3. ICAIE

3. International Conference on Advances and Innovations in Engineering

ISBN: 978-975-394-098-6

SEPTEMBER 25-27, 2025

CONGRESS PROCEEDINGS BOOK

Firat University
Faculty of Engineering / ELAZIĞ

Editor in Chief

Erkut YALÇIN

Editors

Mehmet ARZU
Canan KOÇ
Çağla HOROZ
Keremcan KARAKAYA

Fırat University

International Conference on Advances and Innovations in Engineering

3.ICAIE



Conference Date
25-27 September 2025



15 June 2025
Deadline for abstract submission

30 June 2025
Deadline for full-text submission



www.icaie.org.tr **2025**

Place: Fırat University, Atatürk Culture and congress Center

arala AİL MEGA ALACAKAYA MARBLE

INTRODUCTION

Dear Distinguished Delegate,

The third edition of the International Conference on Advances and Innovations in Engineering (ICAIE) was held between 25-27 September 2025 at Fırat University Faculty of Engineering, Elazığ.

International Conference on Advances and Innovations in Engineering is an international scientific forum of distinguished scholars engaged in scientific, engineering and technological research, dedicated to the furtherance of science, engineering and technology. The academic research conference since its inception is at the cutting edge of international nonprofit scientific, engineering and technological progress to promoting excellence in science.

The conference plays an influential role in science and promotes developments in science, engineering and technology in a wide range of ways. The conference aims to foster research in the area of science and technology and its impact to mainstream human activities. Specifically, it serves as a venue for discussions and exchange of ideas in current issues in science and technology.

All full paper and abstract submissions to the conference are peer reviewed and refereed and evaluated based on originality, research content and correctness, relevance to contributions, and readability. In this content the full paper and abstract submissions are chosen based on technical merit, interest, applicability and how well they fit a coherent and balanced technical program. The accepted papers after rigorous peer reviewing process have been published in the refereed international conference proceedings

Chair of ICAIE Committee
Prof. Dr. Ebru AKPINAR

Honour Committee of the ICAIE

Prof. Dr. Fahrettin GÖKTAŞ, Rector
Prof. Dr. Mehmet YILMAZ, Vice Rector

Committee Members of the ICAIE

Chair of Conference

Prof. Dr. Ebru AKPINAR, Dean

Organizing Committee

Abayhan BURAN, Fırat University (Bioengineering)
Ahmet ÇINAR, Fırat University (Computer Engineering)
Akın ODABAŞI, Fırat University (Metallurgical and Materials Engineering)
Arif GÜLTEN, Fırat University (Electrical and Electronics Engineering)
Arkadiusz DOBRZYCKI, Poznan University
Ayla ÖZER, Mersin University (Chemical Engineering)
Ayşe Nur BAKIRCI
Ayşe Vildan BEŞE, Atatürk University (Chemical Engineering)
A. Garci HERNANDEZ, RWTH Aachen University
Baha Vural KÖK, Fırat University (Civil Engineering)
Buğrahan Fatih FURTANA, Fırat University (Civil Engineering)
Buket ERZEN, Fırat University (Chemical Engineering)
Burçak EBİN, (Chalmers University)
Bünyamin ERGEZİN, Fırat University (Mechatronics Engineering)
Canan KOÇ, Fırat University (Software Engineering)
Celal KISTAK, Fırat University (Mechanical Engineering)
Cenk YANEN, Fırat University (Mechanical Engineering)
Cevdet AKOSMAN, Fırat University (Chemical Engineering)
Çağla HOROZ, Fırat University (Software Engineering)
Çağrı KAYMAK, Fırat University (Mechatronics Engineering)
Deniz ÜNER, Middle East Technical University (Chemical Engineering)
Dursun ÖZER, Fırat University (Chemical Engineering)
Emrah ÇELİK, Fırat University (Metallurgical and Materials Engineering)
Emre GÜÇLÜ, Fırat University (Computer Engineering)
Ercan AYDOĞMUŞ, Fırat University (Chemical Engineering)
Erkin EREN, Fırat University (Civil Engineering)
Erkut YALÇIN, Fırat University (Civil Engineering)
Fatih DEMİR, Fırat University (Software Engineering)
Fatih KAYA, Fırat University (Chemical Engineering)
Fatih ÖZYURT, Fırat University (Software Engineering)
Fethi KAMIŞLI, Fırat University (Chemical Engineering)
Filiz KAR, Fırat University (Chemical Engineering)
Fulvio PARISI, Naples Federico II Via Claudio University
Gamze SAK, Fırat University (Chemical Engineering)
Gamze UYGUR
Gizem Gül KATIRCIOĞLU, Fırat University (Mechanical Engineering)
Gülbeyi DURSUN, Fırat University (Chemical Engineering)
Güngör YILDIRIM, Fırat University (Computer Engineering)
Hamdi Soner ALTUNDOĞAN, Fırat University (Bioengineering)
Hasan YETİŞ, Fırat University (Computer Engineering)
Hossein MOAYEDI, Duy Tan University
Hüseyin KARACA, İnönü University (Chemical Engineering)

İlhan AYDIN, Fırat University (Computer Engineering)
İnanç ÖZGEN, Fırat University (Bioengineering)
İnci Türk TOĞRUL, Batman University (Petroleum and Natural Gas Engineering)
Keremcan KARAKAYA, Fırat University (Artificial Intelligence and Data Engineering)
Kıvanç DOĞAN, Fırat University (Electrical and Electronics Engineering)
Kübra BULUT, Fırat University (Civil Engineering)
Kübra ŞİŞLİOĞLU, Fırat University (Food Engineering)
Loke Kok FOONG, Duy Tan University
Mahmut KAYA, Fırat University (Artificial Intelligence and Data Engineering)
Mansour MOSALLANEZHAD, Shiraz University
Mehmet Ali ERTÜRK, Fırat University (Geological Engineering)
Mehmet ARZU, Fırat University (Artificial Intelligence and Data Engineering)
Mehmet Deniz TURAN, Fırat University (Metallurgical and Materials Engineering)
Mehmet ERDEM, Hacettepe University (Environmental Engineering)
Mehmet KALENDER, Fırat University (Bioengineering)
Mehmet KARAKÖSE, Fırat University (Computer Engineering)
Mehmet KÖKÜM, Fırat University (Geological Engineering)
Mehtap TANYOL, Munzur University (Environmental Engineering)
Melek YILGIN, Fırat University (Chemical Engineering)
Mert İNCİDELEN, Fırat University (Artificial Intelligence and Data Engineering)
Merve YILDIRIM, Fırat University (Electrical and Electronics Engineering)
Mesut GÖR, Fırat University (Civil Engineering)
Mete Onur KAMAN, Fırat University (Mechanical Engineering)
Mohammed Abdullahi Mu'azu, Hafr Al-Batin University
Muhammed KARATON, Fırat University (Civil Engineering)
Muhammad Masood RAFI, NED ve Teknoloji University
Mukaddes KARATAŞ, Fırat University (Chemical Engineering)
Musharraf ZAMAN, Oklahoma University
Mustafa YEGİN, Fırat University (Environmental Engineering)
M. Barbaros DURMUŞ, Firat University (Environmental Engineering)
M. Sefa ÇETİN, Fırat University (Electrical and Electronics Engineering)
Neslihan DURANAY, Fırat University (Chemical Engineering)
Nevin ÇELİK, Fırat University (Mechanical Engineering)
Nicola TARQUE, Pontifical Catholic University
Niyazi Furkan BAR, Fırat University (Computer Engineering)
Nuno MENDES, Minho University
Nurhan ARSLAN, Fırat University (Chemical Engineering)
Özge HANAY, Fırat University (Environmental Engineering)
Pranshoo SOLANKI, Illinois State University
Ragip İNCE, Fırat University (Civil Engineering)
Ramazan ORHAN, Fırat University (Chemical Engineering)
Rouzbeh GHABCHI, South Dakota State University
Seda ARSLAN TUNCER, Fırat University (Software Engineering)
Sibel ASLAN, Fırat University (Environmental Engineering)
Soheil GHAREH, Payame Noor University
Şeyda TAŞAR, Fırat University (Chemical Engineering)
Zeynep KARACA, Fırat University (Computer Engineering)

Scientific Committee

A. Didem KILIÇ, Fırat University
A. Erçin ERSUNDU, Yıldız Technical University
A. Sükran DEMIRKIRAN, Sakarya University
Abdullah Hilmi LAV, İstanbul Technical University
Abdullah SAR, Fırat University
Abdülhakim ZEYBEK, Muş Alparslan University
Abdulkadir SENGUR, Fırat University
Abuzer ÇALIŞKAN, Fırat University
Adem ATMACA, Gaziantep University
Ahmet Arif AYDIN, İnönü University
Ahmet Bedri ÖZER, Fırat University
Ahmet Burak TATAR, Fırat University
Ahmet Can ALTUNIŞIK, Karadeniz Technical University
Ahmet ÇINAR, Fırat University
Ahmet HASÇALIK, Fırat University
Ahmet Koca, Fırat University
Ahmet ORHAN, Fırat University
Ahmet ŞAŞMAZ, Fırat University
Ahmet TÜRK, Manisa Celal Bayar University
Ahmet Top, Fırat University
Ahmet YILDIZ, Fırat University
Akın Odabaşı, Fırat University
Ali ARI, İnönü University
Ali TOPAL, Dokuz Eylül University
Ali Kaya GÜR, Fırat University
Alp Buğra Aydın, Fırat University
Alper K.TANYILDIZI, Fırat University
Arif GÜLTEN, Fırat University
Arzu Fırat ERSOY, Karadeniz Technical University
Arzu Yadigar DURSUN, Fırat University
Aslı GİRAY, Alanya Alaaddin Keykubat University
Atike NAZİK, Çukurova University
Ayberk KAYA, Recep Tayyip Erdoğan University
AYÇA AK, Marmara University
Aydın ÇITLAK, Fırat University
Aydın DİKİCİ, Fırat University
Ayhan AKBAL, Fırat University
Ayhan ORHAN, Fırat University
Aykut ÖZPOLAT, Fırat University
Aynur UÇAR, Fırat University
Ayşe Ruşen DURUCAN, Fırat University
Ayşe BİÇER, Malatya Turgut Özal University
Ayşegül UÇAR, Fırat University
Aytekin ÇELİK, Fırat University
Aziz AKSOY, Malatya Turgut Özal University
Aziz SARAÇOĞLU, Fırat University
Baha Vural KÖK, Fırat University
Bahar DEMİREL, Fırat University
Bahadır YILMAZ, Bursa Technical University
Bahtiyar ÖZTÜRK, Ondokuz Mayıs University
Banu TAŞKAN, Fırat University
Barış KARAKAYA, Fırat University
Betül ÇİÇEK ÖZKAN, Fırat University
Belkiz TORĞUL, Erzurum Technical University

Beyda TAŞAR, Fırat University
Bihter DAŞ, Fırat University
Bilal ALATAŞ, Fırat University
Bircan ÇALIŞIR, Fırat University
Burak KIYAK, Fırat University
Burak ŞENGÖZ, Dokuz Eylül University
Burhan ERGEN, Fırat University
Bünyamin AKGÜL, Fırat University
C. Fahir ARISOY, İstanbul Technical University
Cafer BAL, Fırat University
Calibe KOÇ TAŞGIN, Fırat University
Canan DÜLGER, Izmir University of Economics
Cebeli ÖZEK, Fırat University
Cemal BÖLÜCEK, Balıkesir University
Cengizhan DURUCAN, Fırat University
Cengiz YILDIZ, Fırat University
Cenk YANEN, Fırat University
Cevdet Emin EKİNCİ, Fırat University
Cihangir KALE, Fırat University
Cihan ÖZEL, Fırat University
Çetin GENÇER, Fırat University
Çetin ÖZAY, Fırat University
Davut HANBAİ, İnönü University
Dicle Bal AKKOCA, Fırat University
Duygu KAYA, Fırat University
Ebubekir ERDEM, Fırat University
EMİNE CENGİL, Bitlis Eren University
Emine Işıl TOPAL, Fırat University
Emre TURGUT, Fırat University
Ender SUVACI, Eskisehir Technical University
Engin AVCI, Fırat University
Engin GÜRTEKİN, Fırat University
Engin Ünal, Fırat University
Enis GÜNAY, Erciyes University
Ercan AKSOY, Fırat University
Erdal ÖBEK, Fırat University
Erdal ÖZBAY, Fırat University
Erdinç ARICI, Fırat University
Ergin TAŞKAN, Fırat University
Erhan AKIN, Fırat University
Erkan AKBAL, Fırat University
Erkan DENİZ, Fırat University
Erkan DUMAN, Fırat University
Erkan TANYILDIZI, Fırat University
Erkut SAYIN, Fırat University
Erman ÇELİK, Fırat University
Ertan BÜTÜN, Fırat University
Eser SERT, Malatya Turgut Özal University
Esra Tuğrul TUNÇ, Fırat University
Esra YILDIRIM, Fırat University
Fahri ÖZKAN, Fırat University
Faruk KARACA, Fırat University
Fatih ERTAM, Fırat University
Fatih ÖZKAYNAK, Fırat University
Ferhat UÇAR, Fırat University
Ferkan SİPAHİ, Gümüşhane University

Feyza ALTUNBEY ÖZBAY, Fırat University
Filiz ÖZGEN, Fırat University
Figen BALO, Fırat University
Furkan SARSILMAZ, Fırat University
Galip AYDIN, Fırat University
Gonca ÖZMEN KOCA, Fırat University
Gökay BAYRAK, Bursa Technical University
Gökhan BAŞMAN, ETİ Krom
Gökmen TAYFUR, Izmir Institute of Technology
Gül TOSUN, Fırat University
Güllü BOZTAŞ, Fırat University
Gülşad Uslu ŞENEL, Fırat University
Gülşah ÇAKMAK, Fırat University
Gülşah KARADUMAN, Fırat University
Gültekin GÖLLER, İstanbul Technical University
H. Lütfi YÜCEL, Fırat University
H. Suha AKSOY, Fırat University
H.Soner ALTUNDOĞAN, Fırat University
Hakan ÇELİK, Fırat University
Murat KARACASU, Eskişehir Osmangazi University
Halil İbrahim YAMAÇ, Fırat University
Hanbey HAZAR, Fırat University
Hanifi GÜLDEMİR, Fırat University
Hanzade AÇMA, İstanbul Technical University
Harun TANYILDIZ, Fırat University
Hasan ÇELİK, Fırat University
Hasan GÜLER, Fırat University
Hasan YETİŞ, Fırat University
Haşim PİHTİLİ, Fırat University
Haydar EREN, Fırat University
Hayrettin CAN, Fırat University
Hediye AYDIN, Kütahya Dumlupınar University
Hikmet ESEN, Fırat University
Hülya BİÇER, Kütahya Dumlupınar University
Hümeyra ŞAHİN, Fırat University
Hüseyin ALTUN, Fırat University
Hüseyin ÇİMENOĞLU, İstanbul Technical University
Hüseyin SEVİNÇ, Fırat University
Hüseyin YAĞLI, Gaziantep University
İbrahim BERKAN AYDİLEK, Harran University
İbrahim TÜRKMEN, Balıkesir University
İbrahim TÜRKOĞLU, Fırat University
İhsan DAĞTEKİN, Fırat University
İlhan AYDIN, Fırat University
İlhan Fırat KILINÇER, Fırat University
Kazım TÜRK, İnönü University
Kürşat Esat ALYAMAÇ, Fırat University
Latif ÖZLER, Fırat University
Levent TAŞÇI, Fırat University
Leyla KALENDER, Fırat University
M. Cihat TUNA, Fırat University
Bay Deniz TURAN, Fırat University
M. Emin EMİROĞLU, Fırat University
M. Erbil ÖZCAN, Fırat University
M. Salih KESKİN, Dicle University
Mahmut Temel ÖZDEMİR, Fırat University

Mediha İPEK, Sakarya University
Mehmet ALTUNBEY, Fırat University
Mehmet ARSLAN, Karadeniz Technical University
Mehmet CEBECİ, Fırat University
Mehmet ÇAVAŞ, Fırat University
Mehmet DAŞ, Fırat University
Mehmet ESEN, Fırat University
Mehmet EROĞLU, Fırat University
Mehmet GEDİKPINAR, Fırat University
Mehmet H. KORKUT, Fırat University
Mehmet KAYA, Fırat University
Mehmet POLAT, Fırat University
Mehmet Sait SÖYLEMEZ, Gaziantep University
Mehmet Sinan BİLGİLİ, Yıldız Technical University
Mehtap MURATOĞLU, Fırat University
Melahat BEYARSLAN, Fırat University
Melek URAL, Fırat University
Meral ÖZEL, Fırat University
Mert GÜRTÜRK, Fırat University
Mete KALYONCU, Konya Technical University
Mete Onur KAMAN, Fırat University
Metin GÜRÜ, Gazi University
Miray Çelikbilek ERSÜNDÜ, Yıldız Technical University
Muammer GÖKBULUT, Fırat University
Muammer TÜRKOĞLU, Samsun University
Muhammed ATAR, Fırat University
Muhammet BAYKARA, Fırat University
Muhammed Emre ÇOLAK, Fırat University
Muhammed Fatih TALU, İnönü University
Muhammed Furkan KÜÇÜK, Fırat University
Muhammed TALO, Fırat University
Muhammed ULUCAN, Fırat University
Muhammed YILDIRIM, Malatya Turgut Özal University
Muhammet AYDIN, Fırat University
Muhammet KARATON, Fırat University
Muharrem AKGÜL, Balıkesir University
Muhsin Tunay GENÇOĞLU, Fırat University
Murat AYDOĞAN, Fırat University
Murat BAYDOĞAN, İstanbul Technical University
Murat ELİBOL, Ege University
Murat KARABATAK, Fırat University
Murat KARACASU, Eskisehir Osmangazi University
Murat ŞEN, Fırat University
Murat Yavuz SOLMAZ, Fırat University
MUSA ÇIBUK, Bitlis Eren University
Musa YETKİN, Fırat University
Mustafa AY, Fırat University
Mustafa BOYRAZLI, Fırat University
Mustafa Can BİNGÖL, Burdur Mehmet Akif Ersoy University
Mustafa GÜR, Fırat University
Mustafa İNALLI, Fırat University
Mustafa KANIK, Fırat University
Mustafa KAYA, Fırat University
Mustafa KAPTANOĞLU, Fırat University
Mustafa Selman AYDOĞAN Balıkesir University
Mustafa TUNCER, Kütahya Dumlupınar University

Mustafa TUNÇ, Fırat University
Mustafa TÜRK, Fırat University
Mustafa ULAŞ, Fırat University
Müge Elif FIRAT, Fırat University
Müjdat FIRAT, Fırat University
Müslüm ARKAN, İnönü University
Müslün Sara TUNÇ, Fırat University
Namık AYSAL, İstanbul University
Necip ÜNLÜ, İstanbul Technical University
Nevin ÇELİK, Fırat University
Nida KATI, Fırat University
Nihat KAYA, Fırat University
Nihat TOSUN, Fırat University
Nilüfer KOÇER, Fırat University
Niyazi ÖZDEMİR, Fırat University
Oğuz YAKUT, Fırat University
Oğuzhan HASANÇEBİ, Middle East Technical University
Oğuzhan KELEŞTEMUR, Fırat University
Oktay YİĞİT, Fırat University
Onuralp YÜCEL, İstanbul Technical University
Orhan ATİLA, Fırat University
Orhan ÇAKAR, Fırat University
Orhan YAMAN, Fırat University
Osman YİĞİD, Bitlis Eren University
Ö. Faruk DURSUN, İnönü University
Ömer GÜLER, Munzur University
Ömer KELEŞOĞLU, Fırat University
Ömer Sinan ŞAHİN, Konya Technical University
Ömer ŞAHİN, İstanbul Technical University
Ömer Yavuz BOZKURT, Gaziantep University
Ömür AYDOĞMUŞ, Fırat University
Özal Yıldırım, Fırat University
Özge Erdoğan YAMAÇ, Fırat University
Özge HANAY, Fırat University
Özgür AVŞAR, Eskisehir Technical University
Özgür KARADUMAN, Fırat University
Özgür YILDIZ, Yıldız Technical University
Özkan ÖZDEMİR, Sakarya University of Applied Sciences
Özlem Öztekin OKAN, Fırat University
Özlem TEPE, Fırat University
Pakistanlı TURGUT, İnönü University
Ragıp İNCE, Fırat University
Recep ARTIR, Marmara University
Remzi TÜNTAŞ, Van Yüzüncü Yıl University
Resul ÇÖTELİ, Fırat University
Resul DAŞ, Fırat University
Reşat ÇELİKEL, Fırat University
Sadettin KAPUCU, Gaziantep University
Sencer ÜNAL, Fırat University
Serdar Ethem HAMAMCI, İnönü University
Serkan ERDEM, Fırat University
Servet TUNCER, Fırat University
Servet YILDIZ, Fırat University
Sevcan KÜRÜM, Fırat University
Sibel ASLAN, Fırat University
Sibel KAYGILI, Fırat University

Sinem AKYOL, Fırat University
Soner BUYTOZ, Fırat University
Şaban TANYILDIZI, Fırat University
Şengül DOĞAN, Fırat University
Taha TAŞKIRAN, Ankara Yıldırım Beyazıt University
Tahsin BOYRAZ, Sivas Cumhuriyet University
Taner ALATAŞ, Fırat University
Taner TUNCER, Fırat University
Tuba DEMİR, Fırat University
Turan GÜRGENÇ, Fırat University
Turgay KAYA, Fırat University
Tülay YILDIRIM, Yıldız Technical University
Tülay YILDIZ, Fırat University
Türker TUNCER, Fırat University
Said ERAY, Sivas University of Science and Technology
Seda ARSLAN TUNCER, Fırat University
Sedat SAVAŞ, Fırat University
Selçuk ALEMDAĞ, Gümüşhane University
Selçuk KARATAŞ, Fırat University
U. Teoman AKSOY, Fırat University
Uğur ÇALIGÜLÜ, Fırat University
Utku KÖSE, Süleyman Demirel University
V.Cem BAYDOĞAN, Fırat University
Vedat ÇELİK, Fırat University
Vedat SAVAŞ, Fırat University
Venhar ÇELİK, Fırat University
Veyis SELEN, Fırat University
Yakup DEMİR, Fırat University
Yaman AKBULUT, Fırat University
Yasemin TOPÇUOĞLU, Fırat University
Yasin VAROL, Fırat University
Yaşar NUHOĞLU, Yıldız Technical University
Yavuz EROL, Fırat University
Yılmaz YILDIRIM, Zonguldak Bülent Ecevit University
Yunus Emre BENKLİ, Atatürk University
Yunus ERKUŞ, Fırat University
Yusuf CALAYIR, Yalova University
Yüksel ESEN, Fırat University
Z. Fuat TOPRAK, Dicle University
Zeynep Bala DURANAY, Fırat University
Zülfü GÜROCAK, Fırat University
Zümriye AKSU, Hacettepe University

Session Chairmans

Hamdi Soner ALTUNDOĞAN

Gülbeyi DURSUN

Fatih DEMİR

Filiz KAR

Ragip İNCE

Melek YILGIN

Nevin ÇELİK

Arif GÜLTEN

İnanç ÖZGEN

Muhammet KARATON

İlhan AYDIN

Mehmet Deniz TURAN

Erkut YALÇIN

Mete Onur KAMAN

Özge HANAY

Baha Vural KÖK

Seda ARSLAN TUNCER

Cevdet AKOSMAN

CONTENTS

Prize Collecting Vehicle Routing for Flight Catering Control Optimization 16
A MILP Model for Peer Evaluation Survey Assignment Optimization 23
Airline Food Optimization: Flightmax 27
Innovative Approach in Document Management System Procurement: Artificial Intelligence and Chatbot 34
A Pharmaceutical Stock Forecasting Approach: Comparative Analysis of Statistical and Machine Learning Models 37
Detection of Satellite Network Anomalies With XOR Based Encryption and Transformer Assisted Approach 46
Blockchain Based Method for Secure Data Storage In Digital Identity Systems 52
Sentiment Analysis on Book Reviews Using Transformer-based Approaches 60
Triple Teacher Framework with Exponential Moving Average for Screw Anomaly Detection 66
Quantum Computing in Mobile Networks: A Comparative Analysis 74
A Particle Swarm Optimization Approach to Federated Bearing Fault Diagnosis 82
Sentiment Analysis of Turkish Song Lyrics Using Machine Learning Algorithms 87
An Application of Face Recognition Authenticated Video Recording Process in Embedded System 95
Critical Success Factors of Lean Manufacturing Tools: The DEMATEL Method based pq-Cubic Quasi rung Orthopair Fuzzy Sets 105
Conceptual Integration of Machine Learning in AI-Directed Soil Stabilization for Sustainable Ground Improvement 114
Effect of TiC on Microstructure and Hardness in Laser Cladding with Fe-Cr-C/TiC Powder Mixtures 117
Effect of Surface Roughness on Average Heat Transfer in Jet Impingement Cooling 123
Effect of Chromium and Carbon on Wear Properties of Laser Coated Low Carbon Steels 128
Photochemical Oxidation and Radical-Driven Processes in Hydrometallurgical Leaching: An Overview 134
Effect of Nitriding Duration on Microstructure and Wear Behavior of Two Different Steels 139
A New BERT Based Hate Speech Detection Method Using Contrastive Learning 145
An Autonomous Aerial Surveillance System for Early Detection of Wildfire Anomalies Using Smart Sensor Integration 152
Design and Calculations of Rocket Launching Mechanism Using Centrifugal Technique 161
Effect of Boron on the Microstructure and Wear Behavior in Laser Surface Coating of SAE 4140 Steel 170
Use of *Lavandula angustifolia* M. Plant in Plant Protection 175
Current Development of Soft Grippers Used in Robotic Applications 190
A Hybrid Model Based on Image Processing for Violence Detection 201
A New Record of the Masarine Wasp *Celonites rugiceps* (Hymenoptera: Vespidae, Masarinae) from Bingöl Province, Turkey 211
The Role of Renewable Energy in the Global Energy Transition and Türkiye's Strategic Position 214
From Debris to Production: The Effect of Aggregates Obtained from Earthquake Debris on Concrete Performance 219
A Preliminary Experimental Investigation on the Compressive Strength of Fiber-Reinforced Polylmer (FRP) Reinforced Concrete Specimens 226
Experimental and Analytical Investigation of Flow Coefficient in Triangular Cross-Section Side Weirs 232
Bisphenol A (BPA): Properties, Environmental Impact, Health Risks, and Regulatory Actions 238
Chemical Aspects of the Vanadium (V)-Oxide Extraction Process From Alunite Ore 243

Dye Economy in Endemic and Medicinal Plants: An Analysis Based on *Isatis, Reseda,* and *Rubia* 248
The Use of Algae in Agriculture 254
Development of a Mobile Application Supporting Sustainable Living Habits for Individual Carbon Footprint Tracking 265
Artificial Intelligence-Assisted Image Processing Models for Disease Detection in Wheat Plants 271
An Artificial Intelligence-Based Fraud Detection Approach for Securing Blockchain Networks 277
Autonomous Unmanned Aerial Vehicle Application for Real-Time Wounded Detection Based on YOLOv5 287
Design and Energy Efficiency of Nearly Zero Energy Buildings 295
Detecting Street Signs for Visually Impaired People: A Yolov8 and Federated Learning-Based Approach 302
MediFind: An Interactive Mobile Application for Drug Information Access and Drug Compliance 316
From Tradition to Future: The Economic, Ecological and Cultural Importance of Natural Dye Plants in Elazığ Province 322
Fuzzy Logic Controlled Regenerative Braking for Electric Vehicles: Energy Recovery and Driving Safety 330
A Deep Learning Based Application for Deepfake Video Detection 339
Lignin Recovery and Characterization from the Liquid Obtained by the Organosolv Process of MDF Dust 352
Energy Efficiency İmprovement Based on Smart Grid: A Simulation Assessment 359
Enhanced Semantic Segmentation of Microscopic Rock Images Using DeepLabV3+ with SE Attention and Superpixel-Aided Labeling 364
Production of Functionally Graded Composite Materials by Vacuum Infusion Method and Investigation of Their Impact Behavior 375
Machine Learning-Based Prediction of Key PVT Parameters: A Comparison with Classical Regression 383
The Production of Fibres for Mineral Wool by Utilization of Olivine Extracted as A by-product of Chromite Ore 392
Slope Stability Analysis Using the Limit Equilibrium Method 400
Explainable Diabetes Prediction in a Large-Scale Clinical Dataset Using TabNet and SHAP 409
Modernization of The Refueling Probe With an Electromagnetic Magnet 416
Effects Of Sodium Hydroxide On The Unconfined Compressive Strength Of Low Plasticity Soils 424
Examination Of The Behavior Of Soil Stabilized With Silica Fume 430
Determining the Optimal Stator Slot–Rotor Pole Combination for Brushless DC and Permanent Magnet Synchronous Motor 436
KS-Net: Multi-layer network model for determining the rotor type from motor parameters in interior PMSMs 443
Quantum-Enhanced Algorithmic Frameworks for Autonomous Vehicles: A Systematic Review and Analysis Emerging Approaches 450
Turkish Regulatory QA: A Hybrid BM25 and Pretrained Transformer Retrieval Framework with Label-Guided Question-Answering 460
Mechanical Stirrer-Assisted Production of $FeCl_3$ from by Using HCl and Mill Scale 472
A Hybrid Quantum Federated Learning Approach for Image Classification 476
Electric Field Estimation in High Voltage Ceramic Insulator Using Deep Learning and Random Forest 483
Customer Churn Prediction in Telecom: A Comparative Machine Learning Approach Using IBM Telco Dataset 492
Artificial Intelligence Based Waste Classification and Guidance System 501
Swarm UAV-Assisted System Design for Border Protection and Threat Detection 507
Artificial Intelligence Design Paradigms: Review of AI-Native and Artificial General Intelligence (AGI) Systems 515
Development of a Rheological Model for Drilling Muds and Determining Their Rheological Properties 522

Heat Transfer in A Corrugated Conduit Filled with Porous Material 530
Bellman-Ford and SMOTE Based Optimization Method in Hyperspectral Image Classification 541
Classification of Liver Disease Using Biochemical Data: A Machine Learning Approach 547
Efficiency of Energy Management in Unmanned Aerial Vehicles: An Examination within The Framework of Current Approaches 552
Development of A Docker-Based Demand Forecasting System Integrated with Continuous Integration and Continuous Deployment Automation Using Prometheus and Grafana 560
Pneumonia Detection in Chest X-Ray Images via Blockchain-Enabled Federated Learning 574
A Hybrid Convolutional Neural Networks and Hu Moments-Based Approach for Automated Textile Defect Detection 583
Experimental and Artificial Intelligence Methods for Investigating the Effects of Cell Temperature on Panel Performance in Photovoltaic Solar Panels 594
Investigation of Fracture Parameters of Lime-Pumice Mortars 606
An Alternative SCB Test for Fracture Analysis of Asphalt Concretes 613
Evaluation of the Rheological Behavior of Pure and SBS-Modified 70/100 Batman Bitumen under Dynamic Shear Conditions 620
Effect of Reactive Polymer on High Temperature Performance of Crumb Rubber Modified Bitumen 627
Effect of Isocyanate and Crumb Rubber on Stability Properties of Bituminous Mixtures 636
Effect of Ethylene-Propylene-Diene-Monomer Addition on Rheological Properties of Bitumen 645
Effect of Damping Ratio on Seismic Response of a Reinforced Concrete Building Using Nonlinear Time History Analysis 653
Evaluation of Ground Motion Attenuation Relationships: Case Study of the 2023 Kahramanmaraş Earthquakes in Elazığ, Türkiye 659
Road Surface Crack Detection Using Deep Learning in Smart Cities 668
Barrette Pile Foundations and Application Principles: A Case Study of the Çeşmeli–Kızkalesi Highway 678
Rheological Properties of Permeation Grout Using Sodium Silicate–Boric Acid Mixtures 687
High-Performance Spyware Classification with Wavelet Transform Features 698
MRI-Based Brain Tumor Classification Using Deep Learning Models 706
Disease Prediction Based on Medical Symptoms: A Comparative Study of Machine Learning Models 715
Fuel Optimization in Flight Routes Using Machine Learning and Graph-Based Methods 725

Prize Collecting Vehicle Routing for Flight Catering Control Optimization

Ahmet Cihat BAKTIR[1], Imran TOPRAK [2]

[1,2]Operations Research Directorate, Turkish Technology, Istanbul, Turkey.
[1]abaktir@thy.com, [2]itoprak@thy.com

Abstract
In-flight catering is critical for airlines to enhance the customer experience and satisfaction during travel. To ensure a certain level of quality of catering products, quality control on aircraft at parking positions plays a key role throughout the overall journey. With many flights arriving and departing daily at an airport, the catering products on aircraft require control within equally divided shifts. However, due to limited resources of control teams and vehicles, prioritization among flights is necessary. In this direction, this study addresses the vehicle routing problem (VRP) for quality control of catering products. We formulate this problem as a prize-collecting VRP with time windows (PCVRPTW). To solve this model, we develop both mixed-integer linear programming (MILP) and matheuristic approaches. The main goal is to maximize the number of visited aircrafts prioritizing those with higher priority, subject to operational and business-related constraints. Based on a comprehensive set of experiments, our results show a 20% increase in the number of aircraft visited compared to the current heuristic approaches used by control teams, while ensuring visits to highly prioritized flights. This study provides an efficient solution for optimizing vehicle routes in seconds, enhancing the quality control process for catering products on aircraft.

Keywords: optimization, operations research, airline industry, vehicle routing problem, MILP, catering quality control

1. Introduction

The airline industry transforms its business to provide value-added services for passengers towards enhanced experience, satisfaction, and loyalty. Along with the digital evolution for operational efficiency and increased effort towards efficient revenue management, recent trends in the industry empower airlines to deliver customer-oriented solutions. Therefore, airlines can achieve higher degrees of customer satisfaction, which ultimately leads to higher profits in the long run.

In-flight catering plays a key role in achieving a competitive advantage in the market with regards to customer satisfaction throughout the flight services [1]. Providing high-quality meals, beverages and other products during the flight not only enhances the experience during the journey but also strengthens customer loyalty and enhances the brand reputation [2].

Even though high-quality catering products are ensured by catering companies, management of the products prior to flights is critical, especially for airlines with many flights during the day. In addition to the loading/unloading process of the catering products [3], as the demand for in-flight catering increases with the competition among companies, the catering products and supplies should be controlled by a team before departure or after arrival. Effective control of the products and logistic coordination of the control teams is critical for maintaining a certain level of quality and ensuring the products are installed or uninstalled without any deficiency. Furthermore, these controls are necessary to indicate any risks so that immediate actions can be taken accordingly. Ultimately, actions to mitigate such risks can also avoid financial losses or degradation in customer experience.

However, controlling catering for all flights is not achievable due to restricted resources of control teams and time constraints. Therefore, there should be a prioritization among the flight based on various criteria. On the other hand, even though the prioritization is achieved, the limited time should be utilized efficiently to maximize the number of flights that are controlled by taking the business rules into consideration as well

In this direction, this paper proposes an optimization model to produce optimal routes for control teams at airport apron to achieve maximum benefit in terms of flight prioritization and number of controlled flights. A mixed-integer linear programming (MILP) model is implemented towards this objective. The problem is modelled as a prize-collecting vehicle routing problem with time windows (PCVRPTW) as the priority of a flight is reflected as an optional prize to be collected after ontrol is accomplished. However, due to the complex nature of routing problems with many nodes

[1]Corresponding author

and time windows, achieving optimal solution in a very short time is challenging. Therefore, we also proposed a matheuristic approach, which decreases the solution space for the original MILP model.

The proposed methodology is implemented and evaluated with an extensive set of experiments which are designed according to the flight data of Turkish Airlines. The results depict that even though MILP model is capable of providing high-quality solutions in a short time, matheuristic approach enhances efficiency by achieving better results even in a shorter time.

The rest of the paper is organized as follows. Section 2 introduces the related literature and highlights the contribution of this paper. Section 3 presents the problem definition and then the proposed MILP and matheuristic approaches. Section 4 provides an overview of the experiment design and presents the results. Section 5 concluded the paper along with the future directions in this domain.

2. Literature Review

In the literature, the applications of vehicle routing problem (VRP) vary in terms of problem characteristics and properties. Practically, VRP consists of determining the optimal routes for a fleet of vehicles to visit a set of nodes (i.e., locations) and then return to the base location. Variations of objectives can be defined such as minimizing the cost, maximizing the visited locations and maximizing the total benefit. In this direction, many verticals customize the typical VRP models by integrating the concepts of time windows, capacitated vehicles, prize collection and so on.

The airline industry, as a result of digital transformation in the domain, also integrates VRP solutions to improve the overall efficiency of the operations. In the study made by Li et al. [4], crew pairing problem is modelled as VRP. In this scope, a heuristic algorithm is developed which is then evaluated with data provided by an airline company.

An extension of VRP is applied to aircraft routing problem, which is modeled as VRP with time windows (VRPTW). In the study made by Evler et al. [5], aircraft are considered as vehicles, and each aircraft is enforced to visit each node within a certain time window. As the nodes are specified as flights including the depot (i.e., base), the VRPTW is modeled to assign aircraft to flight routes.

The idea of prize collection can also be integrated into planning operations. In cases where visiting all locations (e.g., customers) is not feasible, the main objective could be to determine a subset of nodes to be visited in order to maximize overall prize. Even though the definition of prize might differ for each problem, the main concept is to maximize the locations which provide higher benefit. There are various studies in literature that accommodate the prize collection VRP (PCVRP) for different applications [6, 7].

Based on an extensive review of the literature, it is observed that prize collecting VRPTW (PCVRPTW) is also modeled to define time windows for each potential customer to be visited. In the study made by Dogan and Alkaya [8], a novel method is provided to solve prize collecting traveling salesman problem with time windows (PCTSPTW). This approach incorporates a two-staged solution, which is composed of a heuristic approach and genetic algorithm. In another study, Gama et al. investigate PCVRTW for multi-agent environments for VRP [9].

To the best of our knowledge, the PCVRPTW model introduced by our paper is a novel approach in literature. Even though there is extensive research on VRPs and their applications in different industries, the specific problem of PCVRPTW in airline industry has received relatively little attention in the literature. A review of existing studies reveals that most research has focused on traditional VRP variants, with limited exploration of the prize-collecting aspect in the context of airline operations. Therefore, the original contribution of this study is to provide an optimization model that solves this problem type in less than 10 seconds for an airline industry case, based on real data.

3. Prize-collecting VRP with Time Windows

In this section, first the business problem is defined explaining why it is formulated as PCVRPTW and why heuristics are used to enhance its performance. Then, full Mixed-Integer Linear Programming (MILP) formulation for PCVRPTW is introduced together with the explanations of each constraint. At the end, matheuristic approach is elaborated which accommodates a heuristic phase before MILP model to decrease the problem size.

3.1. Problem Definition

The proposed model for this business problem generally determines the visit sequence of the control teams with vehicles among hundreds of aircraft at parking positions. In other words, the problem this paper aims to tackle can be classified as a VRP. Apart from that, each aircraft can only be visited within a pre-determined time window depending on their estimated departure or arrival times, which makes this problem a VRP with time windows (VRPTW). As control teams have limited resources, namely the number of available vehicles, visiting all aircraft is impossible. Therefore, prioritization among flights is necessary while deciding on the visiting sequences. To this end, a priority point parameter is defined for each flight leg, which is used for prioritizing a flight over the rest. The vehicle needs to maximize the total priority points of the flights which are visited within its route. The priority points can be considered as prizes, and the problem can be classified as PCVRPTW.

Class of VRPs are categorized as NP-Hard [10], and grows exponentially with increasing number of nodes, in this case flights. Finding optimal or near-optimal solutions in reasonable time is challenging, and so various exact or heuristics methods are developed and applied to handle this highly computationally complex problem [11]. Particularly in this business problem, reaching a satisfying solution in reasonable time is strongly desired, as all operations in apron setting are crucial and dynamic. For this reason, we developed a heuristic approach as a pre-process phase of the proposed MILP model satisfying the business unit needs, which reduces the solution space by a specific ratio depending on flight count. After eliminating a fraction of flights which are less likely to visit in optimal solution, PCVRPTW model can be run to receive a sufficiently good solution within a specific time limit. As the previously introduced approach combines exact optimization method with heuristics, it can be labelled as a matheuristic approach.

3.2. Proposed MILP Optimization Model

As aforementioned, the main objective of this paper is to propose an optimization model that maximizes the total priority point of the flights. In this regard, MILP formulation of the proposed model is presented in this subsection, which covers the sets, parameters, decision variables, objective function and constraints.

In particular, MILP optimization model incorporates four different sets. The set of nodes, which is denoted as N, represents the flights within a specific time period where control teams can potentially visit. Each flight can be categorized as either departure or arrival, where set of departure flights D and arrival flights A are subset of N. As control criteria, time windows and control time might differ among these categories, we define these subsets for more accurate modeling.

Sets and indices:
$\boldsymbol{N} = \{\boldsymbol{0}, \boldsymbol{1}, \dots, \boldsymbol{n}\}$: Set of nodes (flights), where 0 is the base location (depot)
$\boldsymbol{D}$: Set of departure nodes ($\boldsymbol{D} \subset \boldsymbol{N}$).
$\boldsymbol{A}$: Set of arrival nodes ($\boldsymbol{A} \subset \boldsymbol{N}$).
$\boldsymbol{K}$: Maximum number of vehicles.

Correspondingly, we define travel time between two flights (i, j), which is represented as Tij. The concept of time windows is defined through parameter $[Ei, Li]$, where Ei defines the earliest allowable service start time and Li defines the latest allowable service start time. As stated, each flight, depending on the fleet type and category (i.e., arrival or

departure), requires a certain time for controlling service, which is defined by Si. Lastly, according to business rules, there could be a certain lower or upper bound for the proportion of arrival flights among all visited flights. For example, it could be demanded that at least 50% of the visited flights should belong to departure class. Such lower and upper bound parameters are specified as $[LB, UB]$, respectively. Parameter M is used as a large constant for linearization purposes.

Parameters:
$\boldsymbol{Tij}$: Travel time between nodes $\boldsymbol{i}$ and $\boldsymbol{j}$.
$[\boldsymbol{Ei}, \boldsymbol{Li}]$: Time window for node $\boldsymbol{i}$ (earliest and latest allowable service start time).
$\boldsymbol{Si}$: Service time required at node $\boldsymbol{i}$.
$\boldsymbol{Wi}$: Importance weight (prize) for node $\boldsymbol{i}$.
$[\boldsymbol{LB}, \boldsymbol{UB}]$: Lower bound and upper bound for the proportion of arrival flights among all visited flights.
$\boldsymbol{M}$: Large constant for linearization (big-M).

Based on these properties, we define three different types of decision variables. In this direction, xij, which is a binary variable, becomes 1 if any vehicle travels from node i to node j, 0 otherwise. In order to keep track of the visit times and define related constraints, we define another decision variable, $\boldsymbol{ti}$, to determine the time at which service starts at flight $\boldsymbol{i}$. By using this set of information, the last decision variable we use in this model is $\boldsymbol{zi}$, which is a binary variable representing whether node $\boldsymbol{i}$ is visited or not.

Decision variables:
$\boldsymbol{xij} \in \{\boldsymbol{0}, \boldsymbol{1}\}$ **:** 1 if any vehicle travels from node i to node j, 0 otherwise.
$\boldsymbol{t_i}$ **:** Time at which service starts at node i.
$\boldsymbol{zi} \in \{\boldsymbol{0}, \boldsymbol{1}\}$ **:** 1 if node i is visited, 0 otherwise.

Objective Function:

$$\max \sum_{i \in D \cup A} Wizi \quad (1)$$

Constraints:

$$\sum_{j \in N, j \neq i} xij - \sum_{j \in N, j \neq i} xji = 0, \quad \forall i \in D \cup A \quad (2)$$

$$\sum_{j \in D \cup A, j \neq i} xij \leq 1, \quad \forall i \in D \cup A \quad (3)$$

$$\sum_{i \in D \cup A, j \neq i} xij \leq 1, \quad \forall j \in D \cup A \quad (4)$$

$$ti + Si + Tij \leq tj + M(1 - xij), \ \forall i \in N, \forall j \in D \cup A \quad (5)$$

$$Ei \leq ti \leq Li, \quad \forall i \in N \quad (6)$$

$$zi \leq \sum_{j \in N, j \neq i} xji, \quad \forall i \in N \quad (7)$$

$$\sum_{i \in A} zi \leq \left(\sum_{j \in D \cup A} zj\right) UB \quad (8)$$

$$\sum_{i \in A} zi \geq \left(\sum_{j \in D \cup A} zj\right) LB \quad (9)$$

$$xij \in \{0,1\}, \quad zi \in \{0,1\}, \quad \forall i \in N, \forall j \in N \tag{10}$$

Equation 1 represents the objective of the optimization model. For each flight visited by a vehicle, a prize is collected, and total prizes are maximized. Equation 2 ensures the flow conservation, where each visited aircraft (i.e., flight) must be left. If not visited, then that aircraft cannot be left. Equations 3 and 4 limits the number of visits. Each flight can be visited and left at most once. In other words, if multiple vehicles exist, each flight can be visited by at most 1 vehicle. Equations 5 and 6 secures the time window feasibility. Visit times, service times and time windows have to be coherent and subtours must be eliminated. Parameter M should have the lowest possible value to avoid computational complexity. Equation 7 activates the decision variable z. In other words, a node can only be visited if a vehicle enters it. Equations 8 and 9 prioritizes specific flight types. Proportion of arrival controls cannot be less than a pre-specified lower bound and exceed the upper bound. Lastly, Equation 10 includes the integrality constraints.

3.2. Matheuristic

The optimization model proposed in the previous sub-section is capable of providing a high-quality solution for many cases. However, it may suffer from scalability problem as the complexity increases exponentially with the increasing number of flights to be assessed by the model.

To address this issue, a promising approach is to decrease the solution space. In this direction, we propose to develop a heuristic-based pre-processing phase which removes a subset of flights that are relatively assigned with lower priority values. Even though this approach might lead to sub-optimal solution for some cases, it is beneficial for implementing a scalable approach that produces high-quality solution within a shorter time. For each category of flights, we analyze the input set of nodes and then omit a subset of flights based on the business rules and prioritization criteria. Then, the proposed MILP model operates on the $Nh \subset N$, where Nh represents the filtered set of flights.

4. Performance Evaluation and Results

The proposed solution methodologies in this paper are evaluated based on an extensive set of experiments. In order to assess the performance of both MILP model and matheuristic approaches, we parameterized different settings and comparatively analyzed the quality of the solution as well as the solution time.

Both MILP model and matheuristic solution are solved with Gurobi 12.0.1 running on an 8-core CPU with 64 GB main memory. The experiments utilize real flight data provided by Turkish Airlines, collected on different time periods. In order to evaluate the generalizability of the proposed solutions, different set of flights are given as input for each run and the results obtained through experiments are compared in this section. The overview of the parameters utilized throughout the experiments is depicted in Table 1.

Table 1. Experiment settings

Parameter	Values
Run time	10 seconds
Vehicles	[1, 2, 3] vehicles
Number of flight cases	10

30 different experiments are conducted for both the original MILP model and matheuristic approach and their solutions are reported. Comparison tables for each setup (1, 2, and 3 vehicles) can be found in Table 2, Table 3, and Table 4.

Table 2. Experiment results for one vehicle

Scenario Name	Flight Count	MILP Objective Value	Matheuristic Objective Value	Increase in Obj
F1-V1	340	133.05	151.58	13.9%
F2-V1	352	122.01	144.83	18.7%

Table 2. (Continue) Experiment results for one vehicle

F3-V1	317	133.57	131.49	-1.6%
F4-V1	345	145.59	156.78	7.7%
F5-V1	313	131.05	140.13	6.9%
F6-V1	357	124.54	136.28	9.4%
F7-V1	349	125.31	145.04	15.7%
F8-V1	339	121.24	144.76	19.4%
F9-V1	334	150.37	160.07	6.5%
F10-V1	333	132.13	139.21	5.4%

Table 3. Experiment results for two vehicles

Scenario Name	Flight Count	MILP Objective Value	Matheuristic Objective Value	Increase in Obj
F1-V2	340	269.02	282.20	4.9%
F2-V2	352	235.55	269.96	14.6%
F3-V2	317	244.13	246.93	1.1%
F4-V2	345	271.65	291.35	7.3%
F5-V2	313	251.42	260.13	3.5%
F6-V2	357	231.32	250.38	8.2%
F7-V2	349	240.70	249.67	3.7%
F8-V2	339	223.04	270.63	21.3%
F9-V2	334	262.21	276.21	5.3%
F10-V2	333	231.69	240.94	4.0%

Table 4. Experiment results for three vehicles

Scenario Name	Flight Count	MILP Objective Value	Matheuristic Objective Value	Increase in Obj
F1-V3	340	377.89	386.20	2.2%
F2-V3	352	333.24	383.01	14.9%
F3-V3	317	345.73	361.44	4.5%
F4-V3	345	403.39	409.64	1.6%
F5-V3	313	349.82	367.00	4.9%
F6-V3	357	324.90	342.43	5.4%
F7-V3	349	329.90	354.43	7.5%
F8-V3	339	327.10	380.30	16.3%
F9-V3	334	370.39	390.91	5.5%
F10-V3	333	312.73	353.36	13.0%

Based on the comparative analysis, it can be assessed that the proposed matheuristic approach provides better solution than the MILP model. Except for one scenario (F3-V1), matheuristic approach surpasses the MILP approach under a time limit of 10 seconds. On average, the reported objective value is 8.4% is higher in matheuristic method. While the improvement is higher for some cases, such as 21.3% for F8-V2, the improvement can be negligible for another set of cases such as F3- V2. If sums of all 30 objective values are compared, matheuristic method still increases the objective value by 7.8%.

On the other hand, we evaluated the potential impact of the number of flights on the improvement provided by the matheuristic approach in terms of objective value. However, there is not significant evidence that flight counts and

increase in the objective values are correlated. Since the flight data for experiment cases are collected independently, the capability of matheuristic approach is validated as different set of flights with various properties are evaluated systematically.

5. Conclusion

In-flight catering is a critical service provided by airlines that incorporates operations to enhance customer experience as well as brand loyalty. Even though the catering companies execute their processes in this direction, it is an inherent step of the overall process to ensure that the same quality of catering products is maintained end-to-end. Therefore, the catering control teams need to prioritize some flights over others as controlling all of them is not feasible in practice. In this scope, this paper presents an optimization approach to determine the set of flights to be controlled in a sequential manner. For this purpose, we designed this operation as a PCVRPTW and introduced MILP model as the solution approach. In order to assess the scalability of the proposed optimization model, we additionally present a matheuristic approach the integrates a heuristic-based pre-processing approach that reduces the solution space. Based on real data provided by Turkish Airlines, we executed a set of comprehensive experiments and provided the performance evaluation of proposed approaches. It is observed that matheuristic approach provides a higher quality solution within the given run time limit. As a future direction in this domain, application of machine learning (ML) approaches could enhance the intelligence of pre-processing step and implement a better solution for solution space reduction. Apart from that, various metaheuristic methods can be utilized to create and initial solution as warm start for the commercial solver used, rather than an elimination-based approach.

6. References

[1] Messner, W. (2016). The impact of an aircraft's service environment on perceptions of in-flight food quality. Journal of Air Transport Management, 53, 123-130.

[2] An, M., & Noh, Y. (2009). Airline customer satisfaction and loyalty: impact of in-flight service quality. Service Business, 3(3), 293-307.

[3] Sze, S. N., Suk-Fong, A. N., & Chiew, K. L. (2012, September). An insertion heuristic manpower scheduling for in-flight catering service application. In International Conference on Computational Logistics (pp. 206-216). Berlin, Heidelberg: Springer Berlin Heidelberg.

[4] Li, Z., Wang, S., Li, W., & Jiang, H. (2022, December). A VRP-based Approach for the Airline Crew Pairing Problem. In 2022 10th International Conference on Information Systems and Computing Technology (ISCTech) (pp. 309-315). IEEE.

[5] Evler, J., Lindner, M., Fricke, H., & Schultz, M. (2022). Integration of turnaround and aircraft recovery to mitigate delay propagation in airline networks. Computers & Operations Research, 138, 105602.

[6] Long, J., Sun, Z., Pardalos, P. M., Hong, Y., Zhang, S., & Li, C. (2019). A hybrid multi-objective genetic local search algorithm for the prize-collecting vehicle routing problem. Information Sciences, 478, 40-61.

[7] Li, K., & Tian, H. (2016). A two-level self-adaptive variable neighborhood search algorithm for the prize-collecting vehicle routing problem. Applied Soft Computing, 43, 469-479.

[8] Dogan, O., & Alkaya, A. F. (2021, August). A novel method for prize collecting traveling salesman problem with time windows. In International Conference on Intelligent and Fuzzy Systems (pp. 469-476). Cham: Springer International Publishing.

[9] Gama, R., Fuertes, D., del-Blanco, C. R., & Fernandes, H. L. (2024). Multi-Agent Environments for Vehicle Routing Problems. arXiv preprint arXiv:2411.14411.

[10] Lenstra, J., Rinnooy, K., (1981). Complexity of vehicle routing and scheduling problems. Networks 11, 221–227.

[11] El-Sherbeny, N.A., (2010), Vehicle routing with time windows: An overview of exact, heuristic and metaheuristic methods. Journal of King Saud University - Science, 22, 3, 123-131.

A MILP Model for Peer Evaluation Survey Assignment Optimization

Ahmet Zahit ASLAN[1], Mehmet ERTEM[2]

[1,2]Operations Research Directorate, Turkish Technology, İstanbul, Türkiye.
[1]ahmet.aslan@thy.com, [2]mertem@thy.com
[1](0009-0004-9947-7788), [2](ORCID: 0000-0001-5363-3619)

Abstract
Modern airline organizations employ thousands of staff across intricate hierarchical structures. Within this scope, making an effective peer evaluation system is essential for overall performance management and operational efficiency. This study introduces a mixed-integer programming (MILP) model designed to optimize peer evaluation survey assignments within a large-scale airline company. The proposed model solves the complex task of assigning over 10,000 employees to evaluate one another by using both email interaction data and organizational hierarchy information. The model assigns surveys while considering the business constraints such as fairness, number of evaluations, and hierarchical boundaries, enabling more accurate performance assessments. The proposed optimization framework enables airlines to automate and enhance their peer review processes, reducing administrative effort while improving the quality and relevance of evaluation outcomes. The results highlight the potential of data-driven optimization to support fair and effective peer review systems in dynamic organizational settings.

Keywords: optimization, operations research, peer evaluation, assignment problem, MILP, performance review

1. Introduction

Even though the 360-degree performance evaluation process is valuable for organizational development, it causes significant logistical and fairness challenges when managed manually [1]. The optimization model presented in this study addresses these challenges by automating the assignment of evaluators, balancing workloads, respecting organizational hierarchies, and accommodating manual adjustments. By leveraging mathematical programming, this approach not only reduces the administrative burden on HR personnel but also enhances the fairness and effectiveness of the evaluation process. In addition, strict business rules are enforced to maintain assignment consistency. For example, a top manager is not permitted to evaluate an associate who works under another top manager. These rules ensure that evaluations respect organizational hierarchies and prevent inappropriate evaluator-assessed relationships. Moreover, a workload balance and a minimum number of evaluations per employee should also be considered to ensure that everyone receives sufficient feedback for a fair and accurate assessment. Our proposed model makes all evaluation assignments in accordance with business rules and enhances overall survey quality by assigning evaluations to individuals who are familiar with each other.

2. Literature Review

The performance evaluation survey assignment problem is closely related to the Reviewer Assignment Problem (RAP), which is widely studied for assigning reviewers to relevant papers. The core idea of our problem closely parallels the reviewer assignment problem, particularly with respect to workload balancing and relevance constraints. Taylor (2008) addresses the critical challenge of optimally assigning a large volume of conference papers to reviewers, a common bottleneck in the peer-review process for major academic conferences. This model was successfully deployed at CVPR 2006, demonstrating its practicality by assigning over 1100 papers to 500 reviewers in seconds [2]. Stelmakh et al. (2019) proposed the PeerReview4All algorithm, which balances fairness and accuracy in reviewer assignments. Their max-flow–based approach demonstrated a near-optimal solution and outperformed prior methods in both synthetic and real-world peer review experiments [3]. Similarly, Ramotsisi et al. (2022) examined the student-to-project supervisor assignment problem, which also mirrors reviewer and evaluator allocation. They proposed an integer linear programming model that integrates both student and supervisor preferences while balancing workloads. Tested with real data, their approach improved preference satisfaction and reduced coordinator effort by offering a standardized, transparent framework [4].

To sum up, these studies show how optimization-based frameworks can successfully handle complex allocation problems. These studies illustrate that similar approaches can be effectively adopted for performance evaluation survey assignments.

[1]Corresponding author

3. Modeling Framework

A mixed-integer linear programming (MILP) model is used to formulate the proposed problem. To assign employee evaluations in a fair and balanced manner, the model considers organizational hierarchy, workload, and email traffic among employees.

The primary binary decision variable $x_{i,j} \in \{0,1\}$ indicates whether employee i evaluates employee j. To ensure feasibility in edge cases—such as when email communication is limited, or the organization is relatively small— slack variables are added for both the minimum–maximum number of received assessments and the total workload. These slack variables are penalized in the objective function to discourage unnecessary deviations from the specified rules.

Sets:

E: $Set\ of\ employees$

$P \subseteq E \times E$: $Set\ of\ Pairs$

E: $Set\ of\ employees$

$h1(e)$: $First\ level\ hierarchy\ for\ employee\ e \in E$ $h2(e)$: $Second\ level\ hierarchy\ for\ employee\ e \in E$ $h_3(e)$: $Third\ level\ hierarchy\ for\ employee\ e \in E$ $h_4(e)$: $Fourth\ level\ hierarchy\ for\ employee\ e \in E$

Parameters:

$m_{i,j}$: $Mail\ interaction\ between\ employees(i, j) \in P$
α: $Minimum\ percentage\ of\ evaluations\ from\ same\ group$

L: $Lower\ bound\ for\ number\ of\ evaluations$

U: $Upper\ bound\ for\ number\ of\ evaluations$

W: $Maximum\ survey$ workload

Decision Variables:
$x_{i,j} \in \{0,1\}$: $if employee\ i\ evaluates\ employee\ j\ , (i, j) \in P$

$s^e \geq 0$: $slack\ variable\ for\ minimum\ received\ evaluations\ for\ e,\ e \in E$

$s^e \geq 0$: $slack\ variable\ for\ maximum\ evaluations\ for\ e,\ e \in E$

$s^e \geq 0$, $slack\ variable\ for\ workload, e \in E$

$s^e \geq 0$, $slack\ variable\ for\ same\ group\ assignment, e \in E$

Objective Function:

$$\max \sum_{i \in E} \sum_{j \in E:(i,j) \in P} w_h \cdot \frac{m_{i,j} - m_{min}}{m_{max} - m_{min}} \cdot x_{i,j} - \sum_{i \in E} \left(p_{min} \cdot s^i_{min} + p_{max} \cdot s^i_{max} + p_{work} \cdot s^i_{work} \right) \quad (1)$$

Constraints:

Maximum Survey Assignment for Each Employee

$$\sum_{j \in E:(i,j) \in P} x_{i,j} - s^i_{work} \leq W \quad \forall i \in E \quad (2)$$

Minimum Same Hierarchy Assignment Constraint

$$\sum_{i \in h_1(j)} x_{i,j} + \sum_{i \in h_2(j)} x_{i,j} + \sum_{i \in h_3(j)} x_{i,j} + \sum_{i \in h_4(j)} x_{i,j} + s^j_{group} \geq \alpha \sum_{i \in E:(i,j) \in P} x_{i,j} \quad \forall j \in E \quad (4)$$

Lower Bound on Evaluations

$$\sum_{i \in E:(i,j) \in P} x_{i,j} + s^j_{min} \geq L \quad \forall j \in E \quad (5)$$

Upper Bound on Evaluations

$$\sum_{i \in E:(i,j) \in P} x_{i,j} + s^j_{max} \leq \mathrm{U} \quad \forall j \in E \quad (6)$$

Here, the objective function (1) aims to make the evaluation process as effective as possible. It makes the assignments by rewarding assignments that reflect stronger communication ties between employees, with higher weights given to evaluations from closer hierarchy levels. At the same time, it penalizes cases where workload, evaluation limits, or group requirements are not fully met, using slack variables to keep the model flexible. In this way, the objective balances meaningful organizational relationships with fairness and feasibility in the overall assignment. In the objective function, different metrics are combined, including email traffic rewards and penalties. To ensure that these components are comparable and balanced, both the emailing intensity and the penalty values are scaled between 0 and 1. This normalization prevents any single metric from dominating the optimization.

Constraint (2) limits the number of surveys each employee can be assigned. Specifically, it requires that the number of evaluations an employee conducts does not exceed the maximum workload. To ensure model feasibility, the slack variable is added to this constraint to allow limited violations when capacity becomes too restrictive. Constraint (3) enforces a minimum percentage of evaluations to come from individuals within the same hierarchical structure. By summing evaluations across the first, second, third, and fourth hierarchy levels of an employee, the constraint ensures that at α percent of the evaluations come from individuals from the same hierarchy. In the objective function, different hierarchy levels are given different rewards. So that the model is encouraged to assign evaluation pairs starting with colleagues from the closest (first level) hierarchy. Constraint (4) sets a lower bound on the number of evaluations each employee must receive. This ensures that every individual is evaluated at least L times. Similarly, Constraint (5) sets an upper bound on the number of evaluations each employee can receive. This ensures that no individual is evaluated more than U times.

4. Results

The model was implemented using the CBC solver on a system with a 4-core CPU and 16 GB of RAM. Real-world data from Turkish Airlines was employed, consisting of 10,277 employees. Despite the scale and complexity of the problem, the solver successfully found an *optimal solution within 10 minutes*. One particularly interesting finding was that 89% of employees were evaluated by coworkers they had previously interacted with. In practice, achieving such a high rate is typically very challenging because of considering the balance between workload constraints, hierarchical preferences, and fairness in such a big organization. However, the model demonstrated its usefulness by reaching this level while complying with all constraints.

5. Conclusion

This study presented an optimization-based framework for assigning employee evaluations in a large organizational setting. By considering emailing intensity, hierarchical preferences, and workload restrictions, the model balances practical feasibility with organizational goals. Applied to real-world data from Turkish Airlines, the approach demonstrated strong performance. Notably, 89% of employees were matched with colleagues they already knew, a level of alignment that would be extremely difficult to achieve through manual planning. These results highlight the model's potential as a tool for managing large-scale survey assignments, ensuring both efficiency and meaningful evaluation outcomes.

6. References

[1] J. W. Smither, M. London, and R. R. Reilly, "Does performance improve following multisource feedback? a theoretical model, meta-analysis, and review of empirical findings," Personnel Psychology, vol. 58, no. 1, pp. 33–66, Mar. 2005, doi: 10.1111/J.1744-6570.2005.514_1.X.

[2] Taylor, Camillo. (2008). On the Optimal Assignment of Conference Papers to Reviewers.

[3] I. Stelmakh, N. B. Shah, and A. Singh, "PeerReview4All: Fair and Accurate Reviewer Assignment in Peer Review," arXiv preprint arXiv:1806.06237, 2019. [Online]. Available: https://arxiv.org/abs/1806.06237

[4] J. Ramotsisi, M. Kgomotso, and L. Seboni, "An Optimization Model for the Student-to-Project Supervisor Assignment Problem—The Case of an Engineering Department," Journal of Optimization, vol. 2022, Article ID 9415210, pp. 1–11, Oct. 2022, doi: 10.1155/2022/9415210

International Conference on Advances and Innovations in Engineering

Airline Food Optimization: Flightmax

Adem Erdem KURT[1], Sena YILMAZ[2], Emre Atlıer OLCA[3]

[1] Flight Support Solutions, Turkish Airlines Technology, Istanbul, Türkiye.
[2,3]Software Engineering, Engineering and Natural Sciences Faculty, Maltepe University, Istanbul, Türkiye.
[1] erdemkurt@thy.com, [2] ysena8846@gmail.com, [3] emreolca@maltepe.edu.tr
[1](ORCID: 0009-0008-3703-880X), [2](ORCID: 0009-0008-3703-880X), [3](ORCID: 0000-0001-6812-5166)

Abstract
The FlightMax project is a machine learning (ML)-driven system designed to optimize in-flight food loading processes. Airline catering operations often face challenges related to cost efficiency and food waste, and this project leverages artificial intelligence and data analytics to enhance operational performance.The developed machine learning model analyzes historical flight data, passenger numbers, flight duration, departure and arrival locations, seasonal variations, and other operational factors to predict the optimal amount of food for each flight. This approach not only increases operational efficiency but also minimizes food waste.A simple web-based interface has been developed where airline staff can manually enter flight-related information such as passenger count, flight route, and estimated flight duration. Based on the entered data, the system provides a direct prediction of how much food should be loaded onto the flight. This interaction enables quick and informed decision-making without requiring advanced technical knowledge from the user.FlightMax uses a PostgreSQL database to securely store and manage all flight and consumption data, ensuring a robust foundation for future optimizations and in-depth analyses. During development, Rider IDE is utilized for backend development, while DataGrip is preferred for efficient database management.This project offers a practical AI-based solution for the aviation industry, aiming to reduce operational costs while supporting environmental sustainability by minimizing food waste.

Keywords: in-flight catering optimization, meal demand forecasting, machine learning, inventory management, food waste reduction.

1. Introduction

FlightMax is a system designed to optimize in-flight catering processes in the aviation industry. This project focuses on critical objectives such as reducing operational costs, improving resource utilization, and promoting environmental sustainability. In-flight catering planning is crucial not only for customer satisfaction but also for lowering airline costs and minimizing food waste.

According to reports by the International Civil Aviation Organization (ICAO), food waste generated by air transportation amounts to approximately 5.7 million tons annually [1], leading to significant economic costs and environmental impacts. FlightMax addresses this issue by offering an innovative system that optimizes catering requirements.

The core component of the system is its artificial intelligence (AI) module, developed using Python. This module analyzes variables such as flight routes, passenger profiles, flight duration, and historical flight data to calculate the optimal amount of food required. The AI module's analyses prevent overloading and enhance the efficiency of operational processes. All data-driven operations rely on a secure PostgreSQL-based database.

FlightMax is supported by a web-based interface that allows users to input flight information and manage operational processes seamlessly. This interface consists of the following modules:
Dashboard: Users can quickly manage operational processes by viewing flight details, efficiency metrics, and AI-generated recommendations.
Catering Planning Module: Food quantities calculated by the AI can be adjusted and tailored for specific flights within this module.
Analytics and Reporting: Trend analyses and reports based on historical data are generated to evaluate the system's operational success.

This structure not only simplifies the user experience but also enables more effective optimization of catering processes.

[1]Corresponding author

Existing research in this field primarily focuses on predictive models for individual flights or passenger survey-based analyses. For instance, studies on meal planning models emphasize reducing food waste through techniques such as shopping lists, inventory management, and portion control [2]. Additionally, mathematical optimization techniques have been explored to create efficient and balanced meal plans that align with nutritional requirements while minimizing waste [3]. The FlightMax project builds on these concepts, integrating real-time AI- powered food provisioning optimization within the aviation sector.

Another key area of research highlights the application of optimization methods in food processing workflows. Optimization models have proven effective in minimizing waste by analyzing variables such as portion sizes, package dimensions, and expiration dates [4]. Furthermore, research has shown that decision support models significantly enhance operational efficiency in airline catering management [5].FlightMax adopts a similar approach, ensuring precise food loading calculations to minimize both waste and operational inefficiencies.

Furthermore, studies have been conducted on optimizing quality control in the food supply chain. These models monitor factors such as time and temperature to prevent product deterioration and improve logistics efficiency [6]. Recent findings indicate that incorporating AI in the food supply chain enhances quality control and minimizes waste [7]. Additionally, AI-driven optimization has been shown to improve predictive maintenance in food logistics, reducing spoilage risks and ensuring compliance with safety standards [1]. FlightMax aligns with these efforts by integrating quality control measures in its catering optimization processes, ensuring that food remains fresh and consumable while minimizing excess inventory.

Airlines such as Emirates and Etihad Airways have begun leveraging AI-driven food waste management systems. Emirates has set a target of reducing in-flight food waste by 35% [8], while Etihad Airways employs machine learning to analyze uneaten food after flights to enhance future meal planning [9].Additionally, studies on AI-based decision-making models in aviation indicate that intelligent algorithms improve resource allocation and cost management in airline operations [10]. However, these approaches primarily focus on single-flight optimizations, lacking consideration for sequential flight connections. Recent studies suggest that incorporating real- time passenger data into AI-driven models significantly improves forecasting accuracy and reduces meal wastage in airline operations [11]. FlightMax aims to bridge this gap by implementing a holistic, multi-flight optimization strategy.

In conclusion, FlightMax delivers a technological solution that modernizes operational processes in air transportation, reducing costs, optimizing resource usage, and minimizing food waste. By integrating artificial intelligence, data analytics, and user-friendly interfaces, the system represents a significant advancement in sustainable catering management for the aviation industry.

2. Materials and Methods

This section outlines the materials and methodologies employed in the development and implementation of the FlightMax system. The primary goal is to provide a detailed overview of the technical and methodological approaches used to optimize in-flight meal loading processes.

Figure 1 illustrates the operational structure of the FlightMax system and presents a comprehensive overview of the data interactions among all stakeholders. The system collects flight and catering data from external sources such as Airline Company & Kaggle, as well as from internal sources through the Flight Operations Manager. These data are transmitted to the core of the system, FlightMax, where they are analyzed by the Machine Learning Model to generate optimized food load calculations for each flight. The resulting calculations are then fed back into the system to support decision-making by operations managers. Additionally, flight and catering data are stored in the Flight Records and Catering Records databases to ensure data continuity and enable future analytical use. All these data flows are managed by the Admin, who also monitors system feedback. Furthermore, the outcomes of the analytical processes are reported through the Analysis Report unit. Through this structure, FlightMax aims to optimize operational decisions based on data, reduce food waste on flights, and enhance the efficiency of resource utilization. Figure 1 clearly depicts the data flow and functional relationships within this multi-actor system.

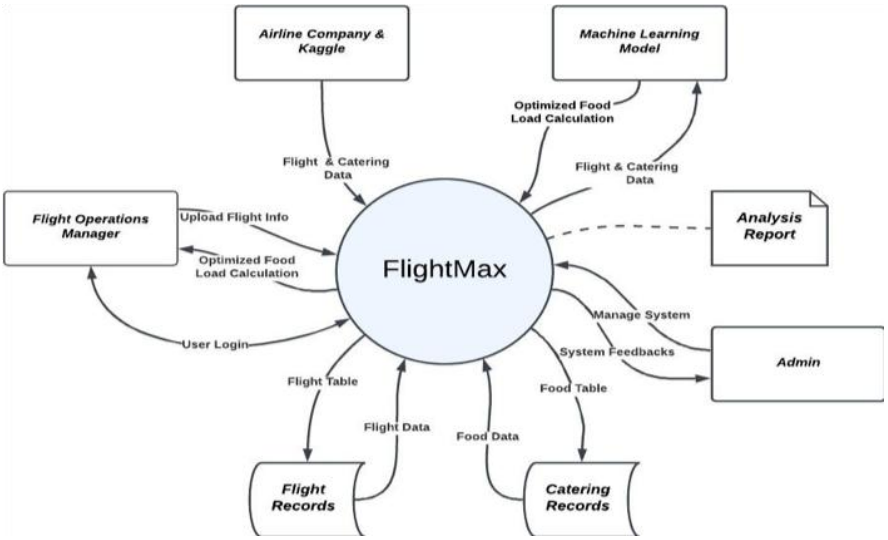


Figure 1. FlightMax context diagram

2.1. Core Functional Modules

The following subsections describe the core modules of FlightMax and their respective functionalities.

2.1.1. Flight Data Processing Module

This module is responsible for retrieving and processing flight-related data to generate meal loading recommendations.

Data Acquisition: The system integrates with airline data sources via APIs to retrieve flight details, including passenger count, route and class distribution.
Data Preprocessing: The retrieved data is cleaned and normalized to ensure consistency across different airline formats.
Real-time Updates: The module continuously synchronizes with airline systems to reflect last-minute passenger changes.

2.1.2. Meal Demand Prediction Module

This module leverages machine learning algorithms to forecast the required number of meals for each flight.

Predictive Model: A regression-based model analyzes historical flight data, passenger trends, and meal preferences to generate accurate predictions.
Seasonal Adjustments: The model incorporates seasonality and special event factors to fine-tune meal estimates.
Dynamic Scaling: Predictions are adjusted dynamically in response to real-time updates from the Flight Data Processing Module.

2.1.3. Inventory and Catering Management Module

This module ensures that the right amount of food is allocated and loaded efficiently.
Stock Tracking: Integration with inventory management systems allows real-time tracking of available meal components.

Order Generation: Based on demand predictions, the system generates procurement orders for required food items.
Waste Reduction Strategies: The module minimizes food waste by optimizing meal distribution based on past consumption patterns [11]. Recent research highlights that targeted meal planning strategies in airline catering can significantly reduce excess inventory and associated costs [12]. A study on cost-efficient airline catering suggests that adjusting meal provisioning based on historical waste data can lead to a significant reduction in operating expenses [13].

2.2. System Architecture

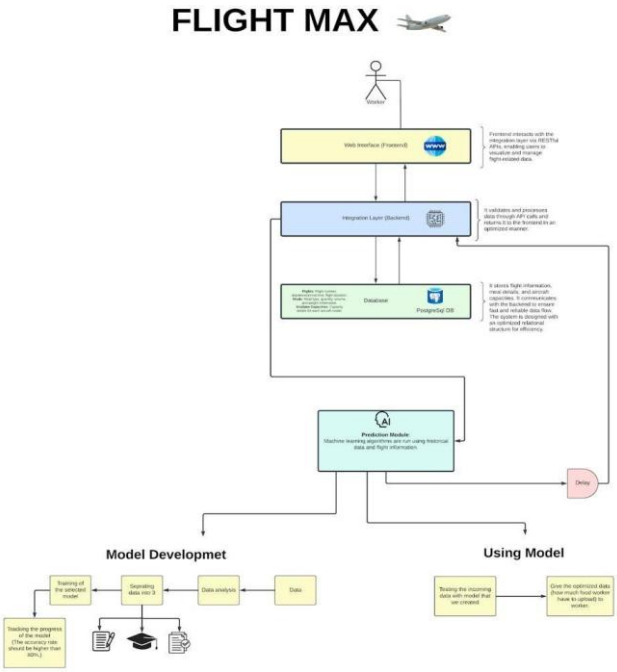


Figure 2. FlightMax high level design

FlightMax follows a multi-layered system architecture designed for scalability, security, and performance. Figure 2 is explained in the following items.

2.2.1. Frontend Layer

Developed using ASP.NET and based on MVC model.
Provides an intuitive dashboard for catering staff to monitor meal loading processes.

2.2.2. Backend Layer

Built using Python for machine learning and .NET for business logic processing.
Handles data analytics, inventory management, and order automation

2.2.3. Database Layer

Uses PostgreSQL to store flight and meal data.
Ensures data security with encryption and role-based access controls.

2.3. Development Environment

The FlightMax system was developed using the following tools and methodologies:
Programming Languages: Python for machine learning, C# for web application development.
Integrated Development Environments (IDEs): Rider for .NET development, Jupyter Notebook for machine learning experiments.
Version Control: Git was used for source code management.

2.4. Data Collection and Management

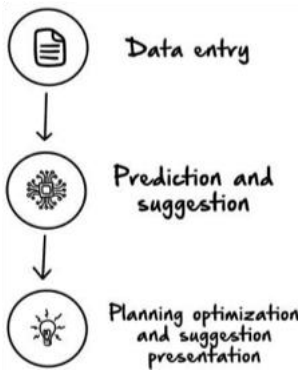


Figure 3. FlightMax dataflow

Historical Flight Data: Collected from airline records to train and validate the meal prediction model. The data comes from kaggle or flight companies.
Real-time Passenger Data: Synchronized with airline APIs to adjust meal orders dynamically.
Stock and Waste Data: Stored and analyzed to optimize future meal planning and reduce unnecessary food waste.
Data Cleaning: Data has to be cleaned and comes to meaningful data.

This methodology ensures that FlightMax operates efficiently, delivering precise meal loading recommendations while reducing food waste and operational costs.

3. Results

The FlightMax project dynamically calculates food loading based on flight type, passenger count, and flight duration, determining the optimal amount of food for each flight. FlightMax prevents unnecessary material loading, thereby increasing operational efficiency. The algorithms used in the system offer flexibility, allowing them to be customized based on different flight routes of airlines, ensuring that the right food and beverage selection is made for each flight. These calculations help prevent errors during the pre-flight preparation process and allow the staff to work faster and more accurately. Additionally, with the system's continuously updated databases and dynamic structure, it can quickly adapt to sudden changes in flights.

The web-based application has been developed to enable staff to make loading plans quickly and accurately. The application features a user-friendly interface that allows staff to easily input flight information, passenger numbers, and special food requests. This makes the loading planning process much faster, significantly reducing the pre-flight preparation time. The fast calculations and accurate data presentation minimize errors and make the system more reliable. As a result, the reduction in food loading errors has made operational processes more efficient.

Interface improvements based on user feedback have further simplified the system's use and increased staff efficiency. The user-friendly design of the application has particularly accelerated the onboarding process for new staff, allowing them to effectively use the system with minimal training.Research on UX optimization in aviation technology suggests that streamlined interfaces reduce operational errors and increase efficiency in complex workflows [14]. Furthermore, adjustments made based on operators' ongoing feedback have made the interface more understandable and accessible. These improvements have reduced time losses before flights and greatly contributed to speeding up operational processes. Additionally, research highlights that catering staff equipped with real-time data analytics tools experience a 20% improvement in decision-making speed and accuracy [13].

In conclusion, the FlightMax project has optimized food loading processes, reduced preparation times for flights, significantly reduced waste, and improved staff efficiency. The developed web-based application allows flight preparations to be carried out more quickly, accurately, and efficiently, increasing workforce productivity, accelerating operational processes, and reducing costs. These achievements will ensure smoother flight preparations in the future and enable the system to provide large-scale efficiency in the long run.

4. Conclusions

As a result, the FlightMax project has optimized food loading processes, making flight preparations faster, more accurate, and more efficient. By performing dynamic calculations based on flight type, passenger count, and flight duration, it has significantly reduced food waste. Thanks to the high accuracy of the calculation algorithms, the error rate among staff has been minimized, and the loading processes have been accelerated. Future studies suggest that integrating AI-driven optimization with blockchain-based inventory tracking could further enhance efficiency and security in airline catering operations [15]. Furthermore, research on sustainable airline catering emphasizes that incorporating regional food sourcing strategies can reduce waste and enhance passenger satisfaction [11]. Studies also indicate that integrating predictive analytics into airline catering logistics can enhance food loading efficiency and reduce overstocking issues [1]. The web-based application, with its user-friendly interface and fast processing capabilities, has enabled staff to plan loadings more effectively.These improvements have reduced time losses in the pre-flight preparation process, increased operational efficiency, and lowered costs. Continuous improvements made based on user feedback have further enhanced the system's efficiency, minimizing potential errors during flight preparations. FlightMax provides a robust infrastructure to manage flight preparation processes more quickly, efficiently, and error-free in the future.

5. Declaration of Interest

The authors declare that there is no conflict of interest regarding the development or presentation of the FlightMax project.

Acknowledgements

The authors acknowledge the support of Maltepe University, particularly the Software Engineering Department. Special thanks are extended to Dr. Emre Atlıer, Head of the Software Engineering Department, for his valuable guidance and supervision throughout the development of the FlightMax Project.

Author Contributions

Sena Yılmaz & Erdem Kurt: Collaboratively designed and developed the FlightMax system, implemented backend functionalities using Python and .NET/C#, conducted data analysis for flight catering optimization, and prepared the technical documentation and ICADA submission materials.

Dr. Öğr. Üyesi Emre Atlıer Olca: Served as the academic advisor, providing continuous feedback, supervision, and ensuring the academic rigor of the project.

6. References

[1] International Air Transport Association, *Global Air Travel Demand Continued Its Bounce Back in 2023*, Jan. 31, 2024. [Online]. Available: https://www.iata.org/en/pressroom/2024-releases/2024-01-31-02/

[2] M. A. van Rooijen, J. C. Gerdessen, G. D. H. Claassen, and S. de Leeuw, “Optimizing household food waste: The impact of meal planning, package sizes, and performance indicators,” *Resour. Conserv. Recycl.*, vol. 205, p. 107559, 2024. [Online]. Available: https://doi.org/10.1016/j.resconrec.2024.107559

[3] M. Türkay, *Optimization*. [Online]. Available: https://ku-people.s3.eu-west-1.amazonaws.com/cdn/files/Mysite/mturkay/indr262/Optimizasyon.pdf

[4] J. R. Banga, E. Balsa-Canto, C. G. Moles, and A. A. Alonso, “Improving food processing using modern optimization methods,” *Int. J. Prod. Econ.*, 2003. [Online]. Available: https://doi.org/10.1016/j.ijpe.2009.11.026

[5] E. E. Nebati, “Havayolu operasyonlarında dayanıklı ekip eşlemesi için bir karar destek modeli önerisi: Pisagor AHP- Pisagor WASPAS yaklaşımı,” *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, no. 81, pp. 92–110, 2024. [Online]. Available: https://doi.org/10.51290/dpusbe.1430329

[6] A. Rong, R. Akkerman, and M. Grunow, “An optimization approach for managing fresh food quality throughout the supply chain,” *Int. J. Prod. Econ.*, vol. 131, no. 1, pp. 421–429, 2011. [Online]. Available: https://doi.org/10.1016/j.ijpe.2009.11.026
[7] J. Zhu, X. Zhao, Y. Sun, *et al*., “Relational data cleaning meets artificial intelligence: A survey,” *Data Sci. Eng.*, 2024. [Online]. Available: https://doi.org/10.1007/s41019-024-00266-7
[8] Ginjer, *Consuming Data: How AI Can Help Airlines Reduce Food Waste*. [Online]. Available: https://ginjer.co.uk/consuming-data-how-ai-can-help-airlines-reduce-food-waste/
[9] Aviation Guide, *Etihad, Lumitics to Use Technology to Reduce Food Wastage on Flights*. [Online]. Available: https://aviationguideem.com/news/etihad-lumitics-to-use-technology-to-reduce-food-wastage-on-flights/
[10] H. Sarker, “AI-based modeling: Techniques, applications and research issues towards automation, intelligent and smart systems,” *SN Comput. Sci.*, vol. 3, no. 1, p. 158, 2022. [Online]. Available: https://doi.org/10.1007/s42979-022-01043-x
[11] İ. Anasız, “Otel işletmelerinde yiyecek-içecek maliyet kontrolü için israfı azaltacak farklı bir uygulama önerisi: ‘Hedef maliyetleme’,” *DergiPark*. [Online]. Available: https://dergipark.org.tr/en/download/article-file/686109
[12] A. Özdemir and E. Yılmaz, “Türkiye'de yiyecek seçiminin bölgelere göre farklılaşması,” *Res. Gate*, 2020. [Online]. Available: https://www.researchgate.net/publication/342568250_Turkiye_de_Yiyecek_Seciminin_Bolgelere_Gore_Farklilasmasi_Variati on_of_Food_Choices_According_to_Regions_in_Turkey

International Conference on Advances and Innovations in Engineering

3.ICAIE

Innovative Approach in Document Management System Procurement: Artificial Intelligence and Chatbot

Erbaycan KILIÇKAYA[1]

[1]Turkish Technology Central MRO Solutions İstanbul, Türkiye
e.kilickaya@thy.com

Abstract
Document Management Systems (DMS) are essential for organizations seeking digital transformation by enhancing efficiency, compliance, and accessibility while minimizing human error. Traditional procurement and implementation methods, however, often face difficulties such as complexity, low user adoption, time losses, and limited automation. This paper proposes an innovative approach to DMS procurement through the integration of Artificial Intelligence (AI) and chatbot technologies. The model demonstrates how AI-driven assistants can contribute to requirement analysis, supplier evaluation, and automated document classification. Moreover, chatbot-based interfaces are emphasized as user-friendly tools that facilitate interaction, provide timely reminders, and improve decision-making processes. The findings indicate potential benefits including cost savings, improved document and asset management, higher accuracy in information retrieval, and increased user satisfaction. By bridging technical systems with human interaction, the study shows how AI and chatbots can transform traditional DMS procurement into a more adaptive, transparent, and intelligent process.

Keywords: document management systems, artificial intelligence, chatbot, procurement, digital transformation

1. Introduction

Document Management Systems (DMS) have become an indispensable part of modern organizations by supporting digital transformation, facilitating asset management, ensuring compliance with legal regulations, and enhancing operational efficiency. As organizations generate and manage increasing volumes of information, effective document handling has a direct impact on decision-making, productivity, and corporate governance. Despite their growing importance, traditional approaches to DMS procurement and implementation are often constrained by significant challenges such as high complexity, insufficient user adoption, time and cost inefficiencies, and limited automation capabilities. These barriers not only delay digital transformation initiatives but also diminish the expected value of DMS investments.

In response to these challenges, innovative methods that leverage emerging technologies are being increasingly explored. Artificial Intelligence (AI) and chatbot solutions offer promising opportunities to improve both the procurement and utilization of DMS by enhancing automation, simplifying interactions, and supporting smarter decision-making. This study introduces an innovative approach that integrates AI-driven assistants and chatbot-based interfaces into the procurement process. The objective is to demonstrate how these technologies can bridge the gap between technical systems and human interaction, enabling organizations to achieve more adaptive, time-efficient, transparent, and intelligent document management strategies.

2. Sectoral Applications of Document Management Systems

Document Management Systems (DMS) have become indispensable tools across various industries, offering structured methods to capture, store, and manage documents throughout their lifecycle. In the healthcare sector, DMS ensures secure handling of patient records while complying with strict data protection regulations. Financial institutions rely on these systems to streamline audit processes and manage vast amounts of transactional data. Similarly, manufacturing and aviation industries benefit from DMS by maintaining accurate technical documentation and operational manuals critical for safety and compliance. Government agencies and public institutions utilize DMS to improve transparency, enable efficient archival processes, and support digital transformation efforts. These sectoral applications demonstrate the versatility of DMS and highlight the increasing need for reliable, accessible, and regulation-compliant document handling solutions.

[1]Corresponding author

3. Why Document Management Systems Alone Are Not Sufficient

Although Document Management Systems (DMS) provide structured storage, version control, and regulatory compliance, they often fall short of meeting the dynamic needs of modern organizations. Traditional DMS platforms primarily focus on static document archiving and retrieval, but they lack the ability to interpret unstructured data, automate decision-making, or provide contextual insights. Users frequently encounter challenges such as time-consuming manual categorization, limited search capabilities in complex datasets, and difficulties in integrating with other enterprise systems. Moreover, as the volume of digital content grows exponentially, these limitations become more evident, creating inefficiencies and reducing the overall value of information assets. Therefore, while DMS offers a strong foundation for document organization, it remains insufficient as a standalone solution for today's data-driven business environments.

4. Why Document Management Systems Alone Are Not Sufficient

The integration of Document Management Systems (DMS) with Artificial Intelligence (AI) and chatbot technologies has transformed conventional document handling into a more intelligent, interactive, and efficient process. AI-powered DMS can automatically classify, tag, and retrieve documents, significantly reducing the need for manual intervention. This not only minimizes human error but also ensures that employees spend less time searching for or organizing files, resulting in considerable time savings and measurable labor-hour efficiency.

Chatbots further enhance this ecosystem by providing users with instant, conversational access to stored information without the need for complex screen designs or training on system navigation. Employees can simply request specific reports, data, or documents through natural language queries, and the chatbot delivers accurate results in seconds. This capability ensures rapid report generation, supports informed decision-making, and promotes accessibility across the organization.

In addition, AI-powered validation mechanisms allow the system to verify document authenticity and accuracy automatically. By cross-checking metadata, version histories, and predefined compliance rules, the system can detect inconsistencies, outdated information, or unauthorized alterations. This ensures that documents used in daily operations are not only quickly accessible but also accurate, reliable, and compliant with organizational standards.

Another key advantage of AI integration is process design based on both metadata and document content. Since the system can access not only the metadata but also the actual document itself, complex workflows can be designed and automated. For example, when a process requires the bulk transfer of a specific set of documents, the system can recognize, group, and process them according to predefined rules without manual intervention. This allows organizations to streamline repetitive document-heavy tasks, optimize workflows, and ensure consistent execution across departments.

Overall, the synergy of AI and chatbots with DMS creates a user-friendly, time-efficient, and error-reducing environment, while also strengthening the integrity of corporate knowledge assets and enabling process-driven document management.

5. Security and Access Control Concerns in AI-Chatbot Integrated Document Management System

While the integration of Document Management Systems (DMS) with Artificial Intelligence (AI) and chatbots offers significant advantages, security and access control remain critical considerations. Especially in environments handling sensitive documents, several risks must be addressed. Unauthorized access can occur if user authentication and authorization mechanisms are not sufficiently robust, allowing individuals without proper permissions to gain access to sensitive information via chatbot or AI interfaces. Misconfigured APIs or shared environments may further increase this risk.

The choice between on-premises and cloud infrastructure also impacts security. On-premises systems provide greater internal control since data is stored locally, offering protection against external attacks, but require continuous maintenance, including server security, physical access control, backups, and patch management. Cloud-based systems introduce additional challenges related to data privacy, data residency, and the shared nature of infrastructure.

Data leakage and system misconfigurations remain key concerns during AI or chatbot integration with DMS. Inaccurate access paths or improper configurations may lead to unauthorized exposure of confidential documents or allow unintended external access to databases. Moreover, AI models themselves can be vulnerable to adversarial attacks or data poisoning, and reverse engineering or model inversion may expose sensitive information embedded in training datasets.

Maintaining data integrity and accuracy is also crucial. Ensuring documents remain unaltered, that version histories are traceable, and that metadata aligns with the document content helps preserve trustworthiness. AI-powered validation mechanisms can assist in this regard, but they must be correctly configured to avoid false approvals.

Finally, compliance with privacy regulations is essential. Standards such as GDPR and KVKK impose strict requirements on how personal data is processed, stored, and accessed. Organizations must implement measures such as data anonymization, encryption, and detailed access logging to ensure AI and chatbot-integrated DMS solutions adhere to regulatory requirements.

6. References

[1] Egnyt "AI Chatbot Security: Understanding Key Risks and Testing Best Practices" https://www.egnyte.com/blog/post/ai-chatbot-security-understanding-key-risks-and-testing-best-practices?utm_source=chatgpt.com (09.05.2025)

[2] Medium "Securing AI workloads in the Cloud and On-prem" https://medium.com/%40cloudwithusama/securing-ai-workloads-in-the-cloud-and-on-prem-53f8d36dc647

[3] Wiz "7 Serious AI Security Risks and How to Mitigate Them" https://www.wiz.io/academy/ai-security-risks?utm_source=chatgpt.com (28.03.2025)

[4] Gocloudforce " Don't Spill the Beans: How Private AI Keeps Your Data Safe" https://gocloudforce.com/private-ai-data-safety/?utm_source=chatgpt.com (21.01.2025)

[5] Ec Council Univercity "How AI-Powered Chatbots Are Becoming Cybersecurity Threats" https://www.eccu.edu/blog/how-ai-powered-chatbots-are-becoming-cybersecurity-threats/ (2025)

[6] OpenAI "Privacy Policy." https://openai.com/tr-TR/policies/privacy-policy/ (27.06.2025)

[7] IBM "What is chat bot" https://www.ibm.com/think/topics/chatbots (15.01.2025)

A Pharmaceutical Stock Forecasting Approach: Comparative Analysis of Statistical and Machine Learning Models

Osman Furkan KÜÇÜK[1], Mehmet KARAKÖSE[2], İlker GÜRELLİ[3], Yaren BAYDENİZ[4], Eray HANOĞLU[5]

[1,2]Department of Computer Engineering, Faculty of Engineering, Fırat University, Elazığ, Türkiye.
[3,4,5]Panates Technology Investment R&D Center, İzmir, Türkiye
[1]210260042@firat.edu.tr, [2] mkarakose@firat.edu.tr, [3] ilker.gurelli@panates.com, [4]yaren.baydeniz@panates.com, [5]eray.hanoglu@panates.com
[1](ORCID: 0009-0002-5594-3009), [2](ORCID: 0000-0002-3276-3788), [3](ORCID: 0009-0006-8117-0475), [4](ORCID: 0009-0002-5625-6550),[5](ORCID: 0000-0001-9491-7028)

Abstract
Pharmaceuticals are undoubtedly one of the most critical components of the healthcare supply chain. Ordering the correct amount of medications is vital for preventing issues such as unnecessary inventory surplus or shortages. Therefore, a robust stock analysis is imperative. In this study, for medication stock forecasting, statistical methods such as Single Exponential Smoothing (SES), Holt, Holt-Winters, SARIMA, and SARIMAX, as well as the Random Forest model are presented in a comparative manner. In addition to forecasting medication stock in a univariate (single-variable) approach, the relationship between disease incidence and drug usage is evaluated through a multivariate (multi-variable) approach. The obtained results clearly show that including external variables in the models significantly increases prediction accuracy compared to using univariate models alone. This study contributes to the field by combining traditional time series models with machine learning approaches and evaluating them under real-world healthcare data conditions. By introducing a multivariate forecasting structure, the study highlights the added value of integrating diagnosis information into drug demand estimation. The results demonstrate the practical applicability of the proposed methods and their potential to support more informed, data-driven decision-making in healthcare supply chain management.

Keywords: drug stock forecasting, time series analysis, statistical models, random forest

1. Introduction

Undoubtedly, one of the most important aspects in healthcare services is to provide patients with the right medication at the right time and in the proper quantity. In addition to ensuring that every element of the healthcare supply chain operates systematically, it is essential that the amount of drugs supplied is planned and sufficient to meet demand [1]. Adequate and timely stocking is also an indicator of the high-quality service provided to patients [2]. However, factors such as demand uncertainty, disease crises, emergencies, and deficiencies in organizational planning negatively affect inventory management in all healthcare organizations, especially hospitals and pharmacies. Incorrect or insufficient stock management can lead to high costs, waste of resources, or drug shortages. Drug shortages may cause treatment interruptions, delays in treatment, or even the absence of vital medications, which directly jeopardizes patient safety [3-4]. This situation is unacceptable in modern healthcare systems.

In traditional drug supply operations, the quantities of medications are mostly determined based on the simple statistical averages of past data and fixed assumptions. These approaches often fail to adequately capture dynamic factors such as sudden changes in demand and seasonality. As a result, pharmacies and hospitals may not be able to meet medication needs, leading to delays in patient treatment or even causing patients to seek drugs from alternative supply sources [5].

As illustrated in Figure 1, pharmaceutical products follow a linear flow from raw material suppliers and manufacturers to patients through distribution platforms, wholesalers, regional distributors, and points of care such as hospitals and pharmacies [6]. At each stage, delays or misestimations in demand can lead to critical shortages or excessive stock. In this context, it is required that records obtained from hospital and pharmacy systems be utilized for proactive stock forecasting. The integration of a Pharmaceutical Stock Forecast Module is deemed essential, as it plays a critical role in reducing forecasting uncertainties and ensuring timely medication availability within the supply chain. Furthermore, it is widely recognized that data-driven and adaptable methods not only offer more competent solutions in today's dynamic supply chain environment, but that advanced AI methods can significantly enhance

[1]Corresponding author

both the reliability and efficiency of healthcare operations, thereby providing robust support for clinical decision processes [7]. These approaches have the ability to quickly identify sudden changes and uncertainties in stock levels, thereby helping to prevent critical issues such as excessive stock or shortages. In particular, machine learning-based models are notable for their capacity to reveal complex relationships and patterns in the data, and to generate more realistic forecasts of future trends.

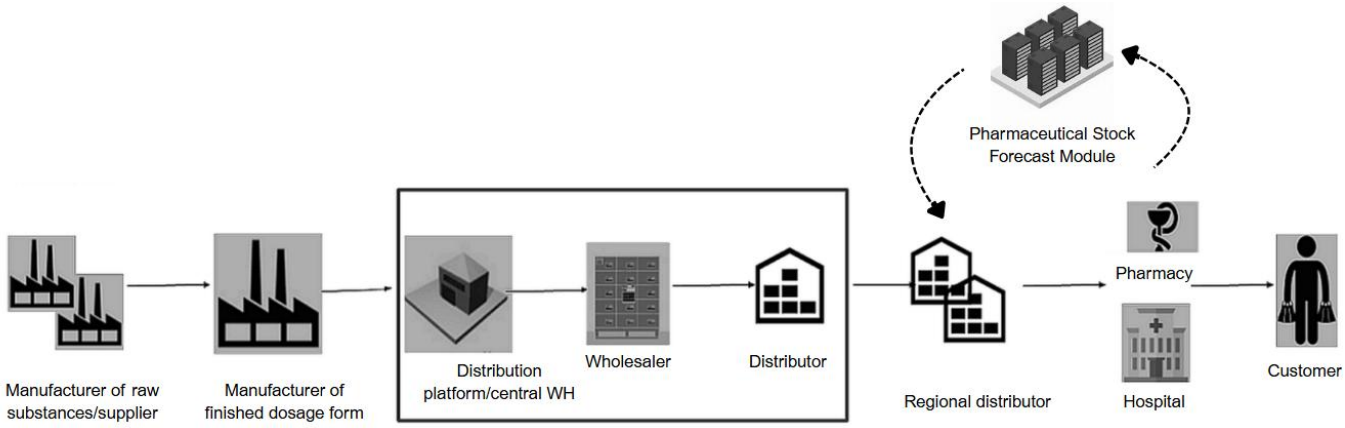


Figure 1. A general flow chart for the pharmaceutical supply chain
(Adapted and extended from Merkuryeva et al. [6])

Queiroz et al. [8] stated that digital supply chain models enhance forecasting capabilities in inventory management. The authors demonstrated that, by integrating technologies such as IoT, cloud computing, big data analytics, and artificial intelligence, a much more dynamic and real-time flow of data is achievable compared to traditional approaches. This increased data fluidity improves inventory predictions by minimizing uncertainties and sudden demand fluctuations.

Baryannis et al. [9] elaborated on how predictive analytics methods support operational efficiency and flexibility. Their study revealed that machine learning and other forecasting techniques are highly effective at detecting uncertainties in inventory management processes and preventing problems associated with excess stock or shortages. They detailed that data-driven models provide higher accuracy and adaptability than conventional methods, thereby enabling enhanced operational performance and cost optimization.

Shanthi and colleagues [2] applied various machine learning methods for hospital inventory management to forecast future drug usage. They trained Random Forest, Decision Trees, and Convolutional Neural Networks (CNN) models using Kaggle datasets. Their study demonstrated that Random Forest enhances generalization ability and reduces overfitting, emerging as the most accurate model for predicting medication demand. In addition, they highlighted that the CNN model was more effective at capturing subtle fluctuations in drug usage levels.

Mousa and Al-Khateeb [10] comprehensively explored the applicability of machine learning and deep learning techniques in drug demand forecasting. In their study, they compared the performance of traditional time series methods with that of artificial intelligence-based models to minimize issues such as overstocking and understocking in medication inventory management. The authors investigated approaches including linear regression, Decision Trees, Support Vector Machines, Random Forest and recurrent neural network (RNN) models based on Long Short-Term Memory (LSTM). Their results confirm the feasibility of using machine learning and deep learning methods for drug demand forecasting while also showing that data limitations and market dynamics can significantly impact model performance.

Raisya et al. [11] proposed the use of the ARIMA (3,1,3) model in the stock estimation of single-use drugs. In the study, historical consumption data from January 2022 to February 2025 of Banda Sakti Health Center were used. As a result of the estimations, it was presented that the model produced reliable outputs and was effective in minimizing

problems such as excess or shortage of stock. It was also emphasized that the ARIMA model provided high accuracy in short-term predictions, thus allowing healthcare institutions to make more precise resource planning.

In another study, Bhat et al. [12] performed demand forecasts using ARIMA, Auto ARIMA, regression, LSTM models and stochastic models using three-year data on Aceclofenac 100 mg drug consumption, lead time, patient numbers and service level obtained from 53 healthcare institutions in India. Their findings showed that regression-based methods provide more reliable predictions in predicting stock shortages, time series models in short-term forecasts, and stochastic models in situations of uncertainty.

Elisawati and colleagues [13] developed a forecasting model using an artificial neural network-based backpropagation algorithm for more accurate planning of medication inventories in a public healthcare center. In the study, the model was trained using 70 rows of real inventory data (including variables such as initial stock, procurement, usage, remaining stock, etc.) from 2019–2020; 42 records were used for training and 28 records for testing. As a result of the forward-backward propagation steps, a prediction accuracy of 88.04% was obtained with 9 hidden layer neurons. These results provide a strong example of the applicability of classical artificial intelligence-based methods for drug inventory management in the healthcare supply chain, even with a limited amount of data.

When the studies are examined, it can be said that machine learning methods provide successful results in drug stock estimation. For this reason, it is believed that machine learning and classical models hold tremendous importance and that their use in the healthcare supply chain is very significant.

The aim of this study is to contribute to the healthcare supply chain by using classical machine learning methods. It is of vital importance that medications, as one of the main components in the supply chain, are ordered in the correct quantity and that needs are accurately determined. We comparatively test the performance of time series analysis methods, including Single Exponential Smoothing (SES), Holt, Holt-Winters, SARIMA, and SARIMAX, along with the Random Forest algorithm. The distinctive feature of our study compared to others is that a forecasting model is not developed solely based on past stock information. Instead, we perform forecasting using two variables by considering that drug usage increases or decreases in parallel with the number of specific diseases or disease groups: past usage records and the number of disease incidences. Additionally, in contrast to other studies that use a a limited set of statistical models, this study presents a comparative evaluation of a wider range of time series forecasting methods. In the following section, the methods employed, application processes, and the obtained results are explained in detail, thereby providing recommendations for inventory management in the healthcare supply chain.

2. Forecasting Models Using Time Series Data for Healthcare

Time series data is the expression of records obtained at specific time intervals, represented by equally spaced and sequentially ordered numbers [14-15]. Time series forecasting, on the other hand, enables the prediction of future trends and patterns using historical data. Accurately determining drug demand in the health supply chain is of great importance for effective stock management. Although many statistical methods and techniques are used in time series forecasting, the number and scope of these methods have expanded with the introduction of machine learning into our lives.

2.1. Single Exponential Smoothing (SES)

The SES method is a statistical method that requires very little computation. SES produces future estimates based on the weighted contribution of the most recent past observation.

$$\hat{y}_{t+1} = \alpha y_t + (1-\alpha)\hat{y}_t \quad (1)$$

In Equation 1 y_t represents the actual value at time t, $\hat{y}_{t+1}$ represents the forecast for time t+1 and α is a parameter defined between 0 and 1.

2.2. Holt and Holt – Winters

The Holt Model also takes into account the trend component observed in the time series, and is an extended version of SES. It provides more accurate estimates for data showing increasing or decreasing trends by updating both the level and the trend over time. The Holt-Winters model is the Holt model with the seasonal component added. This model captures not only the trend but also the periodic (seasonal) fluctuations in the data set.

$$\hat{y}_{t+m} = L_t + mb_t \quad (2)$$

$$\hat{y}_{t+m} = L_t + mb_t + S_{t-L+m}\ mod\ L \quad (3)$$

Equation 2 represents the equations of the Holt model, while Equation 3 shows the equation of the Holt-Winters model. L_t ve b_t represent the level and trend components at time t respectively. S_t denotes the seasonal component, m is the number of periods to be forecasted and L refers to the seasonal period. In both models, all components are updated at each observation.

2.3. SARIMA and SARIMAX

Seasonal Autoregressive Integrated Moving Average (SARIMA) model is an extension of ARIMA model with seasonal features. Its structure includes seasonal difference operation, seasonal autoregressive terms and seasonal moving average features [16]. These additional features are designed to capture the patterns that repeat at certain periods.

$$\Phi_P(L^S)\ .\ \phi_P(L)\ .(1-L)^d\ .(1-L^S)^D y_t = \Theta_Q(L^S)\ .\theta_q(L)\ .\ \epsilon_t \quad (4)$$

In Equation 4, the term y_t represents the actual value at the time t. Each term in this equatio corresponds to a different structural component of the series. The expression $(1-L)^d$ performs differencing to identify the overall trend in the data (e.g, upward or downward movements), while $(1-L^S)^D$ apples seasonal differencing to model repeating seasonal patterns. The component $\phi_P(L)$ and $\theta_q(L)$ indicate the relationships with past values of the data (autoregressive – AR) and past forecast errors (moving average – MA), respectively. $\Phi_P(L^S)$ and $\Theta_Q(L^S)$ are the seasonal counterparts of these terms. Finally ϵ_t denotes the random errors that the model fails to explain. This structure enables SARIMA to generate highly accurate forecasts based on both past trends and seasonal patterns. These features make SARIMA a prominent statistical model, especially in studies where training data is limited [17].

SARIMAX (Seasonal ARIMA with eXogenous variables), an extension of the SARIMA model, additionally incorporates external (exogenous) variables into the calculation. This allows forecasting to be based not only on the historical values of stock data but also on other variables that may influence stock levels, such as diagnosis count in this study. While the structure of the SARIMAX model is largely similar to that of SARIMA, an additional term βX_t is included at the end of the equation to represent the effect of the exogenous variable. Here, X_t denotes the external variable affecting stock levels, and β represents its corresponding coefficient.

2.4. Random Forest

Random Forest is a powerful machine learning method that can capture nonlinear relationships in data beyond classical time series models. In time series data, predictions are made using features obtained from past observations. Each decision tree is trained on a specific subset of data, and the final prediction is calculated as the average of the outputs of all trees. The low risk of overfitting and the lack of time spent on hyperparameter tuning make it the preferred machine learning method for various problems [18].

3. The Proposed Approach for Pharmaceutical Stock Forecasting

In our study, the statistical methods discussed in the previous section and the Random Forest machine learning model were tested for drug stock forecasting in the healthcare supply chain. The forecasting process aims to predict future stock levels based on historical records of diagnosis counts and stock quantities. The models were structured to account for both seasonality and trend components.

In our approach, a dual-pronged method was employed for drug stock forecasting. Specifically, we structured our forecasting workflow into two distinct branches—one utilizing solely historical stock data, and the other incorporating both stock data and diagnosis counts. This dual-branch strategy allowed us to directly compare univariate and multivariate models, demonstrating that the inclusion of external variables significantly enhances prediction accuracy. Figure 2 illustrates a comprehensive diagram of the applied steps and methods.

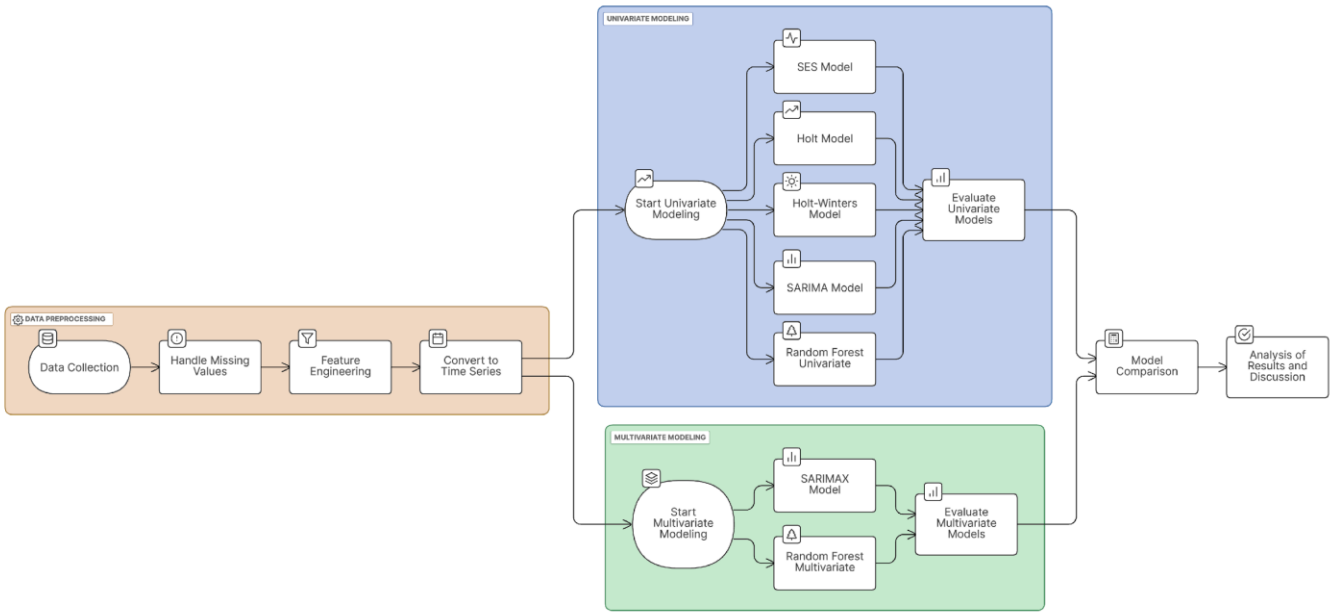


Figure 2. General flow diagram of the implemented data processing, modeling and analysis steps

4. Experimental Results

4.1. Dataset

The dataset was obtained from the Hospital Information Management System (HIMS) module of a private chest diseases hospital. The dataset includes the drug "PIRAZINID 500 MG 50 TABLET", which is related to tuberculosis disease, and includes historical stock records and diagnosis number data. These daily recorded data were used as basic features for stock quantity estimation (Figure 3).

DATE	DRUG_NAME	DIAGNOSIS_CODE	DIAGNOSIS_NAME	PRICE	DIAGNOSIS_COUNT	STOCK_COUNT
2010-01-02	PIRAZINID 500 MG 50 TABLET	A15-A19	Tuberculosis	0.082879	10	40.0
2010-01-03	PIRAZINID 500 MG 50 TABLET	A15-A19	Tuberculosis	0.082879	10	40.0
2010-01-06	PIRAZINID 500 MG 50 TABLET	A15-A19	Tuberculosis	0.082880	1	8.0
2010-01-07	PIRAZINID 500 MG 50 TABLET	A15-A19	Tuberculosis	0.082880	6	20.0
2010-01-08	PIRAZINID 500 MG 50 TABLET	A15-A19	Tuberculosis	0.082880	13	58.0
...	...	...	...	...	...	...
2024-12-27	PIRAZINID 500 MG 50 TABLET	A15-A19	Tuberculosis	4.555100	11	34.0
2024-12-28	PIRAZINID 500 MG 50 TABLET	A15-A19	Tuberculosis	4.555100	11	34.0
2024-12-29	PIRAZINID 500 MG 50 TABLET	A15-A19	Tuberculosis	4.555100	11	34.0
2024-12-30	PIRAZINID 500 MG 50 TABLET	A15-A19	Tuberculosis	4.555100	13	59.0
2024-12-31	PIRAZINID 500 MG 50 TABLET	A15-A19	Tuberculosis	4.555100	11	37.0

5437 rows × 6 columns

Figure 3. Data frame view of the dataset

Records between 2010 and 2025 were obtained and preprocessed. The preprocessing steps are explained in Section 4.2.

4.2. Data Preprocessing

In order to make the data suitable for analysis, pre-processing steps must be applied. First, unused columns (drug name, diagnosis name, diagnosis code, price) were removed, and missing data were detected. A total of 42 unrecorded daily data points were generated using the Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) method. Seasonal features were extracted by deriving day, month, weekday, and quarter information from the date index. These temporal components were transformed into cyclical features by applying sine and cosine transformations (e.g., sin and cos values for day, month, weekday, and quarter). Subsequently, additional features were created by calculating 1, 7, and 30-day lag values, 7 day moving averages, and 7 day moving standard deviations based on the diagnosis and stock quantities. The dataset was scaled to a range between 0 and 1 and sorted chronologically. Finally, 5113 records dated between 2010 and 2024 were used as training data, while 366 records from 2024 to 2025 were reserved for testing.

4.3. Model Setup And Training

As mentioned earlier, the model setup and training was carried out in two distinct steps. In the first step, estimates were made using only stock data. Single Exponential Smoothing, Holt, Holt-Winters, SARIMA, and Random Forest univariate models were applied. In the second step, in order to account for the relationship between diagnosis counts and drug usage, SARIMAX and Random Forest multivariate models were employed. The parameters of each model were optimized with methods such as GridSearch, AIC/BIC criteria and cross-validation, and the performances of the models were compared by creating training and test data sets.

After the training step the forecasting performances were evaluated and compared using the metrics of MSE (Mean Squared Error), RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), and MAPE (Mean Absolute Percentage Error). Table 1 presents the results of the models trained in both univariate and multivariate forms.

Table 1. Error metrics of models

	Model	MSE	RMSE	MAE	MAPE (%)
0	SES	0.005749	0.075822	0.057043	35.525504
1	Holt	0.005789	0.076085	0.056636	34.624921
2	Holt – Winters	0.005459	0.073886	0.055855	33.149436
3	SARIMA	0.006511	0.080691	0.058432	31.409126
4	Random Forest	0.002414	0.049127	0.033546	18.694484
5	SARIMAX (Multivariate)	0.002042	0.045191	0.031365	16.023269
6	Random Forest (Multivariate)	**0.001783**	**0.042221**	**0.028479**	**14.832471**

In models where only stock data was used, similar results were obtained between Single Exponential Smoothing (SES), Holt, Holt-Winters, and SARIMA models. However, when the error metrics of these models were examined, the single-variable version of Random Forest provided the lowest error values. In multivariate models (in the case where the number of diagnoses was used together with stock data), SARIMAX and Random Forest multivariate models were applied. These models demonstrated a noticeable improvement in forecasting performance due to the additional information provided by the exogenous variable. The MAPE value of the SARIMAX model was observed to be 16.02%, while the Random Forest multivariate model achieved a MAPE of 14.83%. These two models stand out for producing realistic results and being more practical to implement compared to deep learning methods. The predictions of both models on the test set are presented in Figure 4. In both cases, Random Forest performed better than the statistical models. The results show that incorporating diagnosis information into the forecasting process significantly improves the accuracy of drug stock predictions.

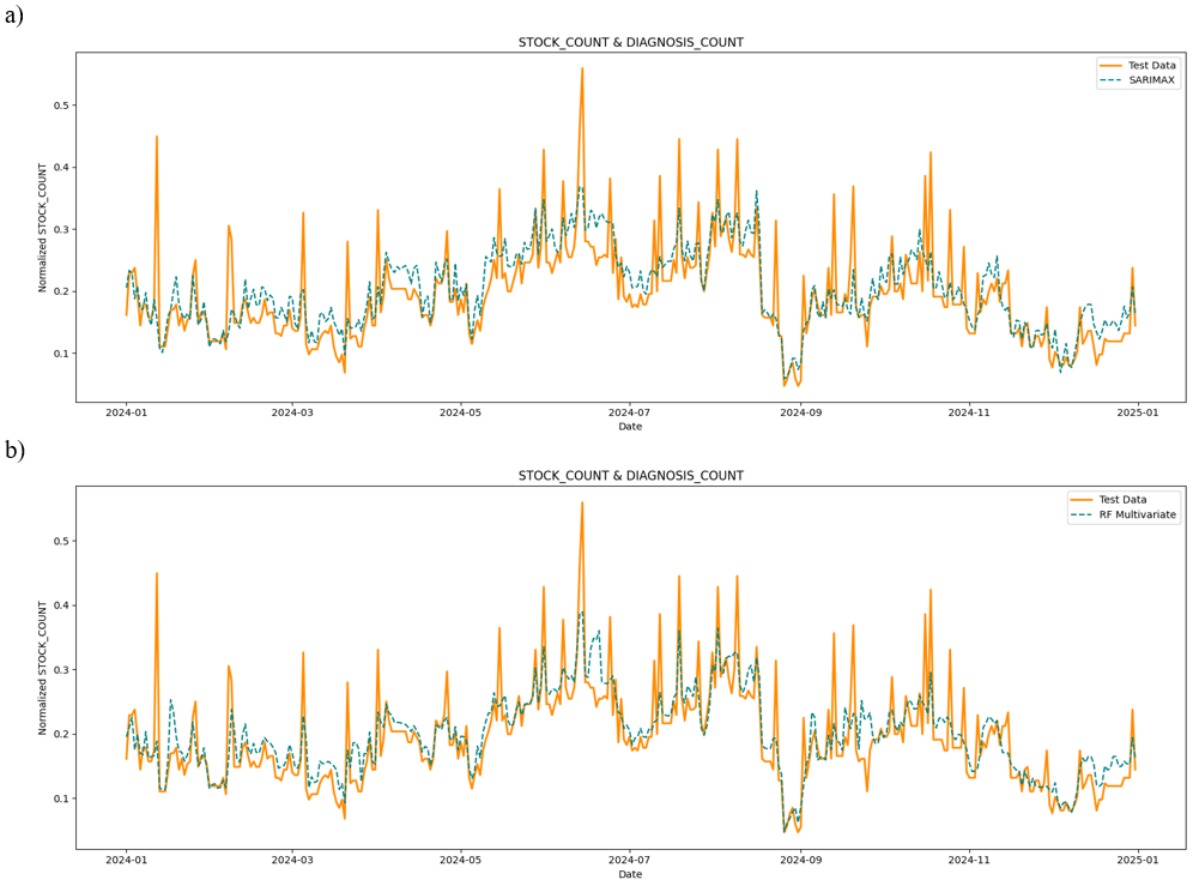


Figure 4. Drug stock forecasting results of multivariate models on the time series
a) SARIMAX, b) Random Forest (Multivariate)

5. Conclusion

In our study, various methods for forecasting future levels of drug stocks in the healthcare supply chain were comparatively examined. Statistical time series models SES, Holt, Holt-Winters, SARIMA, SARIMAX and Random Forest were used to predict the stock change of the drug "PIRAZINID 500 MG 50 TABLET". Model training was performed with only historical stock information and both stock information and Tuberculosis diagnosis frequency variables, respectively. Using only stock information was significantly outperformed by the approach that included both diagnostic and stock data. While SARIMAX and Random Forest showed comparable prediction accuracy in the multivariate case, Random Forest achieved the lowest error values. The ability to realistically predict future stock levels of medications presents a significant contribution in terms of reducing costs and ensuring the continuity of patient care.

In future work, it is planned to include a greater variety of drug types and diseases, allowing for an evaluation that more closely reflects real-world scenarios. The methodological approaches presented in this study provide a solid foundation for developing data-driven decision mechanisms in the healthcare supply chain and offer guidance for future research.

Acknowledgements

This study has been supported by the TUBITAK (The Scientific and Technological Research Council of Turkey) under Grant No: 3230442.

6. References

[1] Fole A, Kulsaputro J, Erniyani, Raodah, Yanasim N. "Application of ABC and EOQ Methods to Improve Control of Patented Medicine Inventory at Pharmacy A". EPI International Journal of Engineering, 7(1), 5-15, 2024.
[2] Darshan.U, Chinmaya.G.P, Harsh Abhinav, S.Varun Kumar, Deepak.R, Dr.Shanthi S. "Predictive Analysis on Medicines Availability in Hospitals Using Machine Learning and Deep Learning Technique". International Journal For Multidisciplinary Research, 7(1), 1-10, 2025.
[3] Said A, Goebel R, Ganso M, Zagermann-Muncke P, Schulz M. "Drug shortages may compromise patient safety: Results of a survey of the reference pharmacies of the Drug Commission of German Pharmacists". Health Policy, 122(12), 1302-1309, 2018.
[4] Castro C, et al. "Logistics reorganization and management of the ambulatory pharmacy of a local health unit in Portugal". Evaluation and Program Planning, 80, 101801, 2020.
[5] Shaban H, Maurer C, Willborn RJ. "Impact of drug shortages on patient safety and pharmacy operation costs". Federal Practitioner, 35(1), 24, 2018.
[6] Merkuryeva G, Valberga A, Smirnov A. "Demand forecasting in pharmaceutical supply chains: A case study". Procedia Computer Science, 149, 3-10, 2019.
[7] Arslanoğlu K, Karaköse M. "An Overview for Trustworthy and Explainable Artificial Intelligence in Healthcare". 28th International Conference on Information Technology (IT), IEEE, 21-24 February 2024.
[8] Queiroz MM, Telles R, Bonilla SH. "Blockchain and supply chain management integration: a systematic review of the literature". Supply Chain Management: An International Journal, 25(2), 241-254, 2020.
[9] Baryannis G, Dani S, Antoniou G. "Predicting supply chain risks using machine learning: The trade-off between performance and interpretability". Future Generation Computer Systems, 101, 993-1004, 2019.
[10] Mousa BA, Al-Khateeb B. "Predicting medicine demand using deep learning techniques: A review". Journal of Intelligent Systems, 32(1), 20220297, 2023.
[11] Kamila R, Abdullah D, Meiyanti R. "Stock Prediction of Single-Use Medicine Using Autoregressive Integrated Moving Average". Proceedings of Malikussaleh International Conference on Multidisciplinary Studies (MICoMS), Lhokseumawe, Indonesia, 20-21 November 2024.
[12] Bhat SS, Srihari VR, Prabhune A, Satheesh SS, Bidrohi AB. "Optimizing Medication Access in Public Healthcare Centers: A Machine Learning Stochastic Model for Inventory Management and Demand Forecasting

in Primary Health Services". 2024 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE), IEEE, 2024.
[13] Elisawati E, Linarta A, Elvaningsih H. "Analysis of Backpropagation Method in Predicting Drug Stock". Sinkron: Jurnal dan Penelitian Teknik Informatika, 6(2), 297-307, 2022.
[14] Liu Z, Zhu Z, Gao J, Xu C. "Forecast methods for time series data: A survey". IEEE Access, 9, 91896-91912, 2021.
[15] Zhang H, Li S, Chen Y, Dai J, Yi Y. "A Novel Encoder-Decoder Model for Multivariate Time Series Forecasting". Computational Intelligence and Neuroscience, 2022(1), 5596676, 2022.
[16] Lim SSY, Phouratsamay SL, Yahouni Z, Gascard E. "Medicine consumption demand forecasting in French hospitals using seasonal auto-regressive integrated moving average (SARIMA) models". 2024 International Conference on Control, Automation and Diagnosis (ICCAD), IEEE, 2024.
[17] Miller JA, Aldosari M, Saeed F, Barna NH, Rana S, Arpınar IB, Liu N . "A survey of deep learning and foundation models for time series forecasting". arXiv preprint arXiv:2401.13912, 2024.
[18] Poyraz İ, Gürhanlı A. "Demand Forecasting with Time Series Analysis Using Drug Sales Data". International Journal of Engineering Research and Applications, 10(7-III), 51–54, 2020.

Detection of Satellite Network Anomalies with XOR Based Encryption and Transformer Assisted Approach

Nida CANPOLAT[1] , Şengül DOĞAN[2], Mehmet KARAKOSE [3], Türker TUNCER[4]

[1,2,4]Department of Digital Forensic Engineering, Faculty of Technology, Firat University, Turkey
[3]Department of Computer Engineering, Faculty of Engineering, Firat University, Elazig, Turkey.
[1]231144110@firat.edu.tr, [2]sdogan@firat.edu.tr, [3]mkarakose@firat.edu.tr, [4]turkertuncer@firat.edu.tr
[1](ORCID: 0000-0002-5262-1509), [2](ORCID: 0000-0001-9677-5684),[3](ORCID: 0000-0002-3276-3788), [4](ORCID:0000-0002-5126-6445)

Abstract
Satellite networks are one of the critical components of modern communication infrastructures. They are used in areas that are particularly important such as defense, weather monitoring, and global data transmission. The vulnerability of these systems, which are so important in communication areas, to cyber attacks makes them quite vulnerable. Therefore this study aims to detect anomalies that may occur in satellite networks. At the same time traditional methods have been exceeded in order to ensure the security of the data. The data to be used was encrypted with an XOR encryption method applied using a random key before analysis. Thus, the proposed model performed the learning process on the encrypted data without directly using the raw data. After the encrypted data was made suitable for the deep learning model, it was classified with a Transformer based architecture. In order to test the success of the model an accuracy of 92.66% was obtained on the test data. This result effectively shows that anomaly detection can be performed with high accuracy without violating the privacy of the data. The developed method reveals that effective analyses can be performed without sharing data especially in systems with high security levels. In addition it enriches the work done in this field by offering a new perspective for artificial intelligence-supported cyber defense applications.

Keywords: satellite networks, cyber attack detection, deep learning, data encryption, transformer architecture

1. Introduction

Satellite networks have become one of the fundamental building blocks of communication in today's globalizing and digitalizing world. These systems play vital roles in many areas such as television broadcasting, GPS based navigation services, internet connection, military communication, meteorological observation and disaster management. Especially in rural areas or in areas with limited infrastructure access, satellites fill a critical gap by providing high reliability and coverage in data transmission, where land-based communication technologies are inadequate. For example, GPS directly rely on satellite-based services in many sectors, from maritime and air transportation to various areas of civil life. In addition, climate change observations through remote sensing, natural resource management and disaster early warning systems are also largely dependent on these technologies. In this context satellite communication infrastructure has an important place in the international arena not only in terms of data transfer but also in terms of strategic and economic superiority. The sustainability accuracy and security of these systems are increasingly important in terms of the independence and cyber sovereignty of countries in the information age.

However the increasing importance of satellite based communication systems has made these systems attractive targets for cyber threats. Satellite networks can be vulnerable to various attack vectors due to both physical infrastructure deficiencies and protocol level weaknesses. These threats include many types of attacks such as jamming, spoofing, unauthorized access, data exfiltration, DoS-DDoS, and system manipulation. For example, if an attacker disrupts or manipulates GPS signals that provide location information, this can lead to serious security problems and even loss of life in highly sensitive sectors such as air transportation. Similarly the seizure of satellite links used in military communications can lead to access to critical operational information or disruption of the chain of command. The potential effects of these attacks are at a level that can threaten not only individuals or institutions, but also national security, public order and economic stability. Therefore satellite networks must not only operate efficiently, but also be strongly protected, and potential threats must be detected at an early stage.

[1]Corresponding author

In this study, an artificial intelligence-supported approach is proposed for early detection of cyber attacks that may occur in satellite networks. While many existing systems process data directly we prioritize protecting the confidentiality of data and designed a system that can detect anomalies without the need for direct access to raw data. For this purpose before the data is analyzed, it is transformed using a randomly generated key with an XOR based encryption algorithm, and then these ciphertexts are processed. This method provides advantages especially in environments where sensitive information is processed and data sharing is not possible due to security policies. Encrypted data was analyzed with a deep learning-based transformer architecture and the accuracy rate of the system was measured as 92.66%. This approach not only provides high accuracy anomaly detection, but also proves that security analysis can be performed without violating data confidentiality. The method we propose has significant potential for restructuring cybersecurity solutions, especially in military and strategic data networks that are not open to the public. Our study aims to contribute to satellite networks becoming more resilient by addressing both security and privacy principles together.

2. Literature Review

This section includes studies on satellite network security.
[1] - Muhammad Usman et al. conducted a study suggesting that the ICMP protocol be used in a more limited way to detect and prevent DOS attacks on satellite networks. As a result of the study DOS attacks on satellite networks were reduced by reducing the use of the ICMP protocol.
[2] - S. Lohani et al. conducted a study using AES and RSA encryption techniques to ensure the security of satellite networks. Thanks to the method used since the data is encrypted and used, preventing unauthorized access has become commonplace.
[3] - N. Koroniotis et al. conducted a study using deep learning and network forensics to detect cyber attacks on satellite networks. They classified the obtained results using the GRU method and achieved an accuracy of 87.93% in the classification result.
[4] - N. Moustafa et al. proposed an architecture called DFSat to detect cyber attacks on satellite networks. As a result of the tests conducted with the proposed recurrent neural network an accuracy rate of 99.66% was obtained.
[5] - V. Gugueoth et al. Deep learning methods classification methods and federated learning methods were used to ensure the security of IOT objects. The results obtained from these methods were compared. The aim here is to find the learning model suitable for IOT networks with different features.
[6] - Z. Zeng et al. used an LSTM based method to detect anomalies on satellite telemetry data. Different data sets were used in the study. An accuracy rate of 93.13% was obtained for one of the data sets used, the VN10 data set.
[7] - E. Hodo et al. Artificial neural networks were used to determine unauthorized access and cyber attacks on IOT networks. A supervised artificial neural network and a multi level detector were used to detect anomalies. An accuracy value of 99.4% was obtained from the data set trained with the proposed method.
[8] - J. Jose et al. used deep neural network methods such as CNN and LSTM on CIC-IDS 2017 dataset to detect anomalies and attacks in IoT systems.
[9] - X. Zhang et al. proposed a CNN-based architecture to detect anomalies in networks. The proposed model was tested on NSl-KDD dataset and an accuracy of 83.31% was achieved.
[10] - R. Uddin et al. presented a distributed architecture implemented in SDN environment to prevent violations in satellite networks. The proposed architecture was based on OpenMined federated learning. With this preferred federated learning, an accuracy value of 79.47% was achieved.

In recent years there have been numerous studies carried out to ensure the security of satellite and IoT based networks. Some of these studies focus on preventing attacks, while others focus on detecting and classifying them. Various approaches such as deep learning, federated learning, encryption methods and statistical analysis have been used, mostly structures such as CNN, LSTM, GRU have been preferred. In some studies, data security has been provided with classical encryption algorithms, while in others, normal data has been directly processed and classified. However a large part of the reviewed literature does not present an effort to evaluate the performance of systems after the confidentiality level of the data has been increased, i.e., in encrypted or masked forms. Especially in critical processes such as attack detection and classification, the need to analyze the data with confidential methods instead of directly

processing it has been largely ignored in the literature. In infrastructures requiring high confidentiality such as satellite networks the need for approaches that will enable effective analysis of the data in its encrypted form draws attention at this point. Therefore, although existing studies provide high accuracy rates, an intrusion detection model that is fully integrated with privacy and security principles has not been developed. This deficiency has been tried to be addressed by our proposed encryption-assisted deep learning based method.

3. Proposed Method

A 5 step architecture is proposed to detect anomalies in satellite networks. The general flow chart of the proposed architecture is shown in Figure 1.

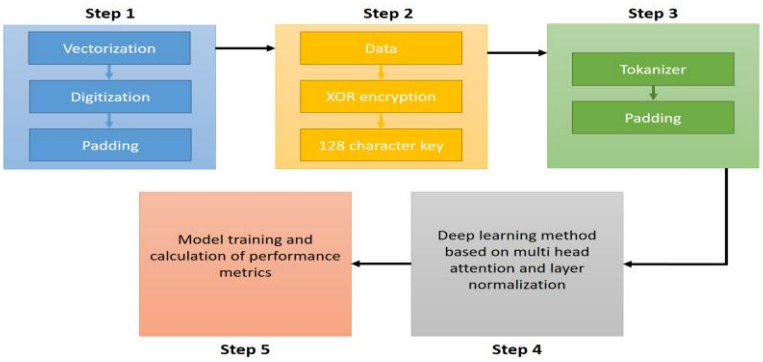


Figure 1. Proposed method steps

Step 1 - Making the Dataset Suitable for Training: In this study the UNSW-NB15 dataset was used. This dataset contains realistic network data created within the Australian Defence Force Academy [11-18]. This dataset is a comprehensive and balanced dataset that includes both normal and attack-based network traffic samples. The dataset used includes various cyber attack types such as DoS, Worms, Shellcode, Fuzzers, Exploits, Generic, Reconnaissance, Analysis, Backdoors and Injection. The UNSW-NB15 dataset is widely preferred in many studies such as network traffic analysis, classification of cyber threats and anomaly detection in the literature due to its various attack types. At this stage the network based data was not directly processed because it was very sensitive and important. The next step was to subject it to an XOR based encryption process.

Step 2 - XOR Based Encryption: As mentioned in the previous step the data transmitted in satellite networks consists of sensitive and important content. Therefore encryption, which is one of the basic components of security, is an important step in this study. In the literature advanced encryption methods such as AES and RSA are generally preferred. However, in this study XOR based encryption, which is a symmetric and lightweight method, was preferred. The main purpose of choosing an XOR based encryption is to hide the data in a way that does not compromise data integrity and to be able to classify it in its encrypted form. The XOR algorithm encrypts each character by subjecting it to a bit level XOR operation with a randomly generated 128 character key. The data obtained as a result of this process moves away from its raw form and turns into a completely different character sequence. Since the corresponding character remains the same length, the length of the sequences will be preserved and the encrypted data can later be given as input to the neural network. In this way the data can be analyzed with methods such as machine learning or deep learning even while it is encrypted.

Step 3 - Data Vectorization and Preprocessing: Data encrypted with XOR is not transferred directly to neural networks. For this the character sequences must first be converted to a numerical form. For all these reasons Tokenizer is used to represent all characters or words of the data with unique indexes. In other words the data to be used is subjected to the vectorization process. The numerical sequences created with the Tokenizer are aligned by zeroing them with the pad_sequences function in order to have a fixed length suitable for the neural network inputs. Thus, shorter sequences are completed with zeros while longer sequences are brought to a limited structure.

Step 4 - Design of Deep Learning Model: The proposed model for detecting anomalies in satellite networks is developed based on the Transformer method. The developed system is based on the multi head attention system. In the input layer of the model, vectorized data is passed through the Embedding layer. In this context, a richer semantic representation is obtained. The Embedding layer allows the model to learn relationships better by representing each character with a vector of a certain size. After this stage, the data is encoded. Each encoder block includes; multi head attention to be able to pay attention to different parts of the input sequence at the same time, layer normalization to stabilize the training process, feed forward network to make the attention results non-linear, and dropout layer to prevent overfitting. In the last stage, the GlobalAveragePooling layer is used to intensify the serial outputs. Finally, two Dense layers are added to the model as output layers. Since the data set here is multi-class, the Softmax activation function is used.

Step 5 - Training and Evaluation of the Model: During model training, 80% of the dataset was divided into training and 20% into test data. During the training phase the batch size was selected as 32 and the model was trained for 5 epochs. The Sparse Categorical Cross-Entropy function, which is suitable for multi-class classification, was used as the loss function. The Adam algorithm was used as the optimization function. After the training was completed, the model's overall accuracy, precision, recall, and F1 score based performance metrics were calculated on the test data.

4. Experimental Results

The model details of which are given in Section 3, was tested on the UNSW-NB15 dataset and its performance metrics were obtained. With the obtained performance metrics it is possible to have an idea about the success of the model from different perspectives. The accuracy, precision, Recall and F1score values of the proposed model are given in Table 1.

Table 1. Performance metrics results

Performance metrics	Percentage rate
Accuracy	92.66
Precision	93.80
Recall	87.86
F1 score	90.25

According to Table 1, the accuracy rate of the proposed model was measured as 92.66%. This value shows that the model can successfully distinguish different data classes thanks to its high success rate. In addition the precision value was obtained as 93.80%. This rate indicates how many of the positive examples predicted by the model are truly positive. In other words the model makes very high accuracy decisions when identifying examples belonging to the positive class. This rate is especially valuable in security-focused applications. The recall value was recorded as 87.86%. This rate shows how accurately the model detects truly positive examples. This obtained value also reveals that an effective performance is exhibited in detecting threats such as cyber attacks. Finally the F1 score was obtained as 90.25%. The F1 score is the harmonic mean of the sensitivity and sensibility values and shows the overall balance performance of the model. This high rate indicates that the model has a balanced performance in terms of both false positive and false negative classifications. As a result Table 1 reveals that the model can provide high accuracy and reliable classification even when working on encrypted data in systems that require secure communication.
A comparison table of studies conducted on satellite and network security is given in Table 2. The methods and results of the studies conducted are included in this table.

Table 2. Comparison with methods in the literature

Study Number	Method	Result
[3]	GRU based deep learning	% 87.93
[4]	RNN based DFSat architecture	% 99.66
[6]	LSTM based architecture	% 93.13
[10]	Federated learning based SDN architecture	% 79.47
Proposed architecture	Encryption + Transformer based DNN	% 92.66

The comparison results show the performance of the proposed method compared to the existing approaches in the literature. Different architectures were used in the reviewed studies. The accuracy rates obtained from these methods vary between 79.47% and 99.66%. While the proposed model increases data security with the encryption technique, it can effectively learn complex relationships in the data thanks to the Transformer based deep learning architecture. In this way 92.66% accuracy rate was obtained. This rate indicates a higher success rate than many studies in the literature. As a result it shows that the method offers a strong solution in terms of both security and attack detection. In addition, it clearly shows that the model can give successful results in real world applications with its stable and balanced attitude.

Table 3 is created to show the advantages and differences of the proposed method compared to the existing methods. The created table includes important features of the methods such as compatibility, confidentiality, accuracy and encryption.

Table 3. Advantages of the proposed method

Feature	Available methods	Proposed method
Data Encryption	Usually not	Yes (XOR based)
Privacy Compliance	Low / medium	High
Deep Learning Method	LSTM / CNN / GRU	Transformer + MHA
Applicability	Standard systems	Security-first systems
Accuracy	Between 79% and 93%	%92.66

According to Table 3, it is seen that the proposed method has superior features in various aspects compared to many studies in the existing literature. First of all in terms of data encryption feature, while a direct encryption approach is not used in the vast majority of existing methods, an XOR based encryption technique is used in the proposed model. The proposed system provides high level of security making it more suitable for use in systems such as satellite networks that contain sensitive data. Instead of approaches such as LSTM, CNN and GRU, which are frequently used in the literature as deep learning methods the Transformer and multi-head attention mechanism are used in the proposed architecture, allowing the model to learn the relationships between the data more effectively. This gives the model a higher accuracy rate. All these features show that the proposed method offers a strong alternative in terms of practical applications.

5. Conclusion

In this study a secure and effective model for detecting cyber attacks in satellite based IoT networks is presented. When the existing literature is examined, it is seen that many studies either focus only on attack detection or develop approaches that only provide data privacy. In order to fill this gap, in our study, both a XOR based encryption technique that increases data security is used and a Transformer-based deep learning architecture that provides high success in attack detection is developed. The success of the model is evaluated with performance metrics. As a result of the evaluation 92.66% Accuracy, 93.80% Precision, 87.86% Recall and 90.25% F1score values are obtained. These results show that the proposed method offers a more balanced and secure performance than existing approaches. In addition the model is quite useful in systems with security requirements, especially thanks to its high compatibility in data privacy. As a result this study makes a significant contribution to strengthening security in satellite based IoT networks.

6. References

[1] M. Usman, M. Qaraqe, M. Rizwan Asghar, and I. S. Ansari, “Mitigating distributed denial of service attacks in satellite networks,” 2020, doi: 10.1002/ett.3936.

[2] S. Lohani, R. J.-2020 I. C. on Emerging, and U. 2020, “Satellite network security,” *ieeexplore.ieee.orgS Lohani, R Joshi2020 Int. Conf. Emerg. Trends Commun. 2020•ieeexplore.ieee.org*, Accessed: Dec. 14, 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9117553/

[3] N. Koroniotis, N. Moustafa, and J. Slay, "A new Intelligent Satellite Deep Learning Network Forensic framework for smart satellite networks," *Comput. Electr. Eng.*, vol. 99, p. 107745, Apr. 2022, doi: 10.1016/J.COMPELECENG.2022.107745.

[4] N. Moustafa, I. Khan, … M. H.-I. T., and undefined 2022, "DFSat: Deep Federated Learning for Identifying Cyber Threats in IoT-based Satellite Networks," *ieeexplore.ieee.orgN Moustafa, IA Khan, M Hassanin, D Ormrod, D Pi, I Razzak, J SlayIEEE Trans. Ind. Informatics, 2022•ieeexplore.ieee.org*, Accessed: Dec. 04, 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9925589/?casa_token=ZxMM81oBNcMAAAAA:_n8WMMSdhvNFSA4IccNtnTeEQD0R3tlCJr8gHrxfkHhv_DQEDeByEv8sjN-TGAkpjAkwVEe0Pcw

[5] V. Gugueoth, S. Safavat, S. S.-I. Express, and undefined 2023, "Security of Internet of Things (IoT) using federated learning and deep learning-Recent advancements, issues and prospects," *Elsevier*, Accessed: Dec. 04, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405959523000334

[6] Z. Zeng, G. Jin, C. Xu, S. Chen, … Z. Z.-I. T. on, and undefined 2022, "Satellite telemetry data anomaly detection using causal network and feature-attention-based lstm," *ieeexplore.ieee.orgZ Zeng, G Jin, C Xu, S Chen, Z Zeng, L ZhangIEEE Trans. Instrum. Meas. 2022•ieeexplore.ieee.org*, Accessed: Dec. 14, 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9716063/?casa_token=731fbRqKAcoAAAAA:cANhUM2KyI2FWizzen7sv4vQLpg0hszMugKB5z4h4makInH3TTSeZ13A3Wo25XLwiCsbF_5mMLg

[7] E. Hodo *et al.*, "Threat analysis of IoT networks using artificial neural network intrusion detection system," *2016 Int. Symp. Networks, Comput. Commun. ISNCC 2016*, Nov. 2016, doi: 10.1109/ISNCC.2016.7746067.

[8] J. Jose and D. V Jose, "Deep learning algorithms for intrusion detection systems in internet of things using CIC-IDS 2017 dataset," *Int. J. Electr. Comput. Eng.*, vol. 13, no. 1, pp. 1134–1141, 2023, doi: 10.11591/ijece.v13i1.pp1134-1141.

[9] X. Zhang, J. Ran, and J. Mi, "An Intrusion Detection System Based on Convolutional Neural Network for Imbalanced Network Traffic," *Proc. IEEE 7th Int. Conf. Comput. Sci. Netw. Technol. ICCSNT 2019*, pp. 456–460, Oct. 2019, doi: 10.1109/ICCSNT47585.2019.8962490.

[10] R. Uddin and S. A. P. Kumar, "SDN-Based Federated Learning Approach for Satellite-IoT Framework to Enhance Data Security and Privacy in Space Communication," *IEEE J. Radio Freq. Identif.*, vol. 7, pp. 424–440, 2023, doi: 10.1109/JRFID.2023.3279329.

[11] N. Moustafa, B. Turnbull, and K. K. R. Choo, "An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4815–4830, Jun. 2019, doi: 10.1109/JIOT.2018.2871719.

[12] N. Koroniotis, N. Moustafa, E. Sitnikova, and J. Slay, "Towards Developing Network Forensic Mechanism for Botnet Activities in the IoT Based on Machine Learning Techniques," *Lect. Notes Inst. Comput. Sci. Soc. Telecommun. Eng. LNICST*, vol. 235, pp. 30–44, 2018, doi: 10.1007/978-3-319-90775-8_3.

[13] N. Moustafa, G. Misra, and J. Slay, "Generalized Outlier Gaussian Mixture Technique Based on Automated Association Features for Simulating and Detecting Web Application Attacks," *IEEE Trans. Sustain. Comput.*, vol. 6, no. 2, pp. 245–256, Feb. 2018, doi: 10.1109/TSUSC.2018.2808430.

[14] M. Keshk, N. Moustafa, E. Sitnikova, and G. Creech, "Privacy preservation intrusion detection technique for SCADA systems," *2017 Mil. Commun. Inf. Syst. Conf. MilCIS 2017 - Proc.*, vol. 2017-December, pp. 1–6, Dec. 2017, doi: 10.1109/MILCIS.2017.8190422.

[15] N. Moustafa, E. Adi, B. Turnbull, and J. Hu, "A New Threat Intelligence Scheme for Safeguarding Industry 4.0 Systems," *IEEE Access*, vol. 6, pp. 32910–32924, Jun. 2018, doi: 10.1109/ACCESS.2018.2844794.

[16] N. Moustafa, G. Creech, and J. Slay, "Anomaly Detection System Using Beta Mixture Models and Outlier Detection," *Adv. Intell. Syst. Comput.*, vol. 710, pp. 125–135, 2018, doi: 10.1007/978-981-10-7871-2_13.

[17] N. Moustafa, G. Creech, and J. Slay, "Flow Aggregator Module for Analysing Network Traffic," *Adv. Intell. Syst. Comput.*, vol. 710, pp. 19–29, 2018, doi: 10.1007/978-981-10-7871-2_3.

[18] N. Moustafa and J. Slay, "A Network Forensic Scheme Using Correntropy-Variation for Attack Detection," *IFIP Adv. Inf. Commun. Technol.*, vol. 532, pp. 225–239, 2018, doi: 10.1007/978-3-319-99277-8_13.

Blockchain Based Method for Secure Data Storage In Digital Identity Systems

Nida CANPOLAT[1] , Hasan YETIS [2], Mehmet KARAKOSE [3]

[1]Department of Digital Forensic Engineering, Faculty of Technology, Firat University, Turkey
[2,3]Department of Computer Engineering, Faculty of Engineering, Firat University, Elazig, Turkey.
[1]231144110@firat.edu.tr, [2]h.yetis@firat.edu.tr, [3]mkarakose@firat.edu.tr
[1](ORCID: 0000-0002-5262-1509), [2](ORCID: 0000-0001-7608-3293),[3](ORCID: 0000-0002-3276-3788)

Abstract
In the digital world security is becoming increasingly critical due to the vulnerability of personal data to cyber attacks. Encrypting data securely provides an effective solution against these threats. Strong encryption algorithms, especially RSA, ensure that only authorized users have access to data. However encryption processes can often be complex and difficult for users to understand. In this project, a user friendly system has been developed where users can securely encrypt and decrypt their passwords. Thanks to key management and export features, users can manage encryption processes in a more controlled and secure manner. Even people without technical knowledge can easily perform these operations thanks to the user friendly interface. In addition private and public keys are stored securely and can be exported when necessary. Blockchain technology has been integrated to further strengthen data security. The decentralized structure of the blockchain allows data to be recorded without modification, thus making encryption keys and encrypted data more secure. Every transaction becomes transparent and verifiable. Key pairs and encrypted data recorded on the blockchain provide control over the history of data. In addition this system is designed to be accessible only with the correct private keys. As a result, blockchain integration reinforces security in the digital environment by increasing the security of encryption and key management processes.

Keywords: blockchain, RSA, encryption, decryption, identity protection

1. Introduction

In the digital age, it is increasingly important for individuals to store their credentials and passwords securely. In traditional password management systems, passwords are stored in central databases, which pose various security threats. Major data breaches and cyber attacks in recent years have increased security concerns about these centralized structures. This has led to decentralized systems and technologies such as blockchain coming to the forefront as more secure identity management solutions. With the proliferation of digital platforms the need for users to securely store and manage their credentials and passwords is increasing day by day. Since users perform various transactions on multiple platforms they use different passwords for each platform. It is of great importance to store these passwords securely. Traditional password management systems usually store passwords in databases connected to a central server, and master passwords are often used to protect passwords. However, centralized systems are vulnerable to cyber attacks, which can lead to major data breaches. Security vulnerabilities and data leaks especially in recent years, have questioned the security of centralized systems and increased the need for alternative security solutions. In traditional encryption methods user passwords are usually stored and managed on central servers. However if malicious attackers infiltrate these systems, theft of passwords can also put users' accounts on other platforms at risk. Moreover although users are constantly advised to create strong and different passwords, remembering and storing these passwords safely poses a major challenge. Therefore the need for decentralized and more secure solutions is increasing.

Today with the rapid increase in digitalization, individuals and institutions store, share and process more data than ever on online platforms. While this digital transformation brings with it great opportunities, it also brings critical issues such as data security and privacy to the agenda. In particular the seizure of personal and sensitive data by unauthorized persons brings with it serious threats such as identity theft, information leakage and financial fraud. For this reason secure data transmission and storage in the digital environment has become one of the most important problems faced by modern information societies.
One of the most effective ways to ensure data security is encryption methods. In particular public key encryption algorithms allow data to be accessible only to authorized persons. The RSA algorithm is one of the most common and reliable methods in this area. RSA works with public and private key pairs created specifically for the user. Thus it provides high security when encrypting and decrypting data. However, in traditional RSA systems, processes such as key management, secure storage and data verification can be complicated for users and can cause security vulnerabilities. At this point Blockchain

[1]Corresponding author

technology offers a new perspective. Blockchain increases security by providing a decentralized, unchangeable and distributed data structure. Thanks to the fact that data is kept in a chain and in cryptographically interconnected blocks, the history of any data can be followed transparently and it becomes almost impossible to change it later. Thanks to these features, blockchain technology is not limited to cryptocurrencies, but is used for secure information sharing in many areas such as identity verification, health data management and smart contracts.

In this project a hybrid system has been developed that combines the RSA encryption method with a user friendly interface, while also using the secure recording and verification opportunities offered by blockchain technology. The aim here is to reduce the need for technical knowledge and enable both individuals and institutions to perform encryption and decryption operations securely and easily, while also increasing the security of key management and data history. Users can access a transparent, verifiable and indelible data security layer by recording the encrypted data they create in the system on the blockchain. As a result the main motivation of this project is to provide an effective solution against increasing digital threats and to reinforce the reliability of encryption technologies with the blockchain. With this approach the concerns of individuals and institutions about data security in the digital environment are reduced; a secure, sustainable and user oriented system design is made possible. In this direction, the system developed aims to be one of the cornerstones of secure digital communication in the future.

2. Literature Review

This section includes a literature review on digital identity and personal information storage.
Md Abdullah Al Mamun and his colleagues presented an approach for a Blockchain-based digital identity system [1]. In this study, a system on Blockchain based digital identity for individuals using bio information is proposed. The proposed digital identity system is implemented using Ethereum smart contract. The aim of this study is to ensure the security of user information. May Htet and his colleagues conducted a study on a Blockchain based digital identity management system [2]. This study addresses the possible use of Blockchain technology for a digital identity management system to solve the passport problem in Myanmar. With the study all passport offices will be connected in a decentralized environment and sufficient passport information of a person can be shared securely among peers. Thus passports cannot be shared illegally. Mehmet Aydar and his colleagues conducted a study on investigating the inefficiencies of traditional identity systems [3]. It was investigated and discussed how the appropriate implementation of Blockchain technology would result in more secure, safe, privacy friendly and remotely accessible identity systems. Francesco Buccafurri and his colleagues conducted a study on integrating digital identity structure into Blockchain technology [4]. The existing problem was addressed by proposing a solution that combines the public digital identity mechanism with Blockchain through identity based encryption. Rogelio Rivera and his colleagues conducted a study on how digital identity can contribute to the smart city environment [5]. In the study a systematic mapping review was conducted to collect all the relevant existing research on digital identity on Blockchain technology applied in the smart city environment. Eranga Bandara and his colleagues conducted a study on Blockchain and self-sovereign identity supported digital identity platform [6]. In this study, a digital identity platform based on Blockchain and self-sovereign identity was proposed to address the problems in centralized identity systems. The proposed method is named "Casper". This application is a mobile identity wallet application based on both Android and IOS that combines Blockchain integration and self sovereign identity based approach. Jagger Bellagarda and his colleagues conducted a study based on NFT to eliminate fraud risks [7]. A Blockchain enabled NFT application is proposed to reduce fraud and provide authenticated digital identity. A system that combines NFTs and verifiable credentials to store the underlying NFT digital assets in a decentralized manner will be discussed for the existing fundamental problems. As a result, this study aims to eliminate the risks of fraud. Xiaohui Yang et al. conducted a study on a zero-knowledge proof based digital identity management scheme in Blockchain technology [8]. This study utilized smart contracts and zero-knowledge proof algorithms to improve the existing blockchain identity claim model to realize identity decoupling. In addition, it is planned to effectively prevent the decryption of attribute ownership. Andrej J. Zwitter et al. conducted a study on universal identity management and the concept of the “self sovereign” individual [9]. This study aims to contextualize recent developments and provide a categorization to frame the general landscape as developments rapidly progress. Guillermo Balastegui García et al. conducted a study on Blockchain based corporate digital identity [10]. This study aims to propose and develop a digital identity for legal entities based on a blockchain network. For this, a methodology based on implementing the same idea of personal digital identity is used.

Additionally the proposal takes into account the Ethereum blockchain, as it is strongly accepted for the deployment of smart contracts.

Although there have been many studies on Blockchain-based digital identity systems in the current literature, these studies generally remain at a conceptual level and do not provide integrated solutions that allow users to manage their passwords on different platforms in a secure, decentralized and user friendly manner. Therefore, an application that both increases security and provides practical use is needed. This project aims to fill this gap and provide a secure digital identity and password management solution supported by Blockchain.

3. Proposed Method

In this study a graphical user interface based application based on the RSA algorithm has been developed in order to securely encrypt, store, decrypt user passwords and perform key management in a reliable manner. The developed system generates a special public and private key pair for each user based on the asymmetric key cryptography widely used today and provides password security through this infrastructure. The application is equipped with a user friendly graphical interface to enable users to easily perform encryption and decryption operations without requiring technical knowledge. The system enables passwords to be encrypted with the RSA algorithm and this encrypted data to be securely stored and decrypted only by the user when necessary. In addition the application can easily import and export key pairs of users. Thus, the key management process can be carried out flexibly and securely. This developed system is not limited to encryption and decryption operations only but also meets critical security requirements such as data integrity, confidentiality and access control; it offers a practical, accessible and powerful solution for secure password management. The visualization of the proposed method is given in Figure 1.

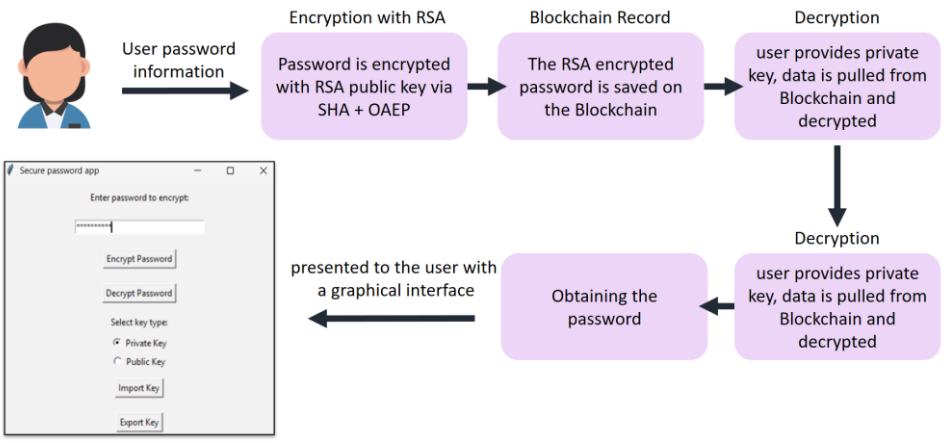


Figure 1. Architecture of the proposed method

Step 1- Determining the username with RSA:
In this step the password information entered by the user is secured with the RSA method, which is a symmetric encryption method [11]. The RSA algorithm has been used to ensure security in many studies [12-14]. In this process, a 2048 bit key pair is generated using the RSA algorithm. This pair consists of two keys, one public and the other private. The user's password information is encrypted with the public key used only for encryption purposes, and the SHA 256 algorithm is used together with the asymmetric encryption padding method OAEP to increase security during this process. In this study OAEP is used as a modern padding method that makes it difficult to break the RSA algorithm [15-17]. The encrypted data is saved to the file system and made ready to proceed to the next step. Thanks to this step, the user's password is protected with strong cryptographic methods so that no third party can access it.

Step 2 – Ensuring Blockchain Integration:
In this step the user password information encrypted with RSA in the previous step is saved to the file system as a file with the extension .bin. This file is later used to integrate with an external Blockchain network. This process, which is performed

with the “save_to_file” function in the application, creates the infrastructure for transferring data to the Blockchain network. Since Blockchain technology ensures that data is stored in an unchangeable and distributed structure, the security of encrypted data is further increased. Since only encrypted data is kept in each block, data confidentiality is protected, while the integrity of the chain prevents unauthorized changes to the data. This structure also increases the continuity of the system and its resistance to attacks without being dependent on a central server. This step of the application provides data production ready for Blockchain integration.

Step 3 – Performing the decryption process:
In order for the user to access the data that they have previously encrypted and stored in the system again, the decryption process is required. This process can only be performed by authorized users who have the private key. In this process, which is carried out through the application, the user first selects the encrypted data and then loads the private key with the .pem extension from the system. While the private key is imported with the “import_key” function in the application, the “decrypt_password” function performs the decryption with the reverse operation of the RSA algorithm. The OAEP scheme and the SHA-256 algorithm are also used to ensure security in this step. When the decryption process is successfully completed, the original version of the password is presented to the user securely through the application interface. Thanks to this structure, only the authorized user can access the data, while data confidentiality and integrity are protected.

Step 4 – Ensuring practicality of use with a user friendly interface:
One of the most important components of the project is that technical operations are carried out through a simple and understandable interface. To ensure this, all operations are presented through a user-friendly graphical interface. This interface, developed for user convenience, allows users to perform encryption, decryption and key generation operations without the need for any technical knowledge. The user can easily create key pairs and start encryption and decryption operations through the buttons on the main screen of the application. During each operation, the user is guided with file selection windows and informative messages. The interface increases ease of use by providing instant feedback on operations to the user.

4. Experimental Results

The graphical user interface obtained as a result of the application of the method detailed in section 3 is presented in Figure 2. This user friendly and functional interface works in integration with the RSA based asymmetric cryptography infrastructure and allows users to securely encrypt and decrypt their passwords. Using the password entry field on the interface, the user can enter the text to be encrypted and use the “Encrypt Password” button; similarly, they can convert previously encrypted data to its original form with the “Decrypt Password” button. The system supports both private and public key management; users can easily perform key creation, import and export operations. The options provided for key type selection increase the security of the encryption process by ensuring that the correct key is used in the operations to be performed. All these operations are performed through a simple graphical interface, thus minimizing the need for technical knowledge and allowing the system to be used effectively by users at all levels.

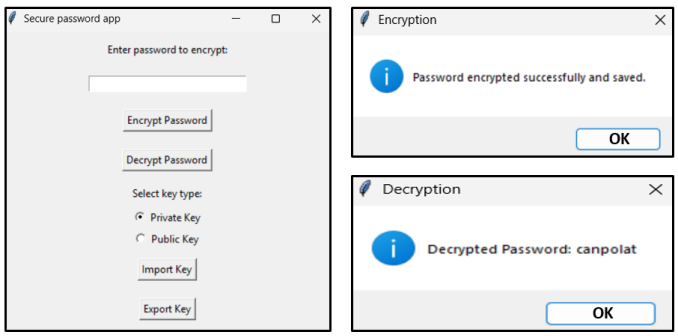


Figure 2. User interface outputs

In this study, a Blockchain based structure was preferred instead of classical centralized systems to store data in the proposed system. The main reason for this is the security, immutability and transparency advantages offered by blockchain. Since data is stored on a single server in classical centralized systems, this center can be targeted and security vulnerabilities such as data modification, deletion or access blocking can occur. In addition access continuity may be interrupted due to connection problems in these systems and users' data history tracking capabilities are limited. In contrast blockchain technology eliminates the centralization problem by allowing data to be stored in a distributed manner on many nodes and enables the system to continue operating even if any node in the network is disabled. Thanks to the immutable structure of blockchain, no manipulation can be made on the data recorded once, and this guarantees data integrity. In addition, since all transactions performed in the system are recorded transparently and chronologically on the network, data movements provide full traceability. For this reason, blockchain-based data storage method was preferred in the proposed system to maximize the security of user passwords. The comparison of blockchain and classical centralized systems is given in Table 1.

Table 1. Comparison of data storage methods

Feature	Classic central system	Proposed Blockchain based system
Data Modifiability	Possible	Impossible
Attack Resistance	Low Probability	High
Continuity of Access	The Connection May Be Interrupted	Uninterrupted
Traceability	Limited	Full Traceability

As seen in Table 1, it clearly reveals the fundamental differences between classical centralized systems and the proposed Blockchain based system. Traditional centralized structures are open to change due to the fact that data is kept in a single center, and the integrity of the system is weaker against attacks. These structures allow data to be easily changed in case a malicious user accesses the central server. In addition in these systems, access continuity cannot be provided in cases such as connection breaks and there is no possibility of full tracking of the history of the data. However, the proposed Blockchain-based system makes it technically impossible to change the data thanks to its distributed structure. In this way, it provides a high level of data integrity and security. Since every transaction on the Blockchain is recorded encrypted and time dependent, full traceability is provided. Each step taken in the system can later be verified. In addition, thanks to the system working over multiple nodes, access is not interrupted even in cases such as connection breaks, and high access continuity is maintained. Considering these aspects, the proposed system has a structure that is superior to classical centralized systems in terms of data security, transparency and sustainability.

In this study RSA is preferred as the encryption algorithm, OAEP is preferred for padding and SHA 256 is preferred to obtain the hash value. In order to clearly demonstrate the advantages of the proposed method, the basic features of various popular algorithms are presented comparatively in Table 2. Key length, padding method and used hash function criteria are included to compare the algorithms. The proposed system is designed by using the RSA algorithm and the OAEP method together. The OAEP method is a padding method generally used with RSA [15]. This method prevents information leakage by adding randomness to the system and provides higher resistance to attacks. In addition, both data integrity and security levels are significantly increased by using the SHA-256 hash function. SHA-256 is a strong and reliable algorithm that is widely preferred in the field of cryptography today [18-20].

Table 2. Comparison of popular encryption techniques
(PKCS = Public-Key Cryptography Standards, ECDSA = Elliptic Curve Digital Signature Algorithm)

Algorithm	Key Length	Padding Process	Hash Function Used
Recommended system (RSA + OAEP)	2048 bit	OAEP	SHA-256
RSA	2048 bit	PKCS	SHA-1
AES-128	128 bit	None	AES-128
ECC (P-256)	256 bit	ECDSA	SHA-2

Considering the technical specifications given in Table 2, the reasons for preferring the proposed system include providing a strong and reliable structure that is suitable for current security needs. The support of the RSA algorithm used in the proposed method with OAEP padding and SHA 256 hash function provides much higher security compared to traditional methods. The use of a 2048 bit long key in the proposed method also provides strong encryption to the system.

In this project, OAEP offers critical advantages, especially in terms of password security, with its ability to prevent information leakage during the randomization and decryption process. In addition, the fact that the system supports both public and private key management increases its flexibility of use and allows users to easily encrypt and decrypt their data. With these aspects, the proposed structure presents a strong and sustainable solution in terms of both technical adequacy and user security. Another reason why the OAEP padding method was preferred in this study is that it eliminates the security vulnerabilities brought by deterministic encryption by ensuring that the same data is encrypted differently each time. Unlike older methods such as PKCS, OAEP is more resistant to chosen ciphertext attacks and provides a cryptographically stronger security foundation. In this way, the system has become more modern, secure and resistant to attacks.

As a result the proposed RSA + OAEP based system provides higher security compared to other methods with both key length, advanced padding method and strong hash function. This structure is very effective in protecting data confidentiality and integrity, especially in critical areas such as password security.

In this study, instead of classical or singular methods, a multi featured method is proposed. The proposed method provides user specific key management, Blockchain-based data storage method, strong encryption security with RSA + OAEP + SHA 256, and easy access with a simple interface that the user can easily use. In order to understand the advantages of the proposed method more clearly, a comparison of the proposed method and other systems used is given in Table 3.

Table 3. Comparison of other systems and the proposed system

Feature	Recommended System	Classic RSA Systems	AES + Central Systems
Key Management	User specific	Shared	Centralized key management
Data Storage	Blockchain based	External file system	Central server
Encryption Security	RSA + OAEP + SHA-256	RSA + PKCS	AES-128
Access Control	Only with private key	Usually the user password	Standard login
User Interface	Graphical interface	Command line or not	Web panel

Considering the methods compared in Table 3, it is seen that the proposed system offers significant advantages in terms of both security and user experience. While classical RSA systems generally use shared key structure and old padding techniques such as PKCS, the system proposed in this study offers private key management for each user and increases security with RSA encryption supported by OAEP. Symmetric encryption methods such as AES are more vulnerable to attack because they usually run on central servers and require central key management. On the other hand, the proposed system makes it impossible to change the data by storing it in a blockchain based structure and provides full traceability. In addition, it increases the accessibility and practicality of the system with the graphical user interface developed so that users can perform transactions without needing technical knowledge. When all these features come together, the proposed system offers a much more functional and sustainable solution compared to classical methods with its secure and user friendly structure.

4. Conclusion

In this study, a blockchain based password encryption and storage system supported by RSA based, OAEP padding method and SHA 256 hash function has been developed in order to increase the security of user passwords. The developed system has a significant advantage over classical centralized systems with its distributed structure, high security level, unchangeable data storage capacity and user-friendly interface. The experimental analysis results and comparative tables clearly showed that the proposed structure has significant advantages in terms of both data security and user experience. Especially thanks to the blockchain integration, the system can protect the integrity of encrypted data, make key

management independent from central authorities and provide user-specific access control. As a result, this study offers an innovative approach that addresses password security in accordance with modern needs with its both technically robust and user-friendly structure.

Acknowledgements

This study was supported by Scientific and Technological Research Council of Türkiye (TUBITAK) under Grant Number 124E001. The authors thank TUBITAK for their support.

5. References

[1] M. A. Al Mamun, S. M. M. Alam, M. S. Hossain, and M. Samiruzzaman, "A Novel Approach to Blockchain-Based Digital Identity System," *Adv. Intell. Syst. Comput.*, vol. 1129 AISC, pp. 93–112, 2020, doi: 10.1007/978-3-030-39445-5_9.

[2] M. Htet, P. T. Yee, and J. R. Rajasekera, "Blockchain based Digital Identity Management System: A Case Study of Myanmar," *Proc. 4th Int. Conf. Adv. Inf. Technol. ICAIT 2020*, pp. 42–47, Nov. 2020, doi: 10.1109/ICAIT51105.2020.9261785.

[3] M. Aydar, S. Ayvaz, and S. C. Cetin, "Towards a Blockchain based digital identity verification, record attestation and record sharing system," Jun. 2019, Accessed: Mar. 12, 2025. [Online]. Available: http://arxiv.org/abs/1906.09791

[4] F. Buccafurri, G. Lax, A. Russo, and G. Zunino, "Integrating Digital Identity and Blockchain," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11229 LNCS, pp. 568–585, 2018, doi: 10.1007/978-3-030-02610-3_32.

[5] R. Rivera, J. G. Robledo, V. M. Larios, and J. M. Avalos, "How digital identity on blockchain can contribute in a smart city environment," *2017 Int. Smart Cities Conf. ISC2 2017*, Oct. 2017, doi: 10.1109/ISC2.2017.8090839.

[6] E. Bandara, X. Liang, P. Foytik, S. Shetty, and K. De Zoysa, "A Blockchain and Self-Sovereign Identity Empowered Digital Identity Platform," *Proc. - Int. Conf. Comput. Commun. Networks, ICCCN*, vol. 2021-July, Jul. 2021, doi: 10.1109/ICCCN52240.2021.9522184.

[7] J. Bellagarda and A. M. Abu-Mahfouz, "Connect2NFT: A Web-Based, Blockchain Enabled NFT Application with the Aim of Reducing Fraud and Ensuring Authenticated Social, Non-Human Verified Digital Identity," *Mathematics*, vol. 10, no. 21, Nov. 2022, doi: 10.3390/MATH10213934.

[8] X. Yang and W. Li, "A zero-knowledge-proof-based digital identity management scheme in blockchain," *Comput. Secur.*, vol. 99, Dec. 2020, doi: 10.1016/J.COSE.2020.102050.

[9] A. J. Zwitter, O. J. Gstrein, and E. Yap, "Digital Identity and the Blockchain: Universal Identity Management and the Concept of the 'Self-Sovereign' Individual," *Front. Blockchain*, vol. 3, 2020, doi: 10.3389/FBLOC.2020.00026/FULL.

[10] G. Balastegui-García, E. M. S. Sabau, A. S. Payá, and H. Mora, "Corporate Digital Identity Based on Blockchain," *Springer Proc. Complex.*, pp. 645–655, 2023, doi: 10.1007/978-3-031-19560-0_55.

[11] X. Zhou and X. Tang, "Research and implementation of RSA algorithm for encryption and decryption," *Proc. 6th Int. Forum Strateg. Technol. IFOST 2011*, vol. 2, pp. 1118–1121, 2011, doi: 10.1109/IFOST.2011.6021216.

[12] J. Wu, Y. Long, Q. Huang, and W. Wang, "Design and Application of IBE Email Encryption Based on Pseudo RSA Certificate," *2016 12th Int. Conf. Comput. Intell. Secur.*, pp. 282–286, 2017, doi: 10.1109/cis.2016.0071.

[13] P. M. Sonsare and S. Sapkal, "Stegano-CryptoSystem for Enhancing Biometric-Feature Security with RSA," *Int. Conf. Inf. Netw. Technol. IPCSIT 2011*, vol. 4, no. January, pp. 196–200, 2011, [Online]. Available: https://www.researchgate.net/publication/228528402

[14] G. Singh, "A Study of Encryption Algorithms (RSA, DES, 3DES and AES) for Information Security," *Int. J. Comput. Appl.*, vol. 67, no. 19, pp. 975–8887, 2013.

[15] Y. Zhong, "An Overview of RSA and OAEP Padding," *Highlights Sci. Eng. Technol.*, vol. 1, pp. 82–86, Jun. 2022, doi: 10.54097/HSET.V1I.431.

[16] B. Chevallier-Mames, D. H. Phan, and D. Pointcheval, "Optimal Asymmetric Encryption and Signature Paddings," *Lect. Notes Comput. Sci.*, vol. 3531, pp. 254–268, 2005, doi: 10.1007/11496137_18.

[17] D. Pointcheval, “OAEP: Optimal Asymmetric Encryption Padding,” *Encycl. Cryptogr. Secur.*, pp. 443–445, Sep. 2005, doi: 10.1007/0-387-23483-7_284.
[18] R. S. Jha, S. Umar, T. Kossi, and O. M. Lamine, “Token Bases Valid and Secure Payment System Using SHA-256,” *Commun. Comput. Inf. Sci.*, vol. 1738 CCIS, pp. 29–38, 2022, doi: 10.1007/978-3-031-23724-9_3/FIGURES/5.
[19] M. L. L. Sai Maneesha, V. D. Nagesh, N. G. Mahalakshmi, V. S. Yamini, and M. Srinivas, “CERTIFICATE VALIDATION WITH SHA,” *Int. J. Eng. Sci. Adv. Technol.*, 2023.
[20] I. Mahfud, P. H. Utomo, and B. Winarno, “Implementation of Rivest-Shamir-Adleman (RSA) Cryptography System with Secure Hash Algorithm (SHA-256) on REST API Authentication Mechanism,” *AIP Conf. Proc.*, vol. 3285, no. 1, Mar. 2025, doi: 10.1063/5.0262777/3340990.

Sentiment Analysis on Book Reviews Using Transformer-based Approaches

Efe Batın SEÇKİN[1] , Mustafa Berkay DEMİR [2], Mehmet ARZU [3] , Mahmut KAYA [4]

[1,2,3,4] Artificial Intelligence and Data Engineering Department, Fırat University, Elazığ, Türkiye.
230291011@gmail.com, 230291022@gmail.com, [3]marzu@firat.edu.tr, [4]mahmutkaya@firat.edu.tr
[1](ORCID: 0009-0001-9235-1718), [2](ORCID: 0009-0001-2472-6071),[3](ORCID: 0000-0001-6610-2788), [4](ORCID: 0000-0002-7846-1769)

Abstract
In this study, a balanced dataset for multi-class sentiment analysis was prepared using user reviews obtained from online book review platforms. The dataset was created manually and contains 4687 examples. The labels are organized into three classes: 1,584 positive, 1,520 negative, and 1,583 neutral. In the study, three different Transformer-based models (RoBERTa, ELECTRA, and Vanilla BERT) were fine-tuned and their performances were compared. According to the experimental results, the RoBERTa model showed the highest performance with an accuracy rate of 88.38%. The ELECTRA model achieved similar success with an accuracy rate of 88%, while Vanilla BERT showed lower performance with an accuracy rate of 78%. These findings reveal that Transformer-based models offer higher success in sentiment analysis tasks compared to traditional methods.

Keywords: sentiment analysis, natural language processing, transformer models, RoBERTa

1. Introduction

Sentiment analysis is an important technique that aims to determine positive, negative, or neutral emotional tendencies by examining text-based data using natural language processing methods. This method is widely used in the evaluation of large-scale text data such as social media posts, customer feedback, product reviews, and news comments. Sentiment analysis plays a critical role not only in commercial fields but also in various disciplines such as public opinion research and the study of social events. Businesses can use insights gained from such analyses to make decisions aimed at increasing customer satisfaction, improving service quality, and shaping marketing strategies more effectively [1].

In recent years, significant advancements have been made in the field of sentiment analysis with the widespread adoption of deep learning methods. Deep learning algorithms have significantly improved the accuracy and reliability of sentiment analysis systems by leveraging their ability to learn complex patterns and contextual relationships within large datasets [2]. Initial studies based on word lists or statistical methods have gradually yielded much more successful results with the introduction of machine learning algorithms (Decision Trees, Support Vector Machines, Naive Bayes, etc.) and later Transformer-based deep learning models (BERT, ELECTRA, RoBERTa, etc.). Transformer architectures provide a distinct advantage, particularly in complex language structures, due to their ability to learn the semantic relationships between words within a text without losing context.

In this study, sentiment analysis was performed on user reviews obtained from online book review platforms; the effects of data preprocessing steps on model performance were examined. Additionally, the comparative performance of Transformer-based methods with traditional machine learning algorithms was evaluated. One of the original aspects of the study is the use of a manually prepared dataset with a balanced class distribution. It is believed that this dataset, due to its inclusion of long and literary language structures, provides a meaningful basis for measuring the performance of models on more complex text structures. Thus, the advantages and limitations of the methods used in the field of sentiment analysis are more clearly revealed.

2. Related Studies

Numerous studies have been conducted in the field of sentiment analysis using different data sets and methods. This section summarizes a few of these studies.

[1]Corresponding author

Shehu [3] performed sentiment analysis on Turkish Twitter data using a polarity lexicon and artificial intelligence methods. The study used supervised and unsupervised learning approaches as well as the SVM algorithm. Özkul [4] examined perceptions of intelligent transportation systems using text mining methods and applied sentiment analysis. This study used the KNN algorithm, SVM, fuzzy logic, and nonlinear support vector machines. Özdeş [5] performed sentiment analysis using big data tools and applied Decision Trees, Logistic Regression, and Naive Bayes algorithms. Demir and Yılmaz [6] examined the effect of extracurricular educational activities on students' attitudes towards science and technology using sentiment analysis methods. Participants were manually selected, a manual dataset was created, and machine learning methods were then applied. Saka and Cömert [7] performed sentiment analysis on airline customer reviews using neural network language models and a universal sentence encoder. In the study, the CNN-LSTM model was trained from scratch; additionally, transfer learning approaches based on NNLM and USE were tested. The highest accuracy rate of 95.63% was achieved with the USE model. Budak [8] examined social media comments and scores of airline companies using data and text mining methods; he used SVM, Naive Bayes, and deep learning algorithms. Uslu and Özmen [9] classified Turkish news texts using machine learning methods; they compared SVM, Random Forest, and Naive Bayes algorithms and obtained the highest accuracy rate of 91% with the Naive Bayes model. Tuzcu [10] classified online user comments using MLP, Naive Bayes, SVM, and logistic regression algorithms. In experiments conducted with Python and RapidMiner tools, MLP yielded the most successful results with an accuracy rate of 89%. Deniz and colleagues [11] classified scientific article abstracts using the kNN algorithm and achieved an accuracy rate of 96.67%.

As seen in these examples, different methods are used in sentiment analysis studies. This study, however, only addresses Transformer-based models and aims to highlight their superior aspects by comparing their performance in sentiment analysis.

3. Methods and Materials

This section details the steps involved in preparing the dataset, training the models, and comparing the models. Presenting the process in a detailed and systematic manner is important for the reliability of the study.

3.1. Data Set Preparation

During the data set creation phase, data collection was first performed, followed by data cleaning and data preprocessing steps. Google Colab was used as the compilation environment for the study. In the initial phase, various web scraping methods were attempted, but no successful results were obtained. Subsequently, manual data extraction was performed using JavaScript via the browser console.

Since some of the collected data was unreadable or written in languages other than English, such data was filtered out during the data cleaning phase. The langdetect library was used for this process. In the next step, punctuation marks were removed as part of data preprocessing, words were separated according to their indices, and the data was made suitable for model training. The resulting dataset contains 4,687 rows. The dataset consists of three columns:

- **reviews:** user reviews,
- **label:** positive, negative, or neutral classification labels,
- **labels**: numerical equivalents of these labels.

The class distribution of the prepared data set is presented in Table 1. As can be seen, there is a balanced distribution among positive, negative, and neutral classes. This balance is important in preventing models from overfitting to a particular class during training.

Table 4. Class-based distribution of the dataset

Label	Number of samples
positive	1584
negative	1520
neutral	1583

To better illustrate the structure of the dataset, the first five sample rows are presented in Table 2. This table contains the sentiment label and numerical values corresponding to each comment.

Table 5. Sample rows from the dataset

review	label	labels
The storyline was deeply moving and kept me engaged until the end.	Positive	1
The characters were poorly developed and the plot felt incomplete.	Negative	2
Reading this book was a frustrating experience with repetitive ideas.	Negative	2
I truly enjoyed the writing style; it was elegant and inspiring.	Positive	1
The book provided information, but without a strong emotional impact.	Nötr	0

3.2. Models Used

3.2.1. RoBERTa

RoBERTa is a Transformer-based model developed by Facebook AI that improves upon BERT's training methods to enhance its performance. RoBERTa is preferred because it achieves high success in both general language understanding tasks and specific natural language processing applications such as sentiment classification. It differs from BERT thanks to long-term training on larger datasets, dynamic masking, and optimization techniques.

Although it uses the same architecture as BERT, it produces stronger language representations by adopting different pre-training strategies. Performance gains are achieved through dynamic masking, sentence batching, large mini-batch usage, and byte-level BPE vocabulary [12]. RoBERTa has achieved high success in GLUE benchmark tasks and outperformed previous models on datasets such as SQuAD and RACE [13].

3.2.2. ELECTRA

Unlike BERT, ELECTRA uses a method based on distinguishing tokens modified by the generator rather than predicting masked words. This approach enables fast and effective language representations to be learned with lower computational costs. ELECTRA consists of two components: a generator and a discriminator. While the generator replaces some words with plausible alternatives, the discriminator learns which words are original and which have been modified.

This method yields more efficient results compared to the traditional Masked Language Model (MLM) approach by processing all input tokens instead of just a small masked subset [14]. Experimental results show that ELECTRA performs better than BERT models of the same size. Furthermore, thanks to its scalability, it can achieve similar or better results with lower computational costs compared to powerful models such as RoBERTa and XLNet [15].

3.2.3. Vanilla BERT

The BERT model developed by Google better understands the meaning of words in texts thanks to bidirectional context learning and plays a fundamental role in many natural language processing tasks. BERT is pre-trained on large unlabeled text corpora and generates powerful language representations by performing tasks such as predicting masked words in sentences and inter-sentence relationships.

The fundamental idea behind BERT is to randomly mask words, enabling the model to learn both the left and right context. This feature provides a more comprehensive understanding of language compared to unidirectional models. The learned language representations can be adapted with additional layers for use in different natural language processing tasks [16]. Its Transformer-based structure provides a significant advantage over traditional methods because it can process relationships between words simultaneously [17].

3.3. Performance Metrics

The performance metrics used in the study, along with their formulas and explanations, are shown in Table 3.

Table 3. Performance metrics used

Metrics	Equation	Definition
Accuracy	Ac=TP/(TP+FN)	It indicates the percentage of correct predictions among all predictions made by the model.
F1Score	F1Score=2TP/(2TP+FN+P)	A balanced measure of sensitivity and recall is used to evaluate the average performance of true positive predictions.
Precision	Precision=TP/(TP+FP)	It shows how many of the examples classified as positive by the model are actually positive.
Recall	Recall=TP/(TP+FN)	It shows that the model correctly classified the true positive examples.

4. Experiment Results and Discussion

In this study, multi-class sentiment analysis was performed on user reviews obtained from online book review platforms using three different Transformer-based models (BERT, RoBERTa, and ELECTRA). The models were trained and evaluated on three different classes: positive, negative, and neutral.

The results show that the RoBERTa and ELECTRA models generally performed well, while the Vanilla BERT model showed lower performance. In particular, it was noted that the Vanilla BERT model had difficulty correctly classifying the positive class. This may be due to the model's inability to learn contextual relationships as effectively as RoBERTa and ELECTRA.

Considering the performance metrics, the RoBERTa model achieved the highest success with an accuracy rate of 88.38% and an F1 score of 89.2%. The ELECTRA model showed a performance close to RoBERTa with an accuracy rate of 88% and an F1 score of 88.3%. The Vanilla BERT model achieved the lowest success with a 78.46% accuracy rate and a 78.39% F1 score. Its low learning success in the negative class, in particular, contributed to these results.

Class-based analyses have revealed that models often conflate neutral classes with positive or negative classes. This may be because neutral comments can contain both positive and negative words due to the structure of language. This situation demonstrates that separating the neutral class is more difficult in sentiment analysis studies.

Overall, the findings once again demonstrate that Transformer-based models are successful in text classification. The high success rates of RoBERTa and ELECTRA show that these models can be used effectively in sentiment analysis. These evaluations provide important clues for selecting the most suitable model for specific datasets and application scenarios.

Table 4. Experimental results

Models	Performance Metrics			
	Accuracy	**Precision**	**F1Score**	**Recall**
RoBERTa	%88.38	%89	%89.8	%89.2
ELECTRA	%88	%88.6	%88.1	%88.3
Vanilla BERT	%78.46	%83	%78.46	%78.39

The results clearly show that RoBERTa achieved the highest success, ELECTRA performed similarly, and Vanilla BERT showed lower success compared to other models. RoBERTa's language modeling capacity and pre-training on a large dataset are considered among the main reasons for this success.

5. Conclusion

In this study, multi-class sentiment analysis (positive, negative, neutral) was performed using user reviews obtained from online book review platforms. Three different Transformer-based architectures (BERT, RoBERTa, and ELECTRA) were trained and then their performances were compared. Standard metrics such as accuracy, precision, sensitivity, and F1 score were used in model evaluation.

According to the results, the RoBERTa model achieved the highest success in terms of both overall accuracy and class-specific F1 scores. The ELECTRA model showed similar performance, while the Vanilla BERT model achieved relatively lower accuracy and F1 scores. In particular, Vanilla BERT's low success in the negative class can be explained by the model's inability to learn contextual relationships sufficiently. Furthermore, the presence of positive or negative elements in neutral comments led to more errors in this class.

These findings demonstrate that Transformer-based models are an effective tool for multi-class sentiment analysis tasks. However, model success is directly influenced by factors such as dataset size, label distribution, language diversity, and preprocessing strategies.

6. References

[1] Cömert Ö, Yücel N. Müşteri Duyarlılığını Keşfetmek İçin Yapay Zeka Destekli Analiz ile Çevrimiçi Ürün İncelemelerinden Anlamlı Bilgiler Elde Etme. Fırat Üniversitesi Mühendislik Bilimleri Dergisi 2023; 35(2): 679-690.

[2] Saka, Semih Osman, and Zafer Cömert. "Sinir Ağı Dil Modelleri ve Evrensel Cümle Kodlayıcı Kullanarak Havayolu Müşteri Yorumlarının Duygu Analizi." Fırat Üniversitesi Mühendislik Bilimleri Dergisi 37.1 (2025): 167-182.

[3] Shehu, Harisu Abdullahi. Kutupsallık sözlüğü ve yapay zeka yardımı ile Türkçe twitter verileri üzerinde duygu analizi. MS thesis. Pamukkale Üniversitesi Fen Bilimleri Enstitüsü, 2019.

[4] Özkul, Mübin. Metin madenciliği yöntemi ile akıllı ulaşım sistemleri algısının duygu analizi ile ölçülmesi. MS thesis. Sosyal Bilimler Enstitüsü, 2019.

[5] Özdeş, Merve. Büyük veri araçlarını kullanarak duygu analizi gerçekleştirimi. MS thesis. Pamukkale Üniversitesi Fen Bilimleri Enstitüsü, 2017.

[6] Demir, Cennet Göloğlu, and Hakan Yılmaz. "Sınıf dışı eğitim faaliyetlerinin öğrencilerin bilim ve teknolojiye yönelik tutumlarına etkisi ve duygu analizi." İnsan ve Toplum Bilimleri Araştırmaları Dergisi 7.5 (2018): 101-116.

[7] Saka, Semih Osman, and Zafer Cömert. "Sinir Ağı Dil Modelleri ve Evrensel Cümle Kodlayıcı Kullanarak Havayolu Müşteri Yorumlarının Duygu Analizi." Fırat Üniversitesi Mühendislik Bilimleri Dergisi 37.1 (2025): 167-182.
[8] Budak, İbrahim. "Veri ve metin madenciliği ile hava yolu işletmelerinin sosyal medya yorum ve skorlarının değerlendirilmesi." (2021).
[9] Uslu, Osman, and Serel Özmen-akyol. "Türkçe haber metinlerinin makine öğrenmesi yöntemleri kullanılarak sınıflandırılması." Eskişehir Türk Dünyası Uygulama ve Araştırma Merkezi Bilişi Dergisi 2.1 (2021): 15-20.
[10] Tuzcu, Seda. "Çevrimiçi kullanıcı yorumlarının duygu analizi ile sınıflandırılması." Eskişehir Türk Dünyası Uygulama ve Araştırma Merkezi Bilişim Dergisi 1.2 (2020): 1-5.
[11] Kılınç, Deniz, et al. "KNN algoritması ve r dili ile metin madenciliği kullanılarak bilimsel makale tasnifi." Marmara Fen Bilimleri Dergisi 28.3 (2016): 89-94.
[12] https://huggingface.co/docs/transformers/model_doc/roberta
[13] https://www.geeksforgeeks.org/machine-learning/overview-of-roberta-model/
[14] https://huggingface.co/docs/transformers/model_doc/electra
[15] https://arxiv.org/abs/2003.10555
[16] https://huggingface.co/docs/transformers/model_doc/bert
[17] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention is All You Need*. Advances in Neural Information Processing Systems (NeurIPS), 30, pp. 5998–6008.

International Conference on Advances and Innovations in Engineering

Triple Teacher Framework with Exponential Moving Average for Screw Anomaly Detection

Emre GÜÇLÜ[1] , İlhan AYDIN[2], Erhan AKIN [3]

[1,2,3]Computer Engineering, Engineering Faculty, Fırat University, Elazığ, Türkiye.
1eguclu@firat.edu.tr, 2iaydin@firat.edu.tr, 3eakin@firat.edu.tr
[1](ORCID: 0000-0002-4566-7517), [2](ORCID: 0000-0001-6880-4935),[3](ORCID: 0000-0001-6476-9255)

Abstract
Screw defect detection is a key factor to characterize product reliability and safety in a manufacturing setting. This work proposes a triple teacher knowledge distillation framework with an additional exponential moving average implementation in order to obtain stable feature representation, as well as accurate anomaly localization. The framework includes three teacher networks and one student network. The pair of teacher networks include a trainable teacher which is trained, as well as a target teacher which derives feature weights from the exponential moving average of the trainable teacher, and a reference teacher that is fixed that helps to ensure consistency during training and representations across teachers. The multi-scale features from the target teacher are pushed through a bottleneck module before being reconstructed by the student. The proposed method is trained through the combination of distillation loss and alignment loss resulting in improved accuracy and improved robustness. The proposed method was explored on the screw category of the MVTec Anomaly Detection dataset. The MVTec dataset contains high-resolution images of screw defects, consisting of normal images and random defect types of a screw. The proposed abnormality detection method attained the pixel-level area under the receiver operating characteristic curve of 91.8 percent. The proposed method depicted strong capability for defect localization, with a reduced number false positive detections, while obtaining more accurate defect boundaries than baseline methods.

Keywords: screw defects, knowledge distillation, exponential moving average, anomaly detection

1. Introduction

Quality control in industrial manufacturing processes plays a critical role in ensuring product reliability and customer satisfaction. Timely detection of surface defects, scratches, cracks, or assembly errors occurring on production lines is of great importance for both reducing costs and increasing production efficiency. Traditional quality control methods mostly rely on manual inspection by the human eye or rule-based computer vision algorithms. However, in high-volume production environments, these methods can lead to time-consuming, costly, and error-prone processes. In particular, screws are fundamental mechanical elements used in industrial assembly processes to join different parts and form a secure connection. The locking mechanism provided by the screw and nut combination ensures that the parts are securely fastened while maintaining the connection over time against mechanical effects such as vibration, load, and pressure [1]. Therefore, timely detection of even the smallest manufacturing defects in screws is critical for the sustainability of both mechanical safety and production quality.

Thanks to their durability and reusability, screws are indispensable in many industries such as machinery manufacturing, automotive, electric motors, aviation, construction, furniture, and household appliances. In industrialized countries with high production capacity, demand for these products is greater, resulting in a larger market size [2].
One of the most significant problems faced by screw manufacturers is surface defects. High production speed on the line, raw material quality, machining tolerances, and variability in coating processes can lead to defects such as cracks, scratches, burrs, dents, coating errors, or deformation [3]. Such defects not only affect aesthetics but also negatively impact the strength and reliability of the connection.

Traditional quality control is often carried out through visual inspection by human operators; however, this method is time-consuming and prone to errors due to operator fatigue [4]. For this reason, in recent years, real-time defect detection systems based on computer vision and artificial intelligence deep learning algorithms have been developed for in-line inspection. Especially models such as YOLO, Faster R-CNN, and Deeplabv3+ integrated with high-resolution cameras minimize human error during production and increase the defect-free product ratio [5],[6],[7].

[1]Corresponding author

Academic studies have shown that combining deep learning with techniques such as transfer learning, domain adaptation (scale-sensitive adaptation), and weakly supervised learning enables high-accuracy defect detection even under different production conditions [8],[9],[10]. For example, transfer learning-based models have been reported to yield faster and more consistent results than human-eye-based methods in detecting loosening or surface defects in fasteners such as screws and nuts [10],[11]. Domain adaptation approaches minimize distribution differences in image data collected from different production lines, improving model generalization capability [12]. This increases quality control efficiency and reduces production costs.

In production lines, the occurrence probability of defective products is generally very low. This causes a serious class imbalance problem in datasets collected for quality control purposes. The small number of rare defect types negatively affects the learning process of deep learning models, increasing the risk of overlooking defective products while correctly classifying defect-free ones [13]. Therefore, research is increasingly focusing on methods that learn only from defect-free data to detect defects [14]. In particular, anomaly detection approaches with autoencoders are trained exclusively on defect-free product data and can identify deviations from the observed normal distribution as defects [15]. Additionally, knowledge distillation techniques enable lightweight and fast models to learn from the knowledge of high-performance deep networks, making it possible to detect defects with high accuracy and low latency in real-time production lines [16]. Recently, hybrid systems combining these two approaches have been developed to address the class imbalance problem effectively [17]. Bergmann et al. [18] used the MVTec AD dataset to generate anomaly maps of surface defects with a deep convolutional autoencoder (CAE), achieving high accuracy in real-time inspection of industrial parts. Zong et al. [19] combined a deep autoencoder with a Gaussian Mixture Model (GMM) to estimate anomaly scores in complex manufacturing data.

The GANomaly model developed by Akçay et al. [20] combined the Generative Adversarial Network (GAN) and autoencoder structure, achieving strong results in visual anomaly detection. To overcome the imbalanced dataset problem, one-class classification and self-supervised methods are also used. Ruff et al. [21] proposed the Deep Support Vector Data Description (Deep SVDD) method, learning only from the normal class to distinguish anomalies. Tack et al. [22] enhanced anomaly detection performance by learning representative features from only defect-free products via self-supervised contrastive learning. Yi and Yoon [23] combined transfer learning and feature embedding strategies for one-class classification. In addition, knowledge distillation-based models can achieve high performance. Salehi et al. [24] used a teacher-student architecture to distill knowledge from high-capacity models into lighter networks, providing low-latency solutions in industrial defect detection.

Synthetic data generation is also frequently used to address the class imbalance issue. Qiu et al. [13] generated synthetic defective images via style transfer to increase the representation power of the defective class during training of a YOLOv8-based model. Niu et al. [25] used GAN-based data augmentation methods to increase defect diversity.

In this study, an EMA-Teacher (Exponential Moving Average Teacher) mechanism is integrated into a student-teacher architecture to enable smooth parameter updates of the teacher model during training. Thus, the teacher model, derived from the parameters of the student model, produces a more stable and generalizable feature space. The proposed method is supported by multi-scale cosine similarity-based distillation loss and tested on the screw category of the MVTec Anomaly Detection (MVTec AD) dataset. Experimental results show that EMA-Teacher integration significantly improves both pixel-level and sample-level anomaly detection performance.

The main contributions of this study can be summarized as follows:

- The parameters of the teacher model are updated from the parameters of the student model with momentum, providing more stable feature representations.
- The consistency between teacher and student features is optimized using cosine similarity loss computed at different resolution levels.
- Experiments on the MVTec AD dataset demonstrate that the EMA-Teacher approach improves AUROC and AUPRO metrics compared to the traditional fixed teacher method.
- The proposed approach offers a lightweight and flexible solution that can be integrated into industrial quality control systems without the need for labeled anomaly data.

2. Materials and Methods

2.1. Overall Architecture

In the student-teacher-based methodology proposed, a pre-trained WideResNet-50 model on ImageNet is utilized as the teacher backbone. Multi-scale feature maps from different intermediate layers of the teacher model are reconstructed by the student model. The student model, sharing the same architectural setup, is trained from an updated teacher model during training with the Exponential Moving Average (EMA) technique. Thus, instead of directly optimizing the parameters of the teacher model, they are updated with a smooth transition using the EMA of the student model parameters. The three steps in the training process are given below. During the first step, feature extraction, both the student and the EMA-teacher models both take in the input image and produce intermediate layer feature maps at different resolutions. During this stage, the teacher model generated features act as target features during training time. During the second stage, feature reconstruction, the student model generated features are reconstructed through an additional Batch Normalization (BN) layer and a decoder network. The decoder network enables learning in the student model by generating representations that are close to the feature maps of the teacher model.

A multi-scale similarity loss is calculated in the final stage. For this purpose, teacher and student features are matched through a cosine similarity-based distillation loss. Feature maps of different resolution levels are upsampled to a uniform size and then pooled to calculate similarity. This approach allows different levels of representations to be compared in a shared space and offers a more robust learning signal. Figure 1 illustrates the overall architecture of the proposed EMA-based triple-teacher knowledge distillation framework for screw anomaly detection.

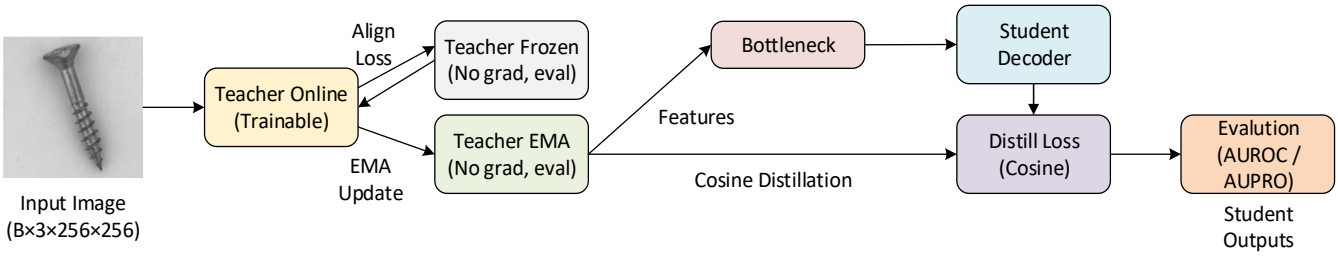


Figure 3. Schematic diagram of the proposed model

During the testing process, anomaly score maps are generated using only the difference of teacher-student feature. In normal regions, teacher and student features are similar, resulting in low scores. In anomalous regions, teacher and student features are different, resulting in high scores. The generated score maps are converted into binary masks (abnormal/normal) with a specified threshold value, and performance is measured in pixel-by-pixel AUROC. The EMA update procedure generates a more stable progression of the teacher model's parameters and assists the student model to acquire more stable and highly generalizable features.

2.2. Architectural Components

The proposed approach is based on an EMA-driven multi-scale feature distillation paradigm. The system consists of three distinct teacher models: Trainable Teacher, Target Teacher and Reference Teacher, and a student model.

The Trainable Teacher (T_θ) is the teacher model optimized during training. Its parameters θ are updated in each iteration through standard backpropagation and an optimizer.

The Target Teacher ($T_{\bar{\theta}}$) is updated from the parameters of the Trainable Teacher using the EMA method and provides the student with stable target representations.

The Reference Teacher ($T_{\theta*}$) is the non-updated (no-grad) fixed copy of the Trainable Teacher and is used as a reference for computing the Alignment Loss.

Given an input image x ∈ $R^{3\times H\times W}$, the multi-scale feature maps produced by these teachers are defined as:

$$Z_{train}(x) = T_{\theta}(x),\ Z_{target}(x) = T_{\bar{\theta}}(x),\ Z_{ref}(x) = T_{\theta*}(x) \quad (1)$$

Each feature set consists of K scales corresponding to different spatial resolutions:

$$Z^{(.)}(x) = \{z^{(1)}(x), z^{(2)}(x), \dots, z^{(K)}(x)\} \quad (2)$$

The multi-scale features from the Target Teacher are compressed by a Bottleneck module B_p:

$$\hat{Z}(x) = B_p\left(Z_{target}(x)\right) \quad (3)$$

The student model (decoder) D_p then reconstructs the teacher features from these compressed representations:

$$\hat{Z}(x) = D_p\left(\hat{Z}(x)\right) \quad (4)$$

The parameters of the Target Teacher are derived from the Trainable Teacher using the Exponential Moving Average (EMA) update, which helps to produce temporally consistent target features. The update rule is given by:

$$\bar{\theta} \leftarrow \tau.\bar{\theta} + (1-\tau).\theta \quad (5)$$

Where, $\bar{\theta}$ Parameters of the Target Teacher, θ Parameters of the Trainable Teacher $\tau\epsilon(0,1)$, EMA decay factor (set to 0.999 in this work). A high τ value results in slow updates, effectively filtering short-term fluctuations in the teacher parameters and ensuring stable target representations for the student.

2.3. Loss Function

Model optimization is performed using a total loss function that integrates two key components, Distillation Loss and Alignment Loss.

Distillation Loss loss measures the similarity between the reconstructed features of the student and the features of the Target Teacher using a multi-scale cosine similarity loss:

$$L_{distill} = \frac{1}{K}\sum_{k=1}^{K}\left[1 - \cos\left(\hat{z}^{(k)}(x), z_{target}^{(k)}(x)\right)\right] \quad (6)$$

This loss enforces consistency between the Trainable Teacher and the Reference Teacher by combining cosine distance and L_1 norm:

$$L_{align} = \frac{1}{K}\sum_{k=1}^{K}\left[\alpha.\left(1 - \cos\left(z_{train}^{(k)}(x),\ \ z_{ref}^{(k)}(x)\right)\right) + \beta.\left|z_{train}^{(k)}(x),\ \ z_{ref}^{(k)}(x)\right|_1\right] \quad (7)$$

The total loss is formulated as follows

$$L_{total} = L_{distill} + \lambda_{align}.L_{align} \quad (8)$$

3. Experimental Results

All experiments were carried out on a workstation equipped with an NVIDIA RTX 3060 GPU, an Intel Core i9-12900K CPU, and 64 GB of system memory. The proposed EMA-based triple-teacher knowledge distillation framework was implemented using PyTorch 2.0 with CUDA 11.8 for GPU acceleration.

We evaluated the method on the screw category of the MVTec Anomaly Detection (MVTec AD) dataset, a widely used benchmark for industrial anomaly detection. The MVTec AD dataset consists of 15 object and texture categories, each containing high-resolution color images of both normal and defective samples. Defects vary in type (scratches, dents, missing parts, discolorations, etc.) and size, making the task challenging. In this study, only the screw category was used for experimentation, while the proposed method remains applicable to all categories.

The training images were resized to 256×256 pixels, and normalization was performed according to ImageNet statistics. Data augmentation techniques such as random rotation, horizontal flipping, and color jittering were applied to improve generalization. The model was trained for 11 epochs using the Adam optimizer with a learning rate of 0.005, $\beta 1$=0.5, and $\beta 2$=0.999. The batch size was set to 16. For the EMA update, the decay factor τ was fixed at 0.999, ensuring slow parameter updates for stable target representations. The loss balancing hyperparameters were set to α=1.0, β=0.1 and λalign=0.1.

For comparison, we implemented a baseline knowledge distillation model without EMA to evaluate the contribution of the EMA mechanism. Both models used identical backbone architectures for fair comparison: WideResNet50 as the encoder in the teacher networks, a bottleneck transformation layer, and a de-WideResNet50 as the student decoder.

3.1. Evaluation Metrics

To comprehensively evaluate anomaly detection performance, we employed three metrics. Pixel-level AUROC measures the ability of the model to localize anomalous pixels in the image, where a higher value indicates better pixel-wise discrimination between normal and anomalous regions. Sample-level AUROC evaluates image-level binary classification performance by determining whether an image contains any defect. Pixel-level AUPRO computes the Area Under the Per-Region Overlap curve, quantifying the proportion of correctly localized anomaly pixels normalized by the defect region size. In this evaluation framework, Pixel-level AUROC emphasizes localization accuracy, Sample-level AUROC assesses the overall anomaly detection capability, and Pixel-level AUPRO is particularly sensitive to the completeness of defect coverage, especially for small or irregular anomalies.

3.2 Experimental Results on MVTec AD – Screw Category

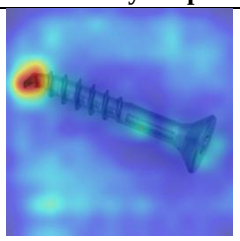
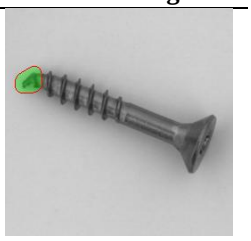
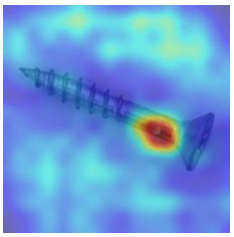
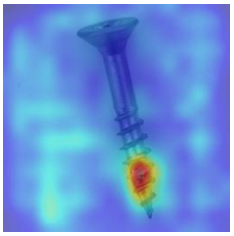
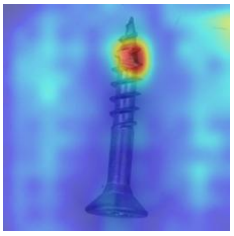
The experimental evaluations were conducted on the MVTec Anomaly Detection (MVTec AD) dataset, which consists of high-resolution images of industrial products. This dataset contains both normal and defective samples for each category, along with binary ground truth masks indicating the pixel-level locations of defects.

In this study, the screw category was selected as the test case. This category includes various types of defects such as scratches, dents, and contamination. All test images were resized to a resolution of 256 × 256 pixels to match the model's input size. They were then normalized using the mean and standard deviation values applied during the model's training phase. These preprocessing steps help stabilize the statistical distribution of the input data, thereby enhancing the consistency of model predictions.

After preprocessing, each RGB test image was passed through the encoder–decoder-based anomaly detection architecture. The encoder (WideResNet50) extracted multi-scale feature maps from the input image, which were then processed through a batch normalization layer. The decoder reconstructed these features, highlighting the differences between the learned "normal" appearance and the actual image.

As a result of this process, an anomaly map was generated for each pixel, representing the probability of being anomalous. The model produced low scores for normal regions and high scores for defective regions. Table 1 presents the anomaly maps obtained from sample images in the dataset.

Table 6. Visual comparison of the input image, generated anomaly map, and the highlighted defective region

Input Image	Anomaly Map	Defective Region
		
		
		
		

The anomaly maps obtained from the model were first normalized to the range [0, 1]. Then, to identify only high-confidence defect regions, percentile-based thresholding was applied. In this study, the top 5% of pixel values (95th percentile and above) were considered as defects. The resulting binary masks were filtered using a morphological opening operation to remove small and isolated noise, followed by morphological dilation to preserve the integrity of defect boundaries. The refined masks were overlaid onto the original RGB image in two layers. The interior of defect regions was filled with semi-transparent green, allowing background details to remain visible while emphasizing the defect area. The boundaries of these filled regions were outlined in red with a thickness of 3 pixels, making defect edges sharp and clearly visible. The results are presented in Table 1.

The model achieved a Pixel-level AUROC of 91.8%, indicating a strong ability to localize defective regions at the pixel level. The Sample-level AUROC reached 98.9%, demonstrating near-perfect binary classification capability for distinguishing between defective and non-defective samples. Furthermore, the Pixel-level AUPRO score was 86.0%, reflecting the method's robustness in accurately covering the entire defect area, even for small or irregularly shaped anomalies. These results confirm that the proposed approach is effective in both precise defect localization and overall anomaly detection accuracy.

Table 7. Quantitative results on the MVTec AD (*screw*) dataset

Pixel-level AUROC	91.8%
Sample-level AUROC	98.9%
Pixel-level AUPRO	86.0%

Table 3 shows that our method achieves the highest AUROC (91.8%) in screw anomaly detection, outperforming ADPS [26] by 2.3% and exceeding other methods such as Pull & Push [28], SSAD-Pseudo [29], and IAD-CLIP [30] by notable margins.

Table 8. Comparison of AUROC scores for screw defect detection between the proposed method and existing approaches.

Reference	**AUROC (Screw)**
Pull & Push [27]	85.6%
SSAD-Pseudo [28]	74.1%
IAD-CLIP [29]	77.8%
ADPS [26]	89.5%
Ours	91.8%

Acknowledgements

This work was supported by The Scientific and Technological Research Council of Türkiye (TÜBİTAK) under Project No. 125E031.

4. References

[1] Fzang, E., Li, S., Liu, Z., Wang, Y., & Tao, T. (2024, December). Research on Real-Time Inspection Scheme of Tunnel Lining Surface Defects Based on YOLOv10. In *International Conference on Artificial Intelligence and Autonomous Transportation* (pp. 18-27). Singapore: Springer Nature Singapore.

[2] Chen, P., Ma, Z., Xu, C., Zhang, M., Li, H., Zheng, K., & Jin, Y. (2024). Scale-aware domain adaptation for surface defects detection on machine tool components in contaminant measurements. IEEE Transactions on Instrumentation and Measurement.

[3] Xu, Y., Luo, Y., Majeed, A. P. A., Liu, X., Zhu, Y., & Chen, W. (2024, August). Ball Screw Drive Surface Defect Model Based on Transfer Learning Approach. In *International Conference on Intelligent Manufacturing and Robotics* (pp. 451-457). Singapore: Springer Nature Singapore.

[4] Tin, D. D. MealySup: A Multi-loss and weakly supervised learning approach for fine-grained object recognition.

[5] Li, C., Yan, H., Qian, X., Zhu, S., Zhu, P., Liao, C., ... & Li, X. (2023). A domain adaptation YOLOv5 model for industrial defect inspection. *Measurement*, *213*, 112725.

[6] Güçlü, E., Aydın, İ., Şahbaz, K., Akın, E., & Karaköse, M. (2021). Demiryolu bağlantı elemanlarında bulunan kusurların YOLOv4 ve bulanık mantık kullanarak tespiti. *Demiryolu Mühendisliği*, (14), 249-262.

[7] Guclu, E., & Akin, E. (2025). Enhanced defect detection on steel surfaces using integrated residual refinement module with synthetic data augmentation. Measurement, 250, 117136.

[8] Choi, J., Ha, M., & Lee, J. G. (2023). Transfer learning-based object detection model for steel structure bolt fastening inspection. *Applied Sciences*, *13*(17), 9499.

[9] Alif, M. A. R., Hussain, M., Tucker, G., & Iwnicki, S. (2024). Boltvision: A comparative analysis of cnn, cct, and vit in achieving high accuracy for missing bolt classification in train components. *Machines*, *12*(2), 93.

[10] Huang, J., Liu, J., Gong, H., Gong, H., & Deng, X. (2023). CDMTNet: a novel transfer learning model for the loosening detection of mechanical structures with threaded fasteners. *Structural Health Monitoring*, *22*(6), 3840-3855.

[11] Chen, Y., Wei, J., Gao, Z., Li, W., & Wu, J. (2024). Multibolt looseness monitoring of steel structure based on multitask active sensing method and substructure cross-domain transfer learning. Structural Health Monitoring, 23(6), 3486-3504.
[12] Ramasso, E., Teloli, R. D. O., & Marcel, R. (2024). On the condition monitoring of bolted joints through acoustic emission and deep transfer learning: generalization, ordinal loss, and super-convergence. *Structural Health Monitoring*, 14759217241259668.
[13] Qiu, S., Cai, B., Wang, W., Wang, J., Zaheer, Q., Liu, X., ... & Peng, J. (2024). Automated detection of railway defective fasteners based on YOLOv8-FAM and synthetic data using style transfer. *Automation in Construction*, *162*, 105363.
[14] Hridoy, M. W., Rahman, M. M., & Sakib, S. (2024). A framework for industrial inspection system using deep learning. *Annals of Data Science*, *11*(2), 445-478.
[15] Prunella, M., Scardigno, R. M., Buongiorno, D., Brunetti, A., Longo, N., Carli, R., ... & Bevilacqua, V. (2023). Deep learning for automatic vision-based recognition of industrial surface defects: A survey. *IEEE Access*, *11*, 43370-43423.
[16] De Donato, L., Flammini, F., Marrone, S., Mazzariello, C., Nardone, R., Sansone, C., & Vittorini, V. (2022). A survey on audio-video based defect detection through deep learning in railway maintenance. *IEEE Access*, *10*, 65376-65400.
[17] Alif, M. A. R., Hussain, M., Tucker, G., & Iwnicki, S. (2024). Boltvision: A comparative analysis of cnn, cct, and vit in achieving high accuracy for missing bolt classification in train components. *Machines*, *12*(2), 93.
[18] Bergmann, P., Fauser, M., Sattlegger, D., & Steger, C. (2019). MVTec AD--A comprehensive real-world dataset for unsupervised anomaly detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 9592-9600).
[19] Zong, B., Song, Q., Min, M. R., Cheng, W., Lumezanu, C., Cho, D., & Chen, H. (2018, February). Deep autoencoding gaussian mixture model for unsupervised anomaly detection. In *International conference on learning representations*.
[20] Akcay, S., Atapour-Abarghouei, A., & Breckon, T. P. (2018, December). Ganomaly: Semi-supervised anomaly detection via adversarial training. In Asian conference on computer vision (pp. 622-637). Cham: Springer International Publishing.
[21] Ruff, L., Vandermeulen, R., Goernitz, N., Deecke, L., Siddiqui, S. A., Binder, A., ... & Kloft, M. (2018, July). Deep one-class classification. In International conference on machine learning (pp. 4393-4402). PMLR.
[22] Tack, J., Mo, S., Jeong, J., & Shin, J. (2020). Csi: Novelty detection via contrastive learning on distributionally shifted instances. *Advances in neural information processing systems*, *33*, 11839-11852.
[23] Yi, J., & Yoon, S. (2020). Patch svdd: Patch-level svdd for anomaly detection and segmentation. In Proceedings of the Asian conference on computer vision.
[24] Salehi, M., Sadjadi, N., Baselizadeh, S., Rohban, M. H., & Rabiee, H. R. (2020). Multiresolution knowledge distillation for anomaly detection. 2021 IEEE. In *CVF conference on computer vision and pattern recognition (CVPR)* (pp. 14897-14907).
[25] Niu, S., Li, B., Wang, X., & Lin, H. (2020). Defect image sample generation with GAN for improving defect recognition. *IEEE Transactions on Automation Science and Engineering*, *17*(3), 1611-1622.
[26] Xing, P., Tang, H., Tang, J., & Li, Z. (2025). ADPS: Asymmetric distillation postsegmentation for image anomaly detection. IEEE Transactions on Neural Networks and Learning Systems, 36(4), 7051–7063.
[27] Zhou, Q., He, S., Liu, H., Chen, T., & Chen, J. (2023). Pull & Push: Leveraging differential knowledge distillation for efficient unsupervised anomaly detection and localization. IEEE Transactions on Circuits and Systems for Video Technology, 33(5), 2176–2189.
[28] Shah, R. A., Urmonov, O., & Kim, H. (2025). Self-supervised image anomaly detection through diverse pseudo anomaly insertion. IEEE Access, 13, Article 3580864.
[29] Li, Z., Ge, Y., Li, Q., & Meng, L. (2024, November 26–30). IAD-CLIP: Vision-language models for zero-shot industrial anomaly detection. In Proceedings of the 2024 International Conference on Advanced Mechatronic Systems

Quantum Computing in Mobile Networks: A Comparative Analysis

Çağatay Umut ÖĞDÜ[1], Mehmet KARAKÖSE[2]

[1,2]Firat University, Computer Engineering Department, 23119 Elazig, Türkiye.
1cagatayogdu@gmail.com, 2mkarakose@gmail.com
[1](ORCID: 0009-0004-1697-4392), [2](ORCID: 0000-0002-3276-3788)

Abstract
In this paper, we systematically examine the convergence of quantum communication with 6G mobile networks in terms of security, timing, and management dimensions. Quantum applications (Shor, Grover) provide a path to public-key security, mitigating the risk of "listen-in-the-move" schemes, enabling transitions to both physical quantum security (QKD/QSDC) and post-quantum cryptography (PQC). Methodologically, we compare the current literature with the deterministic transport (TSN/DetNet, advanced URLLC) capabilities of 6G, analyzing the operations of quantum networks (entanglements, teleportation, repeaters). Our contribution is of three types: (i) we present a holistic reference architecture that unifies QKD/QSDC, deterministic classical transport, PQC, and software-defined slicing; (ii) We present design/practical facilities for multipath routing (Q-CAST Mobility) and widely reliable roaming (REDP) in mobile scenarios; (iii) We compile integration policies for the quantum routing, deterministic classical transport, and PQC triangle. The findings demonstrate that teleportation is a factor for the classical side channel of deterministic 6G transport; when considered alongside field demonstrations of QKD/QSDC, it appears that latency and throughput targets are achievable in mobile quantum services. A feasible roadmap for operators, including cryptographic transformation, inventory, and slotted storage, is proposed.

Keywords: 6G, mobile networks, quantum networks, quantum key distribution, deterministic networks.

1. Introduction

Quantum information processing and communication methods have the potential to radically transform the architecture of communication networks thanks to their computational models that fundamentally challenge the security assumptions of classical cryptography. Shor's algorithm, which solves prime factorization and discrete logarithm problems in polynomial time, and Grover's method, which provides square-root speedup in unverified search, introduce new threats, especially for public-key schemes (RSA, ECC) and key lookup/hash attacks [1]–[3]. In this context, considering the quantum secure communication capabilities together with the end-to-end deterministic latency and high reliability targets of 6G, the design of classical/quantum hybrid networks has become a critical research area [4], [12]. Figure 1 shows the distribution of studies on quantum computing and quantum networks in the context of mobile networks/5G–6G over the years. As can be seen, there has been a significant increase in recent years; This trend confirms that classical/quantum hybrid architectures, with their deterministic latency and high reliability targets for 6G, have taken a central position in the research agenda.

On the one hand, physics-based quantum security mechanisms, such as quantum key distribution (QKD) [1] and direct quantum secure communication (QSDC), offer different design considerations to mitigate the risk of "harvest-now, decrypt-later" behavior. QSDC, when transporting messages directly through a quantum channel, has been analyzed for information-theoretic security even with finite block sizes; practical demonstrations have been reported, reaching 100 km over fiber and kbps in multi-user setups [7]–[9]. The extension of this security principle to high-integrity structures such as blockchains is increasingly mature in the literature, with examples of quantum-safe signatures, voting/consensus, and optical network-based applications [5], [6]. On the other hand, the end-to-end operation of quantum networks (QNs) relies on classical transport planes at a significant scale for operations such as fragile entanglement distribution, clearing/purification, and error correction. The classical two-bit requirement of quantum teleportation may necessitate further reduction of deterministic latency limits beyond 5G, particularly for services such as blind quantum computing (BQC), to levels of hundreds of microseconds targeted by 6G [12]. Therefore, the evolution of URLLC and the tight integration of 6G with deterministic wired technologies such as TSN/DetNet are triggering new use cases that enable mobile and non-terrestrial (satellite/underwater) quantum applications [12].

[1]Corresponding author

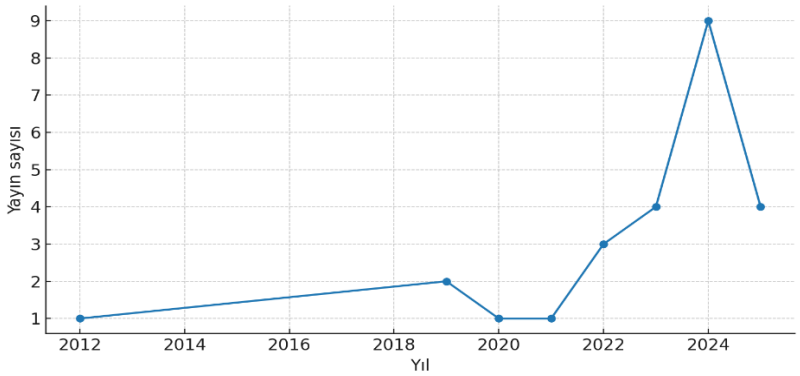


Figure 1. Annual number of publications in the literature on quantum computing in mobile networks (2012–2025).

On the one hand, physics-based quantum security mechanisms, such as quantum key distribution (QKD) [1] and direct quantum secure communication (QSDC), offer different design considerations to mitigate the risk of "harvest-now, decrypt-later" behavior. QSDC, when transporting messages directly through a quantum channel, has been analyzed for information-theoretic security even with finite block sizes; practical demonstrations have been reported, reaching 100 km over fiber and kbps in multi-user setups [7]–[9]. The extension of this security principle to high-integrity structures such as blockchains is increasingly mature in the literature, with examples of quantum-safe signatures, voting/consensus, and optical network-based applications [5], [6]. On the other hand, the end-to-end operation of quantum networks (QNs) relies on classical transport planes at a significant scale for operations such as fragile entanglement distribution, clearing/purification, and error correction. The classical two-bit requirement of quantum teleportation may necessitate further reduction of deterministic latency limits beyond 5G, particularly for services such as blind quantum computing (BQC), to levels of hundreds of microseconds targeted by 6G [12]. Therefore, the evolution of URLLC and the tight integration of 6G with deterministic wired technologies such as TSN/DetNet are triggering new use cases that enable mobile and non-terrestrial (satellite/underwater) quantum applications [12].

At the upper layers of quantum networks, routing and slicing problems that allocate limited resources for multiple concurrent requests become prominent. Protocols such as Q-CAST, which are sensitive to concurrent multiple sessions and proposed for fixed topologies, aim to achieve high entanglement performance with ideas such as multipath, wide path, and local state change, while topology volatility poses additional challenges for mobile quantum nodes (UAV, vehicle) [15]. In this context, new mobility-focused proposals that adapt the hybrid (intra-region proactive, inter-region reactive) approach of ZRP to the Q-CAST framework and improve performance with multipath/adaptive updating are inspired by classical MANET protocols such as AODV/OLSR and quantum MANET studies such as QRS and MFI-EDA [13]–[19]. Simultaneously, approaches such as entanglement-assisted network architectures and reliable entanglement distribution protocols (REDP) for reliability and scalability of entanglement distribution are taking an increasingly central role in the 6G context [13], [14].

The transition to quantum-resistant classical cryptography is critical to the security of the 6G ecosystem. NIST's FIPS 203 (ML-KEM), FIPS 204 (ML-DSA), and FIPS 205 (SLH-DSA) standards, published in 2024, provide references for implementing PQC in key components such as certificates, TLS/IPsec, EAP-TLS, and SUCI protection [10]. Documents such as GSMA's guidelines and ETSI TR 103 619 also outline the inventorying, risk analysis, gradual transition, and crypto agility processes for operators [11], [20]. Consequently, 6G-quantum convergence is shaped on four pillars: (i) physics-based quantum secure communication (QKD/QSDC), (ii) deterministic classical transport (TSN/DetNet and advanced URLLC), (iii) quantum-resistant cryptography (PQC), and (iv) slicing with software-defined management (SD-QN, 6G slice-service) [6], [12]–[14], [20]. This paper aims to address the co-design of these components and, in particular, the systematic integration of the quantum routing–deterministic classical transport–PQC triangle in mobile quantum networks. As shown in Table 1, recent work focuses on three axes: (i) QKD/QSDC-based physical security, (ii) 6G/TSN-DetNet-supported deterministic classical transport, (iii) transition to PQC and quantum-supported network protocols (e.g., mobile QN routing and slicing).

Table 1. Comparison of important works on quantum computing and quantum networks in mobile networks

Study	Purpose of the Study	Method Used	Benefits
[6]	Propose a QSDC-based quantum blockchain scheme; realize identification, encryption, and consensus with QSDC, and give an optical network-based implementation.	Architecture consisting of SDN controller, QSDC network node and quantum-secure blockchain node; examining the effects on resource usage through simulation.	It quantifies the resource utilization advantages of QSDC-enabled blockchain and presents a practical implementation with optical network components available today.
[12]	Demonstrating how "6G and Beyond (6GB)" networks can support quantum networks with deterministic classical connectivity and slicing.	Case/use case focused conceptual framework; mapping 6GB's timing/synchronization and end-to-end deterministic latency capabilities to QN requirements.	It is based on the assumption that the control/operation planes of QNs will remain essentially classical and that 6GB provides tangible benefits to QNs with synchronization/location and deterministic latency.
[13]	To compile a comprehensive review of the mechanics, enabling technologies, challenges, and research directions of entanglement-assisted quantum networks.	Survey/literature review; mapping the future integration topics of quantum networks for the 6G context.	With 6G, the hybrid (classical + quantum) interface provides critical axes such as repeaters, entanglement distribution and error correction in a collective framework, guiding subsequent design/protocol studies.
[14]	Designing a reliable entanglement distribution protocol (REDP) for large-scale quantum networks.	Protocol design; network-level modeling of entanglement distribution with a focus on large scale and reliability.	We address the fundamental bottleneck in 6G-quantum convergence by proposing a concrete design approach for scalable and reliable entanglement deployment in wide-area quantum networks.
[15]	We propose a mobility-focused adaptation of Q-CAST (Q-CAST Mobility) for mobile quantum nodes.	Hybrid (proactive+reactive) multipath routing inspired by ZRP; simulation-based evaluation considering entanglement generation/purification/exchange.	Throughput superiority over QRS and MFI-EDA, especially at high mobility, and robust entanglement maintenance with source/route adaptation.
[20]	Discuss the mandatory role of PQC for 6G mobile networks; examine the threats, algorithm families, and quantum-safe transition steps based on inventory.	Cryptographic inventory + step-by-step migration approach; analytical review summarizing NIST FIPS 203/204 and different PQC approaches (grid, code, hash, isogeny, etc.).	Provides a migration roadmap for 5G/6G; embodies the "harvest-now, decrypt-later" risk and urgency of migration; highlights the need for crypto-agility.
[21]	To compile the role of quantum networks (QKD, entanglement, etc.) in the 6G ecosystem and summarize the potential for integration with 6G.	Literature review/survey: technologies, architectures (hybrid classical-quantum, quantum internet, satellite-based), security and applications.	It systematically outlines the rationale for quantum networks for 6G, their architectural options, and challenges (scalability, hardware limitations, cost).
[22]	Summarize the current development of quantum computers and quantum neural networks; discuss performance and research directions.	Literature review; descriptive analysis gathering QNN concepts and applications.	It presents future research directions by compiling the intersection of quantum computing with artificial intelligence and QNN/optimization performance.

2. Conceptual Framework: Mobile Networks and Quantum Technologies

This section conceptually and architecturally frames the intersection of mobile networks and quantum technologies to establish the common ground we will use for the remainder of this paper. We first clarify terminology and scope, outlining the layers of the 5G/6G architecture. We then discuss the fundamental building blocks of quantum communication (QKD, QSDC, entanglement distribution/purification, quantum repeaters, and the classical two-bit requirement of teleportation), along with their latency and interoperability constraints.

2.1. Mobile networks and Quantum Computing: Definitions and Background

Mobile networks are end-to-end managed communication systems where access, transport, and core network layers are integrated with a service-based architecture [23]. The 5G System (5GS) architecture is defined by the separation of user and control planes, network slicing, policy/charging functions, and service interfaces [24], and as this structure evolves towards 6G, it aims for higher prevalence, reliability, and delay predictability [25]. While 3GPP's core architecture definition for 5GS (TS 23.501) and the service requirements in cyber-physical control applications (URLLC; TS 22.104) provide the technical foundations of this framework, the ETSI MEC (Multi-access Edge Computing) ecosystem that enables processing close to the application and the IETF Deterministic Networking (DetNet) approach that sets upper bounds for delay/queuing/jitter in the transport plane play a critical role in adapting mobile networks to scenarios that require "deterministic" operation. When these components are taken together, the mobile network is defined as a holistic service platform with not only wireless access but also edge computing, synchronization/positioning and transport plane guarantees.

Quantum computing is a computing paradigm that processes information on quantum bits (qubits) instead of classical ones. Properties of qubits, such as superposition and entanglement, enable quantum circuits and algorithms that promise exponential gains for certain problems [26]; the standard reference work in the field systematically introduces these concepts through qubit gates, measurement, and circuit models. The quantum teleportation protocol, which uses entanglement as a "source," lays the foundation for quantum communication by demonstrating that a quantum state can be transferred to distant ends using two classical bits. However, the noisy and mesoscale (NISQ) nature of current devices means that fully fault-tolerant universal machines are still in the research phase, and quantum advantages are particularly sought in certain optimization, simulation, and learning tasks in the near future [27]. This table also outlines the technical limits that will determine how quantum will be implemented in practice at the security (e.g., quantum key distribution/new cryptography), communication, and computation levels.

2.2. 6G/Mobile Network Architecture and Deterministic Transport

The distinguishing feature of 6G is that it introduces deterministic behavior into the access-transport-core chain, making end-to-end latency and reliability predictable. In ranges that 5G cannot reach (e.g., long-distance sub-ms latency targets and non-terrestrial links), the 6G/6GB vision explicitly targets deterministic low latency support over a wide area, with time synchronization and planned redundancy [12]. Towards this goal, the coupling of quantum applications to the classical transport layer is provided by standards such as TSN (L2) and DetNet (L3); the 6G architecture aims to integrate these two approaches into the transport plane, providing end-to-end guarantees for quantum workloads as well as for very sensitive traffic such as holograms and tactile internet [12]. Architectural realities on the quantum side reinforce this requirement: due to the no-cloning and fragile nature of qubits, quantum state transport is practically performed via teleportation; While entanglement is distributed over quantum connections, Bell measurement results are necessarily transmitted as two bits over the classical network. Therefore, it is particularly emphasized in the literature that the control/operation planes of quantum networks will remain "predominantly classical" and that quantum implementations (e.g., BQC) will impose strict delay budgets on the classical channel [12]. Figure 2 shows the entanglement generation, repeaters, and classical information flow between Alice and Bob nodes together; the deterministic properties of 6G transport are necessary to keep the reliability and delay of this classical flow within bounds [12].

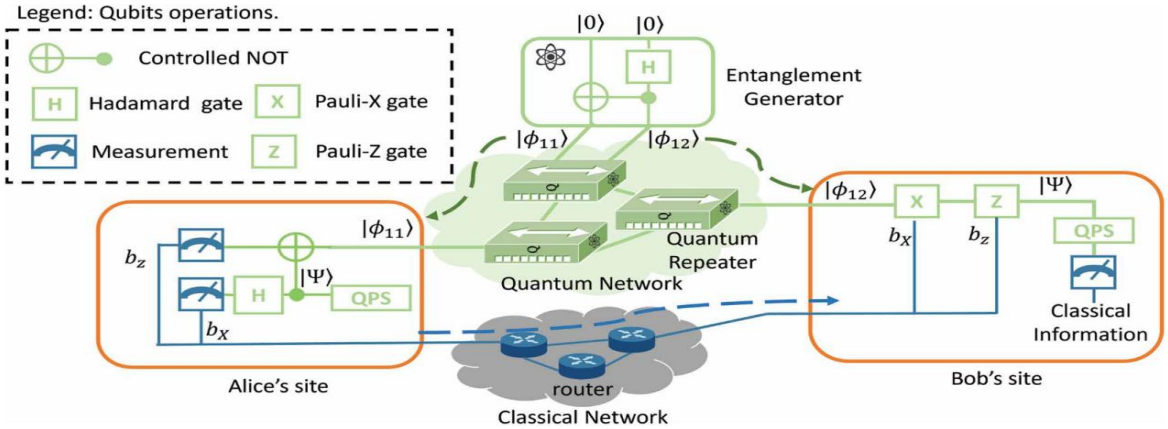


Figure 2. Quantum network model and qubit teleportation; flow of measurement bits supported by classical network [12].

This body of literature is reinforced by two additional lines: (i) a recent review summarizing the 6G context highlights that the integration of quantum communication with classical infrastructure is essential for secure and scalable architectures, and that hybrid classical-quantum architectures are the dominant option in 6G [21]; in the same source, entanglement-assisted networks and protocols such as reliable entanglement distribution (REDP) are cited as fundamental building blocks to address the performance/reliability limits of quantum transport [13], [14]. (ii) 6G-assisted PoP findings quantitatively demonstrate that quantum applications can impose low latency requirements that 5G cannot achieve in the classical channel; the same study points out that repeater experimentation operating in a 50 km channel on the telecom wavelength is progress, but challenges on the transport and device side remain [12]. These two lines confirm that 6G's deterministic transport (including TSN/DetNet) plays a co-factor role for quantum services to actually work.

2.3. Quantum Communication Components and Network Operations

Quantum communication is built on three fundamental pillars underlying the end-to-end architecture that will serve the application layer: (i) quantum key distribution (QKD) and entanglement-based communication, (ii) network operations based on entanglement purification exchange and quantum repeaters, and (iii) quantum memories and limited coherence times. In the context of 6G, these pillars constitute the core of hybrid classical-quantum networks, carrying together the objectives of security, latency, and scalability [21].

The BB84 flow summarized in Figure 3 is based on the principles of randomly encoding individual photons in one of two orthogonal bases, performing measurements on the same random base at the receiver, and comparing the bases in the open channel, converting only the matching measurements into a "filtered key" [28]. The entanglement-based E91 family, on the other hand, structures key generation through shared entanglement and is a natural extension of the architecture [29]. Recent reviews of the 6G ecosystem place QKD and entanglement-based communication among the core technologies of quantum networks and key security enablers for 6G, emphasizing the need for quantum repeaters for range extension.

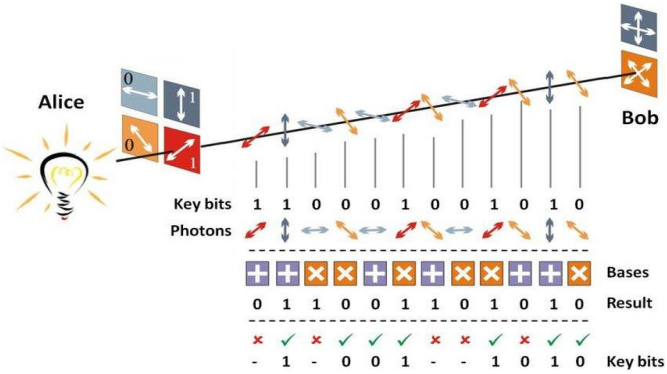


Figure 3. Simplified operation of the BB84 key distribution protocol.

Unlike QKD, Direct Quantum Secure Communication (QSDC) carries the message directly over the quantum channel, thus offering a security framework that does not rely on key negotiation [15]. A recent study details the implementation of an optical network consisting of application blockchain node components for a QSDC-based distributed blockchain and demonstrates that the architecture can be reconfigured with resource monitoring over SDN. The same study presents the privacy bitrate-distance curve for a multi-intensity QSDC protocol over a finite block size, assuming a 1 GHz repetition rate; practical performance records report 100 bps over 100 km over fiber and ~1 kbps per edge for 15 users over 40 km [15]. These findings demonstrate that QSDC can provide kbps-class secure transmission over tens of kilometers with today's infrastructure and can be connected to real-world network operations with SDN.

Entanglement distribution, purification, repeaters, and teleportation. To create an end-to-end “virtual” connection in quantum networks, entanglement is first generated in neighboring links; then, entanglement swapping is used to extend the entanglement between nodes, and purification is applied to address fidelity drops. Routing and resource management embed these physical operations into the network workflow: for mobile scenarios, the proposed Q-CAST Mobility runs entanglement swapping as a phase of path establishment, evaluates performance metrics (reliability, latency, throughput) in this context, and re-optimizes by monitoring link lifetimes (T_life). For global scale, the performance of quantum repeaters is decisive, and repeater chains extend distance by swapping and purification; however, efficiency and memory limitations are still major obstacles. The practical method for state migration is teleportation: entanglement is shared over the QN, the results of the Bell measurement are transmitted over the classical network, and the state is re-established by corrections at the receiver; The classical side channel is mandatory due to fragile qubits and non-replicability, and the delay of this channel directly affects the operability of quantum services.

Quantum memory and coherence time. Entanglement generation, purification, and multi-hop exchange chains involve a series of classical-quantum interactions, such as measurement, feedback, and resetting. The total latency of these interactions rivals the coherence times of the states waiting at intermediate nodes; therefore, network protocols track qubit lifetimes (T_life) as an explicit state variable and use this quantity in path selection/refresh. Current frameworks emphasize that quantum memory efficiency and lifetime are among the most critical bottlenecks for practical deployment. Interoperability with classical transport. Today's architectural reality is that control/operation planes will remain predominantly classical; while entanglement distribution is handled via QN, the classical information required for teleportation must be transported with deterministic low latency. Therefore, the integration of quantum networks with 6G transport technologies is a prerequisite not only for cost and prevalence but also for protocol-level functionality.

3. Conclusion

This study systematically outlines the point where quantum algorithms' disruptive impact on classical cryptography meets 6G's end-to-end deterministic latency and high reliability objectives. As the literature demonstrates, the Shor and Grover-based threat model expands not only public-key schemes but also key lookup and hash-based attack surfaces, making it necessary to rethink mobile networks for the quantum era. In parallel, the 6G/6GB vision's goal of providing deterministic low latency across a wide area through time synchronization and planned redundancy makes the classical side channel required by quantum services (e.g., two-bit measurement results of teleportation) a design prerequisite for carrying within delay/jitter budgets. In this context, the study summarizes 6G-quantum convergence in four pillars (QKD/QSDC based physical security, deterministic transport with TSN/DetNet and advanced URLLC, transition to NIST approved PQC, and management with SDN/slicing) and emphasizes the need for the joint design of the "quantum routing-deterministic classical transport-PQC" triangle in mobile quantum networks.

Our comparative analysis highlights both the maturing security principles and the practical bottlenecks of the field. The demonstrated transmission capability of QSDC as well as QKD over 100 km of fiber and the performance reaching kbps scale in multi-user installations confirm that “keyless” information theoretic security-based applications (e.g., blockchain, quantum-safe signatures, and consensus) can be integrated with today’s available optical components. Furthermore, while the establishment of end-to-end quantum connections is shaped by the efficiency of entanglement generation, clearing, and purification steps, and the limited coherence times of quantum memories, the deterministic behavior of the classical transport accompanying these operations is decisive for functionality. Therefore, it is assumed that TSN/DetNet supported 6G transport plays a co-factor role in terms of real-time operability of quantum applications; in mobile scenarios, entanglement continuity and throughput can be increased with multipath and hybrid proactive/reactive updates such as Q-

CAST Mobility; In the broader context, the findings that REDP-like approaches offer tangible contributions to reliable/scalable entanglement deployment stand out. Furthermore, to mitigate the "listen-then-decipher" risk, the gradual transition to PQC families standardized by NIST in FIPS 203/204/205 requires a full-stack cryptographic inventory and agility, extending from certificates to TLS/IPsec and access protocols. Considering all these axes together, the benefit of this work is that it provides a viable reference framework for mobile quantum networks that combines physics-based security, deterministic classical transport, PQC migration, and software-defined management in the same design space. Priority directions for future work include end-to-end experiments that combine deterministic latency targets with quantum memory coherence times, near-field evaluations of multi-session entanglement routing in mobile/topology-agitated environments, and quantifying the middleware impact of PQC migration in the 6G core/access stack.

In conclusion, this work provides a concrete roadmap for the field by integrating security, timing, and management in a common design space at the intersection of 6G and quantum communication; the proposed framework brings together physics-based security (QKD/QSDC), deterministic classical transport (TSN/DetNet and advanced URLLC), gradual transition to post-quantum cryptography, and software-defined/slicing-based management in mobile quantum networks into a single and consistent architecture.

4. References

[1] Bennett Charles H., and Gilles Brassard,"Quantum cryptography: Public key distribution and coin tossing.", Theoretical computer science, 2014.

[2] Shor, Peter W., "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer.", SIAM review, 1999.

[3] Grover, Lov K., "A fast quantum mechanical algorithm for database search.", Proceedings of the twenty-eighth annual ACM symposium on Theory of computing, 1996.

[4] Satoshi Nakamoto, "Bitcoin: A peer-to-peer electronic cash system.", 2008.

[5] Evgeniy O. Kiktenko, et al, "Quantum-secured blockchain.", Quantum Science and Technology, 2018.

[6] Z.-Z. Sun et al., “Quantum Blockchain Relying on Quantum Secure Direct Communication Network,”, IEEE Internet of Things Journal, vol. 12, no. 10, pp. 14375–14386, 2025.

[7] D. Pan et al., “The evolution of quantum secure direct communication: On the road to the qinternet,” IEEE Commun. Surveys & Tutorials, vol. 26, no. 3, 2024.

[8] Haoran Zhang, et al, "Realization of quantum secure direct communication over 100 km fiber with time-bin and phase quantum states.", Light: Science & Applications, 2022.

[9] J. Wu, G.-L. Long, and M. Hayashi, “Quantum secure direct communication with private dense coding using a general preshared quantum state,” Phys. Rev. Applied, 2022.

[10] NIST, “FIPS 203 (ML-KEM), FIPS 204 (ML-DSA), FIPS 205 (SLH-DSA),” Aug. 2024.

[11] J. Prados-Garzón, T. Taleb, L. Chinchilla-Romero, and M. Shokrnezhad, “Deterministic 6GB-Assisted Quantum Networks With Slicing Support,” IEEE Network, 2023.

[12] Z. Li et al., “Entanglement-assisted quantum networks: Mechanics, enabling technologies, challenges, and research directions,” IEEE Commun. Surveys & Tutorials, 2023.

[13] L. Chen et al., "REDP: Reliable Entanglement Distribution Protocol Design for Large-Scale Quantum Networks," in IEEE Journal on Selected Areas in Communications, vol. 42, no. 7, pp. 1723-1737, July 2024, doi: 10.1109/JSAC.2024.3380101.

[14] D. Abreu, A. Pimentel and A. Abelém, "Q-CAST Mobility: Multipath Routing Protocol for Mobile Quantum Networks," 2025 International Wireless Communications and Mobile Computing (IWCMC), Abu Dhabi, United Arab Emirates, 2025, pp. 1366-1371, doi: 10.1109/IWCMC65282.2025.11059645.

[15] Zygmunt Haas, "The zone routing protocol (ZRP) for ad hoc networks.", Internet Draft draft-zone-routing-protocol, 1998.

[16] Charles Perkins, Elizabeth Belding-Royer, and Samir Das. “Ad hoc on-demand distance vector (AODV) routing”, Optimized Link State Routing Protocol, 2003.

[17] Ling Zhang, and Qin Liu, "Optimisation of the routing protocol for quantum wireless Ad Hoc network.", IET Quantum Communication, 2022.

[18] Little Lion Scientific, "Energy efficient routing in quantum flying ad hoc network (q-fanet) using mamdani fuzzy inference enhanced dijkstra's algorithm (mfi-eda).", Journal of Theoretical and Applied Information Technology, 2024.
[19] Sogo Pierre Sanon, and Hans D. Schotten, "Securing Mobile Networks in the Quantum Era: Imperative Role of Post-Quantum Cryptography.", 2025 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit), IEEE, 2025.
[20] L. Qian and J. Zhao, "Quantum Networks in 6G Communications: Technologies, Challenges, and Applications," 2025 International Conference on Quantum Communications, Networking, and Computing (QCNC), Nara, Japan, 2025, pp. 701-703, doi: 10.1109/QCNC64685.2025.00119.
[21] Z. Huang, L. Qian and D. Cai, "Analysis on the recent development of quantum computer and quantum neural network technology," 2022 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), Dalian, China, 2022, pp. 680-684, doi: 10.1109/ICAICA54878.2022.9844614.
[22] Alliance, N. G. M. N., "5G white paper.", Next generation mobile networks, white paper, 2015.
[23] J. Ping, "Network Resource Model for 5G Network and Network Slice," in Journal of ICT Standardization, vol. 7, no. 2, pp. 127-140, 2019, doi: 10.13052/jicts2245-800X.724.
[24] P. Vetter, "An Updated Research Vision of the Next Generation Network After Several Years on the 6G Journey," in IEEE BITS the Information Theory Magazine, vol. 3, no. 1, pp. 5-13, March 2023, doi: 10.1109/MBITS.2023.3344377.
[25] Nielsen, M. A., & Chuang, I. L., "Quantum computation and quantum information", Cambridge university press, 2010.
[26] Cheng B., Deng X. H., Gu, X., He, Y., Hu, G., Huang, P., ... & Yu, D., "Noisy intermediate-scale quantum computers.", Frontiers of Physics, 2023.
[27] Bennett C. H., & Brassard G., "Quantum cryptography: Public key distribution and coin tossing.", Theoretical computer science, 2014.
[28] Ekert A. K., "Quantum cryptography based on Bell's theorem.", Physical review letters, 1991.

A Particle Swarm Optimization Approach to Federated Bearing Fault Diagnosis

Musa YENİLMEZ[1] , Ilhan AYDIN[2]

[1,2]Department of Computer Engineering, Firat University, Elazig, Turkey.
[1]myenilmez@firat.edu.tr, [2]iaydin@firat.edu.tr
[1](ORCID: 0009-0006-1722-3050), [2](ORCID: 0000-0001-6880-4935)

Abstract
Electric motors constitute a critical component in numerous industrial and commercial systems due to their capability to efficiently convert electrical energy into mechanical work with minimal energy loss and low fault incidence. Their applications span a broad spectrum including industrial fans, pumps, transportation, power tools, medical devices, smart home and household appliances. However even with relatively low failure probabilities bearing related faults can cause substantial downtime and financial losses if not detected promptly. In this study presents a Particle Swarm Optimization enhanced Federated Learning framework for efficient and privacy-preserving bearing fault classification. In the proposed approach multiple clients independently train local models on their proprietary datasets without transferring raw data thereby ensuring data confidentiality. The global model is iteratively aggregated on a central server with PSO employed to optimize hyperparameters and model weights leading to faster convergence and improved classification performance. Experimental evaluation conducted on the Case Western Reserve University bearing fault dataset demonstrates that the proposed particle swarm optimization integrated federated learning method achieves superior accuracy compared to conventional FL approaches. The results confirm its capability to enable early fault diagnosis mitigate potential breakdowns and reduce operational losses in real world electric motor applications.

Keywords: federated learning, particle swarm optimization, bearing fault diagnosis, electric motors, fault classification

1. Introduction

Electric motors are indispensable components in modern industrial, commercial and domestic applications due to their capability to efficiently convert electrical energy into mechanical power with high reliability and low operational costs[1]. Their widespread adoption spans diverse sectors including manufacturing, transportation, agriculture, medical equipment, smart home technologies and household appliances. Despite their robust design and generally low failure rates electric motors remain susceptible to mechanical and electrical faults particularly in bearings which constitute a significant portion of motor failures. Undetected bearing faults can lead to unplanned downtime, increased maintenance costs and severe disruptions in critical operations.

Traditional bearing fault diagnosis methods often rely on centralized machine learning models that require aggregating data from multiple facilities into a single repository. While this approach can improve diagnostic accuracy by leveraging diverse datasets it introduces several critical challenges such as data privacy concerns, regulatory compliance constraints and the high communication costs associated with transmitting large volumes of raw sensor data. These limitations have motivated the exploration of Federated Learning(FL) a distributed machine learning paradigm that enables collaborative model training without direct sharing of raw data[2][3]. FL ensures that data remains localized at the client side with only model parameters exchanged thereby enhancing privacy and compliance with industrial data governance policies[4].
However federated learning models face several optimization challenges, including slow convergence rates, susceptibility to local minima and performance degradation in heterogeneous data environments. To address these limitations recent research has focused on integrating metaheuristic optimization algorithms into FL workflows[5]. Among these Particle Swarm Optimization (PSO) has emerged as a promising candidate due to its simplicity, computational efficiency and strong global search capabilities. Inspired by the collective behavior of biological swarms PSO dynamically adjusts candidate solutions based on both individual and collective performance making it particularly effective for tuning hyperparameters and refining model weights in nonconvex optimization problems.

In this study we propose a PSO enhanced Federated Learning framework for bearing fault classification in electric motors. The proposed method leverages FL to enable decentralized training across multiple clients ensuring data privacy while employing PSO to optimize model parameters at the global aggregation stage thereby improving accuracy and

[1]Corresponding author

convergence speed. We validate the effectiveness of the proposed approach using the Case Western Reserve University (CWRU) bearing fault dataset[6], which contains vibration signals under various load conditions. Experimental results demonstrate that the integration of PSO with FL significantly enhances classification performance compared to conventional FL approaches achieving higher accuracy and robustness in real world industrial scenarios.

2. Background

2.1. Federated Learning

Federated learning is a decentralized machine learning paradigm that enables multiple clients to collaboratively train a shared model without transferring their raw data to a central server[7][8]. Instead each client trains the model locally on its own dataset and sends only the model parameters or gradient updates to the server. The server aggregates these updates typically using algorithms such as Federated Averaging (FedAvg) and sends the updated global model back to the clients for further training. This iterative process continues until the model converges.

One of the key advantages of FL is data privacy preservation as sensitive or proprietary datasets remain on the client side reducing the risk of data leakage and ensuring compliance with regulations such as GDPR or HIPAA. FL also enables collaboration across heterogeneous data sources which improves the generalization capability of the model by incorporating diverse patterns and conditions from multiple environments.

2.2. Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a population based stochastic optimization algorithm inspired by the social behavior of bird flocks. It was introduced by Kennedy and Eberhart in 1995 as a method to efficiently explore and exploit search spaces for finding optimal or near optimal solutions to complex problems[9].

In PSO a swarm of candidate solutions referred to as particle navigates the search space by updating their positions and velocities according to two main influences:

Cognitive component: Each particles tendency to return to its own best known position (*personal best*, pBest).
Social component: Each particles tendency to move toward the best known position found by the entire swarm (*global best*, gBest). The velocity(v_i) (1) and position(x_i) (2) updates at iteration t are governed by the following equations:

$$v_i^{(t+1)} = w \times v_i^{t} + c_1 \times r_1 \times (pBest_i - x_i^{t}) + c_2 \times r_2 \times (gBest - x_i^{t}) \quad (1)$$

$$x_i^{(t+1)} = x_i^{t} + v_i^{(t+1)} \quad (2)$$

where w is the inertia weight controlling exploration versus exploitation c_1 and c_2 are acceleration coefficients and r_1 and r_2 are random values between 0-1.

When integrated with Federated Learning PSO can be used to fine-tune global model parameters during the aggregation phase enhance convergence speed and improve overall classification accuracy [10]. Particularly in heterogeneous data environments such as those encountered in industrial fault diagnosis.

3. The Proposed Approach

3.1. Data Acquisition and Preprocessing

In this study Case Western Reserve University(CWRU) dataset is used. The CWRU dataset is a comprehensive dataset of vibration signals frequently used as a reference for diagnosing bearing faults. This dataset was obtained by Case Western Reserve University's School of Engineering using experimental test benches. Faults were deliberately applied to bearing components (inner race, ball, or outer race) using electro-erosion (EDM) and vibration data were recorded at various engine

loads. CWRU dataset link(https://engineering.case.edu/bearingdatacenter/download-data-file)[11]. All features of dataset is seen on Table 1.

Table 1. Dataset Description

Feature	Description
Bearing Locations	Drive end, Fan end
Fault Types	Inner race fault, Outer race fault (3 positions), Ball fault, Normal
Fault Creation Method	Electric Discharge Machining (EDM)
Fault Sizes	0.007 in, 0.014 in, 0.021 in
Operating Speeds	1797, 1772, 1750, 1730 RPM
Motor Loads	0, 1, 2, 3 horsepower
Sampling Frequencies	12 kHz, 48 kHz
Sensors	Accelerometers mounted on drive and fan ends

This dataset consist of float values in time domain as MATLAB(.mat) format. Recorded viration data is converted into image dataset with time and amplitude attribute. Than sampled 1000 images for each fault types. Finally dataset divided into 2 parts for training(%80) and testing(%20). An example of new image dataset is seen below Figure 1.

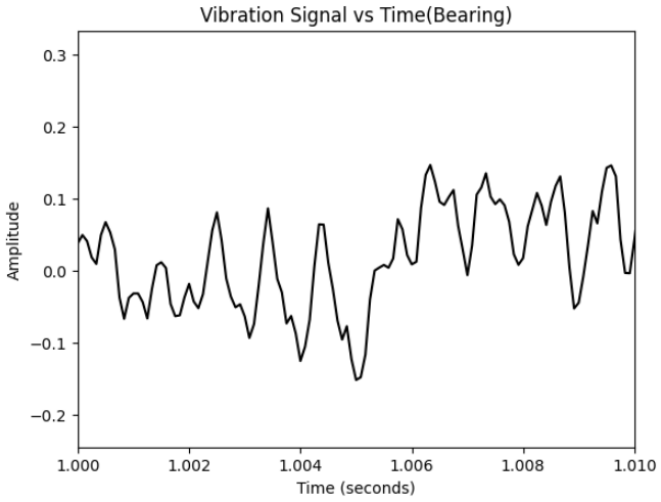


Figure 4. Example of vibration signal converted into image

CWRU experimental setup is seen on Figure 1. This setup have the dynamometer on the right, a torque transducer/encoder in the middle and 3 hp motor left. This is most used dateset by the researcher for fault diagnosis.

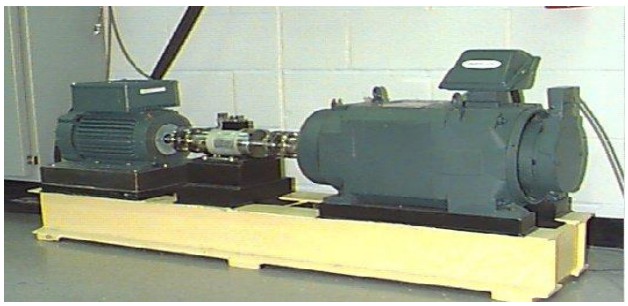


Figure 5. Experimental setup of CWRU dataset[12]

3.2. Federated Learning and Particle Swarm Optimization for Model Aggregation

Incorporating PSO into FL aims to enhance the aggregation process by optimizing the global model parameters rather than relying solely on simple averaging. Each particle in PSO can represent a potential global model configuration. Local model updates from clients are used to evaluate the fitness of each particle, for instance, based on validation accuracy or loss. PSO iteratively adjusts the global model weights to find an optimal aggregation that maximizes performance while respecting client contributions. Our proposed method firstly trainnig personal model after senndig parameters to server. After parameter optimization happening on server with PSO. Then best parameters sending to clients. The block diagram of proposed approach is seen on Figure 2.

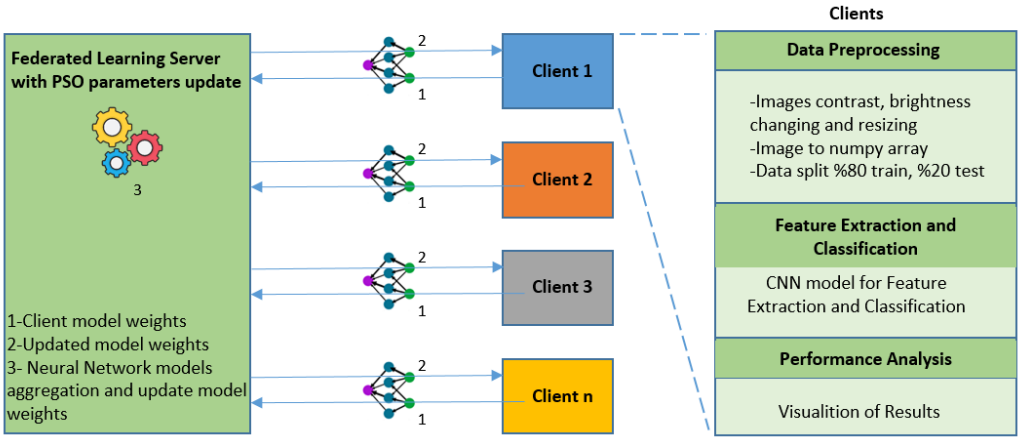


Figure 6. Block diagram of proposed approach

4. Experimental Results

The proposed Particle Swarm Optimization–enhanced Federated Learning framework was evaluated using the CWRU bearing fault dataset. The dataset was preprocessed by transforming raw vibration signals into image based representations enabling the model to capture discriminative temporal and frequency domain features. During training %80 of the dataset was used for model development while %20 was reserved for testing to ensure unbiased evaluation. Results showed that the integration of PSO into the federated aggregation process significantly improved both convergence speed and final classification accuracy when compared with baseline federated averaging methods. The performance results obtained from the experiments performed using the proposed method and the cwru dataset are shown in Table 2. CWRU dataset and FL approach without PSO results taken our another study[13].

Table 2. CWRU dataset classification results

Base Model	FL without PSO Accuracy(%)	FL with PSO Accuracy(%)
Xception	97.19	97.65
VGG16	**99.06**	**99.47**
VGG19	98.69	99.03
ResNet50	97.19	97.24
MobileNet	92.56	94.14
MobileNetV2	88.06	88.91
DenseNet169	96.88	97.45
Inception V3	97.81	98.32

5. Conclusions

In this study introduced a federated learning framework enhanced with Particle Swarm Optimization to address the challenges of bearing fault diagnosis in electric motors. By leveraging federated learning the approach ensured data privacy

and compliance with industrial data governance policies while PSO played a key role in refining the aggregation process and accelerating model convergence. The experimental evaluation conducted on the CWRU dataset demonstrated that the proposed method achieved superior fault classification accuracy and robustness compared to conventional federated learning techniques. A success rate of 99.47% was achieved with optimum parameters and VGG16 based neural network integradted PSO. These outcomes highlight the framework capability to enable early and reliable detection of bearing faults thereby minimizing unplanned downtime and financial losses in industrial environments.

6. References

[1] H. Zhang and B. Wang, "Supervised Contrastive Learning for Electric Motor Bearing Fault Detection," *2024 Int. Conf. Electr. Mach. ICEM 2024*, 2024, doi: 10.1109/ICEM60801.2024.10700325.

[2] A. Sen, S. H. Heng, and S. C. Tan, "A comprehensive Review of Cryptographic Techniques in Federated Learning for Secure Data Sharing and Applications," *IEEE Access*, 2025, doi: 10.1109/ACCESS.2025.3593953.

[3] G. Chandu, T. Karthik, and B. Parag, "Federated Learning for Distributed IoT Security: A Privacy-Preserving Approach to Intrusion Detection," *IEEE Access*, 2025, doi: 10.1109/ACCESS.2025.3592481.

[4] G. Y. Huang and C. H. Lee, "Federated learning architecture for bearing fault diagnosis," *Proc. 2021 Int. Conf. Syst. Sci. Eng. ICSSE 2021*, pp. 408–411, Aug. 2021, doi: 10.1109/ICSSE52999.2021.9538492.

[5] Z. Li, H. Li, and M. Zhang, "Hyper-parameter Tuning of Federated Learning Based on Particle Swarm Optimization," *Proc. 2021 7th IEEE Int. Conf. Cloud Comput. Intell. Syst. CCIS 2021*, pp. 99–103, 2021, doi: 10.1109/CCIS53392.2021.9754676.

[6] K. K. Raj, S. Kumar, R. R. Kumar, and M. Andriollo, "Enhanced Fault Detection in Bearings Using Machine Learning and Raw Accelerometer Data: A Case Study Using the Case Western Reserve University Dataset," *Inf.*, vol. 15, no. 5, 2024, doi: 10.3390/info15050259.

[7] O. R. A. Almanifi, C. O. Chow, M. L. Tham, J. H. Chuah, and J. Kanesan, "Communication and computation efficiency in Federated Learning: A survey," *Internet of Things*, vol. 22, p. 100742, Jul. 2023, doi: 10.1016/J.IOT.2023.100742.

[8] S. P. Karimireddy, N. R. Veeraragavan, S. Elvatun, and J. F. Nygård, "Federated Learning Showdown: The Comparative Analysis of Federated Learning Frameworks," *2023 8th Int. Conf. Fog Mob. Edge Comput. FMEC 2023*, pp. 224–231, 2023, doi: 10.1109/FMEC59375.2023.10305961.

[9] J. Kennedy, R. Eberhart, and bls gov, "Particle Swarm Optimization".

[10] G. G. Devarajan, S. M. Nagarajan, A. Daniel, T. Vignesh, and R. Kaluri, "Consumer Product Recommendation System Using Adapted PSO with Federated Learning Method," *IEEE Trans. Consum. Electron.*, vol. 70, no. 1, pp. 2708–2715, Feb. 2024, doi: 10.1109/TCE.2023.3319374.

[11] "Case Western Reserve University |Download Data File." [Online]. Available: https://engineering.case.edu/bearingdatacenter/download-data-file

[12] W. A. Smith and R. B. Randall, "Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study," 2015, doi: 10.1016/j.ymssp.2015.04.021.

[13] M. Yenilmez and I. Aydin, "Federated Learning-Based Bearing Fault Classification," *2024 Int. Conf. Innov. Intell. Informatics, Comput. Technol. 3ICT 2024*, pp. 202–206, 2024, doi: 10.1109/3ICT64318.2024.10824453.

Sentiment Analysis of Turkish Song Lyrics Using Machine Learning Algorithms

Merve ÇALOĞLU[1] , Recep GÜLTER[2], Mehmet ARZU[3], Mahmut KAYA[4], Muhammed Emre ÇOLAK[5]

[1,2,3,4,5] Yapay Zeka ve Veri Mühendisilği Bölümü, Fırat Üniversitesi, Elazığ, Türkiye.
[1]230291031@firat.edu.tr, [2]23029108@firat.edu.tr, [3]marzu@firat.edu.tr, [4]mahmutkaya@firat.edu.tr, [5]memrecolak@firat.edu.tr
[1](ORCID: 0009-0007-7039-8674), [2](ORCID: 0009-0002-3610-5608), [3](ORCID: 0000-0001-6610-2788),[4](ORCID: 0000-0002-7846-1769), [5](ORCID: 0000-0003-2317-2385)

Abstract
Emotion classification in Turkish texts has become an increasingly important research topic in the fields of natural language processing and machine learning in recent years. In particular, the analysis of emotionally charged texts such as song lyrics can provide valuable contributions to recommendation systems, psychological assessment tools, and user-focused digital applications. In this study, a classification model was developed to predict individuals' moods (happy, sad, angry) based on song lyrics. Approximately 3,250 Turkish song lyrics collected via the Genius API were manually labeled and processed through various preprocessing steps. Both classical machine learning algorithms (Logistic Regression, Naive Bayes, Support Vector Machines, Random Forest, Decision Trees, K-Nearest Neighbors) and pre-trained transformer-based models (BERTurk, ELECTRA, DistilBERT) were evaluated on the obtained dataset. To improve model performance, early stopping, K-Fold cross-validation, and majority voting methods were applied. Experimental results were analyzed using basic classification metrics such as accuracy and F1 score; it was observed that transformer architectures, which have a high ability to capture contextual relationships, offer higher accuracy rates compared to classical methods. The findings reveal that context-sensitive models are an effective alternative for sentiment classification in morphologically rich languages and suggest that this approach could contribute to the development of systems that better evaluate the structural and semantic features of language.

Keywords: emotion classification, song lyrics, turkish natural language processing, transformer models

1. Introduction

Music is one of the most universal ways for humans to express their emotions and share them with others. Lyrics deepen this experience; they contain expressions that reflect a person's inner world and often have multiple layers of meaning. Today, listeners form a strong emotional connection not only with melodies but also with lyrics, and sometimes even find their own moods reflected in these lyrics.

Therefore, automatically detecting emotions in song lyrics is not just an academic research topic; it can be a functional tool in many areas, from music recommendation systems to digital therapy applications. Sentiment analysis is a natural language processing (NLP) method used to identify and classify emotional expressions in texts. While these techniques are primarily used in more everyday and direct texts such as social media posts, product reviews, and news content, they can also yield important data about people's emotional states. However, when it comes to artistic texts such as song lyrics, its use becomes more complex. The main reason for this is that metaphors, ambiguous expressions, and cultural references make interpretation difficult. Indeed, the limited number of studies on Turkish song lyrics in the literature highlights the need for innovative approaches in this field.

Turkish, with its agglutinative structure and rich morphological features, requires special attention in natural language processing. The large number of morphological elements added to word roots, flexible syntax, and context-dependent meaning changes can reduce the success of traditional machine learning methods. However, in recent years, pre-trained, context-sensitive transformer-based models developed specifically for Turkish (e.g., BERTurk, ELECTRA) have been effective in overcoming many of these limitations.

In this study, classical machine learning methods (e.g., Random Forest, Naive Bayes) and deep learning approaches (e.g., BERTurk, ELECTRA) were evaluated together to predict individuals' moods (happy, sad, angry) based on Turkish song lyrics. with the aim of contributing to sentiment analysis studies on Turkish music data in terms of both methodology and

[1]Corresponding author

application. Our study proposes an integrated approach that can improve sentiment classification accuracy, particularly in texts with high metaphor density and strong cultural context.

2. Related Studies

Erşahin and Uçkan (2018) performed root extraction using the Zemberek library on a dataset consisting of approximately 300 songs, and then extracted n-gram and TF-IDF (Term Frequency-Inverse Document Frequency) representations. The Naive Bayes and Support Vector Machine (SVM) algorithms were tested on these representations, with the highest success rate of 43.7% accuracy and 46.9% F1 score coming from Naive Bayes. However, these low figures suggest that frequency-based methods struggle to capture relationships, especially in context-rich texts [1].

More recently, Dağdeviren (2023) developed the BERTurk model in his work published on the HuggingFace platform. Despite working with a small training set, the model's accuracy approached 90%. This suggests that transformer architectures can be more effective than classical methods when considering the contextual complexity of Turkish [2].

Different approaches are also seen in English song lyrics. Rösch and Rauch (2025) used the concept of "emotional arc" to examine changes in emotions over time in songs and related them to popularity data. With this approach, emotion analysis was not reduced to classification alone; the time dimension was also part of the study [3]. On the other hand, Agrawal and his team (2022) classified song lyrics using XLNet; their findings revealed that contextual methods yielded better results than acoustic-based approaches [4]. Another Turkish-focused study was conducted by Altınel Girgin and colleagues (2024). In this study, dictionary-based, classical, and transformer methods were compared. The findings revealed that the problems caused by the morphological richness of Turkish can be managed more easily, especially in transformer architectures [5].

Similarly, Mutlu and Özgür (2022) developed BERT-based models for goal-oriented sentiment analysis and shared an open-source dataset. The models were able to capture changes in sentiment expressions depending on goals [6]. Additionally, Zümberoğlu and his team (2025) examined the BERTurk, ELECTRA, and BERTweet models based on Turkish social media posts. The study focused on the effects of data quality and structure on performance; it also noted that transformer-based models are more advantageous in complex sentences [7].

Classic methods have also been tested on different datasets recently. For example, Durak and Hark (2024) compared Naive Bayes, SVM, and XGBoost algorithms with TF-IDF representations on e-commerce reviews. In the experiments, SVM achieved approximately 84% accuracy, while XGBoost exceeded 85%. This result suggests that classical methods may still be a strong alternative in certain situations [8]. Kılıçer and Samlı (2023), on the other hand, focused on positive–negative–neutral classification; they reported that the neutral class was difficult to distinguish due to low sentiment intensity and that classical methods performed poorly in this class [9].

Finally, Arzu and Aydoğan (2023) compared various transformer architectures such as BERTurk, ELECTRA, DistilBERT, and ConvBERTurk on a large-scale dataset consisting of balanced Turkish user reviews called TRSAv1. The highest accuracy rate achieved (83.05%) was observed in the ConvBERTurk mC4 (uncased) model; this finding once again confirms that transformer architectures offer high contextual performance in Turkish sentiment classification tasks [10].

3. Material and Method

3.1. Data Set

The dataset used in this study consists of Turkish song lyrics, which were automatically collected via the Genius API. The dataset, which contains a total of 3,250 song lyrics, was selected from Turkish popular music genres and works by various artists. Each data row represents an entire song or a meaningful section of a song.

The dataset has been manually labeled into three basic emotion categories: happy, sad, and angry. During the labeling process, the emotional tone, context, and thematic structure of the song lyrics were taken into consideration. Care was

taken to include a similar number of examples in each category, thus obtaining a balanced dataset. Figure 1 shows the distribution of emotion categories in the dataset.

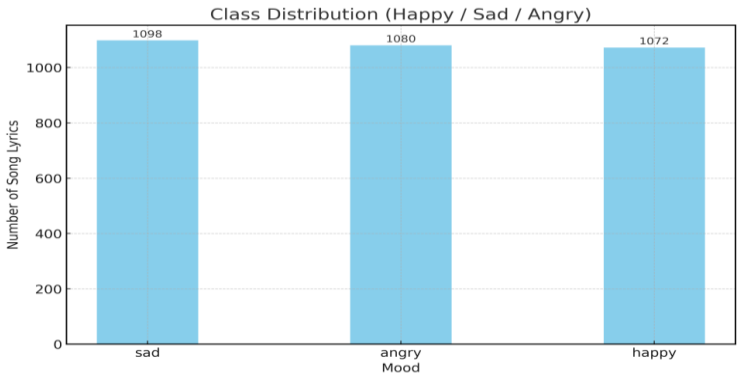


Figure 1. Distribution of sentiment classes in the dataset.

The dataset consists of five columns, each containing various information about song lyrics: the artist column contains the artist who performs the song, the song_title column contains the title of the song, the lyrics column contains the raw (unprocessed) lyrics, the mood column contains the emotion class to which the song belongs, and the clean_lyrics column contains the simplified lyrics that have undergone preprocessing steps. Figure 2 shows an example of a section from the dataset.

artist	song_title	lyrics	clean_lyrics	mood
Onur Can Özcan	Hırka	Sen dağıttın bak Ben topluyorum beni Madem anl...	sen dağıttın bak ben topluyorum beni madem anl...	üzgün
Nükhet Duru	Cambaz	1 ContributorCambaz Herkes toplanmış Bayram ye...	herkes toplanmış bayram yerinde geziniyorlar e...	mutlu
Simge	Ne Zamandır	Ne Zamandır Eskisi gibi değil ne bakışın ne gü...	zamandır eskisi değil bakışın gülüşün artık is...	mutlu
mor ve ötesi	Park	dün neler mi kaybettin belki zamanın yok şimdi...	dün neler mi kaybettin zamanın yok şimdi giden...	mutlu
Dolu Kadehi Ters Tut	Karanlık	Ey protest halkım! Bana ne yaptın? İçinizden h...	ey protest halkım bana yaptın içinizden hiçbir...	üzgün

Figure 2. Sample dataset

3.2. Method

This section provides a detailed explanation of the preprocessing steps applied to the dataset, feature extraction methods, models used, and experimental process. Common classification metrics such as accuracy and F1-score were used to evaluate model performance.

All experimental studies were conducted using the Python programming language and Google Colab, a cloud-based development environment.

3.2.1. Data Preprocessing

Before training the model, a series of preprocessing steps were performed on the song lyrics in the dataset. Punctuation marks and special characters in the text were removed, all letters were converted to lowercase, and unnecessary spaces were removed. Then, stop words were removed.

Stop words are words that frequently appear in a language but do not have a clear meaning on their own. Examples of such words in Turkish include “bir (one),” “ve (and),” “ile (with),” “gibi (like/as),” and “için (for). Therefore, a list of stop words appropriate for Turkish was prepared for this study, and these words were then removed from the dataset.

3.2.2. Feature Extraction and Vectorization

In order for text data to be processed using machine learning, it must first be converted into a numerical representation. In this study, the TF-IDF (Term Frequency–Inverse Document Frequency) approach was preferred for classical methods. The mere fact that a word appears frequently in a text does not necessarily mean that it is important. TF-IDF comes into play here: while calculating the frequency of a word, it also considers its distinctiveness in other documents within the dataset. This suppresses frequently used but meaningless words, making important words more prominent.

The preprocessed song lyrics (clean_lyrics) were vectorized using the TF-IDF method and each example was converted into a feature vector of approximately 5,000 dimensions. These numerical representations were used as input data for classical supervised learning algorithms such as Logistic Regression, Naive Bayes, Random Forest, K-Nearest Neighbor (KNN), SVM, and Decision Trees.

Transformer-based models, on the other hand, do not require an external vectorization step. These types of models can process text directly thanks to their own tokenizer structures and contextual embedding layers; they can automatically learn the semantic relationships between words through deep learning layers. Transformer architectures such as BERTurk, DistilBERT, and ELECTRA used in this study convert raw text into multidimensional contextual representations by passing it through their own tokenizer processes. Therefore, external feature extraction methods such as TF-IDF have not been applied to transformer models.

3.2.3. Models Used

In this study, both classical machine learning algorithms and deep learning-based transformer architectures were applied to predict individuals' moods from Turkish song lyrics.

Classic Machine Learning Models

Classic machine learning algorithms are computationally efficient and interpretable models that have been used in text classification problems for a long time [11]. These models work on data converted into numerical form using representations such as TF-IDF. However, their ability to learn contextual relationships is limited in texts with high contextual meaning density, such as song lyrics. Therefore, they were used in this study primarily to conduct a comparative evaluation and observe their performance compared to transformer-based architectures.

Logistic Regression: A linear statistical model commonly used in classification problems. The model estimates the probability of an example belonging to a particular class based on given features. While highly effective in binary classification, it can also be used in multi-class classification. It can achieve successful results in high-dimensional problems when used in conjunction with representations such as TF-IDF [12].

Naive Bayes: This method, based on Bayes' theorem, relies on probability calculations. It generates predictions using word frequency information and assumes that variables are independent of each other. Thanks to its simple structure, it is fast and can be easily applied to large data sets [13].

Decision Trees: This is an algorithm that classifies data by dividing it into branches and is easy to understand visually. It has a clear and intuitive structure. However, when used alone, it may carry the risk of overfitting [14].

Random Forest: This is one of the ensemble methods created by combining multiple decision trees instead of using them alone. Each tree learns from a randomly selected subset of the training data. When making predictions, the results produced by these trees are combined; the final decision is made by majority vote or average. This approach produces more stable results than a single decision tree and increases generalization capacity [15].

K-Nearest Neighbor: When classifying a new example, it looks at the k nearest neighbors in the training set. Its working principle is based on distance measurements. It is a simple method that does not require parameters. However, its performance may weaken in cases where the data is very large or sparse [16].

Support Vector Machines: Aims to separate data using a hyperplane that provides the widest margin between classes. It can deliver effective results in high-dimensional data problems, especially with linear kernel functions. It can demonstrate strong overall performance even with small data sets [17].

Transformer-Based Deep Learning Models

The Transformer architecture was developed in 2017 by researchers at the University of Toronto and Google Brain and introduced in a paper titled "Attention is All You Need" [18]. This architecture uses a self-attention mechanism to learn the relationships between words in input texts. Thus, meaningful representations are created by taking into account the context of each word in relation to other words in the sentence.

Transformer-based models are effectively used in many tasks in the field of natural language processing, such as sentiment analysis, text classification, text summarization, and question answering, thanks to their ability to produce outputs with high contextual awareness [18], [19].

In this study, pre-trained transformer architectures specific to Turkish were preferred, and context-aware classification was targeted. In model selection, different variations of popular Turkish language models such as BERTurk, DistilBERT, and ELECTRA were evaluated; different training strategies such as early stopping, manual K-Fold cross-validation, and HuggingFace Trainer API were applied. In addition, the BERT and ELECTRA models were combined using a majority voting-based ensemble method and tested.

BERTurk: A BERT model specifically designed for Turkish. The "cased" version takes case sensitivity into account, and the model works with a WordPiece vocabulary consisting of approximately 32,000 tokens. Provided by HuggingFace, this model delivers effective results in contextually analyzing Turkish morphological structures [20].

ELECTRA: Offers a structure similar to BERT but uses a "word replacement detection" strategy instead of masking. This method enables more efficient learning. The Turkish-adapted version is case-sensitive and is generally preferred in projects that require fast training with limited data [21].

DistilBERT: A model created by compressing information obtained from the original BERT model, with a more compact structure. It operates with fewer parameters and lower computational costs; however, it may perform less well than BERT in some tasks. It is advantageous in real-time or resource-constrained environments [22].

BERTurk + ELECTRA Ensemble: Classification outputs from BERT and ELECTRA models are combined using majority voting to create a hybrid structure. This method aims to achieve more consistent results by balancing the biases of individual models [23].

In order to prevent overfitting, training was stopped after 3 epochs. During the training phase, the data was divided into K parts and the model was retrained for each k-fold. A one-time training/validation split was performed using HuggingFace's Trainer interface.

4. Results and Discussion

In this study, sentiment classification experiments were conducted using both classical machine learning algorithms and transformer-based deep learning models on a dataset consisting of 3,250 Turkish song lyrics. The results are presented in Table 1, Table 2.

As shown in Table 1, among classical machine learning algorithms, the highest accuracy rate of 62.15% was obtained with the Logistic Regression model. Naive Bayes (61.85%), Random Forest (61.54%), and SVM (60.92%) models showed similar accuracy levels. In these models, the "angry" class was generally classified more accurately, while declines were observed in the "sad" class. The Decision Tree algorithm remained below average with 49.23% accuracy, while the lowest success rate was 34.92% for the KNN algorithm. As can be seen in Table 1, the KNN model made almost no accurate predictions for the "sad" class. This highlights the limitations of methods that cannot model contextual relationships.

Table 1. Success metrics for classical machine learning models

Model	Accuracy	Precision	Sensitivity	F1-Score
KNN	0.34	0.33	0.35	0.27
Decision Tree	0.49	0.50	0.49	0.49
SVM	0.60	0.61	0.61	0.61
Random Forest	0.61	0.62	0.61	0.62
Naive Bayes	0.61	0.63	0.62	0.62
Logistic Regression	0.62	0.63	0.62	0.62

Since classical models rely on vector representations of texts such as TF-IDF, they cannot fully capture the context and cannot directly model subtle details such as word order or emotional tone. Therefore, as shown in Table 1, their success has been limited in song lyrics that require contextual analysis.

As shown in Table 2, the highest accuracy among transformer-based models was achieved by the BERT model trained with the HuggingFace Trainer API. This model outperformed all other models with 73.38% accuracy and a F1 score of 0.736. In particular, the 87% precision and 83% F1 score in the "anger" class demonstrate that the model effectively utilizes contextual awareness.

As can be seen in detail in Table 2, the manually applied 5-fold stratified K-fold cross-validation strategy provided a more stable structure for testing the model's generalizability. In this method, the data set was divided into different subsets, the model was retrained in each layer, and accuracy values were calculated for each. Thus, the model's performance was observed not only on a single data split but across the entire dataset. This approach reduced the risk of overfitting and more reliably demonstrated the model's overall performance level. With an average accuracy of 71.38% and high class-based F1 scores, the K-Fold method emerged as an effective strategy, particularly for datasets with class imbalances.

The version of the BERT architecture with early stopping also delivered balanced results with 70% accuracy. The ELECTRA model showed moderate performance with 68% accuracy and an F1 score of 0.68, while DistilBERT produced the lowest results within the transformer family with 59% accuracy and an F1 score of 0.59. This can be attributed to the model's limited number of parameters and contextual representation.

In addition, the ensemble structure, which combines the BERT (early stopping) and ELECTRA (tuned) models using majority voting, achieved 69% accuracy, particularly standing out with a recall value of 73% in the "happy" class. This hybrid structure provided a more balanced distribution across classes and produced more consistent classification outputs by reducing the biases of individual models.

Table 2. Performance metrics of transformer-based models

Model	Accuracy	Precision	Sensitivity	F1-Score
DistilBERT	0.59	0.62	0.59	0.59
ELECTRA	0.68	0.69	0.68	0.68
BERTurk (Standart)	0.69	0.69	0.69	0.69
BERTurk+ELECTRA (Ensemble)	0.69	0.70	0.69	0.69
BERTurk(Early Stopping)	0.70	0.71	0.70	0.70
BERTurk (K-Fold)	0.71	0.73	0.71	0.71
BERTurk (Trainer API)	0.73	0.74	0.73	0.73

When Table 1, Table 2 are evaluated together, it can be seen that transformer-based models achieve significantly higher success rates compared to classical methods. While classical models mostly remain in the 60% range, transformer models exceed 70%. The highest accuracy was measured at 73.38% in the BERT (Trainer API) model, while the lowest accuracy was 34.92% in the KNN model. In terms of training strategies, the Trainer API and K-Fold methods stood out in terms of both performance and generalizability.

The findings show that transformer architectures perform much better than classical methods in context-rich texts. In particular, BERT-based models have been able to produce more successful results when appropriate training strategies are applied. However, the highest accuracy rate achieved was still only 73%. This situation is related more to the semantic and contextual difficulties inherent in sentiment classification than to the inadequacy of the models.

Interpreting song lyrics is quite difficult because they contain metaphors, expressions with multiple meanings, cultural references, and emotional transitions. In such cases, how the emotion is perceived may vary depending on the context of the text, the style of expression, and the reader's own interpretation. The agglutinative structure of Turkish and its flexible sentence structure also complicate matters. Therefore, the accuracy rates obtained should not be viewed as merely a technical result. Considering the complex structure of language and the fact that emotions can be interpreted in different ways, these findings appear to be consistent in the context of existing linguistic and emotional diversity.

5. Conclusions

In this study, sentiment classification was performed on a dataset specifically created from Turkish song lyrics. The performance of different modeling approaches was compared. The experiments revealed that classical machine learning algorithms are useful to a certain extent, but they are limited in capturing contextual relationships. Transformer-based models (e.g., BERTurk and ELECTRA), on the other hand, achieved higher accuracy values by more effectively processing the morphological structure and context-based features of Turkish.

The findings show that context-sensitive methods are more reliable than traditional models for song lyrics containing metaphors, polysemy, and cultural references. On the other hand, it was also observed that accuracy rates are limited to a certain level due to the interpretive nature and uncertainty of sentiment classification.

Future research plans include creating broader and more balanced datasets, including different music genres in the analysis, and testing the impact of multilingual models on Turkish. Additionally, developing hybrid methods that can utilize contextual information more efficiently is emerging as an important research direction.

6. References

[1] Erşahin Y, Uçkan H. “Music emotion classification for Turkish songs using lyrics”. *Pamukkale University Journal of Engineering Sciences*, 24(7), 1284–1290, 2018.

[2] Dağdeviren E. “Turkish lyrics sentiment classification with BERTurk”. https://huggingface.co/emre-dagdeviren/turkish-lyrics-sentiment (30.06.2025).

[3] Rösch M, Rauch L. “Emotional arcs in song lyrics and success prediction”. *Journal of Cultural Economics*, (In Press), https://doi.org/10.1007/s10824-025-09444-z, 2025.

[4] Agrawal A, Singh R, Yadav A. “Context-aware sentiment classification of English song lyrics using XLNet”. *Expert Systems with Applications*, 196, 116599, 2022.

[5] Altınel Girgin AB, Gümüşçekiççi G, Birdemir NC. “Turkish sentiment analysis: A comprehensive review”. S*igma Journal of Engineering and Natural Sciences*, 42(4), 1292–1314, 2024.

[6] Mutlu H, Özgür A. “A dataset and BERT-based models for targeted sentiment analysis on Turkish texts”. arXiv preprint, https://arxiv.org/abs/2201.11419, 2022.

[7] Zümberoğlu KB, Dik SZ, Karadeniz BS, Sahmoud S. “Towards better sentiment analysis in the Turkish language: Dataset improvements and model innovations”. *Applied Sciences*, 15(4), 2150,2025.

[8] Durak MH, Hark C. “Makine öğrenmesi yöntemlerinin duygu okuma yeteneği: E-ticaret yorumlarını anlamak”. *Journal of Computer Science*, 9(2), 133–141, 2024.

[9] Kılıçer S, Samlı R. “Türkçe ürün yorumları üzerine makine öğrenmesi algoritmaları ile duygu analizi”. *Veri Bilim Dergisi*, 6(2), 15–23, 2023.
[10] Arzu Z, Aydoğan H. “Türkçe metinlerde transformers tabanlı duygu analizi: TRSAv1 veri seti ile kapsamlı bir karşılaştırma”. *Sosyal Bilimler Teknoloji Dergisi*, 5(1), 45–61, 2023.
[11] Sebastiani F. “Machine learning in automated text categorization”. *ACM Computing Surveys (CSUR)*, 34(1), 1–47, 2002.
[12] Hosmer DW, Lemeshow S, Sturdivant RX. *Applied Logistic Regression*. Vol. 398, New York, USA, John Wiley & Sons, 2013.
[13] McCallum A, Nigam K. “A comparison of event models for Naive Bayes text classification”. *AAAI-98 Workshop on Learning for Text Categorization*, Madison, Wisconsin, USA, 752(1), 41–48, 1998.
[14] Quinlan JR. “Induction of decision trees”. *Machine Learning*, 1(1), 81–106, 1986.
[15] Breiman L. “Random forests”. *Machine Learning*, 45(1), 5–32, 2001.
[16] Cover TM, Hart PE. “Nearest neighbor pattern classification”. *IEEE Transactions on Information Theory*, 13(1), 21–27, 1967.
[17] Cortes C, Vapnik V. “Support-vector networks”. *Machine Learning*, 20(3), 273–297, 1995.
[18] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. “Attention is all you need”. *Advances in Neural Information Processing Systems*, 30, Long Beach, California, USA, 2017.
[19] Devlin J, Chang MW, Lee K, Toutanova K. “BERT: Pre-training of deep bidirectional transformers for language understanding”. *NAACL-HLT*, Minneapolis, Minnesota, USA, 2019.
[20] Schweter S. “BERTurk: BERT models for Turkish”. https://github.com/stefan-it/turkish-bert (30.06.2025).
[21] Clark K, Luong MT, Le QV, Manning CD. “ELECTRA: Pre-training text encoders as discriminators rather than generators”. arXiv preprint arXiv:2003.10555, 2020.
[22] Sanh V, Debut L, Chaumond J, Wolf T. “DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter”. arXiv preprint arXiv:1910.01108, 2019.
[23] Zhou ZH. *Ensemble Methods: Foundations and Algorithms*. Boca Raton, Florida, USA, CRC Press, 2012.

An Application of Face Recognition Authenticated Video Recording Process in Embedded System

İbrahim Fettah DEMİRBAĞ[1], Barış KARAKAYA[2]

[1, 2]Department of Electrical-Electronics Enginering, Faculty of Engineering, Fırat University, Elazığ, Türkiye
[1]231113110@firat.edu.tr, [2]bkarakaya@firat.edu.tr
[1](ORCID: 0009-0002-3535-3183), [2](ORCID: 0000-0001-7995-3901)

Abstract
In this study, a face recognition application is realised using a deep learning model on FPGA and Jetson platform. The model has a total of 5 classes and a new face data set of the user was created in addition to the face images of the people obtained from the existing data set in the literature. In addition, three separate classes belonging to cat, dog and outdoor environment were generated in order to prevent false recognition due to possible different creatures or environments in the test data set to be obtained from the camera in real time. Finally, the aim of the study is that the system automatically starts video recording when the user's face data is recognised over 80% correctly. Thus, user-authenticated video recording can be realised. The study is carried out by training with the DenseNet model installed on both Kria FPGA and Jetson Orin Nano boards. A data set consisting of 3908 test, 13689 training and 6740 validation data was used for training the model. In this dataset, there are dogs-vs-cats, human-faces, house-rooms-streets-image-dataset datasets available in the literature and open to general use, and the user's unique dataset named ibrahim. The performances of this face recognition and image registration process on FPGA and Jetson boards are analysed and compared. As a result of the study, the class accuracies obtained on the Jetson platform are cat: 96.97%, dog: 94.02%, human: 99.83%, ibrahim: 99.74%, none: 99.57%, while class accuracies for Kria FPGA are cat: 95.13%, dog: 86.34%, human: 98.48%, ibrahim: 85.74%, none: 98.45%. According to the results obtained, the artificial intelligence model worked with higher accuracy on a single board computer than on FPGA, this is because the quantization and compile processes required to run the deep learning model in the FPGA environment causes the accuracy of artificial intelligence to decrease.

Keywords: deep neural network; embedded system; face recognition; FPGA

1. Introduction

Thanks to technological developments in hardware and software, the use of artificial intelligence in embedded systems is increasing day by day. Artificial intelligence models that used to require large desktop computers to run can now be run on embedded systems the size of a credit card. Embedded systems are used in a wide range of systems designed for various purposes, such as security cameras, medical equipment, and drones. The aim of this study is to design a security system that performs identity detection using SBC (single board computer) and FPGA (field programmable gate array) based embedded systems and records camera footage when the correct person is identified, and to compare the accuracy performance of this system on FPGA and SBC. Referring to previous studies, FPG-AI, a technology-independent toolchain that automates the loading of CNNs (Convolutional Neural Networks) onto FPGAs, has been introduced to enable a broad user base without hardware knowledge to run CNNs on FPGAs with shorter development times [1]. This study compares the performance of Convolutional Neural Networks (CNN) running on different hardware platforms such as GPU (Graphic Processing Unit), FPGA, and CPU (Central Processing Unit) using various tools. This study was conducted using BNN-PYNQ, which provides ready-to-run code and layers for quantized CNNs and fully connected neural networks on FPGA and CPU. These networks were then copied using TensorFlow and optimized for use on GPU with TensorRT at FP32, FP16, and INT8 precisions. The results show that the FPGA outperforms the GPU by 100 times in inference speed for CNN networks and by 1000 times for fully connected networks. In terms of power efficiency, the FPGA also outperformed the GPU. This study concludes that an FPGA should be preferred for a performance-driven artificial neural network application. However, the GPU offers easier usability thanks to the many tools and frameworks available. If ease of implementation and high design flexibility are priorities, the GPU is recommended [2]. FINN and Vitis AI development environments are two important tools for running DNN (Deep Neural Network) applications within FPGA. The performance of the FPGA-based DNN acceleration application developed using these tools was analyzed in comparison with the NVIDIA Jetson Nano Developer Kit and a high-performance desktop computer equipped with an Intel Core i7-11700K processor, both of which have similar power budgets. The results show that the FINN-based accelerator is significantly superior to Vitis AI, and the comparison with the Jetson Nano and desktop computer also yielded results in

[1]Corresponding author

favor of the FINN-based accelerator. This demonstrates that even when automation frameworks are used, DNN accelerators running on FPGAs can still provide significant performance and energy gains compared to GPUs and CPUs [3]. With the proliferation of edge computing platforms and the entry of new companies into this field, it is becoming increasingly difficult to decide which platform to choose in a given situation. These systems are typically equipped with a wide variety of hardware architectures and hardware acceleration technologies, which can lead to confusion when it comes to integrating them into larger designs as hardware accelerators. Therefore, this study focuses on performance and power consumption analyses conducted on Kria KV260, Jetson Nano, and RTX 3060. Two neural network models trained for image recognition tasks such as ResNet-50 and YOLO were used in this study. The findings reveal that the FPGA-based platform offers superior efficiency in terms of inference rate per watt compared to other platforms. This enables identity verification using AI-powered face recognition systems in security cameras [4]. A new hybrid pipeline architecture is proposed to accelerate the DenseNet model, which operates with binary weights, on the FPGA. This design, which combines two typical architectures, aligns well with the unique structure of the DenseNet model. Application results show that the proposed architecture can operate at a high speed of 889 frames per second (FPS) at 200 MHz and achieves a performance-efficiency of up to 62.4%, which surpasses most similar studies [5]. The performance of deep learning-based navigation models was evaluated on a new platform that combines a four-wheel drive RC vehicle with Jetson Xavier NX for autonomous driving. Focusing on data collection and model training with a single-camera system, obstacle avoidance and path following were the primary objectives. The study compared the training time, accuracy rate, obstacle avoidance efficiency, and path following capabilities of the ResNet18, ResNet34, and DenseNet121 models. The results highlight the computational advantages of the Jetson Xavier NX over the Jetson Nano, demonstrating significant improvements in real-time processing and decision-making capabilities. The integration of Jetson Xavier NX with an RC vehicle not only enhances overall performance but also offers an innovative approach to autonomous vehicle development. The findings provide important insights into optimizing deep learning models and hardware configurations, contributing to the improvement of autonomous navigation [6].

The Xilinx Kria KV260 development platform, based on the Kria K26 system-on-module (SoM), examines the effectiveness and dynamic distribution of hardware-accelerated applications. The platform supports the Python-based PYNQ framework and offers a high degree of flexibility with support for custom bit streams (overlays). The demonstration highlighted the system's reconfigurability and overall ease of use, along with machine learning (ML) algorithms that consume low resources [7]. In this study, the DenseNet201 artificial intelligence model, which was trained in a desktop environment using a data set specially prepared for this study, was run on two embedded systems based on FPGA and SBC, and the accuracy performances of both embedded systems were compared. The Jetson Orin Nano development board developed by Nvidia was used as the SBC. Thanks to the GPU inside this card, artificial intelligence models trained in a desktop computer environment can be run on the Jetson card without the need for additional processing. The total accuracy of the trained model running on Jetson was observed to be 98.026%. The Kria KV260 development board developed by AMD was used as the FPGA. Thanks to the DPU artificial intelligence accelerator IP Core inside this card, artificial intelligence models can be run on the FPGA without using hardware design languages such as VHDL or Verilog. However, if an artificial intelligence model is to be run on the DPU IP Core (Deep Neural Processing Intellectual Property) inside the FPGA, the artificial intelligence model trained on the computer must undergo quantization and compilation processes. After undergoing quantization and compilation processes, the artificial intelligence model experiences a slight decrease in the accuracy of some classes, which causes the artificial intelligence model running on Kria to have lower accuracy than the artificial intelligence model running on Jetson. As a result, the total accuracy of the trained model running on Kria was observed to be 92.828%.

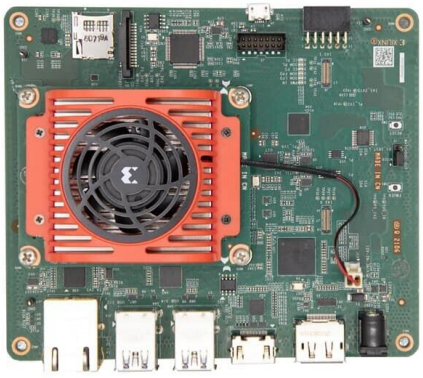


Figure 1a. Kria KV260



Figure 1b. Jetson Orin Nano

2. Method

For training the artificial intelligence model, the Vitis-AI development environment, a Docker-based development environment developed by AMD (Advanced Microdevices), was used [8]. This development environment enables us to develop artificial intelligence models that can be run on FPGAs and embedded systems. The artificial intelligence model trained in this environment is then quantized and compiled within the same environment to make it executable on the DPU IP Core within the FPGA. However, the quantization and compilation processes slightly reduce the accuracy of the artificial intelligence model running on the FPGA, and the results are more evident when compared to the Jetson card, which does not require these processes.

The quantization process converts the 32-bit floating-point weights and activations in the artificial intelligence model to INT8 fixed-point integers, reducing the computational complexity of the artificial intelligence model while compromising as little as possible on the prediction accuracy of the artificial intelligence model. The fixed-point network model requires less memory bandwidth, operates faster, and provides higher power efficiency compared to the floating-point model [9].
The compilation process converts the previously quantized artificial intelligence model into a highly efficient instruction set and data flow model. The compiler is used to create an internal computation graph as an intermediate representation layer. This internal graph consists of independent control and data flow representations. After this process, the compiler performs a series of optimization operations on the artificial intelligence model. For example, if the convolution operator comes before the normalization operator, batch normalization operations are combined with convolution [10].
The TensorFlow library, an open-source machine learning and deep learning library developed by Google for artificial intelligence training, has been selected. This library can be used on various platforms and is particularly suitable for large-scale machine learning projects. The TensorFlow library is typically used to create, train, and deploy deep learning models. Additionally, the TensorFlow library can be run within the Vitis-AI development environment without any additional intervention.

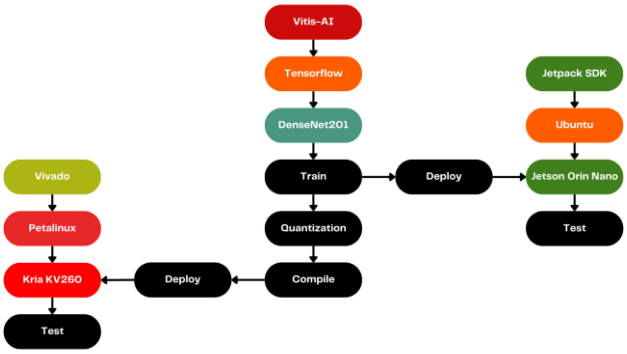


Figure 2. Pipeline

The DenseNet201 convolutional neural network model was selected for artificial intelligence training [11]. The most distinctive feature of the DenseNet CNN architecture, which stands for Densely Connected Convolutional Networks and consists of 201 layers, is that instead of passing the output from each layer only to the next layer, it passes it to all subsequent layers, thereby establishing much denser connections. This allows each layer to benefit from the outputs of previous layers, resulting in better learning performance. To explain in more detail, DenseNet uses "dense connections" for each layer, meaning that the feature map from one layer is transferred to both the next layer and all other layers. This structure allows the network to carry more information at each layer. DenseNet structurally includes special layers called "bottleneck layers" and "transition layers." Bottleneck layers increase the depth of the network while reducing the computational load. Transition layers reduce the size of the network, thereby reducing computational intensity and managing the learning process.

If we look at the advantages of the DenseNet model one by one, since the inputs in each layer of the model are taken from the previous layers, the network can be made deeper with fewer parameters, which makes the network more efficient. During backward propagation, gradients are transmitted directly and more effectively, enabling the training process to be faster and more stable. Dense connections allow the network to learn faster and be more resistant to overfitting. Deep learning models such as DenseNet are typically pre-trained on large datasets and can then be fine-tuned for a specific task or used for another task [12], [13], which shortens the learning process. This process is called transfer learning and means that the knowledge a model has learned in one training can be used in another training. A pre-trained DenseNet model can be customized for a new training through fine-tuning, which enables the model to perform better on datasets prepared for specific purposes. Due to these features, the DenseNet deep learning model can be successfully trained using transfer learning with datasets prepared for specific purposes in various applications [14], [15], [16]. The trained model can be deployed on devices such as computers or embedded systems and used with high accuracy performance.

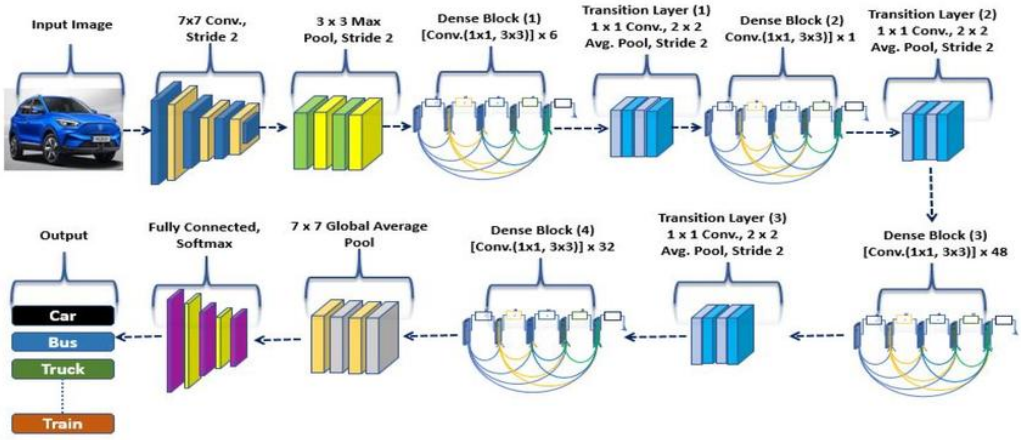


Figure 3. DenseNet201 architecture (feature extraction block) [17].

There are five classes in the dataset. These classes are named cat, dog, human, ibrahim, and none in alphabetical order. Three publicly available datasets and one dataset prepared specifically for this study were used in the dataset. The publicly available datasets are the dogs-vs-cats [18], human-faces [19], and house-rooms-streets-image-dataset [20] datasets obtained from the Kaggle platform. The dataset prepared specifically for this study is the ibrahim dataset, which consists of 636 test, 3389 training, and 1767 validation data points (this dataset contains facial data belonging to the corresponding author who prepared the study). To explain the other datasets, the cat dataset consists of 760 test, 2,170 training, and 1,332 validation; the dog dataset consists of 820 test, 2,680 training, and 1,280 validation; the human dataset consists of 528 test, 2,216 training, and 1,068 validation; and the none dataset consists of 1,164 test, 3,234 training, and 1,293 validation.

The entire dataset consists of 3,908 test, 13,689 training, and 6,740 validation data points. The image data in the dataset was resized to 128x128 pixels and normalized before the training process. Figure 4 shows some of the images

in the dataset arranged in alphabetical order from top to bottom. The main reason for using multiple classes in the dataset is to enable the distinction between positive and negative cases. In addition, in machine learning, especially in binary classification problems, data is generally divided into two classes: positive class and negative class, so that the model can make accurate predictions.

Positive Class: This is the class that the model must predict and classify correctly. For example, in disease diagnosis, it is “detecting the presence of the disease.” Another example is detecting that the person visible in the camera is named Ibrahim. In short, it is the target class that the model must predict correctly.

Negative Class: This is the opposite class that the model must predict and classify correctly. For example, in disease diagnosis, “detecting the absence of the disease.” Another example would be determining that the person seen on camera is not named Ibrahim but is another person, or that there is no person on camera and only an empty environment.

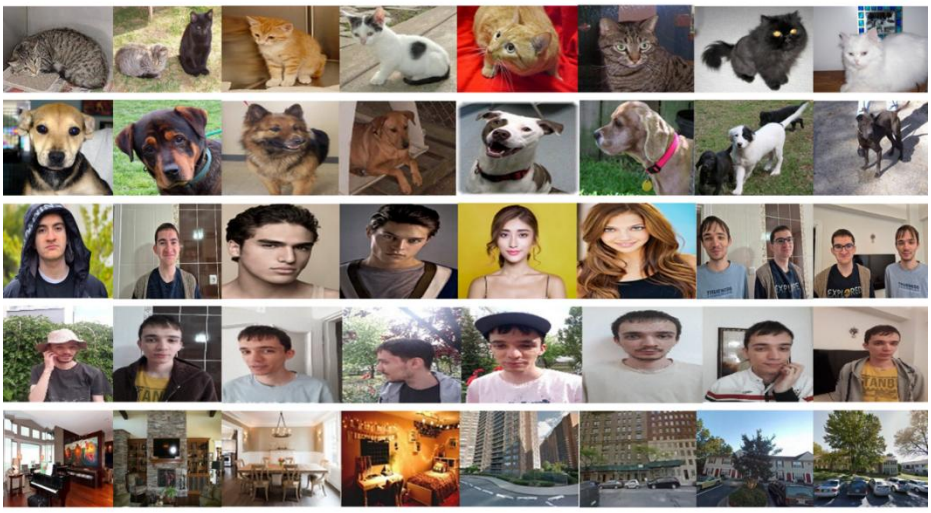


Figure 4. Dataset

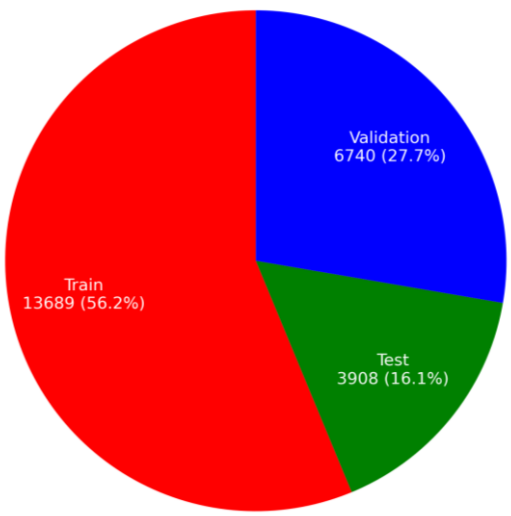


Figure 5a. Distribution of test, training, and validation groups in dataset

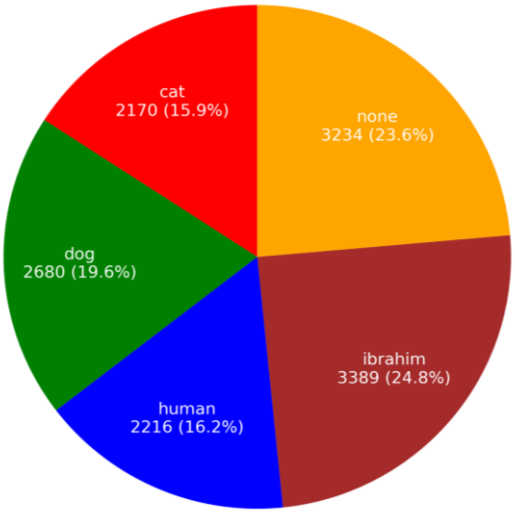


Figure 5b. Distribution of classes in train group

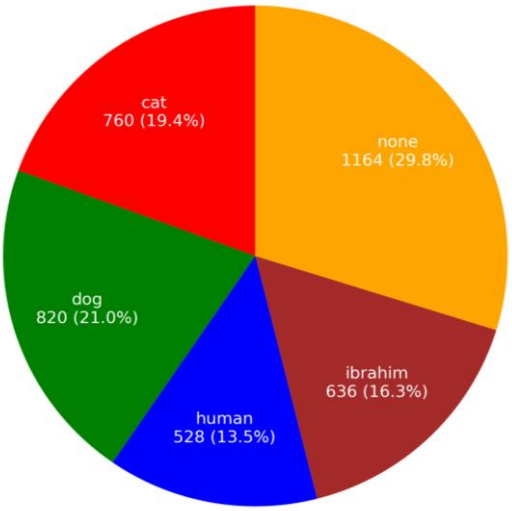


Figure 5c. Distribution of classes in test group

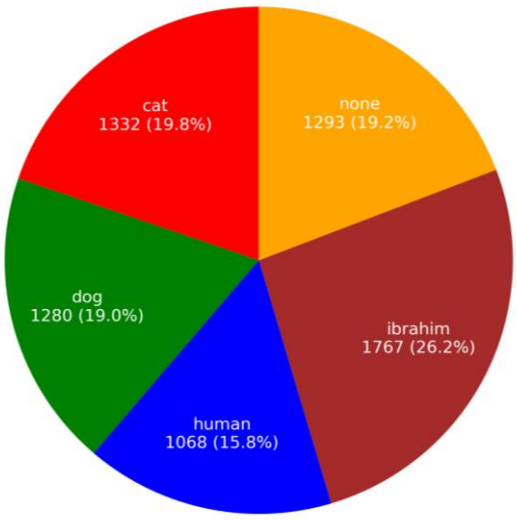


Figure 5d. Distribution of classes in validation group

Figure 5a shows the three main classes that make up the data set: the Train class, the Test class, and the Validation class. The dataset has been carefully divided into three subsets—training (56.2%), validation (27.7%), and testing (16.1%)—to improve the model's generalization ability and reduce the risk of overfitting. During this division process, care was taken to keep the data set balanced by considering the class distributions of the examples in each subset. As shown in Figures 5b, 5c, and 5d, these classes are grouped under five categories: cat, dog, human, ibrahim, and none. Both absolute distribution and proportional balance were sought among these classes, with particular attention paid to minimizing differences between classes in the training and validation sets. During the model training process, positive and negative class separation was also addressed, and efforts were made to prevent the model from becoming biased toward certain classes by maintaining diversity in class representation.

If we examine Figures 6a and 6b, we can see that training accuracy started at approximately 88% and reached 99.9% at the end of the 20th epoch. Validation accuracy generally remained above 95% and, despite the sudden drop observed in the 8th epoch, recovered quickly and followed a parallel trend with training accuracy. Similarly, training loss has steadily decreased and reached a minimum level. Validation loss showed a temporary spike at the 8th epoch but decreased again in the following epochs and progressed steadily. These results demonstrate that the model has successfully adapted to the training data and achieved effective generalization on the validation data. The balanced distribution of classes during training, the proportional presentation of positive and negative examples, and the application of techniques such as early stopping played an important role in preventing overfitting. Possible reasons for the temporary spike observed in the 8th epoch include the random application of normalization and data augmentation to the validation data, as well as the saving of model weights as the “best model” and reloading them at this point. Such operations can cause short-term fluctuations during the validation phase.

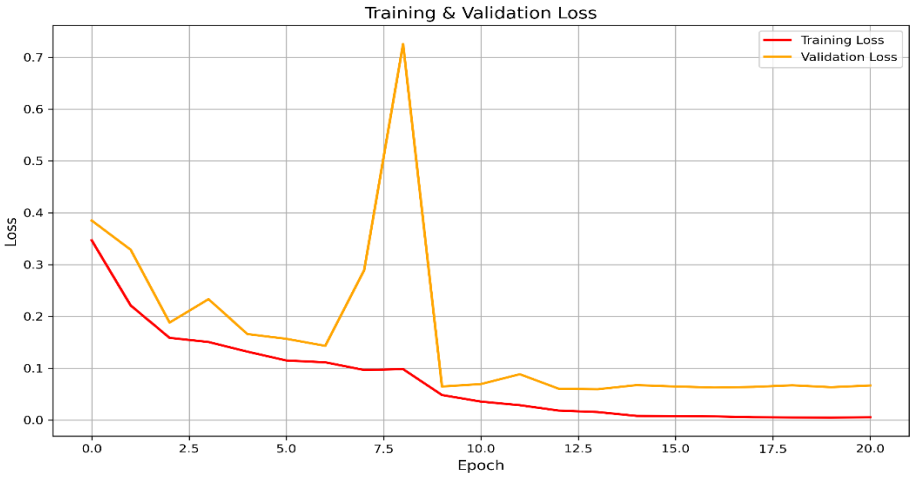


Figure 6a. Train & validation loss

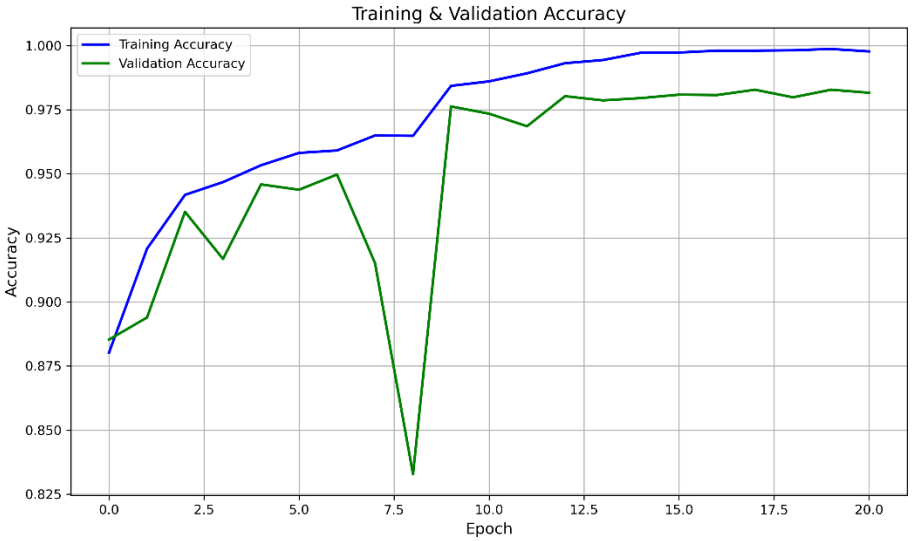


Figure 6b. Training & validation accuracy

3. Results

According to the test results conducted on both embedded systems, the Jetson Orin Nano card demonstrated higher accuracy performance than the Kria KV260 card. The class accuracies for the Jetson Orin Nano card were observed as follows: cat: 96.97%, dog: 94.02%, human: 99.83%, ibrahim: 99.74%, none: 99.57%. The overall accuracy of the Jetson Orin Nano card was observed as 98.026%. For the Kria KV260 card, the class accuracies were observed as follows: cat: 95.13%, dog: 86.34%, human: 98.48%, ibrahim: 85.69%, none: 98.81%. The total accuracy of the Kria KV260 card was observed as 92.828%.

When examining other basic evaluation metrics such as Precision, Recall, and F1-score in Figure 6a and Figure 6b, Jetson Orin Nano has Accuracy: 0.98 Macro F1-score: 0.98 Weighted F1-score: 0.98, while Kria KV260 has Accuracy: 0.93 Macro F1-score: 0.92 Weighted F1-score: 0.93 The Jetson platform consistently achieved higher values. When comparing class-level performance, the F1-Score values on the Jetson Orin Nano were observed to be 0.96 for cat, 0.96 for dog, 0.99 for human, 0.99 for ibrahim, and 0.99 for none. The F1-Score values on the Kria KV260 were observed to be 0.91 for cat, dog 0.90, human 0.91, ibrahim 0.92, and none 0.99. The Jetson Orin Nano stands out with higher F1-score values across all classes. In particular, noticeable differences have emerged in human recognition and data tagged with specific individuals (ibrahim). These findings reveal that the Jetson platform is more successful in distinguishing between classes that are more complex or have similar characteristics.

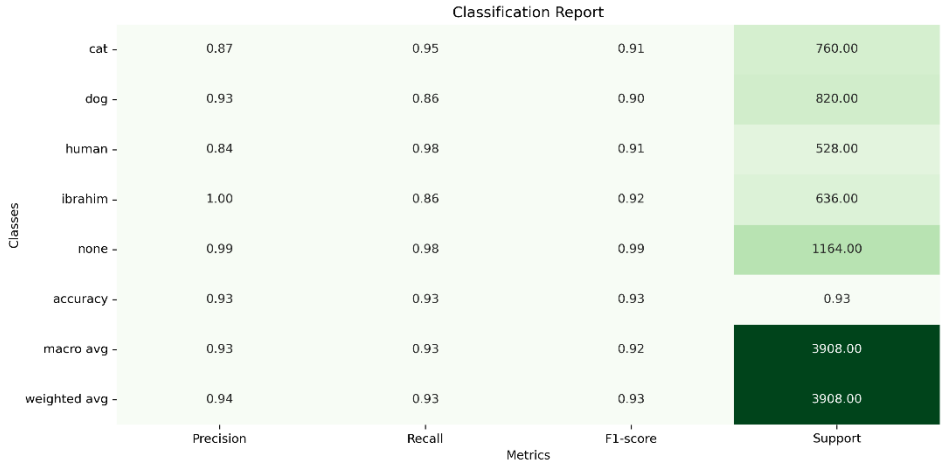


Figure 7a. Precision, recall, f1-score, and support metrics for the kria KV260 card

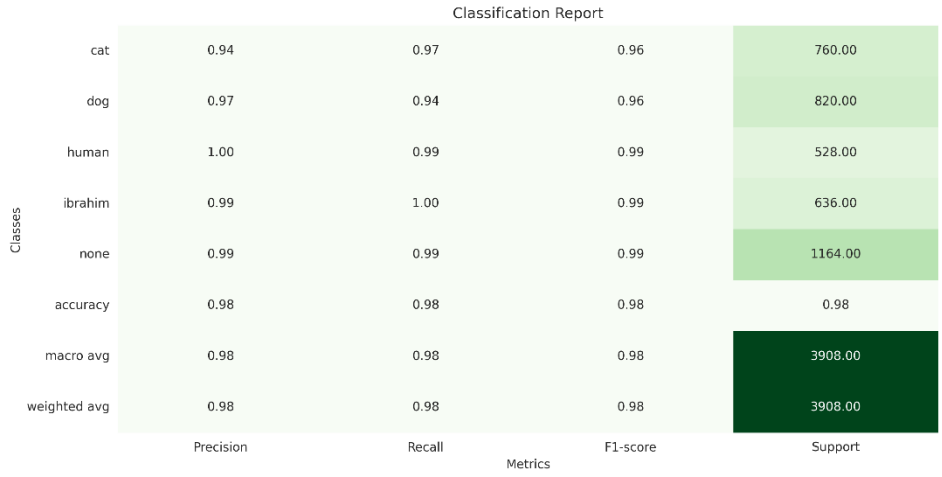


Figure 7b. Precision, recall, f1-score, and support metrics for the jetson orin nano card

A similar trend was observed in analyses conducted using complexity matrices. Jetson Orin Nano produced fewer cross-class errors. For example, on the Jetson platform, only 1 out of 636 examples belonging to the “ibrahim” class was misclassified into another class, while on the Kria platform, this number increased to 61. Similarly, while Kria misclassified 85 examples in the “dog” class, Jetson misclassified only 24.

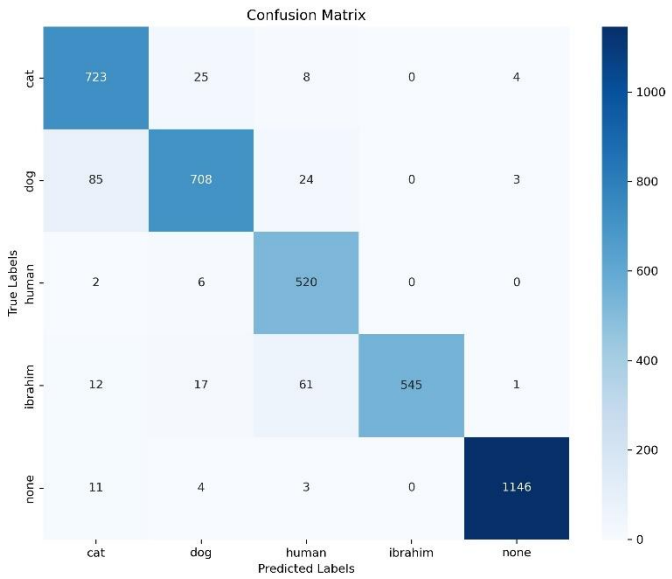


Figure 8a. Kria KV260 confusion matrix

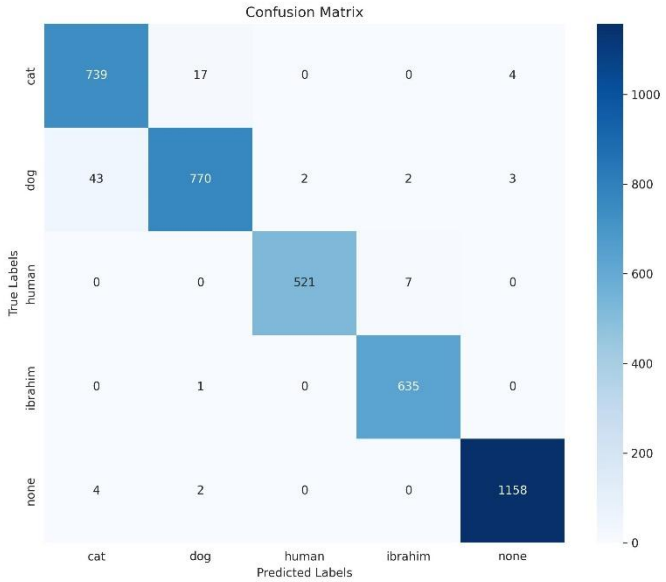


Figure 8b. Jetson orin nano confusion matrix

Based on these results, the accuracy of the artificial intelligence model running on Kria was observed to be lower than that of the artificial intelligence model running on the Jetson card. If we examine these results in more detail, the artificial intelligence model running on the Kria card has undergone quantization and compilation processes, which convert the 32-bit floating-point weights in the DenseNet201 model to INT8 fixed-point weights. If the convolution operators in the DenseNet201 model come before the normalization operator, batch normalization operations are combined with convolution. These processes enable the artificial intelligence model to consume fewer system resources, allowing it to run faster, but they also cause a decrease in the accuracy of the artificial intelligence model. Therefore, when looking at the test results obtained within Kria, the accuracy of some classes is close to that obtained on the Jetson card, while the accuracy of some classes falls short of the class accuracy obtained on the Jetson card. Since the DenseNet201 model running on the Jetson card was not subjected to quantization and compilation processes, the class accuracy of the DenseNet201 model obtained in the test conducted on the Jetson card was higher than the model accuracy obtained in the test conducted on the Kria card. In light of the results obtained here, it has been observed that the artificial intelligence model trained must undergo additional processing before being loaded onto the embedded system, and that the data sets used for positive and negative classification affect the accuracy of the artificial intelligence model. Finally, when comparing the accuracy of artificial intelligence models running on FPGA and SBC, it is observed that the artificial intelligence performance of SBCs is higher than that of FPGAs, and training and loading artificial intelligence models for SBCs is easier than for FPGAs. Therefore, it is expected that SBCs will be used more than FPGAs in practical applications. However, thanks to technological developments, the artificial intelligence performance of FPGAs is approaching that of SBCs day by day, and FPGAs, which previously required a long software development process to run artificial intelligence models, no longer require this long development process. AI models that are trained, compiled, and quantized and run within an FPGA can now achieve nearly the same accuracy performance as an AI model running on an SBC.

4. References

[1] Pacini T, Rapuano E. and Fanucci L. FPG-AI: A Technology-Independent Framework for the Automation of CNN Deployment on FPGAs. in *IEEE Access*, vol. 11, pp. 32759-32775, 2023, doi: 10.1109/ACCESS.2023.3263392. https://ieeexplore.ieee.org/abstract/document/10089388 (08.07.2025)

[2] Pettersson L. Convolutional neural networks on FPGA and GPU on the edge: A comparison. (2020). https://www.diva-portal.org/smash/record.jsf?dswid=1558&pid=diva2%3A1447576 (08.07.2025)

[3] Hamanaka F, Odan T, Kise K. and Chu T. V. An Exploration of State-of-the-Art Automation Frameworks for FPGA-Based DNN Acceleration. in *IEEE Access*, vol. 11, pp. 5701-5713, 2023, doi: 10.1109/ACCESS.2023.3236974. https://ieeexplore.ieee.org/abstract/document/10017269 (08.07.2025)

[4] Lizano S. A. and Westerlund T. Comparison of edge computing platforms for hardware acceleration of AI: Kria KV260, Jetson Nano and RTX 3060. *J. Edge Comput.* 15.3 (2024): 123-135. https://www.utupub.fi/handle/10024/177876 (08.07.2025)

[5] Kalaphotas S, Flamis G, Kitsos P. Efficient Edge-AI Application Deployment for FPGAs (2022) https://www.mdpi.com/2078-2489/13/6/279 (09.07.2025)

[6] Zeng S, Huang Y. A Hybrid-Pipelined Architecture for FPGA-based Binary Weight DenseNet with High Performance-Efficiency *2020 IEEE High Performance Extreme Computing Conference (HPEC)*, Waltham, MA, USA, 2020,pp. 1-5, doi: 10.1109/HPEC43674.2020.9286185. https://ieeexplore.ieee.org/abstract/document/9286185 (09.07.2025)

[7] Ma S, Karkal P. S, Ma J, Alam F. A Comparison of the Performance of ResNet and DenseNet for Obstacle Avoidance and Path Following on a Mobile Robot Using Embedded GPUs *2024 IEEE International Conference on Agents (ICA)*, Wollongong, Australia, 2024, pp. 7-12, doi: 10.1109/ICA63002.2024.00009.https://ieeexplore.ieee.org/abstract/document/10807413 (09.07.2025)

[8] Xilinx "Vitis-AI Documentation" https://xilinx.github.io/Vitis-AI/3.5/html/index.html (12.07.2025)

[9] Xilinx "Vitis-AI Quantization Tool Documentation" https://xilinx.github.io/Vitis-AI/3.5/html/docs/workflow-model-development.html#model-quantization (12.07.2025)

[10] Xilinx "Vitis-AI Compilation Tool Documentation" https://xilinx.github.io/Vitis-AI/3.5/html/docs/workflow-model-development.html#model-compilation (12.07.2025)

[11] Huang G, Liu Z, Maaten L, Weinberger K. Q. Densely Connected Convolutional Networks (2016) https://arxiv.org/abs/1608.06993 (13.07.2025)

[12] Salim F, Saeed F, Basurra S, Qasem S. N, Al-Hadhrami T. Densenet-201 and Xception Pre-Trained Deep Learning Models for Fruit Recognition (2023) https://www.mdpi.com/2079-9292/12/14/3132 (13.07.2025)

[13] Zhang Z, Liang X, Dong X, Xie Y, Cao G. A Sparce-View CT Reconstruction Method Based on Combination of DenseNet and Deconvolution in *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1407-1417, June 2018, doi: 10.1109/TMI.2018.2823338. https://ieeexplore.ieee.org/abstract/document/8331861 (13.07.2025)

[14] Hasan N, Bao Y, Shawon A, Huang Y. DenseNet Convolutional Neural Networks Application for Predicting COVID-19 Using CT Image *SN COMPUT. SCI.* 2, 389 (2021). https://link.springer.com/article/10.1007/s42979-021-00782-7 (13.07.2025)

[15] Dalvi P. P, Edla D. R, Purushothama B. R. *SN COMPUT. SCI.* 4, 214 (2023). Diagnosis of Coronavirus Disease From Chest X-Ray Images Using DenseNet-169 Architecture https://link.springer.com/article/10.1007/s42979-022-01627-7 (13.07.2025)

[16] Liao T, Li L, Ouyang R, Lin X. Classification of asymmetry in mammography via the DenseNet convolutional neural network *European Journal of Radiology* Open,Volume 11, 2023, 100502, ISSN 2352-0477 https://www.sciencedirect.com/science/article/pii/S235204772300028X (13.07.2025)

[17] Akhtar M. J, Mahum R, Butt F. S, Amin R. A Robust Framework for Object Detection in a Traffic Surveillance System (2022) https://www.researchgate.net/publication/364647881_A_Robust_Framework_for_Object_Detection_in_a_Traffic_Surveillance_System (14.07.2025)

[18] Kaggle "Dogs-vs-Cats Dataset" https://www.kaggle.com/c/dogs-vs-cats/data (15.07.2025)

[19] Kaggle "Human-Faces Dataset" https://www.kaggle.com/datasets/ashwingupta3012/human-faces (15.07.2025)

[20] Kaggle "House-Rooms-Streets-Image-Dataset" https://www.kaggle.com/datasets/mikhailma/house-rooms-streets-image-dataset (15.07.2025)

Critical Success Factors of Lean Manufacturing Tools: The DEMATEL Method based pq-Cubic Quasi rung Orthopair Fuzzy Sets

Veysel Çoban[1]

[1,2]Department of Industrial Engineering, Engineering Faculty, Bilecik Seyh Edebali University, Bilecik, Turkiye.
[1]veysel.coban@bilecik.edu.tr
[1](ORCID: 0000-0002-7885-1935)

Abstract
Organizations that produce similar products at similar prices use different tools to increase their profits In a global competitive environment. The most widely used tool with economies of scale is lean production. Lean manufacturing is an important managerial program that aims to increase productivity and efficiency by eliminating physical and informational waste. This study aims to define the causality relationships between Lean manufacturing tools and to determine the importance values of the tools. Seven of the widely used Lean Manufacturing tools are determined by the opinions and evaluations of experts. A fuzzy Decision Making Trial and Evaluation Laboratory (DEMATEL) is used to determine the weight values of lean production tools and to define the cause-effect relationships between them. Pairwise comparisons between tools are made with the p,q-cubic quasi-rung orthopair fuzzy set (p,q-CQOFS) method to increase the evaluation and representation abilities of decision makers. This research contributes to the organizational operability of lean manufacturing tools by providing insights into the critical tools for lean manufacturing adoption. In addition, the DEMATEL based pq-cubic quasi rung orthopair fuzzy sets approach used in the evaluation of the vehicles is an innovative aspect of the study.

Keywords: lean manufacturing, dematel, p,q-cubic quasi-rung orthopair

1. Introduction

Lean production, based on the Toyota production system, is a management approach that focuses on eliminating waste. Lean production is a system that aims to integrate people with technical structures [1]. This structure, which integrates systems with people, focuses on reaching the highest production levels with the lowest waste, lowest labor and time loss. Lean production, which has proven its success in Toyota, emerges with waste, cost and performance improvement activities in different sectors. Lean production applies improvements with small and continuous steps spread over a long period of time instead of high-cost large investments [2]. Success story Lean production also presents disappointing samples because of application difficulties and performance results that do not meet expectations. Therefore, companies that fail in application and continuous improvement with lean production can choose to return to traditional methods [3]. Lean production requires cultural change that can realize transformations in managerial understanding and system applications. The main reasons for failures are seen as the non-spreading the lean philosophy to employees and the non-supporting lean approach by senior management. Understanding and implementing the Lean philosophy among employees at all levels can be achieved by identifying comprehensive lean practices and uncovering the causal relationships between these practices.

Decision-making methods are used to define acceptable lean practices for all production stages [1], [4]. In this study, the Decision Making Trial and Evaluation Laboratory (DEMATEL), based on pq-cubic quasi-rung orthopair fuzzy sets (p, q-CQOFS), is used as a decision-making tool. DEMATEL is a tool used to determine the importance levels of factors and visualize causal relationships between factors. p, q-cubic quasi-rung orthopair fuzzy sets (p, q-CQOFS) identify and incorporate uncertain and fuzzy information into the evaluation process. p, q-CQOFS represents the degrees of membership and non-membership of an element in a set, as well as the degree of uncertainty regarding an element's membership in a set [5], [6].

The study aims to identify the importance of lean manufacturing practices and their interrelationships within the context of these core implementation tools. Experts identify the core Lean application factors by their evaluations, and the developed method defines the priorities and interrelationships for establishing a sustainable lean manufacturing culture. The study's original contributions include the DEMATEL method, based on the p, q-CQOFS method, and its use in evaluating lean manufacturing factors. Therefore, the study consists of two main steps: the development of the DEMATEL

[1]Corresponding author

method based on p, q-CQOFS and the identification of critical lean factors in the sustainable continuous improvement program of this developed method.

The subsequent sections of the paper are organized as follows: Section 2 discusses the objectives and methods through a literature review. Section 3 describes the steps of the proposed decision-making method. Section 4 applies the developed method to identifying critical lean factors. Section 5 discusses the results and offers recommendations for future research.

2. Literature Review

The main topics covered in the literature review are as follows: the relationships among lean production methods, the DEMATEL decision-making method, and the p,q-CQOFS method. A review of existing studies on these topics highlights the originality of the proposed method and the importance of the results to be obtained during the implementation phase. The success of lean applications is measured by the use of the right tools in the right place and at the right time and their impact on performance. The concept of Lean Manufacturing, the equivalent of the Toyota Production System, was first named by James Womack and others in the book "The Machine That Changed the World" [7]. Lean production is defined as the elimination of non-value-added activities from the process and the inclusion of value-added activities [3]. Non-value-added activities are considered waste and do not contribute to customer product evaluation. Overproduction, inventory, waiting, transportation, overprocessing, rework, and unnecessary movements are the primary waste categories [8]. Value-added activities, which should be included at every stage of the production process from design onward, include defining product value, creating and developing the existing value stream map, establishing a traction system, and continuous improvement. 5S, Jidoka, Kaizen, Just-in-Time Production (JIT), Kanban, Single-Minute Die Change (SMED), and Visual Management are the fundamental and widespread tools of lean practices. Establishing and maintaining a lean production system requires management ownership and employee acceptance. Therefore, management and employees must be fully familiar with lean tools and establish the correct relationships between them.

Factors affecting the sustainability of lean practices can be classified under managerial, organizational and cultural main headings [9]. The creation and development of a lean system requires management support and a transformation of employees' ideas and culture. The lean philosophy takes root through employee empowerment, proper guidance, and the maintenance of their lean aspirations through motivating factors [10], [11]. The number of variables expanded under these main headings should be reduced to a smaller number in order to simplify and understand the study. The factors considered in this study are shown in Table 1.

Table 1. Factors affecting the lean practices

Factor Codes	Factors	Definition Sources
FC1	Top Management Support	Top management provides physical and intellectual support in the institutional establishment and development of lean activities. This support should be felt by employees. Top management's training and motivational plans enable employees to be more effective in lean efforts and facilitate the implementation of lean processes.
FC2	Employee Participation	Lean implementation processes are embraced and guided by employees. Employees should accept and actively use lean tools to improve production processes. Employees should be decision-makers in setting, planning, and monitoring lean implementation goals.
FC3	Training	Training aims to disseminate lean tools and implementation processes to all levels of employees. Employee development expands the corporate competency pool and ensures that every employee participates in all lean tasks.

Table 1. (Continue) Factors affecting the lean practices

FC4	Motivaton	Tangible and intangible motivators are used to encourage and maintain employee participation in lean practices. Incentives such as recognizing and celebrating successful employees and paying bonuses are also used.
FC5	Communication	Communication is the establishment of appropriate channels for the corporate dissemination of lean practices and the provision of positive and negative feedback. The participation and suggestions of employees at all levels can be collected and processed through communication channels to achieve lean goals.

This study examines five critical success factors affecting the organizational success of lean manufacturing. These factors aim to include employees at all levels in lean practices and ensure their successful implementation. Management should embrace the lean philosophy and seek the most appropriate ways to communicate the lean approach to employees. Employees should be informed about lean tools, developed through training, and provided with experience through practices. Incentives and communication channels should be established to ensure employee involvement in lean practices and their collaborative efforts.

3. Methodology

p, q Quasi-rung Orthopair Fuzzy Sets

This section discusses cubic sets and p, q-quasi-rung orthopair fuzzy sets ($p, q-QOFS$). A cubic sets (C) [12], [13] are defined in F that is a non-empty set as;

$$\xi = \{t, C(t), v(t) | t \in F\} \tag{1}$$

where $\xi(t) = [C^L(t), C^U(t)]$ is the interval-valued fuzzy set in F and $v(t)$ is a fuzzy set. Cubic set is abbreviated as $\xi = \langle C, v \rangle$[5], [6].

A p, q-QOFS (ς) is defined as $\varsigma = \{f, v_\varsigma(f), \psi_\varsigma(f) | f \in F\}$ where $v_\varsigma(f), \psi_\varsigma(f) : F \to [0,1]$ that they represent support and support against membership grade of element $f \in F$, $\left(v_\varsigma(f)\right)^p + \left(\psi_\varsigma(f)\right)^q \preccurlyeq 1$ and $p, q \geqslant 1$. $\left(v_\varsigma(f),\right), \psi_\varsigma(f))$ is a $p, q-QOF$ number and defined as $\varsigma = \left(v_\varsigma, \psi_\varsigma\right)$. Hesitancy degree of ς is calculated as

$$\pi_\varsigma = \sqrt[s]{1 - \left(v_\varsigma(f)\right)^p - \left(\psi_\varsigma(f)\right)^q} \tag{2}$$

where s is the least common multiple of p and q.

p, q-cubic quasi-rung orthopair fuzzy set ($p, q-CQOFS$) represents the membership degrees with a p, q-interval-valued quasi-rung orthopair fuzzy ($p, q-IVQOF$) value and a p, q-quasi-rung orthopair fuzzy (p, q-QOF) value [6], [14]. Λ as a $p, q-CQOFS$ over $f \in F$ is defined as

$$\Lambda = \{f, L(f), \gamma(f) | f \in F\} \tag{3}$$

where $L(f)$ is represented with p, q-interval-valued quasi-rang orthopair fuzzy set as $L(f) = \{f, [v_L^l(f), v_L^u(f)], [\psi_L^l(f), \psi_L^u(f)]\}$ and p, q-interval-valued quasi-rang orthopair fuzzy value for each $f \in F$ as $\gamma(x) = \{f, v_\gamma(f), \psi_\gamma(f)\}$ [6]. pq-CQOFS Λ over a set F is defined as

$$\Lambda = \left\{\left(\left([v_L^l(f), v_L^u(f)], [\psi_L^l(f), \psi_L^u(f)]\right), \left(v_L(f), \psi_L(f)\right)\right) : f \in F\right\} \quad (4)$$

Also the basic relationships are defined as: $0 \preccurlyeq v_L^l(f) \preccurlyeq v_L^u(f) \preccurlyeq 1, 0 \preccurlyeq \psi_L^l(f) \preccurlyeq \psi_L^u(f) \preccurlyeq 1$ and : $0 \preccurlyeq \left(v_L^u(f)\right)^p + \left(\psi_L^u(f)\right)^q \preccurlyeq 1$.

The score and accuracy functions of $C = \left\{\left(\left([v_L^l, v_L^u], [\psi_L^l, \psi_L^u]\right), (v_L, \psi_L)\right)\right\}$ is defined, respectively, as [6], [14]

$$S(C) = \left(\frac{(v_L^l)^p + (v_L^u)^p - (\psi_L^l)^q - (\psi_L^u)^q}{2}\right) + (v_L^p - \psi_L^q) \quad (5)$$

where $-2 \leq S(C) \leq 2$.

$$A(C) = \left(\frac{(v_L^u)^p + (v_L^u)^p + (\psi_L^u)^q + (\psi_L^u)^q}{2}\right) + (v_L^p + \psi_L^q) \quad (6)$$

where $0 \leq A(C) \leq 2$.

The basic algebraic calculations for two pq-CQROFNs $C_1 = \left\{\left(\left([v_{L_1}^l, v_{L_1}^u], [\psi_{L_1}^l, \psi_{L_1}^u]\right), \left(v_{L_1}, \psi_{L_1}\right)\right)\right\}$and $C_2 = \left\{\left(\left([v_{L_2}^l, v_{L_2}^u], [\psi_{L_2}^l, \psi_{L_2}^u]\right), \left(v_{L_2}, \psi_{L_2}\right)\right)\right\}$are as follows [5], [6], [12], [14]:

$$C_1 \oplus C_2 = \left(\left(\left[\left(\frac{1}{2}\left(\left(v_{L_1}^l\right)^p + \left(v_{L_2}^l\right)^p\right)\right)^{\frac{1}{p}}, \left(\frac{1}{2}\left(\left(v_{L_1}^u\right)^p + \left(v_{L_2}^u\right)^p\right)\right)^{\frac{1}{p}}\right], \left[\left(1 - \frac{1}{2}\left(\left(\psi_{L_1}^u\right)^q + \left(\psi_{L_2}^u\right)^q\right)\right)^{\frac{1}{q}}, \left(1 - \frac{1}{2}\left(\left(\psi_{L_1}^l\right)^q + \left(\psi_{L_2}^l\right)^q\right)\right)^{\frac{1}{q}}\right]\right), \left(\left(\frac{1}{2}\left(\left(v_{L_1}\right)^p + \left(v_{L_2}\right)^p\right)\right)^{\frac{1}{p}}, \left(1 - \frac{1}{2}\left(\left(\psi_{L_1}\right)^q + \left(\psi_{L_2}\right)^q\right)\right)^{\frac{1}{q}}\right)\right) \quad (7)$$

$$C_1 \otimes C_2 = \left(\left(\left[\left(1 - \frac{1}{2}\left(\left(v_{L_1}^u\right)^p + \left(v_{L_2}^u\right)^p\right)\right)^{\frac{1}{p}}, \left(1 - \frac{1}{2}\left(\left(v_{L_1}^l\right)^p + \left(v_{L_2}^l\right)^p\right)\right)^{\frac{1}{p}}\right], \left[\left(\frac{1}{2}\left(\left(\psi_{L_1}^l\right)^q + \left(\psi_{L_2}^l\right)^q\right)\right)^{\frac{1}{q}}, \left(\frac{1}{2}\left(\left(\psi_{L_1}^u\right)^q + \left(\psi_{L_2}^u\right)^q\right)\right)^{\frac{1}{q}}\right]\right), \left(\left(1 - \frac{1}{2}\left(\left(v_{L_1}\right)^p + \left(v_{L_2}\right)^p\right)\right)^{\frac{1}{p}}, \left(\frac{1}{2}\left(\left(\psi_{L_1}\right)^q + \left(\psi_{L_2}\right)^q\right)\right)^{\frac{1}{q}}\right)\right) \quad (8)$$

$$\lambda C = \left(\left(\left[\lambda^{\frac{1}{p}} v_L^l, \lambda^{\frac{1}{p}} v_L^u\right], \left[\lambda\left(1 - \left(\psi_L^u\right)^q\right)^{\frac{1}{q}}, \lambda\left(1 - \left(\psi_L^l\right)^q\right)^{\frac{1}{q}}\right]\right), \left(\lambda^{\frac{1}{p}} v_L, \left(\lambda\left(1 - \psi_L^q\right)\right)^{\frac{1}{q}}\right)\right), 0 \leq \lambda \leq 1 \quad (9)$$

$$C^\lambda = \left(\left(\left[{v_L^l}^\lambda, {v_L^u}^\lambda\right], \left[\left(1 - \left(1 - \left(\psi_L^l\right)^p\right)^\lambda\right)^{\frac{1}{q}}, \left(1 - \left(1 - \left(\psi_L^u\right)^p\right)^\lambda\right)^{\frac{1}{q}}\right]\right), \left(\left(1 - \left(1 - v_L^p\right)^\lambda\right)^{\frac{1}{p}}, \psi_L^\lambda\right)\right) \quad (10)$$

The aggredgated values of pq-CQROFNs $(C_1, C_2, \ldots, C_m)$ are obtained with the p, q-cubic quasi rung orthopair fuzzy weighted averaging (p, q-CQOFWA) as follow:

$$p,q-CQOFWA\ (C_1,C_2,\dots.,C_m)=\left(\left(\left[\left(\sum_{i=1}^{m}\left(w_i\left(v_{L_1}^{l}\right)^{p}\right)\right)^{\frac{1}{p}},\left(\sum_{i=1}^{m}\left(w_i\left(v_{L_1}^{u}\right)^{p}\right)\right)^{\frac{1}{p}}\right],\left[\left(1-\sum_{i=1}^{m}\left(w_i\left(1-\left(\psi_{L_1}^{l}\right)^{q}\right)\right)\right)^{\frac{1}{q}},\left(1-\sum_{i=1}^{m}\left(w_i\left(1-\left(\psi_{L_1}^{u}\right)^{q}\right)\right)\right)^{\frac{1}{q}}\right]\right),\left(\left(\sum_{i=1}^{m}\left(w_i\left(v_{L_i}\right)^{p}\right)\right)^{\frac{1}{p}},\left(1-\sum_{i=1}^{m}\left(w_i\left(1-\left(\psi_{L_i}\right)^{q}\right)\right)\right)^{\frac{1}{q}}\right)\right) \quad (11)$$

where $W=(\omega_1,\omega_2,\dots\omega_m)^T$ is a weight vector of C_i and $w_i\in[0,1]$, $\sum_{i=1}^{m}w_i=1\,;i=1,2,,\dots,m$

Fuzzy DEMATEL

Fuzzy DEMATEL is a decision model applied to analyze the causal relationships between factors. The DEMATEL method creates a structural model of the system based on the cause and effect relationships between factors [4]. The DEMATEL model's application steps are as follows [4], [15], [16], [17]:

Step 1: Formation of the decision-making expert team.

Step 2: The factors are identified through the literature review and expert opinions, and experts create the initial direct relation fuzzy matrix ($\tilde{Z}^k$) for the factors using the linguistic terms (Table 2).

$$\tilde{Z}^k=\begin{bmatrix}0 & \tilde{Z}_{12}^k & \cdots & \tilde{Z}_{1m}^k\\ \tilde{Z}_{21}^k & 0 & \cdots & \tilde{Z}_{2m}^k\\ \vdots & \vdots & \cdots & \vdots\\ \tilde{Z}_{m1}^k & \tilde{Z}_{m2}^k & \cdots & 0\end{bmatrix}$$

where $\tilde{Z}_i$represents the pairwise comparisons of experts, k is experts' number $(k=1,\dots,e)$ and $\tilde{Z}_{ij}^k=\left(\left\{\left(\left(\left[v_{ij_{L_k}}^{l},v_{ij_{L_k}}^{u}\right],\left[\psi_{ij_{L_k}}^{l},\psi_{ij_{L_k}}^{u}\right]\right),\left(v_{ij_{L_k}},\psi_{ij_{L_k}}\right)\right)\right\}\right)$

Table 2. Fuzzy membership functions of linguistic terms

Linguistic terms	Values	Fuzzy scales
No influence	0	(([0,0.1], [0.7,0.9]), (0.9,0.1))
Very low influence	1	(([0.1,0.3], [0.5,0.7]), (0.7,0.3))
Low influence	2	(([0.3,0.5], [0.3,0.5]), (0.5,0.5))
High influence	3	(([0.5,0.7], [0.1,0.3]), (0.3,0.7))
Very high influence	4	(([0.7,0.9], [0,0.1]), (0.1,0.9))

Step 3: Aggregated direct relation fuzzy matrix ($\tilde{Z}$) is obtained using $p,q-CQOFWA$ operation.
Step 4: Normalization of the aggregated direct-relation fuzzy matrix.

$$\tilde{\alpha}=\tilde{Z}_{ij}=\left(\left(\left[\sum_{j=1}^{m}v_{ij}^{l},\sum_{j=1}^{m}v_{ij}^{u}\right],\left[\sum_{j=1}^{m}\psi_{ij}^{l},\sum_{j=1}^{m}\psi_{ij}^{u}\right]\right),\left(\sum_{j=1}^{m}v_{ij},\sum_{j=1}^{m}\psi_{ij}\right)\right) \quad (12)$$

$$\tilde{\beta}_{v^l}=\sum_{i=1}^{m}\sum_{j=1}^{m}v_{ij}^{l},\tilde{\beta}_{v^u}=\sum_{i=1}^{m}\sum_{j=1}^{m}v_{ij}^{u} \quad (13)$$

$$\tilde{\beta}_{\psi^l}=\sum_{i=1}^{m}\sum_{j=1}^{m}\psi_{ij}^{l},\tilde{\beta}_{\psi^u}=\sum_{i=1}^{m}\sum_{j=1}^{m}\psi_{ij}^{u} \quad (14)$$

$$\tilde{\beta}_{v}=\sum_{i=1}^{m}\sum_{j=1}^{m}v_{ij},\tilde{\beta}_{\psi}=\sum_{i=1}^{m}\sum_{j=1}^{m}\psi_{ij} \quad (15)$$

The normalized aggregated direct relation matrix ($\tilde{X}$) is obtained as:

$$\tilde{X} = \begin{bmatrix} \tilde{X}_{11} & \tilde{X}_{12} & \cdots & \tilde{X}_{1m} \\ \tilde{X}_{21} & \tilde{X}_{22} & \cdots & \tilde{X}_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ \tilde{X}_{m1} & \tilde{X}_{m2} & \cdots & \tilde{X}_{mm} \end{bmatrix}$$

where $\tilde{X}_{ij} = \frac{\tilde{z}_{ij}}{\tilde{\beta}} = \left(\left[\frac{v_{ij}^{l}}{\tilde{\beta}_{v^l}}, \frac{v_{ij}^{u}}{\tilde{\beta}_{v^u}} \right], \left[\frac{\psi_{ij}^{l}}{\tilde{\beta}_{\psi^l}}, \frac{\psi_{ij}^{u}}{\tilde{\beta}_{\psi^u}} \right], \left(\frac{v_{ij}}{\tilde{\beta}_{v}}, \frac{\psi_{ij}}{\tilde{\beta}_{\psi}} \right) \right)$

Step 5: The total-relation matrix ($\tilde{T}$) is obtained on the normalized aggregated direct relation matrix ($\tilde{X}$) by converging $\lim_{w\to\infty} \tilde{X}^{W} = 0$.

$$\tilde{T} = \lim_{w\to\infty} (\tilde{X}^{1} + \tilde{X}^{2} + \cdots + \tilde{X}^{W}) \tag{16}$$

$$\tilde{T} = \begin{bmatrix} \tilde{t}_{11} & \tilde{t}_{12} & \cdots & \tilde{t}_{1m} \\ \tilde{t}_{21} & \tilde{t}_{22} & \cdots & \tilde{t}_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ \tilde{t}_{m1} & \tilde{t}_{m2} & \cdots & \tilde{t}_{mm} \end{bmatrix}$$

where $\tilde{t}_{ij} = \left(\left[v_{ij}^{l}{}'', v_{ij}^{u}{}'' \right], \left[\psi_{ij}^{l}{}'', \psi_{ij}^{u}{}'' \right], \left(v_{ij}'', \psi_{ij}'' \right) \right)$.

$$\text{Matrix} \left[v_{ij}^{l}{}'' \right] = \tilde{X}_{v^l} \times \left(pI - \tilde{X}_{v^l} \right)^{-1}, \text{Matrix} \left[v_{ij}^{u}{}'' \right] = \tilde{X}_{v^u} \times \left(pI - \tilde{X}_{v^u} \right)^{-1} \tag{17}$$

$$\text{Matrix} \left[\psi_{ij}^{l}{}'' \right] = \tilde{X}_{\psi^l} \times \left(qI - \tilde{X}_{\psi^l} \right)^{-1}, \text{Matrix} \left[\psi_{ij}^{u}{}'' \right] = \tilde{X}_{\psi^u} \times \left(qI - \tilde{X}_{\psi^u} \right)^{-1} \tag{18}$$

$$\text{Matrix} \left[v_{ij}'' \right] = \tilde{X}_{v} \times \left(pI - \tilde{X}_{v} \right)^{-1}, \text{Matrix} \left[\psi_{ij}'' \right] = \tilde{X}_{\psi} \times \left(qI - \tilde{X}_{\psi} \right)^{-1} \tag{19}$$

where I is the unit matrix.

Step 6: The vector *D* and vector *R* are calculated to obtain the cause and effect diagram.

$A(\widetilde{D}_i) + A(\tilde{R}_i)$ is the horizontal axis vector of the cause-effect diagram, representing the criteria's prominence. $S(\widetilde{D}_i) - S(\tilde{R}_i)$ describes the criteria's cause-effect relationship on the vertical axis. Criteria above the horizontal axis ("Relationship" positive) are the cause criteria, while criteria below the horizontal axis ("Relationship" negative) are the effect criteria. Score and accuracy values of $\widetilde{D}_i$ and $\tilde{R}_i$ are calculated as [6], [14], [18], [19]:

$$S(\widetilde{D}_i, \tilde{R}_i) = \left(\frac{\left(v_{ij}^{l}{}'' \right)^{p} + \left(v_{ij}^{u}{}'' \right)^{p} - \left(\psi_{ij}^{l}{}'' \right)^{q} - \left(\psi_{ij}^{u}{}'' \right)^{q}}{2} \right) + \left(v_{ij}^{''\,p} - \psi_{ij}^{''\,q} \right) \tag{20}$$

$$A(\widetilde{D}_i, \tilde{R}_i) = \left(\frac{\left(v_{ij}^{l}{}'' \right)^{p} + \left(v_{ij}^{u}{}'' \right)^{p} + \left(\psi_{ij}^{l}{}'' \right)^{q} + \left(\psi_{ij}^{u}{}'' \right)^{q}}{2} \right) + \left(v_{ij}^{''\,p} + \psi_{ij}^{''\,q} \right) \tag{21}$$

4. Application and Results

This section describes an example application of the proposed new p,q-QOFS DEMATEL method. In the example application, the proposed method is applied to determine the interdependence of lean manufacturing methods. Three experts with different importance values ($\omega = (0.35, 0.20, 0.45)$) evaluate the methods using Table 2. Expert assessments are combined using the p,q-CQOFWA method and displayed in the aggregate direct relation matrix (Table 3). The normalized aggregate direct relation matrix (Table 4) and the total relation matrix (Table 5) are obtained, respectively, according to the p,q-QOFS DEMATEL application steps. Finally, the cause-effect values of the lean factors are calculated (Table 6) and the causal diagram showed in Figure 1. $p = 4$ and $q = 5$ are selected in the implementation steps of proposed p,q-QOFS DEMATEL method.

Table 3. Aggregated direct relation matrix

$\tilde{Z}$	FC1	FC2	FC3	FC4	FC5
FC1	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0.7,0.9], [0,0.1]), (0.1,0.9))	([0.56,0.75], [0.1,0.29]), (0.28,0.76))	([0.7,0.9], [0,0.1]), (0.1,0.9))	([0.63,0.83], [0.09,0.26]), (0.25,0.83))
FC2	([0.7,0.9], [0,0.1]), (0.1,0.9))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0.5,0.7], [0.1,0.3]), (0.3,0.7))	([0.5,0.7], [0.1,0.3]), (0.3,0.7))	([0.7,0.9], [0,0.1]), (0.1,0.9))
FC3	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))
FC4	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))
FC5	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))

Table 4. The normalized aggregated direct relation matrix

$\tilde{X}$	FC1	FC2	FC3	FC4	FC5
FC1	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0.7,0.9], [0.0.1]), (0.1,0.9))	([0.56,0.75], [0.1,0.29]), (0.28,0.76))	([0.7,0.9], [0,0.1]), (0.1,0.9))	([0.63,0.83], [0.09,0.26]), (0.25,0.83))
FC2	([0.7,0.9], [0.0.1]), (0.1,0.9))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0.5,0.7], [0.1,0.3]), (0.3,0.7))	([0.5,0.7], [0.1,0.3]), (0.3,0.7))	([0.7,0.9], [0,0.1]), (0.1,0.9))
FC3	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))
FC4	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))
FC5	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))	([0,0.1], [0.7,0.9]), (0.9,0.1))

Table 5. Total relation matrix

$\tilde{T}$	FC1	FC2	FC3	FC4	FC5
FC1	([0.001,0.004], [0.012,0.011]), (0.014,0.003))	([0.035,0.028], [0,0.001]), (0.002,0.022))	([0.029,0.024], [0.002,0.004]), (0.004,0.019))	([0.036,0.028], [0,0.001]), (0.002,0.022))	([0.033,0.026], [0.001,0.003]), (0.004,0.021))
FC2	([0.035,0.027], [0,0.001]), (0.002,0.022))	([0.001,0.004], [0.012,0.011]), (0.014,0.003))	([0.026,0.022], [0.002,0.004]), (0.005,0.017))	([0.026,0.022], [0.002,0.004]), (0.005,0.018))	([0.036,0.028], [0,0.001]), (0.002,0.022))
FC3	([0,0.003], [0.012,0.011]), (0.014,0.002))	([0,0.003], [0.012,0.011]), (0.014,0.002))	([0,0.003], [0.012,0.011]), (0.014,0.003))	([0,0.003], [0.012,0.011]), (0.014,0.003))	([0,0.003], [0.012,0.011]), (0.014,0.003))
FC4	([0,0.003], [0.012,0.011]), (0.014,0.002))	([0,0.003], [0.012,0.011]), (0.014,0.002))	([0,0.003], [0.012,0.011]), (0.014,0.003))	([0,0.003], [0.012,0.011]), (0.014,0.003))	([0,0.003], [0.012,0.011]), (0.014,0.003))
FC5	([0,0.003], [0.012,0.011]), (0.014,0.002))	([0,0.003], [0.012,0.011]), (0.014,0.002))	([0,0.003], [0.012,0.011]), (0.014,0.003))	([0,0.003], [0.012,0.011]), (0.014,0.003))	([0,0.003], [0.012,0.011]), (0.014,0.003))

Table 6. Cause-effect values of the lean factors

	D	R	S(D − R)	A(D + R)
FC1	([0.02,0.02], [0.99,0.99]), (0.01,0.99))	([0.02,0.01], [0.99,0.99]), (0.01,0.99))	0,005	3,813
FC2	([0.02,0.02], [0.99,0.99]), (0.01,0.99))	([0.02,0.01], [0.99,0.99]), (0.01,0.99))	0,002	3,816
FC3	([0,0], [0.99,0.99]), (0.01,1))	([0.02,0.01], [0.99,0.99]), (0.01,0.99))	−0,039	3,866
FC4	([0,0], [0.99,0.99]), (0.01,1))	([0.02,0.02], [0.99,0.99]), (0.01,0.99))	−0,045	3,860
FC5	([0,0], [0.99,0.99]), (0.01,1))	([0.02,0.02], [0.99,0.99]), (0.01,0.99))	−0,048	3,857

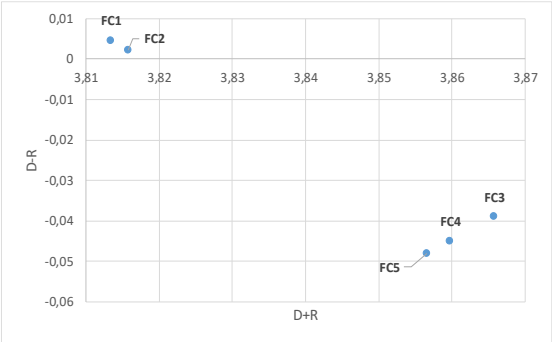


Figure 1. Causal diagram of lean factors

Figure 1 shows that FC1 (Top Management Support) and FC2 (Employee Participation) are main cause criteria (located over D − R axis). Factors FC1 and FC2 directly impact the other factors. Top management support and employee participation in lean practices help shape training, implement motivational tools, and guide lean implementation processes effectively through appropriate communication channels. Guidance on lean factors with top management support ensures more efficient employee and production systems. Top management's decision-making on the right training and motivation tools helps address deficiencies and maintains employee engagement. Employee participation is crucial to the implementation of lean processes. It directly impacts training, motivation, and communication, and determines how these factors relate to lean processes. On the other hand, FC3, FC4 and FC5 are main effect criteria (located under the D − R axis). These criteria are influenced and shaped by the behavior of FC1 and FC2 criteria. Training, motivation, and communication factors are the result of senior management decisions and employee responses. According to the horizontal axis of the cause-effect diagram, the criteria have similar importance values. Furthermore, the FC3 criterion appears as the criterion with the highest importance. The results of the application demonstrate that decision-makers' assessments are helpful in the decision-making process. Experts consider the consistency of the relationships between factors and the validity of the factors' impact levels to be acceptable.

5. Conclusion

Complex problems are solved by bringing together different opinions and suggestions on the common denominator. Fuzzy decision-making methods can be used to incorporate uncertainty and instability into the decision-making process. In this study, the DEMATEL decision-making method based on the p,q-cubic quasi-rung orthopair fuzzy set (p,q-CQOFS) is proposed. The proposed method is applied to evaluate the factors influencing the implementation of lean manufacturing methods. Five factors are considered as effective factors in the application of lean tools. Factors are identified as cause and effect factors by applying the proposed method. FC1 (Top Management Support) and FC2 (Employee Participation) are identified as root causes, while FC3 (Training), FC4 (Motivation), and FC5 (Communication) are identified as root causes. Furthermore, the proposed method also evaluates the importance of factors. While there is no significant difference in importance among the factors, FC3 emerges as the most important factor. The application results of the proposed method are evaluated through discussions with experts. Experts agree that the results reflect consistent opinions. Future study is planned to expand the range of linguistic terms in the p,q-CQOFS - DEMATEL method and expand its application area with new examples.

Acknowledgements

The study is not supported by any institution, project, person, etc.

6. References

[1] Kassem B, Callupe M, Rossi M, Rossini M, Portioli-Staudacher A. "Lean 4.0: a systematic literature review on the interaction between lean production and industry 4.0 pillars," J. Manuf. Technol. Manag., vol. 35, no. 4, pp. 821–847, 2024.
[2] Singh H, Singh B. "Industry 4.0 technologies integration with lean production tools: a review," TQM J., vol. 36, no. 8, pp. 2507–2526, 2024.
[3] Bai C, Ahmet S, Sarkis J. "Investing in lean manufacturing practices: an environmental and operational perspective," Int. J. Prod. Res., vol. 57, no. 4, pp. 1037–1051, Feb. 2019, doi: 10.1080/00207543.2018.1498986.
[4] Ferrazzi M, Costa F, Frecassetti S, Portioli-Staudacher A. "Unlocking synergies in lean manufacturing for enhanced environmental performance: a cross-sector investigation through fuzzy DEMATEL," Clean. Logist. Supply Chain, vol. 15, p. 100219, Jun. 2025, doi: 10.1016/j.clscn.2025.100219.
[5] Amin F, Rahim M, Ali A, Ameer E., "Generalized cubic Pythagorean fuzzy aggregation operators and their application to multi-attribute decision-making problems," Int. J. Comput. Intell. Syst., vol. 15, no. 1, p. 92, 2022.
[6] Chu Y-M, Garg H, Rahim M, Amin F, Asiri A, Ameer E."Some p, q-cubic quasi-rung orthopair fuzzy operators for multi-attribute decision-making," Complex Intell. Syst., vol. 10, no. 1, pp. 87–110, 2024.
[7] Womack J P. "A measure of lean [lean production]," Manuf. Eng., vol. 85, no. 4, pp. 6–7, 2006.
[8] Rossini M, Costa F, Tortorella G L, Valvo A, Portioli-Staudacher A. "Lean Production and Industry 4.0 integration: how Lean Automation is emerging in manufacturing industry," Int. J. Prod. Res., vol. 60, no. 21, pp. 6430–6450, 2022.
[9] Alefari M, Salonitis K, Xu Y. "The Role of Leadership in Implementing Lean Manufacturing," Manuf. Syst. 40 – Proc. 50th CIRP Conf. Manuf. Syst., vol. 63, pp. 756–761, Jan. 2017, doi: 10.1016/j.procir.2017.03.169.
[10] Mangaroo-Pillay M, Coetzee R. "A systematic literature review (SLR) comparing Japanese Lean philosophy and the South African Ubuntu philosophy," Int. J. Lean Six Sigma, vol. 13, no. 1, pp. 118–135, 2022.
[11] Bortolotti T, Boscari S, Danese P. "Successful lean implementation: Organizational culture and soft lean practices," Int. J. Prod. Econ., vol. 160, pp. 182–201, Feb. 2015, doi: 10.1016/j.ijpe.2014.10.013.
[12] Mahmood T, Mehmood F, Khan Q. "Cubic hesitant fuzzy sets and their applications to multi criteria decision making," Int. J. Algebra Stat., vol. 5, no. 1, pp. 19–51, 2016.
[13] Rong Y, Yu L, Niu W, Liu Y, Senapati T, Mishra A R. "MARCOS approach based upon cubic Fermatean fuzzy set and its application in evaluation and selecting cold chain logistics distribution center," Eng. Appl. Artif. Intell., vol. 116, p. 105401, 2022.
[14] Liu P, Ali A, Rehman N, Parveen M. "Artificial intelligence in medical practice: The CRITIC-TOPSIS method based on λ(pq)-cubic quasi rung orthopair fuzzy robust aggregation operators and their applications," Inf. Sci., vol. 710, p. 122076, Aug. 2025, doi: 10.1016/j.ins.2025.122076.
[15] Nezhad M Z, Nazarian-Jashnabadi J, Rezazadeh J, Mehraeen M, Bagheri R. "Assessing dimensions influencing IoT implementation readiness in industries: A fuzzy DEMATEL and fuzzy AHP analysis," J. Soft Comput. Decis. Anal., vol. 1, no. 1, pp. 102–123, 2023.
[16] Abdullah L, Zulkifli N, Liao H, Herrera-Viedma E, Al-Barakati A. "An interval-valued intuitionistic fuzzy DEMATEL method combined with Choquet integral for sustainable solid waste management," Eng. Appl. Artif. Intell., vol. 82, pp. 207–215, Jun. 2019, doi: 10.1016/j.engappai.2019.04.005.
[17] Kiani Mavi R, Standing C. "Critical success factors of sustainable project management in construction: A fuzzy DEMATEL-ANP approach," J. Clean. Prod., vol. 194, pp. 751–765, Sep. 2018, doi: 10.1016/j.jclepro.2018.05.120.
[18] Mirdad W K, Eseonu C I. "A cause-effect strategy map for lean process transformation," Int. J. Syst. Syst. Eng., vol. 8, no. 2, pp. 121–146, 2017.
[19] Costa F, Lispi L, Staudacher A P, Rossini M, Kundu K, Cifone F D. "How to foster Sustainable Continuous Improvement: A cause-effect relations map of Lean soft practices," Oper. Res. Perspect., vol. 6, p. 100091, Jan. 2019, doi: 10.1016/j.orp.2018.100091.

Conceptual Integration of Machine Learning in AI-Directed Soil Stabilization for Sustainable Ground Improvement

Abiola Ayopo ABIODUN[1]

[1]Department of Computer Engineering, Faculty of Engineering and Architecture, Cyprus West University, Famagusta North Cyprus via Mersin 10, Turkey.
[1]a.abiodun@cc.cwu.tr
[1](ORCID: 0000-0003-1655-2312)

Abstract
Sustainable infrastructure development requires innovative solutions to overcome the persistent drawbacks of traditional soil stabilization. Cement- and lime-based methods, although widely used, are carbon intensive and contribute significantly to greenhouse gas emissions, contrary to the global objective of sustainable development. Alternative low-carbon techniques such as electrokinetic, bio-based, and agro-industrial-based stabilization techniques have shown promise, but their wider acceptance is still limited by three critical weaknesses: (i) lack of predictive models to guide treatment mechanisms, (ii) limited scalability in field applications, and (iii) unpredictable performance in heterogeneous soils. This study proposes the conceptual integration of machine learning in AI-directed soil stabilization for sustainable ground improvement. The central research problem is the lack of adaptive, data-driven systems to predict soil responses and optimize stabilization strategies in real time. The objectives are threefold: (1) to create a conceptual framework integrating real-time monitoring with machine learning algorithms, (2) to optimize treatment parameters such as duration, soil-binder mix, environmental factors control, and energy consumption by supporting the decisions with artificial intelligence, and (3) to recommend a closed loop adaptive system that continuously refines predictions using field data. By addressing the identified gaps, the proposed framework will move geotechnical engineering towards sustainability, resilience, scalability, wider application. This conceptual study will guide to achieve reduction of reliance on carbon-intensive methods, minimizes energy consumption, and improves the reliability of soil treatment under different environmental conditions. The study concludes with a recommendation for a roadmap to move from trial-and-error stabilization procedures to intelligent ground-improvement systems with AI capability as a transformative route to sustainable infrastructure.

Keywords: ai-directed soil stabilization, machine learning, sustainable ground improvement, adaptive systems, predictive modeling, geotechnical engineering.

1. Introduction

The global push for sustainable, resilient, and low-carbon infrastructure has placed civil and geotechnical engineering at the forefront of innovation. Ground improvement, particularly in problematic soils such as expansive and soft soils, is essential to ensure the stability and durability of modern infrastructure. However, conventional stabilization methods, mainly cement and lime, although technically reliable, are contributing significantly to CO_2 emissions and undermining the global climate commitments.

Alternative low carbon methods, including electrokinetic stabilization, have shown considerable potential. However, their wider acceptance is constrained by persistent limitations, performance differences between different soils, difficulties in predicting treatment results and difficulties in optimising processes for practical field use. These gaps create a research gap which needs to be addressed in order to move from carbon-intensive to sustainable soil improvements.

Artificial Intelligence (AI) and Machine Learning (ML) offer a unique opportunity to overcome these problems. While artificial intelligence has transformed predictive modelling and optimization across engineering disciplines, its application to soil remediation is still under-explored. This paper proposes a conceptual framework for AI-driven AI-directed soil stabilization and shows how ML can be integrated to close existing gaps, optimize decision making and open scalable and sustainable geotechnical solutions.

[1] Corresponding author

2. Methodological Framework

Although conceptual, this work offers a structured framework-building approach that encapsulates general low carbon soil stabilization methods that can be optimized, scaled up, adapted, and applied to a wide range of soils to achieve the most effective and efficient soil-based improvement. These techniques include electrokinetic treatment, bio-based techniques include microbial-induced calcite precipitation (MICP), enzyme-induced calcite precipitation (EICP), and biopolymers; industrial by-product-based techniques include alkali-activated binders (geopolymers), silica fume, rice husk ash stabilization, biochar, and agricultural waste ashes (pine cone ash, coconut shell ash). Also to be considered are mechanical or physical approaches such as soil freezing (artificial ground freezing with renewable energy sources) and fiber reinforcement (natural fibers of jute, coir, hemp, sisal); emerging low-carbon hybrid methods such as CO2 sequestration in soils or mineral carbonation, magnesium-based binders (MgO cement), and phytostabilization (using plants or roots). These techniques, and their integration in emerging applications of artificial intelligence, will be highly innovative, sustainable, and rewarding as unique conventional and hybrid soil-stabilization techniques in geotechnical engineering. The proposed framework consists of four interconnected components:

Data collection from real-time field and laboratory measurements of parameters such as pH, conductivity, moisture content, pore pressure, ion migration and exchange combined with historical data sets to create reliable training databases. Machine learning models for predictive algorithms including artificial neural networks, support vector machines, and ensemble methods, to predict soil responses such as strength gain, settlement reduction, and durability under varying stabilization conditions.

Optimization engine for decision-making, based on artificial intelligence, dynamically adjusting treatment parameters, including duration, treatment setup, binder type to minimize energy consumption and maximize stabilization efficiency. Feedback loop for building, understanding and applying a self-learning cycle that continuously updates predictive models using new field data, increasing their reliability and scalability over time. This framework is assessed conceptually in terms of sustainability goals, technological readiness and scalability in the real-world applications.

3. Conceptual Results

Integration of artificial intelligence (AI) and machine learning (ML) in soil stabilization is expected to bring several transformational benefits:

- Predictability of robust models reduces uncertainty by predicting soil behavior under different treatment strategies.
- Sustainability of optimized design reduces the use of cement and lime, reduces CO_2 emissions, and supports the global objective of carbon neutrality.
- Energy efficiency in adaptive control minimizes treatment time and electrical input, improving the overall performance of electrokinetic processes.
- Adaptability of the framework applies to diverse soils and climate conditions, improving the resilience of infrastructure.
- Practical applications include the stabilization of soft clay soils in transport corridors, the mitigation of expansive soil in urban development, and the remediation of contaminated soils.

Collectively, these results demonstrate how AI can shift ground improvement from trial-and-error practice to an intelligent, data-driven discipline.

4. Conclusions

This study presents a conceptual framework for soil stabilization in the field of artificial intelligence and addresses the key gaps that have limited the widespread adoption of low carbon alternatives. By integrating machine learning and real-time monitoring, the framework provides predictive accuracy, optimized treatment mechanisms and adaptive control capabilities.

Although conceptual, the framework is a direct response to the pressing need for sustainable and scalable solutions in geotechnical engineering. Future work should focus on developing experimental databases, training predictive algorithms, and conducting pilot field tests to validate the proposed system.

The conceptual innovation outlined here is a decisive step towards the transition of soil improvement from conventional, carbon-intensive methods to AI-enabled, sustainable and resilient infrastructure solutions, placing geotechnical engineering at the heart of global innovation for sustainable development.

Effect of TiC on Microstructure and Hardness in Laser Cladding with Fe-Cr-C/TiC Powder Mixtures

Asuman Merve EKİCİ[1], Mehmet EROĞLU[2], M. Sabri GÖK[3], Rıdvan YAMANOĞLU[4]

[1,2]Metallurgical and Materials Engineering, Engineering Faculty, Fırat University, Elazığ, Turkey
[3]Mechanical Engineering, Architecture and Design Faculty, Bartın University, Bartın, Turkey
[4]Metallurgical and Materials Engineering, Engineering Faculty, Kocaeli University, Kocaeli, Turkey
[1]232146101@firat.edu.tr, [2]meroglu@firat.edu.tr, [3]msabrigok@bartin.edu.tr, [4]ryamanoglu@kocaeli.edu.tr
[1](ORCID: 0009-0005-0602-7305), [2](ORCID:0000-0002-5097-1943), [3](ORCID:0000-0002-8214-2250), [4](ORCID:0000-0002-4661-8215)

Abstract
Laser surface cladding is an advanced surface modification technique employed to enhance the mechanical and chemical properties of metallic substrates. In this process, a metallic or ceramic-reinforced powder is melted and fused onto the substrate surface using a high-energy laser beam, resulting in a metallurgically bonded coating layer with refined microstructure and improved performance characteristics. In this study, it was aimed to produce Fe-Cr-C-based powders with eutectic chemical composition via atomization, for use in laser beam coating of steels. By adding varying amounts of TiC powders to these base powders, the goal was to obtain coatings with high hardness and wear resistance. In the experimental study, following the production of the required Fe-C-Cr powders, laser coating studies were carried out. TiC powders were added to the Fe-C-Cr eutectic powders at 10% and 20% by weight in separate trials, and the effects on the coating microstructure and hardness were investigated. In the coating microstructure, TiC precipitates were observed in addition to the primary austenite and eutectic (austenite + M_7C_3) phases. As a result, it was found that increasing the TiC content in the Fe-Cr-C based powder mixtures led to increased hardness in the coatings.

Keywords: laser coating, tic, fe-cr-c alloy, microstructure, hardness

1. Introduction

Surface engineering plays a critical role in enhancing the performance and lifespan of engineering components operating under severe service conditions. Among various surface modification techniques, laser cladding has emerged as a highly efficient method due to its ability to produce metallurgically bonded coatings with minimal dilution, low porosity, and excellent control over microstructure. This technique involves the deposition of a material onto a substrate using a high-energy laser beam, which melts the coating material and a thin layer of the base material to create a strong metallurgical bond. In recent years, laser cladding has gained considerable attention in applications where enhanced surface hardness, wear resistance, and corrosion resistance are required, particularly in the automotive, aerospace, and tooling industries [1-4].

High chromium and carbon-containing steel powders are frequently used in laser cladding processes due to their ability to form hard carbides and provide excellent resistance to wear and corrosion. The presence of chromium contributes to the formation of stable chromium carbides and enhances oxidation and corrosion resistance. Additionally, the carbon content promotes the formation of hard phases such as Cr_7C_3, which significantly improve surface hardness and wear behavior [5,6]. However, despite these favorable properties, the performance of such coatings can be further improved by incorporating ceramic reinforcements.

Titanium carbide (TiC) is one of the most widely used ceramic particles for reinforcing metal matrix composites due to its exceptional hardness, high melting point, excellent thermal stability, and good chemical compatibility with ferrous alloys. When added to laser cladding powders, TiC particles can serve as nucleation sites during solidification, refine the microstructure, and enhance hardness and wear resistance through dispersion strengthening. Furthermore, TiC particles are known to form strong interfacial bonds with the metal matrix, which contributes to load transfer and improved mechanical performance of the coating [7, 8].

[1]Corresponding author

In this study, a high-chromium and high-carbon steel powder was mixed with varying amounts of TiC (10 wt.%, and 20 wt.%) and used to laser clad the surface of a low-carbon steel substrate. The objective was to investigate the effect of TiC content on the microstructure, hardness, and wear behavior of the laser-cladded layers. By varying the TiC content, the influence of ceramic reinforcement on the formation of hard phases, grain refinement, and the resultant mechanical properties of the coatings was systematically evaluated. The findings of this research are expected to contribute to the development of advanced wear-resistant coatings for industrial applications, providing insights into the optimization of composition and processing parameters in laser cladding technology.

2. Material and Method

In the experimental part of the study, St 37 steel was used for surface coating purposes, and the chemical composition of this steel is given in Table 1. The Fe-Cr-C powders used in the coatings were produced by gas atomization. A SEM image of the powder cross-section is shown in Figure 1. The chemical composition of the powder is given in Table 2, and the particle size range is between 50–150 microns. The dimensions of the steel samples used are 100 mm × 20 mm × 10 mm. For laser cladding of the surface, prior to the process, the substrate surface was subjected to surface preparation steps including coarse sanding, sanding with 120-grit flap discs, and finally sanding with 240-mesh sandpaper.

In the coatings, TiC powders were separately added to the Fe-Cr-C eutectic powders at ratios of 10 wt. %, and 20 wt.%. The particle size of the TiC powders was -75 microns. To prepare the powder mixtures, the powders were first kept at 200°C for 1 hour, then weighed using a precision balance and subjected to mixing in a ball mill. In the ball mill, the ball-to-powder ratio was set at 2:1, the rotation speed at 100 rpm, and the mixing time at 2 hours. After the powder mixtures were prepared, they were mixed with a binder and applied onto the surfaces of the substrates to form a layer with a height of 3 mm. Once the drying process of the powders on the sample surfaces was completed, the laser cladding process was carried out.

During the melting process, a power of 1.5 kW and a scanning speed of 5 mm/s were used. The tracks were applied side by side with 50% overlap. After the laser coating process, samples were extracted from the coated region for microstructural examination, microhardness measurements, and XRD analysis, and were subjected to the necessary preparation procedures. Microhardness measurements were carried out starting from the coated surface toward the substrate material under a load of 200 g.

Table 1. Chemical composition of St 37 steel (wt.%)

C	Si	Mn	P	S	Fe
0,130	0,210	0,620	0,003	0,003	Kalan

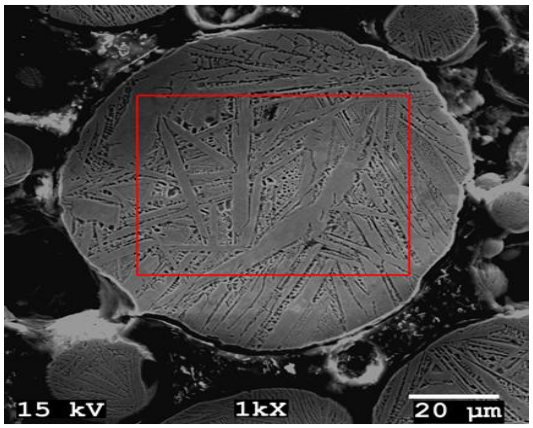


Figure 1. Cross-sectional sem image of the fe-cr-c powder

Table 2. Chemical composition of fe-cr-c alloy powder (wt.%)

C	Si	Mn	P	S	Cr	Mo	Fe
3,251	1,156	0,865	0,002	0,002	23,431	0,214	Kalan

3. Results

3.1. Microstructural Analysis Results

The microstructures obtained from the samples are shown in Figure 2. As an example, Figure 3 presents the result of a line EDS analysis taken from the lower region of the sample containing 10 wt.% TiC. As can be seen from the provided microstructure images, a good metallurgical bond was achieved between the coating and the substrate material, with no observable defects such as voids. In the samples without TiC addition, the microstructure mainly consists of austenite and a eutectic structure (austenite + M_7C_3), whereas in the TiC-added samples, the presence of TiC precipitates is also evident. This can also be seen by XRD analysis given in Figure 4 for 10 wt. % TiC coating. When comparing the samples containing 10 wt.% and 20 wt.% TiC, it is naturally observed that the amount of TiC precipitates increases with increasing TiC content. During laser irradiation, partial melting of TiC particles occurs along with the melting of the matrix [1, 8]. Upon cooling, the matrix formation starts with the solidification of primary austenite, followed by eutectic formation between the austenite dendrites as solidification progresses. Furthermore, undissolved TiC particles are also observed in the sample containing 20 wt.% TiC.

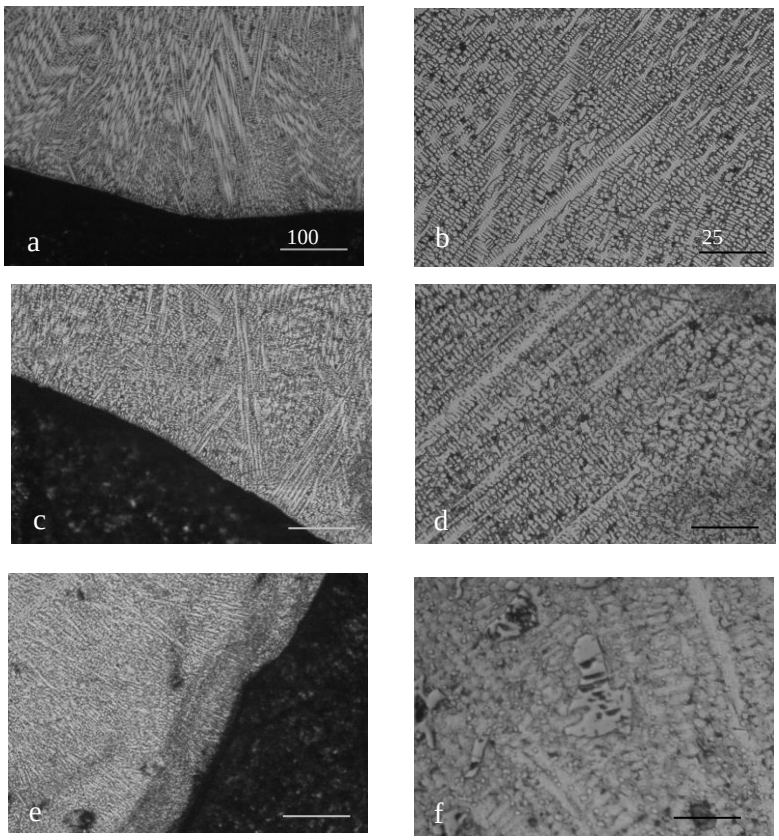


Figure 2. Optical microscope microstructure images taken from the middle region of coatings (a, b: 0 wt.%TiC, c, d: 10 wt.%TiC, and e,f: 20 wt.%TiC)

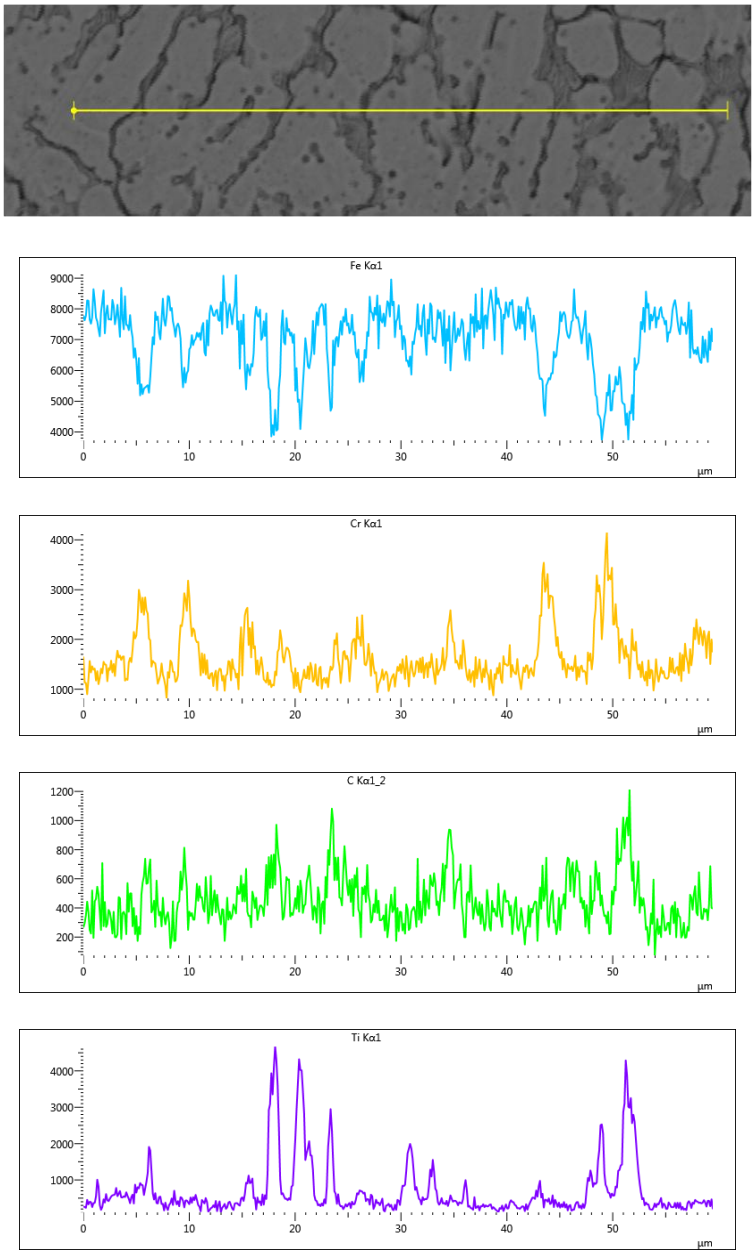


Figure 3. A line EDX analysis taken from the lower region of the sample containing 10 wt.% TiC

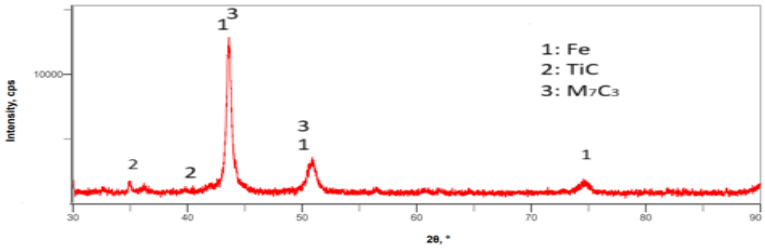


Figure 4. XRD patterns obtained from sample with 10 wt.% TiC

3.2. Hardness results

The maximum hardness values obtained from the coatings are given in Figure 5. As can be seen from these values, the hardness of the coatings increases with the increasing TiC content. This increase in hardness is attributed to the rise in the amount of the TiC particles in the coating [1, 8].

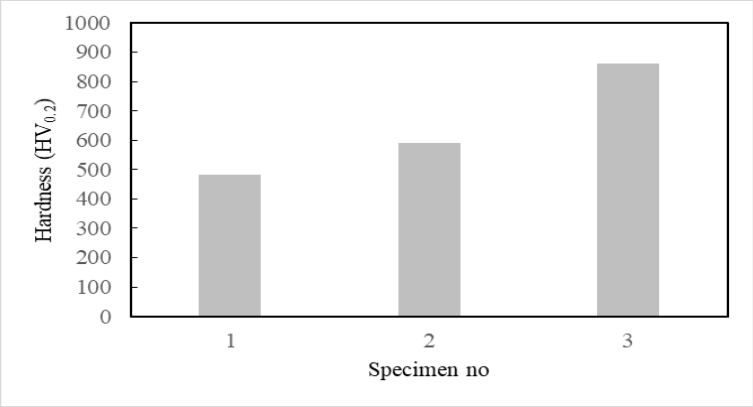


Figure 5. Maximum hardness values

4. Conclusions

In this study, a low carbon steel was coated using a laser beam with a powder mixture containing TiC powder of 10 wt. % and 20 wt.%, The microstructure and hardness of the coatings were examined. The general conclusions drawn from the study are summarized below:

- The coating microstructure changed from a primary austenite + eutectic structure to austenite + eutectic structure + TiC.
- TiC particles mostly dissolved and reformed for the 10 wt.%TiC.
- As the TiC content in the coating increased, the hardness also increased.
- Based on the obtained results, this study appears to have high applicability for components used in the wear industry.

Acknowledgment

This study was supported by Scientific and Technological Research Council of Turkey (TUBITAK) under the Grant Number 223M569. The authors thank to TUBITAK for their supports. This article is a part of the current Master of Science study of Asuman Merve Ekici.

5. References

[1] Zhang H.F, Wang L , Zhang S, C.L. Wu C.L, C.H. Zhang G H, X.Y. Sun X.Y, Chen J. “An investigation on wear and cavitation erosion-corrosion characteristics of the TiC modified Fe-based composite coating via laser cladding. *Journal of Materials Research and Technology*, 26, 8440-8455, 2023.

[2] Yan H, Zhang J, Zhang P, Yu Z, Li C, Xu P, Yunlong Lu. “Laser cladding of Co-based alloy/TiC/CaF2 self-lubricating composite coatings on copper for continuous casting mold”. Surf Coating Technol. 232, 362, 2013.

[3] He X, Xu X, Ge X, Tao J, Zhong Y, Liu Q, Wu G. “F101 Ni-based coating containing La_2O_3 by laser cladding on TC4 titanium alloy”. Rare Metals and Materials Engineering. 46 (4), 1074-1079, 2017.

[4] Zhang H, Zhang S, Wu C, Zhang C, Bai X, Sun X, Chen J. “Microstructural evolution and properties of novel ferritic stainless steel-mox prepared by laser melting deposition”. Journal of Thermal Spray Technology. 32(7), 2236-2249, 2023.

[5] Zhang J, Wei J, Wei S, Huang Z, Wei W, Liujie Xu. "Effect of cooling conditions on microstructure evolution and wear behavior of high chromium cast iron hardfacing layer". Materials Letters. 314, 131417, 2022.
[6] Tippayasam C, Taengwa C, Palomas J, Siripongsakul T, Thaweechai T, Kaewvilai A. "Effects of flux-cored arc welding technology on microstructure and wear resistance of Fe-Cr-C hardfacing alloy". Materials Today Communications. 35, 105569, 2023.
[7] Al Mangour B, Grzesiak D, Jenn M. "Selective laser melting of TiC reinforced 316L stainless steel matrix nanocomposites: influence of starting TiC particle size and volume content". Materials & Design. 104, 141-151, 2016.
[8] Zhang Z, Wang X, Zhang Q, Liang Y, Ren L, Li X. "Fabrication of Fe-based composite coatings reinforced by TiC particles and its microstructure and wear resistance of 40Cr gear steel by low energy pulsed laser cladding". Optics & Laser Technology. 119, 105622, 2019.

Effect of Surface Roughness on Average Heat Transfer in Jet Impingement Cooling

Celal KISTAK[1] , Ali TAŞKIRAN[2] , Nevin ÇELİK[1]

[1]Mechanical Engineering Department, Engineering Faculty, Fırat University, Elazığ, Türkiye.
[2]Mechanical Engineering Department, Engineering Faculty, Fırat University, Elazığ, Türkiye.
[3]Mechanical Engineering Department, Engineering Faculty, Şırnak University, Şırnak, Türkiye.
[1*]ckistak@firat.edu.tr, [2]nevicelik@firat.edu.tr, [3]ataskiran@sirnak.edu.tr
[1](ORCID: 0000-0003-4621-5405), [2](ORCID: 0000-0002-0578-5711),[3](0000-0003-2456-5316)

Abstract
Jet impingement cooling is widely employed in industrial applications requiring high heat dissipation due to its superior thermal performance. The present study experimentally investigates the influence of surface roughness on local and average heat transfer from a moving flat plate subjected to a circular air jet. Experiments were conducted for Reynolds numbers in the range of 7000–20000. A stepper motor was used to control the target plate motion during jet impingement, while surface temperature distributions were obtained using an infrared thermal camera. The acquired data were processed to determine Nusselt number distributions under varying surface roughness conditions. The results indicate that increased surface roughness enhances turbulence intensity in the impingement region, leading to improvement in heat transfer. Furthermore, the enhancement effect was found to be more pronounced at higher Reynolds numbers, demonstrating the critical role of surface roughness in optimizing jet cooling performance.

Keywords: jet impingement cooling, surface roughness, moving plate, heat transfer enhancement, reynolds number, nusselt number

1. Introduction

Efficient heat transfer underpins technologies ranging from electronics thermal management to energy conversion systems. Among enhancement strategies, jet impingement is distinguished by its ability to deliver very high local heat-transfer coefficients with compact hardware and modest pumping power. Decades of research have clarified the roles of Reynolds number, nozzle–plate spacing, and nozzle geometry; systematic reviews and canonical studies provide correlations that remain the baseline for single and multiple jets on smooth, stationary targets [1–3].

Although most investigations address stationary target plates, a smaller—but practically important—literature considers moving surfaces. In a pioneering study, Subba Raju and Schlünder impinged a normal air jet on a heated, continuously moving belt and reported peak heat-transfer coefficients roughly 1.5–2× larger than those on a stationary plate, with strong sensitivity to belt speed and geometric spacing [4]. Later analyses of slot-jet impingement over translating walls corroborate that surface motion reshapes the stagnation and wall-jet regions and can elevate local and area-averaged Nusselt numbers [5].

Beyond jet parameters and target motion, engineered surface roughness offers a powerful, design-controllable lever. Roughened targets—ranging from dimples to pin/mini-pin fins and patterned ribs—disrupt near-wall layers, promote reattachment, and typically increase local and average Nusselt numbers relative to smooth baselines (with pressure-loss trade-offs that must be managed) [6–8]. For example, detailed maps on dimpled targets have established how dimple depth, pitch, and placement relative to the stagnation zone govern augmentation [6,7], while computations and entropy-generation analyses on model rough surfaces further explain the associated transport mechanisms [8]. Collectively, these results suggest that combining surface motion with tailored roughness is a promising route to compact, high-performance designs.

This study experiments on an air jet impinging on a heated moving surface, with and without engineered roughness, to quantify how Reynolds number and roughness topology interact to set local and global heat-transfer performance. A bespoke facility is constructed to emulate realistic operating conditions and to expose the coupled fluid-dynamic and thermal mechanisms governing enhancement.

[1]Corresponding author

2. Materials and Methods

This experimental study examines the influence of target-surface roughness on convective heat transfer performance in a jet-impingement cooling configuration applied to a heated, moving plate. The investigation specifically focuses on quantifying how varying levels of surface roughness affect the heat transfer enhancement relative to a smooth surface benchmark. For comprehensive assessment, a smooth target plate was tested under identical operating conditions to serve as a baseline reference. In addition to comparisons with the smooth–moving configuration, performance was also evaluated, where relevant, against the theoretical limit of a smooth–stationary target, enabling a broader contextual understanding of the impact of both motion and surface topology on heat exchange.

All experiments were carried out using a custom-designed test rig located in the Fluid Mechanics Laboratory, Department of Mechanical Engineering, Faculty of Engineering, Fırat University. The experimental setup was developed and operated with infrastructural and financial support from the Fırat University Scientific Research Projects Coordination Unit (FUBAP). The facility allowed for precise control of flow parameters and surface motion, ensuring high repeatability and reliability in data acquisition across multiple surface configurations and Reynolds number regimes.

2.1. Experimental Setup

The jet-impingement cooling test rig used in this study consists of several integrated subsystems designed to simulate and control high-fidelity convective cooling over a heated moving surface. Central to the system is a variable-speed axial blower fan equipped with a closed-loop speed and flow-rate controller, enabling precise adjustment of the jet velocity and Reynolds number. To ensure uniform and fully developed flow at the nozzle exit, the airflow is first passed through an upstream honeycomb structure and a set of flow straighteners, followed by a PVC ducting line that conditions and directs the flow toward the nozzle section.

The jet is delivered through an interchangeable circular nozzle, with fixed geometry for the primary configuration of this investigation. The nozzle is aligned perpendicularly above the test surface to ensure axisymmetric jet impingement. The target plate module, positioned directly beneath the nozzle, is either smooth or roughened, depending on the experimental case, and is uniformly heated to maintain a constant wall temperature throughout testing. This setup enables controlled comparison of heat transfer behavior across different surface topographies under consistent thermal boundary conditions.

To replicate realistic conditions involving surface motion, the plate is mounted on a motor–belt–pulley drive system, which imparts linear translation at a controlled speed. The motion control system includes rotational encoder feedback, ensuring accurate synchronization and repeatability in plate velocity.

A detailed representation of the experimental setup is provided in Figure 1., which shows both front and top views of the assembled apparatus, including key components and measurement points. A photograph of the actual assembled rig is presented in Figure 2., offering visual confirmation of the experimental infrastructure and component arrangement.

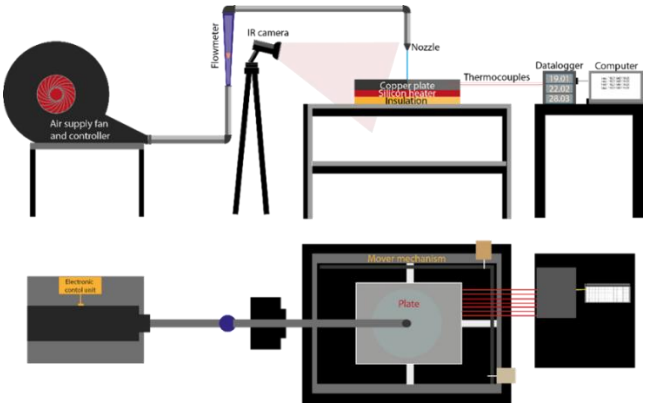


Figure 1. Schematic view of experimental setup

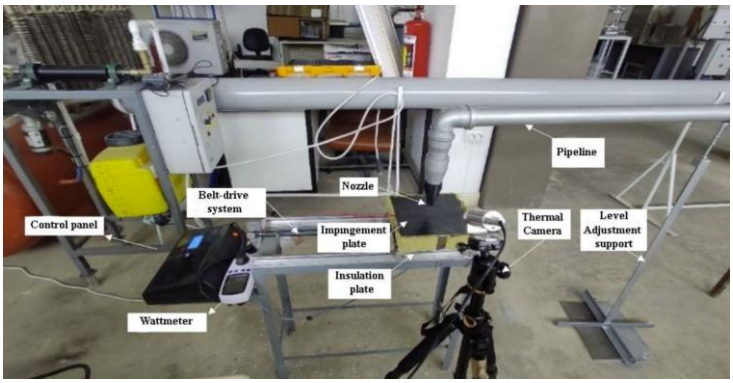


Figure 2. Photo of experimental setup

2.2. Roughened Surfaces and Parameters

To systematically isolate and quantify the effect of surface roughness on heat transfer performance, interchangeable roughness modules were employed on the target plate. These modules were designed to maintain identical material properties and thermal boundary conditions, ensuring that only the geometric surface texture varied across test cases. This modular approach enabled controlled experimentation and direct comparison between different roughness configurations and the smooth baseline.

Surface roughness was characterized by a single geometric parameter, the roughness pitch (denoted as s), defined as the center-to-center spacing between adjacent roughness elements. Three distinct roughness levels were investigated, corresponding to:

Rough 1: s = 45 mm

Rough 2: s = 30 mm

Rough 3: s = 15 mm

These pitch values were selected to span a representative range of practical surface textures while preserving manufacturing simplicity and repeatability. The roughness elements were uniformly distributed across the surface and maintained a consistent height and shape across all modules. By systematically varying the pitch while holding other conditions constant, the study aimed to elucidate the relationship between roughness scale and heat transfer enhancement under jet-impingement cooling with a moving target.

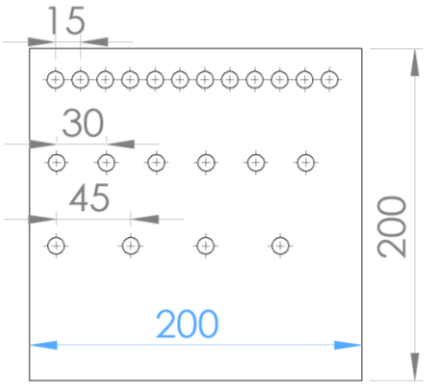


Figure 3. Layout of roughness on the plate

The smooth plate (s=0) served as the reference. All modules were mounted on the same base material so that thermal conductivity and thickness remained constant; thus, the observed differences can be attributed predominantly to the roughness geometry.

3. Results and Discussions

As illustrated in Figure 4., the area-averaged Nusselt number (Nū) exhibits a clear and monotonic increase with the Reynolds number (Re) across all investigated surface conditions, reflecting the fundamental dependency of convective heat transfer on inertial flow characteristics. The data further reveal a consistent and hierarchical enhancement in heat transfer performance as surface roughness increases, with the ranking preserved at each Reynolds number:
smooth < rough 1 (s = 45 mm) < rough 2 (s = 30 mm) < rough 3 (s = 15 mm).

At Re = 7000, the smooth surface yields an average Nusselt number of approximately 97.8, serving as the baseline for comparison. In contrast, rough surfaces rough 1, rough 2, and rough 3 produce values of 101.2, 104.2, and 107.9, respectively. These figures correspond to relative enhancements of ~3.5%, ~6.6%, and ~10.4% compared to the smooth plate, indicating that the introduction of surface roughness contributes meaningfully to convective heat transfer even at lower Reynolds regimes.

As Reynolds number increases to Re = 20000, this trend not only persists but becomes more pronounced. Here, the smooth plate reaches a Nu of 153.3, while rough 1, rough 2, and rough 3 achieve significantly elevated values of 160.2, 164.4, and 172.6, corresponding to heat transfer enhancements of ~4.5%, ~7.3%, and ~12.6%, respectively. These results collectively demonstrate that surface roughness induces a systematic and reproducible augmentation in heat transfer efficiency across the full operating range of Reynolds numbers considered.

Notably, the slope of the Nū–Re curves for all surfaces remains nearly invariant, implying that the dependency of Nusselt number on Reynolds number is fundamentally unaffected by surface roughness. On log–log axes, these curves align closely with a single power-law relationship characterized by an exponent near 0.43, reinforcing the observation that roughness elevates the magnitude of heat transfer without altering its scaling behavior with flow inertia.

Furthermore, a more detailed inspection reveals subtle distinctions in the behavior of rough surfaces at higher Reynolds numbers. The gap between rough 2 and rough 3 widens slightly with increasing Reynolds number, indicating that the most aggressive roughness profile (rough 3) derives greater benefit under conditions of intensified convective momentum. This phenomenon is likely attributable to enhanced turbulence production and wall-jet thinning, which act synergistically to improve convective transport mechanisms. In contrast, the gap between rough 1 and rough 2 remains relatively constant, suggesting a more linear scaling of heat transfer enhancement within that roughness regime.

Overall, these findings underscore the dual effect of surface roughness: while it substantially increases baseline convective performance, it does so by vertically shifting the Nu–Re trend rather than by modifying its sensitivity to flow conditions. This behavior is indicative of a passive augmentation strategy that retains the flow's underlying hydrodynamic structure while leveraging increased surface area, flow disturbance, and local mixing for improved heat exchange..

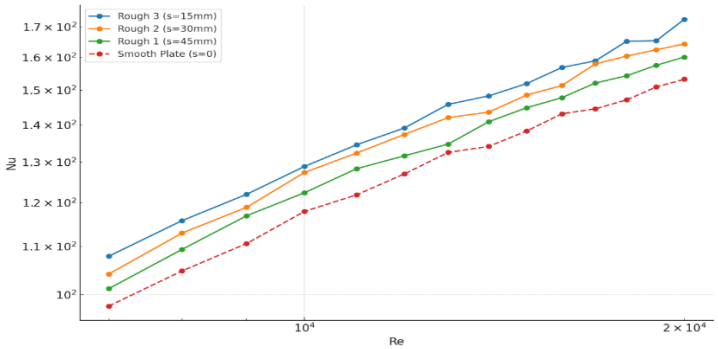


Figure 4. Nu as a function of Re for jet impingement on a moving plate: smooth and three differrent roughness

Figure 4. illustrates the variation of area-averaged Nusselt number (Nu) with Reynolds number (Re) for four different surface conditions: a smooth plate and three roughened surfaces with decreasing roughness pitch values of 45 mm, 30 mm, and 15 mm. The data clearly show that the addition of surface roughness systematically enhances convective heat transfer, with a consistent ranking across the entire range of Re:
Smooth < Rough 1 (s = 45 mm) < Rough 2 (s = 30 mm) < Rough 3 (s = 15 mm).

This enhancement is primarily driven by roughness-induced disruption of the near-wall flow region. The rough elements disturb the thermal boundary layer, promoting stronger mixing and broadening the region of high heat transfer near the stagnation zone. Additionally, they elevate the wall-jet plateau, leading to a sustained increase in Nu. As the pitch decreases, the roughness features more effectively destabilize the boundary layer and lead to earlier reattachment of the jet, especially in the case of Rough 3. This results in a larger and mildly Re-dependent gain in Nu, especially at higher flow rates.

Despite the increasing Nu values with finer roughness, all curves follow a similar slope, indicating that the scaling of Nu with Re is nearly identical for all surfaces. On log-log axes, this corresponds to a single power-law trend with an exponent close to 0.43. This suggests that roughness acts as a multiplicative gain in Nu rather than changing the fundamental dependence on Re.

From an engineering design perspective, this behavior supports a clear two-step optimization strategy. First, the Re should be selected based on system or process constraints. Then, roughness pitch can be adjusted to achieve additional, predictable gains in heat transfer performance. The consistent and monotonic improvement in Nu with decreasing roughness pitch makes it possible to fine-tune surfaces for thermal efficiency without modifying the flow regime.

It should be noted, however, that surface roughness may also introduce additional flow resistance, increasing pressure drop. Therefore, future work should include pressure-drop measurements to assess the thermal-hydraulic performance in a more comprehensive way. Calculating the overall efficiency or performance factor would provide insight into the trade-off between heat transfer enhancement and increased pumping power. Moreover, exploring the effects of unsteady jet conditions or alternative nozzle-to-plate spacings may reveal additional performance benefits when combined with surface roughening.

4. References

[1] Jambunathan, K., Lai, E., Moss, M.A., Button, B.L., “A review of heat transfer data for single circular jet impingement,” International Journal of Heat and Fluid Flow, 13(2), 1992, ss. 106–115.

[2] Beitelmal, A.H., Saad, M.A., Patel, C.D., “Analysis of an impinging two-dimensional jet,” ASME Journal of Heat Transfer, 128(3), 2006, ss. 307–318.

[3] Sagot, B., Antonini, G., Hasnaoui, M., Lebouché, M., “Jet impingement heat transfer on a flat plate at a constant wall temperature,” International Journal of Thermal Sciences, 47(12), 2008, ss. 1633–1641.

[4] Subba Raju, K.S., Schlünder, E.U., “Heat transfer between an impinging jet and a continuously moving belt,” Heat and Mass Transfer (Wärme- und Stoffübertragung), 10, 1977, ss. 165–173.

[5] Chattopadhyay, H., Saha, S.K., “Turbulent heat transfer over a moving plate due to an impinging slot jet,” Proc. 14th International Heat Transfer Conference (IHTC-14), 2010.

[6] Ekkad, S.V., Kontrovitz, D., Han, J.C., “Jet impingement heat transfer on dimpled target surfaces,” International Journal of Heat and Mass Transfer, 45(20), 2002, ss. 4297–4307.

[7] Kanokjaruvijit, K., vd., “Heat transfer correlations of perpendicularly impinging jets on dimpled surfaces,” International Journal of Heat and Mass Transfer, 53(15–16), 2010, ss. 3205–3215.

[8] Xu, P., Qiu, S., Liu, X., “Heat transfer and entropy generation in air-jet impingement on model rough surfaces,” International Communications in Heat and Mass Transfer, 75, 2016, ss. 8–14.

Effect of Chromium and Carbon on Wear Properties of Laser Coated Low Carbon Steels

Ceren ÖZCAN[1], Mehmet EROĞLU[2]

[1,2]Metallurgical and Materials Engineering, Engineering Faculty, Fırat University, Elazığ, Turkey
[1]cerenozcan182@gmail.com, [2]meroglu@firat.edu.tr
[1](ORCID: 0009-0005-7920-6471), [2](ORCID:0000-0002-5097-1943)

Abstract
Owing to its distinct advantages, laser technology has recently been adopted in the surface coating industry and is anticipated to gain increasing significance in the near future. In this study, the surface of a low-carbon steel substrate was coated using a laser beam and extra high-carbon ferrochrome (EHCFeCr) powder mixture. By varying the amount of EHCFeCr in the powder mixture, a progressive increase in chromium and carbon content within the coating was achieved. Following the coating operations, microstructural analyses were carried out on the treated regions, hardness measurements were performed, and abrasive wear tests were conducted. The results indicated that a defect-free coating was successfully obtained via laser processing, and that increased chromium and carbon content led to enhanced hardness and wear resistance. Based on the findings, the application of such coatings is recommended for relevant industrial use.

Keywords: laser coating, microstructure, hardness, wear

1. Introduction

Wear is a critical factor that significantly affects the performance and operational lifespan of mechanical components. It is estimated that approximately %85-90 of machinery failures are attributable to wear-related damage. Among the various wear mechanisms, abrasive wear is the most severe, often resulting in considerable maintenance costs due to the need for frequent repairs and component replacements. This type of wear is particularly detrimental to the functional elements of agricultural and mining equipment, which are frequently exposed to abrasive media during operation. One of the most effective strategies for enhancing the wear resistance and overall durability of such components is the application of hardfacing layers [1-3].

Laser cladding is an advanced surface engineering technique that enables the formation of thick coatings with a metallurgical bond to the substrate. This process employs a high-energy-density laser beam as a heat source to melt both the cladding material and a thin layer of the substrate, resulting in a strong, adherent coating. Compared to conventional surface modification methods, laser cladding offers several significant advantages, including a metallurgical interface between the coating and the base material, a highly refined microstructure, low dilution, and a narrow heat-affected zone (HAZ). In recent years, considerable research has focused on optimizing processing parameters for various powders and materials to enhance the performance of hardfacing coatings, particularly in wear-resistant applications [4, 5].

Carbides are refractory compounds that exhibit a combination of desirable properties, including high hardness, a certain degree of plasticity, and good wettability with metallic binders. This unique set of characteristics makes them attractive candidates for use in protective coatings designed to enhance resistance against thermal, corrosive, and mechanical wear. As a result, carbides are extensively employed in the fabrication of ceramic-based coatings on metallic substrates through conventional hardfacing methods such as reactive sputter deposition, manual shielded metal arc welding, electroslag remelting, and gas tungsten arc welding, all aimed at improving wear resistance [1,6,7]. The wear resistance of Fe–C–Cr alloys is strongly influenced by the volume fraction, size, and distribution uniformity of the carbides formed within the microstructure. Typically, these alloys exhibit a microstructure comprising martensite, retained austenite, and eutectic carbides. Superior wear resistance is often achieved when a well-developed eutectic structure is combined with the presence of primary carbides, which contribute to enhanced hardness and abrasion resistance [1, 8].
In this study, mixtures of pure iron powder and ultra-high carbon ferrochrome (UHCFeCr) were prepared with 20 wt. %, 40 wt. %, and 60 wt. % additions of UHCFeCr, and laser cladding was performed on low-carbon steel

[1]Corresponding author

substrates. The aim is to investigate how varying UHCFeCr contents influence the microstructure evolution**,** phase formation**,** hardness, and wear behavior of the laser-clad coatings. It is hypothesized that increasing the UHCFeCr content will result in a higher volume fraction of chromium-rich carbides, thereby improving the mechanical and tribological performance of the coatings. This research contributes to the growing body of knowledge on the use of ferroalloys in laser cladding and seeks to offer a practical, cost-efficient solution for the surface enhancement of low-carbon steels in industrial applications such as tooling, mining, and manufacturing equipment, where wear and corrosion resistance are critical requirements.

2. Material and Method

2.1. Powders Used and Mixing Ratios

Prior to laser cladding, powder mixtures with three different compositions containing EHCFeCr were prepared for the coating process. For this purpose, EHCFeCr was added to iron powder at ratios of 20 wt%, 40 wt. % and 60 wt.%, along with 2 wt.% ferrosilicon and 1 wt.% ferromanganese. The chemical compositions of the ferroalloys used are presented in Tables 1, 2, and 3, respecively. The particle size of the iron powder was in the range of 50–60 microns, while the particle sizes of the other powders were between 50–150 microns. The sample coding is provided in Table 4.

Table 1. Chemical analysis of EHCFeCr powder

Element	Cr	C	Si	P	S	Fe
wt.%	64-70	8-10	Max.1.5	Max.0.04	Max.0.06	Rest

Table 2. Chemical analysis of ferrosilicon powder

Element	Cr	C	Si	P	S	Al	Fe
wt.%	Max.0.6	Max. 0.1	65-69	Max.0.04	Max.0.015	Max.1	Rest

Table 3. Chemical analysis of ferromanganese powder

Element	Mn	C	Si	P	S	Fe
wt.%	78-84	0.5-1	Max. 1.8	Max.0.024	Max.0.02	Rest

Table 4. Sample coding and powder compositions (wt.%)

Sample no	Iron powder	EHCFeCr	Ferrosilicon	Ferromanganese
1	77	20	2	1
2	57	40	2	1
3	37	60	2	1

2.2. Laser cladding process

Laser cladding operations were carried out on low-carbon steel specimens with dimensions of 20 × 50 × 5 mm. The chemical composition of the substrate material is presented in Table 5. Prior to the cladding process, the surfaces of the substrates were initially rough-ground, followed by sanding with 120-grit flap wheels and finally with 240-mesh water-resistant abrasive paper to ensure a clean and uniform surface. The powder mixtures listed in Table 4 were blended with a small amount of organic binder to form a paste, which was then applied onto the substrate surfaces in a uniform layer with a height of approximately 3 mm. These coated specimens were subsequently dried and used for the laser cladding process.

Table 5. Chemical composition of the substrate material (wt.%).

C	Si	Mn	P	S	Fe
0.15	0.25	0.52	0.0021	0.0035	Kalan

2.3. Metallographic Procedures

Following the laser cladding process, the specimens were sectioned, and cross-sectional samples were prepared for microstructural examination. These samples were mounted in bakelite, then subjected to standard grinding and polishing procedures. After polishing, the surfaces were etched with a 3% nital solution to reveal the microstructure. Microstructural evaluations were carried out using both optical microscopy and scanning electron microscopy (SEM). Representative micrographs were captured, and elemental analyses were performed using spectrometer and energy-dispersive X-ray spectroscopy (EDX). In addition, X-ray diffraction (XRD) analysis was conducted on the coatings to identify the phases present in the clad layers.

2.4. Hardness Measurements

Vickers hardness measurements were performed on the prepared samples starting from the top surface of the coating toward the base material. Measurements were taken at 200 μm intervals under a load of 200 grams.

2.5. Abrasive Wear Tests

After surface grinding and polishing of the coatings, the samples were subjected to abrasive wear testing using the pin-on-disc method. Prior to testing, each specimen was weighed. Following the wear test, the samples were weighed again to determine the weight loss. SiC waterproof sandpaper with 60 grit was used in the test. The tests were conducted under loads of 10 N and wear path was 300 m.

3. Results

3.1. Microstructure

The optical microscope images obtained as a result of the microstructure examinations are presented in Figure 1.

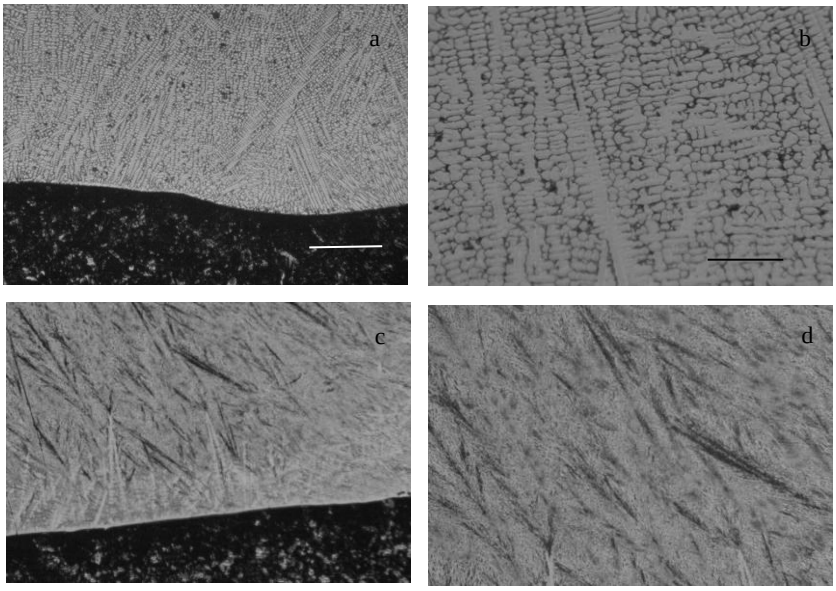


Figure 1. Optical microscope microstructure images taken from the coatings (a,b: sample 1, c,d: sample 2, and e,f: sample 3)

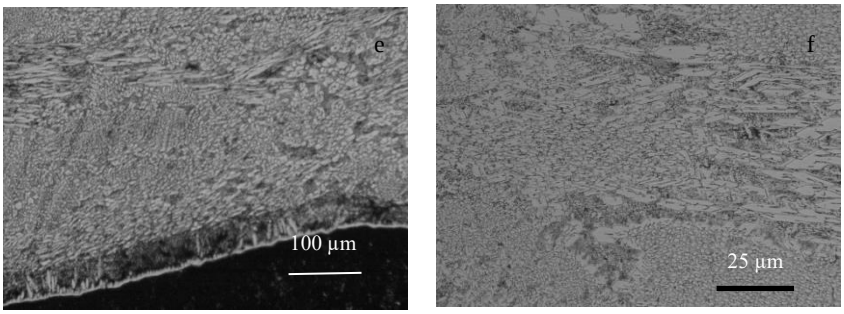


Figure 1. (Continue) Optical microscope microstructure images taken from the coatings (a,b: sample 1, c,d: sample 2, and e,f: sample 3)

As can be seen from these images, when 20 wt.% ferrochromium is used, the microstructure consists of austenite and eutectic (austenite + M_7C_3). The XRD result for the sample 3 is given in Figure 2, and the chemical analysis for Cr and C is presented in Table 5. It is stated that the analysis and microstructure are consistent. As the amount of ferrochromium increases, the amounts of carbon and chromium transferred to the coating also increase, which changes the type and amount of phases formed in the microstructure. In sample 2, the microstructure consists of eutectic structure, while sample 3 consists of primary M_7C_3 carbides with small amount of eutectic structure. This situation is consistent with the chromium and carbon content [1, 8].

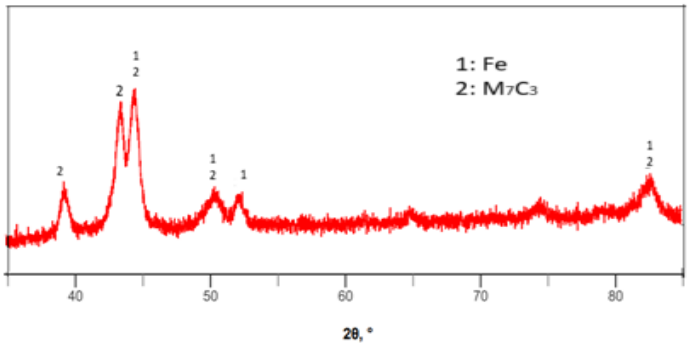


Figure 2. XRD patterns obtained from sample 3

Table 6. Amounts of Cr and C elements in the coatings (wt.%)

Specimen no	Cr	C
1	9.81	1.2
2	21.28	2,9
3	36,23	4,5

3.2. Hardness results

The maximum hardness values obtained from the coatings are given in Figure 3. As can be seen from these values, the hardness of the coatings increases with the increasing EHCFeCr content. This increase in hardness is attributed to the rise in the amount of the M_7C_3 phase formed in the coating [1, 8].

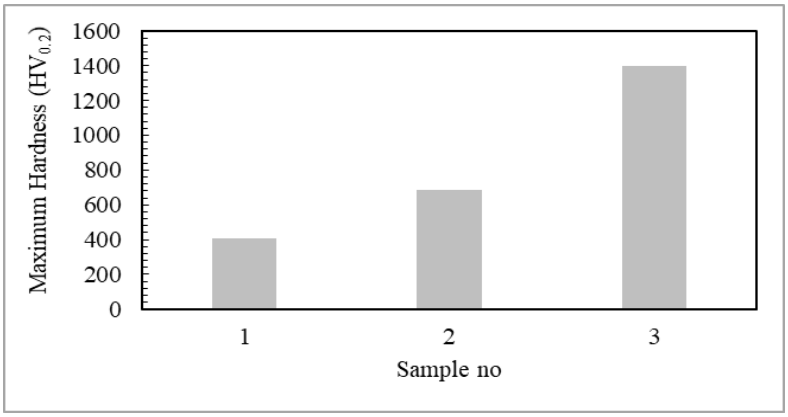


Figure 3. Maximum hardness values

3.3. Wear Test Results

The results of the mass loss measurements from the samples are presented in Figure 4. As can be seen from this figure, the mass losses decrease as the EHCFeCr content increases. This behavior is attributed to the presence of M_7C_3 phases in the coating. The amount of M_7C_3 carbide increases with the EHCFeCr content, and the resulting increase in surface hardness protects the surface against abrasive particles [8].

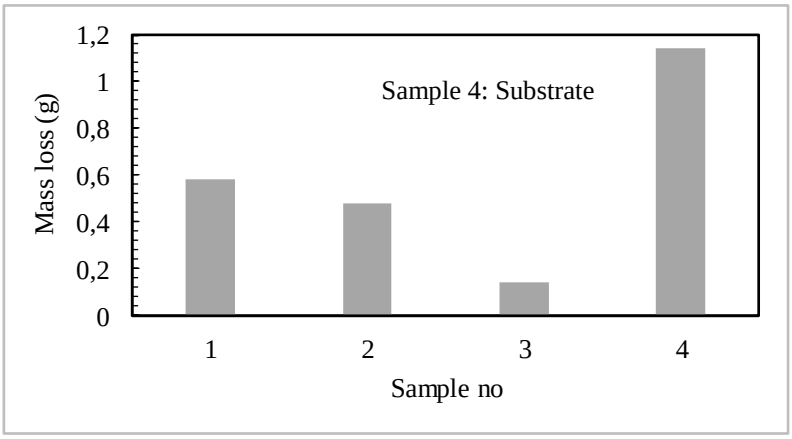


Figure 4. Mass losses of the clad coatings and steel substrate

4. Conclusions

In this study, a low carbon steel was coated using a laser beam with a powder mixture containing EHCFeCr powder of 20 wt.%, 40 wt. %, and 60 wt. %. The microstructure, hardness, and wear properties of the coatings were examined. The general conclusions drawn from the study are summarized below:

- The coating microstructure changed from a primary austenite + eutectic structure to M_7C_3 with small amount eutectic structure.
- As the M_7C_3 content in the coating increased, the hardness also increased.
- As the M_7C_3 content in the coating increased, the wear resistance improved.
- Based on the obtained results, this study appears to have high applicability for components used in the wear industry.

Acknowledgment

This work is supported by the Management Unit of the Scientific Research Projects of Fırat University (FUBAP) (Project Number: MF.25.87). This article is a part of the current Master of Science study of Ceren Özcan.

5. References

[1] Jankauskas V, Katinas E, Laskauskas A, Antonov M, Varnauskas V, Gedzevicius I, Alekneviciene V. "Effect of electrode covering composition on the microstructure, wear, and economic feasibility of Fe-C-Cr manual arc-welded hardfacings". *Coatings*, 10(3), 294, 2024.

[2] Holmberg K, Kivikytö-Reponen P, Härkisaari P, Valtonen K, Erdemir A. "Global energy consumption due to friction and wear in the mining industry". *Tribology International*, 115, 116–139, 2017.

[3] Sadeghi F, Najafi H, Abbasi A. "The effect of Ta substitution for Nb on the microstructure and wear resistance of an Fe–Cr–C hardfacing alloy". *Surface & Coatings Technology*, 324, 85–91, 2017.

[4] Wu Q, Li W, Zhong N, Wang G, Wang H. "Microstructure and wear behavior of laser cladding VC–Cr7C3 ceramic coating on steel substrate". *Materials & Design*, 49, 10–18, 2013.

[5] Liu ZD, Zhang XC, Xuan FZ, Wang ZD, Tu ST. "In situ synthesis of TiN/Ti3Al intermetallic matrix composite coatings on Ti6Al4V alloy". *Materials & Design*, 37, 268–273, 2012.

[6] Pan CG, Wang HC, Zhou JL. "Plasma cladding of Ni60–SiC cermet coating on hot forging die". *Materials Science and Engineering A*, 457, 319–324, 2007.

[7] Lin YC, Lin YC, Chen YC. "Evolution of the microstructure and tribological performance of Ti–6Al–4V cladding with TiN powder". *Materials & Design*, 36, 584–590, 2012.

[8] Yang J, Tian J, Hao F, Dan T, Ren X, Yang Y, Yang Q. "Microstructure and wear resistance of the hypereutectic Fe–Cr–C alloy hardfacing metals with different La2O3 additives". *Applied Surface Science*, 289, 437–444, 2014.

Photochemical Oxidation and Radical-Driven Processes in Hydrometallurgical Leaching: An Overview

Zeynel Abidin SARI[1], M. Deniz TURAN[2], Hasan NİZAMOĞLU[3], Şeyda BALIKÇI[4], Hakan AYHAN[5]

[1]Department of Metallurgy-Iskenderun Vocational School of Higher Education, Iskenderun Technical University, Hatay 31200, Turkey
[2,3,4]Department of Metallurgy and Materials Engineering, Fırat University, 23119 Elaziğ, Turkey
[5] Department of Map and Cadastre-Pertek Sakine Genç Vocational School, Munzur University, 62500 Tunceli, Turkey
[1]zeynel.sari@iste.edu.tr, [2]mdturan@firat.edu.tr, [3]hsnizamoglu@gmail.com, [4]seydabalikci@gmail.com, [5]hayhan@munzur.edu.tr
[1](ORCID: 0000-0001-5932-2141), [2](ORCID: 0000-0002-2136-1425), [3](ORCID: 0000-0002-1801-7156), [4](ORCID: 0000-0001-5632-6219), [5](ORCID: 0000-0001-5849-5243)

Abstract
The integration of photochemical and photocatalytic processes into hydrometallurgical leaching has emerged as a promising approach for enhancing metal recovery from low-grade ores and secondary resources. Conventional pyrometallurgical methods, while efficient, pose significant environmental and economic challenges, motivating the exploration of alternative, sustainable techniques. Photochemical methods utilize ultraviolet (UV) and vacuum UV (VUV) radiation to generate highly reactive species, including hydroxyl (OH·), superoxide ($O_2^{\cdot -}$), hydroperoxyl (HO_2·), and ozone (O_3), which accelerate redox reactions and improve dissolution kinetics. UV-C and VUV lamps, alone or combined with oxidants like H_2O_2 and O_3, have been shown to efficiently produce these reactive species under ambient conditions, reducing energy consumption and chemical oxidant usage. Beyond sulfide minerals, photochemical approaches offer applications in sphalerite oxidation, processing of copper slags, and metal plating waste valorization. The effectiveness of these methods relies heavily on reactor design, light intensity, and wavelength selection. Overall, photochemical leaching presents a sustainable, energy-efficient, and environmentally friendly strategy for improving metal extraction and addressing limitations of conventional hydrometallurgical processes.

Keywords: photochemical oxidation, ultraviolet, radical, leaching.

1. Introduction

The rapid advancement of industrialization and the increasing emphasis on sustainable production have necessitated a critical reassessment of conventional metal recovery techniques. Although pyrometallurgical processes have historically been preferred due to their high efficiency, they are associated with substantial drawbacks, including elevated energy consumption, complex process control, high capital investment, and the emission of environmentally harmful gases. Consequently, the tightening of environmental regulations has compelled the investigation and development of alternative, more environmentally benign methodologies [1]. Hydrometallurgical methods are particularly advantageous for the processing of low-grade ores and secondary metal resources due to their lower energy consumption, simpler process control, and environmentally friendly characteristics. However, the direct leaching of certain sulfide ores and complex secondary wastes often presents significant challenges, as slow dissolution kinetics and limited metal recovery efficiencies have prompted researchers to explore novel process enhancement strategies [2]. In this context, photochemical reactions and photocatalytic processes have emerged as a promising alternative for enhancing metal dissolution in hydrometallurgical leaching operations. These systems, activated by UV or visible light, promote the generation of highly reactive oxidizing species in the solution—such as hydroxyl radicals, superoxide radicals, and ozone—thereby accelerating redox reactions. In particular, the hydroxyl radical, due to its exceptionally high oxidation potential, can engage in rapid and non-selective reactions with both inorganic and organic compounds, offering significant potential to improve the dissolution kinetics in leaching processes [3]. The advantages offered by photochemical and photocatalytic methods have not been confined to environmental remediation technologies; they have also begun to attract considerable interest in metallurgical applications. The utilization of photochemically generated reactive species within leaching systems not only enhances metal recovery efficiencies but also enables the operation of processes under lower energy requirements. Consequently, these approaches open new avenues for reducing environmental impacts and improving the economic performance of metal extraction processes.

In conclusion, the integration of photochemical reactions into hydrometallurgical leaching processes is regarded as a promising approach, particularly for the processing of sulfide ores and secondary metal resources. In this regard,

[1]Corresponding author

understanding the reaction mechanisms of photo-activated species will provide significant contributions to the development of future environmentally friendly and sustainable metal production technologies

2. Classification of UV Radiation and Its Photochemical Activity

UV radiation is one of the most critical light sources employed in photochemical processes. Due to its high energy density and the quantum energy per photon, UV radiation plays a pivotal role in initiating numerous chemical and photochemical reactions. Positioned between visible light and X-rays in terms of wavelength, UV radiation spans the 100–400 nm range, which enables its application across a wide variety of fields [4]. UV radiation, which induces different interactions depending on its energy level, can be classified into four main groups based on wavelength. UVA (315–400 nm) possesses lower energy and generally exhibits surface-level effects, being referred to as long-wavelength UV. UVB (280–315 nm) carries higher energy and plays a significant role not only in biological systems but also in chemical transformations. UVC (200–280 nm), due to its short wavelength, is primarily employed in disinfection applications and facilitates radical formation, thereby enhancing the efficiency of oxidative processes. Additionally, the vacuum ultraviolet (VUV, 100–200 nm) region cannot propagate in the atmosphere and is therefore only usable under vacuum conditions; however, it provides sufficiently high energy to break chemical bonds, playing a critical role in the generation of reactive oxygen species (ROS) [5, 6]. These different UV regions particularly promote the formation of species with high oxidation potential, such as hydroxyl radicals ($OH^{\cdot}$), superoxide anions ($O_2^{\cdot -}$), and ozone (O_3) [7]. These reactive species are utilized not only in environmental remediation technologies but also in hydrometallurgical applications, including metal ion oxidation, enhancement of dissolution kinetics, and photocatalytic precipitation. Therefore, the spectral classification of UV radiation, the energy associated with each wavelength range, and the capacity of this energy to interact with chemical systems are of critical importance for both process design and industrial-scale applications.

3. Photophysical Behaviors and Photochemical Transformations

Photochemical reactions are processes initiated by the absorption of light, leading to permanent transformations in the chemical structure of molecules. At the core of these reactions lies the excitation of molecular electronic states through the energy carried by photons. When a molecule absorbs a photon, its electrons are promoted to higher energy levels, rendering the molecule unstable in its excited state. These unstable species subsequently undergo various chemical reactions, resulting in the formation of new products. In this respect, photochemical processes differ from purely photophysical events, as they involve bond cleavage, radical generation, and the formation of new bonds. Photochemical reactions are generally classified into primary and secondary stages. In the primary stage, the absorption of photons by a molecule promotes it to an excited state. In the secondary stage, the excited molecule interacts with other species, generating radicals or ionic intermediates. These intermediates, due to their high reactivity, reshape the chemical equilibrium of the system and play a critical role in the progression of reaction mechanisms.

3.1. Formation Mechanisms of Reactive Species

The most significant feature of photochemical reactions is their ability to generate reactive oxygen species and radical species with high oxidation potentials. Among these species, the most notable are [8]:

- **Hydroxyl radical ($OH^{\cdot}$):** With a redox potential of 2.8 V, it is one of the strongest known oxidants. It reacts rapidly with organic molecules and metal sulfide compounds.
- **Superoxide radical ($O_2^{\cdot -}$):** Formed through the one-electron reduction of oxygen, it plays an important role in initiating chain reactions.
- **Ozone (O_3):** Acts both as a direct oxidant and as a source of hydroxyl radicals through photolysis.
- **Hydroperoxyl radical ($HO_2^{\cdot}$):** Although a moderate oxidant, it is effective in sustaining radical chain mechanisms.

These radical species provide significant advantages in leaching processes by enhancing the dissolution rate of metal ions.

3.2. VUV Lamp-Induced Radical Production

Vacuum ultraviolet radiation, owing to its high energy, is particularly effective in the photolysis of water molecules. The absorption of VUV photons leads to the following reaction:

$H_2O + h\nu(VUV) \rightarrow OH^{\cdot} + H^{\cdot}$

In this process, hydroxyl radicals are directly generated. Moreover, VUV irradiation can also interact with dissolved oxygen in the solution, triggering the formation of ozone (O_3) and superoxide radicals ($O_2^{\cdot -}$). Due to these characteristics, VUV lamps are considered highly valuable in advanced oxidation processes as well as in the leaching of metal oxides and sulfide minerals.

3.3 Radical Activation by UV-C Lamps

UV-C radiation (200–280 nm), particularly mercury vapor lamps operating at 254 nm, is widely employed on an industrial scale. UV-C, either alone or in combination with oxidants (such as H_2O_2 and O_3), generates radicals with very high efficiency. The most well known mechanisms are as follows [9]:

Photolysis of hydrogen peroxide:

$H_2O_2 + h\nu(UV{-}C) \rightarrow 2OH^{\cdot}$ (1)

Photolysis of ozone:

$O_3 + h\nu(UV{-}C) \rightarrow O_2 + O(1D)$ (2)

$O(1D) + H_2O \rightarrow 2OH^{\cdot}$ (3)

Through these mechanisms, UV-C radiation accelerates the oxidation of metal ions and enhances the dissolution kinetics in leaching solutions. It plays a particularly critical role in overcoming passivation layers and improving the extraction efficiency of sulfide minerals (e.g., ZnS, $CuFeS_2$) during leaching.

4. Application of Photochemical Methods in Hydrometallurgical Leaching Processes

Hydrometallurgical methods have emerged as a more economical and environmentally friendly alternative to pyrometallurgical techniques for the processing of low-grade ores and metallurgical residues. However, the efficiency of hydrometallurgical leaching processes is often limited, particularly in the dissolution of sulfide minerals. Minerals such as chalcopyrite ($CuFeS_2$), sphalerite (ZnS), and galena (PbS) only partially dissolve in acidic media, and enhancing their solubility typically requires the use of oxidants, elevated pressure, and high temperatures. In this context, the integration of photochemical processes into leaching technologies has emerged as a promising novel approach.

4.1. Fundamentals of Photochemical Processes

Photochemical reactions are based on the interaction of electromagnetic radiation—particularly in the UV and VUV regions—with water or solution components, leading to the formation of radical species. These radicals, due to their high oxidation potentials, can significantly accelerate the dissolution of minerals that are otherwise considered inert.

For example, the exposure of pure water to vacuum ultraviolet (VUV) radiation at a wavelength of 185 nm undergoes photolysis as follows:

$H_2O + h\nu(185\ nm) \rightarrow OH^{\cdot} + H^{\cdot}$ (4)

The hydroxyl radical generated in this process is one of the strongest oxidants known, with a standard oxidation potential of 2.80 V. Similarly, the radiation emitted by UV-C lamps (254 nm) is particularly effective in the photolysis of peroxides and chlorinated compounds.

$H_2O_2 + h\nu(254\ nm) \rightarrow 2OH^{\cdot}$ (5)

Furthermore, in the presence of dissolved oxygen in the solution, superoxide ($O_2^{\cdot -}$) and hydroperoxyl ($HO_2^{\cdot}$) species are formed.

$$O_2^{\cdot -} + H^+ \rightarrow HO_2^{\cdot} \quad (6)$$

As a result of these chain reactions, both oxidizing species ($OH^{\cdot}$, $HO_2^{\cdot}$, H_2O_2) and reducing species ($H^{\cdot}$, e^-_{aq}) are generated in the system.

4.2. Photochemical Oxidation in the Leaching of Sulfide Minerals

A major challenge in the leaching of sulfide minerals is the formation of passivation layers on the mineral surface. For instance, during the leaching of chalcopyrite, CuS or elemental sulfur may form, creating a protective film on the surface. In a photochemical environment, radicals such as $OH^{\cdot}$ and $HO_2^{\cdot}$ can disrupt these passivation layers, thereby facilitating dissolution.

The possible photochemical oxidation mechanism for chalcopyrite can be summarized as follows:

Surface Oxidation:

$$CuFeS_2 + (OH^{\cdot}, h\nu) \rightarrow Cu^{2+} + Fe^{2+} + S^0 \quad (7)$$

Oxidation of Elemental Sulfur (S):

$$S^0 + 2OH^{\cdot} \rightarrow SO_2 + H_2O \quad (8)$$

Oxidation of SO_2 in Aqueous Medium:

$$SO_2 + OH^{\cdot} \rightarrow HSO_3^- \quad (9)$$

$$HSO_3^- + OH^{\cdot} \rightarrow SO_4^{2-} + H^+ \quad (10)$$

As a result of this process, both Cu^{2+} and Fe^{2+} ions are released into the solution, while sulfur is completely converted to sulfate ions (SO_4^{2-}). This effectively overcomes the partial dissolution issues commonly encountered in conventional leaching methods.

4.3 Advantages of Photochemical Processes

- **High oxidation power:** $OH^{\cdot}$ and similar radicals exhibit stronger oxidizing ability than conventional oxidants (e.g., Fe^{3+}, O_2).
- **Removal of passivation layers:** Elemental sulfur and other intermediate products are fully converted to sulfate through photo-oxidation.
- **Environmentally friendly approach:** Consumption of chemical oxidants is reduced, relying instead on UV radiation and water-based radicals.
- **Energy efficiency:** Oxidation can occur at ambient temperature and atmospheric pressure.

Photochemical methods are not limited to the dissolution of sulfide minerals; they can also be applied in arsenic mobilization from arsenopyrite, oxidation of refractory gold ores, and treatment of metal plating industry wastes. However, due to the extremely short lifetimes of radical species, reactor design, selection of light intensity and wavelength, and overall process engineering are critical factors in ensuring efficiency.

5. Conclusions

In this study, the applicability of photochemical approaches in hydrometallurgical leaching processes has been investigated. The findings indicate that radicals generated by UV and VUV irradiation at specific wavelengths can oxidize

sulfide minerals, thereby enhancing metal dissolution. In particular, oxidation reactions driven by these radicals improve the dissolution kinetics of sulfide minerals, leading to increased metal recovery. Moreover, photochemical methods have been observed to offer environmental and economic advantages by reducing energy consumption and the need for chemical oxidants. In process design, parameters such as light intensity, wavelength selection, and reactor configuration have been identified as critical factors directly affecting efficiency.

Overall, photochemical leaching methods provide a sustainable and energy-efficient strategy for metal extraction, overcoming the limitations of conventional hydrometallurgical techniques. These approaches hold significant potential for the processing of both low-grade ores and secondary metal resources.

6. References

[1] Norgate, T. E., Jahanshahi, S., & Rankin, W. J. (2007). Assessing the environmental impact of metal production processes. *Journal of cleaner production*, *15*(8-9), 838-848.

[2] Chagnes, A. (Ed.). (2020). *Advances in Hydrometallurgy*. MDPI.

[3] Miklos, D. B., Remy, C., Jekel, M., Linden, K. G., Drewes, J. E., & Hübner, U. (2018). Evaluation of advanced oxidation processes for water and wastewater treatment–A critical review. *Water research*, *139*, 118-131.

[4] Tufail, A., Al-Rifai, J., Price, W. E., Van De Merwe, J. P., Leusch, F. D., & Hai, F. I. (2024). Elucidating the performance of UV-based photochemical processes for the removal of trace organic contaminants: Degradation and toxicity evaluation. *Chemosphere*, *350*, 140978.

[5] Gomez-Lopez, V. M., Koutchma, T., and Linden, K. (2012). Ultraviolet and Pulsed Light Processing of Fluid Foods, Novel Food Processing Technology, pp. 405-421

[6] López-Malo, A., Palou, E. (2005). Ultraviolet Light and Food Preservation, Novel Food Processing Technology, pp. 405-421

[7] Loraine, G. A., & Glaze, W. H. (1993). Destruction of vapor phase halogenated methanes by means of ultraviolet photolysis. In *Proceedings of the Industrial Waste Conference-Purdue University* (Vol. 47, pp. 309-309). Mosby Year Book Europe LTD.

[8] De Jager, T. L., Cockrell, A. E., & Du Plessis, S. S. (2017). Ultraviolet light induced generation of reactive oxygen species. *Ultraviolet light in human health, diseases and environment*, 15-23.

[9] Beltrán, F. J., Chávez, A. M., Jiménez-López, M. A., & Álvarez, P. M. (2024). Kinetic modelling of UVC and UVC/H_2O_2 oxidation of an aqueous mixture of antibiotics in a completely mixed batch photoreactor. *Environmental Science and Pollution Research*, *31*(43), 55222-55238.

Effect of Nitriding Duration on Microstructure and Wear Behavior of Two Different Steels

Emir KARAÇAM[1], Amir MOTALLEBZADEH[2], Mehmet EROĞLU[3]

[1,2]Şenmak Makine, Akçaburgaz Mah.Hadımköy Yolu Cad. No: 94 Esenyurt, İstanbul, Türkiye.
[3]Metallurgical and Materials Engineering, Engineering Faculty, Fırat University, Elazığ, Turkey
[1]emirkaracam@senmak.com, [2]amir@senmak.com, [3]meroglu@firat.edu.tr
[1](ORCID: 0009-0000-3809-0476), [2](ORCID: 0000-0001-6753-9316),[3](ORCID: 0000-0002-5097-1943)

Abstract
Nitriding is a surface hardening technique applied by altering the chemical composition of the surface to both increase the surface hardness of steels and to preserve their surface properties under specific temperatures without degradation. In this study, the effect of nitriding duration on the nitrided layer thickness, microstructure, hardness, and abrasive wear resistance was investigated for 1.8519 and 1.8550 steels, which are commonly used in the polymer processing industry. In the experimental work, the selected steels were subjected to gas nitriding at two different temperatures with three different holding times. After nitriding, the thicknesses of the nitrided layers were measured, the microstructures were examined using a scanning electron microscope (SEM), and abrasive wear tests were conducted. As a result, it was observed that the nitrided layer thicknesses, hardness, and wear resistance varied depending on the type of steel, temperature and duration time.

Keywords: nitriding, microstructure, hardness, wear

1. Introduction

Many metallic materials used in industrial applications are exposed to various mechanical, thermal, and chemical effects during service. These unfavorable conditions, particularly wear, fatigue, and corrosion occurring at the surface, significantly reduce the lifespan of components. Therefore, various surface modification techniques are employed to improve the surface properties of engineering materials such as steels. One of the most widely used techniques is nitriding, a thermochemical surface hardening method, favored for its ability to increase surface hardness at relatively low temperatures [1-5].

Gas nitriding is a diffusion-based nitriding process that uses nitrogen-carrying gases such as ammonia (NH_3). During the process, nitrogen atoms diffuse into the steel surface, forming hard nitride phases (typically γ'-Fe_4N and ε-$Fe_{2-3}N$) and a nitrogen diffusion zone. This hard layer not only enhances the wear resistance of the surface but also improves fatigue life and corrosion resistance. Moreover, as the process is conducted at relatively low temperatures (typically between 500–520 °C), the risk of distortion in components is minimal, making it especially suitable for precision parts. The efficiency of the gas nitriding process depends on many parameters such as the chemical composition of the steel, nitriding temperature and duration, gas composition, and atmosphere control. In particular, steels containing alloying elements such as Al, Cr, Mo, and V tend to form fine and dispersed alloy nitrides in the diffusion zone, contributing to a significant improvement in mechanical performance [6-10].

In recent years, gas nitriding has been extensively applied to various alloy steels used in high-performance applications such as gears, shafts, and powertrain components. Among these, AISI 1.8519 (31CrMoV9) and AISI 1.8550 (34CrAlNi7) steels are of particular interest due to their distinct alloying compositions and nitriding responses. AISI 1.8519 is a low-alloy steel containing chromium and molybdenum, which are strong nitride-forming elements. These elements contribute to the formation of fine and stable alloy nitrides in the diffusion zone, improving wear resistance and hardness [11]. AISI 1.8550, on the other hand, is an aluminum- and nickel-alloyed steel specifically developed for nitriding applications. Aluminum, being a potent nitride former, enhances the formation of a thick and uniform compound layer, while nickel improves toughness and structural stability [12]. Wear is one of the most critical damage mechanisms limiting the service life of machine components, molds, gears, and piston rods. The hard nitride layers formed on the surface through nitriding significantly improve resistance against abrasive and adhesive wear. Recent studies have shown that the thickness, microhardness, and phase composition of the nitrided layer play a decisive role in wear behavior.

[1]Corresponding author

In this study, gas nitriding was applied as a surface modification technique to steel samples, and the resulting nitride layer's microstructure, hardness, and wear properties were investigated. The aim is to evaluate the effects of different nitriding parameters on wear resistance and contribute to the optimization of the process for industrial applications.

2. Material and Method

In this study, AISI 1.8519 and AISI 1.8550 steels were subjected to gas nitriding. The chemical compositions of these steels are presented in Table 1 and Table 2.

Table 1. Chemical composition of AISI 1.8519 steel (wt.%)

C	Si	Mn	P	S	Cr	Mo	V	Ni	Fe
0,32	0,21	0,55	0,021	0,029	2,51	0,21	0,18	-	Rest

Table 2. Chemical composition of AISI 1.8550 steel (wt.%)

C	Si	Mn	P	S	Cr	Mo	Ni	Al	Fe
0,35	0,32	0,61	0,021	0,025	1,72	0,23	1,1	0,9	Rest

The steels whose chemical compositions are presented in Tables 1 and 2 were subjected to gas nitriding after undergoing the necessary surface preparation processes. The parameters of gas nitriding are shown on Table 3. After the gas nitriding process, the samples were sectioned, subjected to standard metallographic procedures, and subsequently examined under scanning electron microscope. Microhardness measurements were carried out starting from the coated surface toward the substrate material under a load of 300 g. In the abrasive wear test, the pin-on-disc method was used, and wear losses were determined under load of 30 N.

Table 3. Gas nitriding parameters

Specimen	Temperature	Time	Additional Treatment
AISI 1.8519	515	120	-
AISI 1.8519	480	140	Treated
AISI 1.8519	480	160	Treated
AISI 1.8550	515	120	-
AISI 1.8550	480	140	Treated
AISI 1.8550	480	160	Treated

3. Results

3.1. Microstructure analysis results

The SEM images obtained from the samples are presented in Figure 1, 2 and 3. As can be clearly seen from these microstructural images, nitride layers have formed on the substrate materials. As also evident from the images, and consistent with the literature [3,4]. As an example, elemental mapping and line scanning EDX obtained from the sample 1.8519 nitrided for 120 hours are given in Figure 4 and 5, respectively.

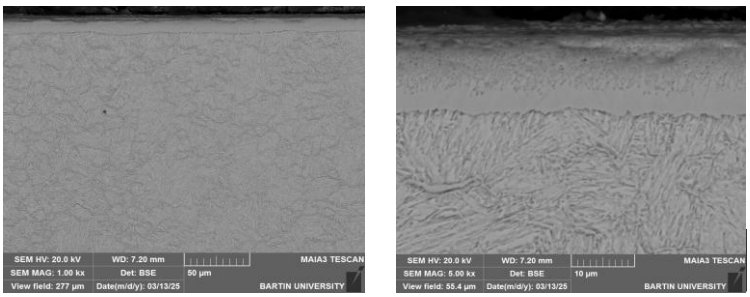


Figure 1. (Continue) SEM images of the nitride layer formed after 120 hours (a, b : AISI 1.8519 and c, d : AISI 1.8550)

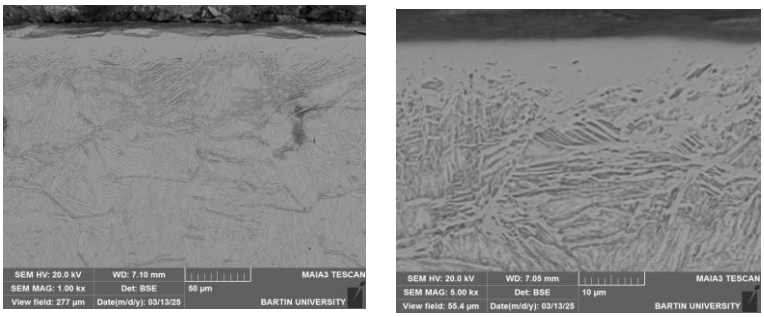


Figure 1. SEM images of the nitride layer formed after 120 hours (a, b : AISI 1.8519 and c, d : AISI 1.8550)

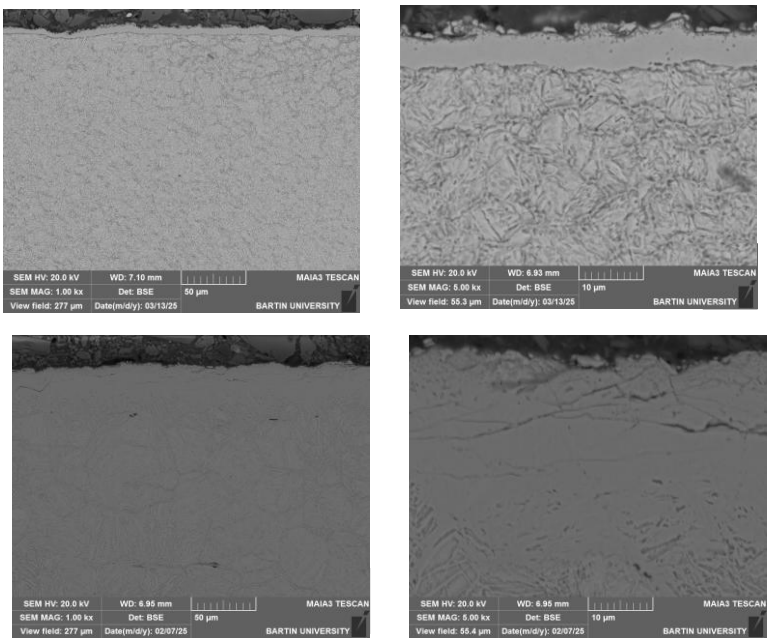


Figure 2. SEM images of the nitride layer formed after 140 hours (a, b: AISI 1.8519 and c, d: AISI 1.8550)

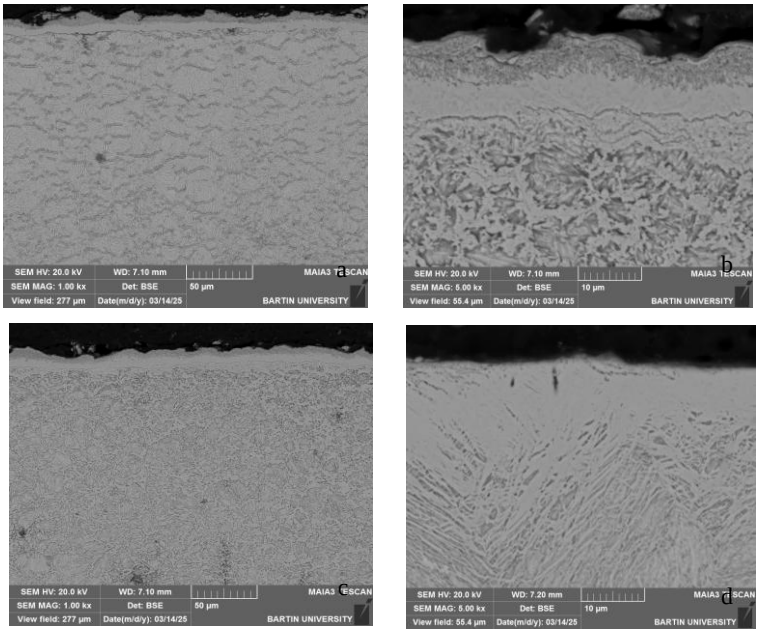


Figure 3. SEM images of the nitride layer formed after 160 hours (a, b: AISI 1.8519 and c, d: AISI 1.8550)

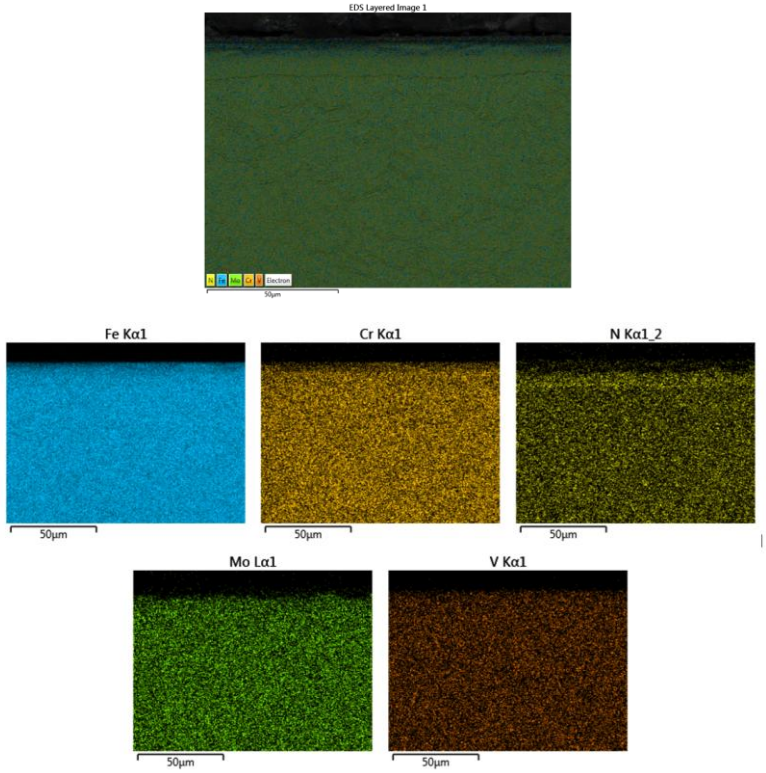


Figure 4. Elemental mapping obtained from the sample AISI 1.8519 nitrided for 120 hours.

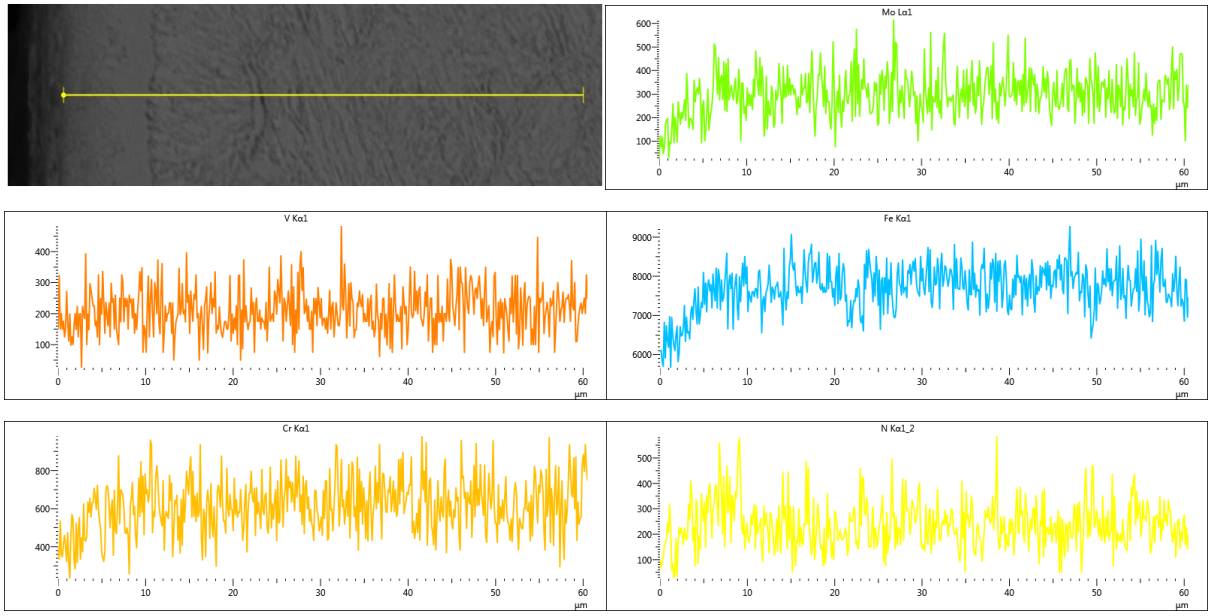


Figure 5. Line scanning obtained from the sample AISI 1.8519 nitrided for 120 hours

During the microstructural examinations, the total thicknesses of the nitride layers were also measured and are presented in Table 4. As can be seen from this table, the layer thickness changed with the temperature and holding time. Nitriding is a thermochemical process based on the diffusion of nitrogen atoms. Considering that diffusion depends on temperature and time, it can be seen that the results obtained are consistent with the literature [11].

Table 4. Nitrided layer thickness values (mm)

Steel	120 h	140 h	160 h
AISI 1.8519	0,62	0,58	0,59
AISI 1.8550	0,63	0,60	0,65

3.2. Microhardness Results

The maximum hardness values obtained from the samples are given in the Table 5. As can be seen from this table, the maximum hardness values showed differences depending on temperature and holding time. The highest hardness values were obtained for the 1.8550 steel. This is attributed to the alloying elements contained in this steel [12].

Table 5. Maximum hardness values (HV)

Steel	120 h	140 h	160 h
1.8519	916	901	887
1.8550	1102	1017	1127

3.3. Wear Test Results

The mass losses obtained from the wear test are given in Table 6. As can be seen from this table, when comparing the 120 and 160-hour durations, it is observed that the amount of wear increases with the nitriding time. This is attributed to the increase in the brittle white layer that forms with longer nitriding durations. When the 140-hour duration is evaluated, the removal of the nitride layer on the surface, leaving only the diffusion layer, results in a slight decrease in wear. When the two steels are compared with each other, the best results were obtained for the 1.8550 steel. This can be associated with the higher hardness of the nitrides formed by the alloying elements present in this steel [12].

Table 6. Weight loss values (g)

Specimen	120 h	140 h	160 h
AISI 1.8519	0.1789	0.3793	0,2612
AISI 1.8550	0.1682	0.2755	01921

4. Conclusions

In this study, two different steels were subjected to gas nitriding at the same temperature but for different durations. The relationships between microstructure, hardness, and wear were investigated, and the following general conclusions were drawn:

- Nitride and diffusion layers were formed during the nitriding process.
- The 1.8550 steel exhibited the highest hardness.
- Wear resistance decreased with increasing nitriding duration.
- The 1.8550 steel showed the highest wear resistance.
- Removal of the white layer from the surface before the wear test increased the amount of wear.
- Based on the results of this study, it is recommended to investigate the effect of nitriding duration on corrosion resistance in future work.

5. References

[1] Balasubramanian K, Khare S.V, Gall D. "Energetics of point defects in rocksalt structure transition metal nitrides: thermodynamic reasons for deviations from stoichiometry". *Acta Materialia*, 159, 77–88, 2018.

[2] Zhang Z, Li X, Dong H. "Plasma-nitriding and characterization of FeAl40 iron aluminide". Acta Materialia, 86, 341–351, 2015.

[3] Tong W, Tao N, Wang Z, Lu J, Lu K. "Nitriding iron at lower temperatures". Science, 299, 686–688, 2003.

[4] Steiner T, Mittemeijer E. J. "Alloying element nitride development in ferritic fe-based materials upon nitriding: a review". *Journal of Materials Engineering and Performance*, 25, 2091–2102, 2016.

[5] Kiełbasa K, Pelka R, Arabczyk W. "Studies of the kinetics of ammonia decomposition on promoted nanocrystalline iron using gas phases of different nitriding degree". *The Journal of Physical Chemistry A*, 114 (13), 4531–453, 2010.

[6] Bartłomiej W, Arabczyk W. "Investigation of nitriding and reduction processes in a nanocrystalline iron-ammonia-hydrogen system at 350 ◦C". *Physical Chemistry Chemical Physics*, 17 (31), 20185–20193, 2015.

[7] Arabczyk W, Skulmowska K, Pelka R, Lendzion-Bielun Oscillatory Z. "Mechanism of α-Fe(N)↔γ'-Fe_4N phase transformations during nanocrystalline iron nitriding". *Materials*, 15 (3), 1006, 2022.

[8] Moszynski D, Moszynska I, Arabczyk W. "Iron nitriding and reduction of iron nitrides in nanocrystalline Fe-N system". *Materials Letters*, 78, 32–34, 2012.

[9] Baranowska J, Szczecinski K, Wysiecki M. "Increasing of gas nitriding kinetics via surface pre-treatment". *Surface and Coatings Technology*, 151-152, 534–539, 2002.

[10] Wang B, Fu W, Dong F, Jin G, Feng W, Wang Z, Sun S. "Significant acceleration of nitriding kinetics in pure iron by pressurized gas treatment". *Materials & Design*, 85, 91–96, 2015.

[11] M. Sommer, Epp J, Steinbacher M, Fechte-Heinen H, Hoja S. "Variation of the compound layer structure by controlled gas nitriding and nitrocarburizing". *Heat Treatment of Materials*, 77 214-227, 2022.

[12] Szittnick A, Krull H.-G, Gaber C, Demuth M, Hochenauer C, Schwarz S. "Impact of process gases on scale formation when heating tool steels". *Practical Metallography*, 62 (4), 258-271, 2025.

A New BERT Based Hate Speech Detection Method Using Contrastive Learning

Turan Göktuğ ALTUNDOĞAN[1] , Mehmet KARAKÖSE [2]

[1]Computer Engineering Department, Engineering Faculty, Manisa Celal Bayar University, Manisa, Türkiye.
[2]Computer Engineering Department, Engineering Faculty, Firat University, Elazığ, Türkiye.
[1]turan.altundogan@cbu.edu.tr, [2]mkarakose@firat.edu.tr
[1](ORCID: 0000-0002-8677-3105), [2](ORCID: 0000-0002-3276-3788)

Abstract
Today, millions of people use social networks. The widespread and free use of social networks has made it necessary to monitor the content of posts for hate speech. Many hate speech detection methods utilizing language models with transformer architectures exist in the literature. These approaches utilize frozen large language model embeddings and a combination of different neural classifiers to perform classification. In this study, the BERT encoder was fine-tuned using a contrastive learning strategy. The feature extraction performance of the fine-tuned BERT encoder is presented using various data visualization and evaluation techniques. Then, this trained encoder was used as the backbone of a different classifier architectures to perform hate speech classification. The proposed method clearly outperforms existing studies in the literature with its phase utilizing contrastive learning strategies (F1-score: 98.5%).

Keywords: hate speech detection, large language models, transformer, contrastive learning

1. Introduction

Social networks are communication platforms that utilize web technologies, allowing people to share their thoughts freely and easily. Social network X (formerly Twitter), one of the most frequently used platforms, has over 600 million users. This technological advantage is also frequently exploited by those who exploit hate speech and harassment. To address this problem, many approaches focused on automatic hate speech detection have been developed. Hate speech classification has been achieved with success rates exceeding 90%, particularly with the development of large language models using Transformer architectures.

Different methods have been developed for hate speech detection. For example, in a study using GloVe embeddings and CNN-based models, an F1-score metric yielded 92% accuracy [1]. Another approach, using autoencoder-based feature extraction followed by an SVM classifier, achieved 95% accuracy [2]. Furthermore, neural network architectures based on the attention mechanism have achieved 90% accuracy in hate speech detection [3].

In a method using BERT encoder-based embeddings in a Bi-LSTM classifier, the success rate remained at 80% [4]. In another study conducted on multilingual data, the combination of MuRIL embedding and SVM (RBF) yielded the best results, achieving 96% accuracy in Malayalam, 72% in Tamil, and 66% in Kannada [5]. A K-means-based method both detected and categorized hate speech, but its F1-score was reported as 54% [6].

Other studies conducted on Twitter data have also yielded remarkable results. A deep learning-based model developed using Bag of Words and TF-IDF features achieved approximately 92% accuracy [7]. Furthermore, n-gram-based feature extraction was performed on a comprehensive dataset specifically created for Twitter, testing classifiers such as Logistic Regression, SVM, and Random Forest, with the highest success achieved by the Logistic Regression model [8].

The HCovBi-Caps model, a combination of CNN, Bi-GRU, and Capsule Network, was proposed. This model, powered by GloVe embeddings, performed multi-layer feature extraction and achieved an F1-score of 93%, demonstrating one of the most successful results in the literature [9]. Another study proposed a hybrid method focused on hate speech detection in imbalanced datasets [10]. This study combined local context information with CNN layers and sequential dependencies with RNN layers, developing a two-layer hybrid CNN-RNN architecture. The proposed model demonstrated successful results with an F1-score of 91.4% in experimental evaluations.

Finally, a pattern-based embedding approach was developed in another study [11]. Tweet representations enriched with POS labels and sentiment scores were transferred to a Transformer-based model via a two-stage cross-attention structure.

[1]Corresponding author

The proposed method achieved an F1-score of 88% in experimental evaluations, making it one of the most powerful methods in the literature. Figure 1 shows a block diagram representing the methods of current literature studies focused on hate speech detection.

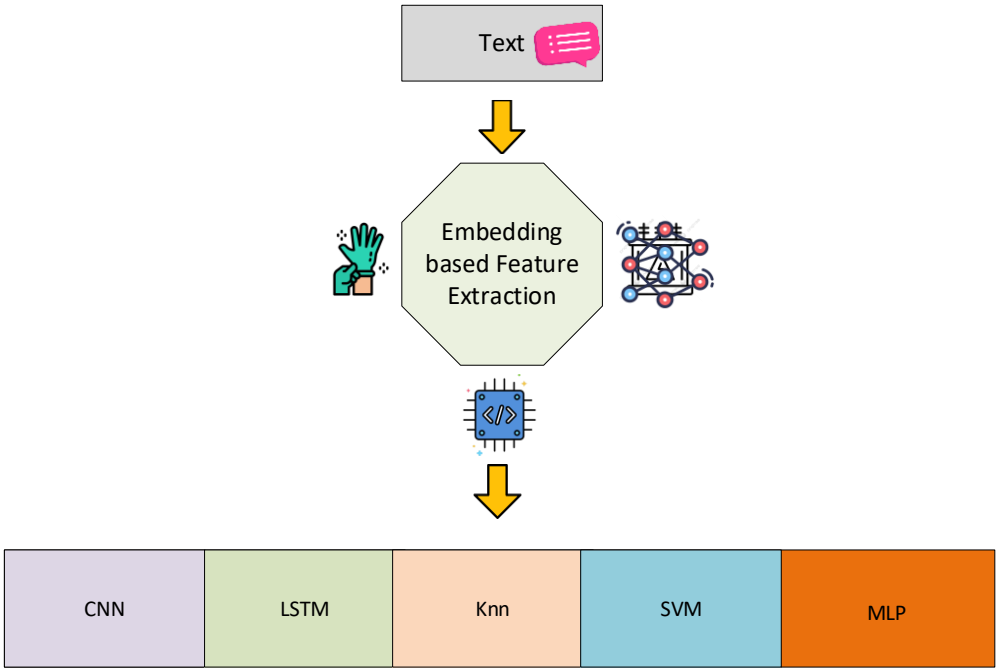


Figure 1. Block diagram of methods proposed for hate speech detection in the literature.

When examining the studies in the literature [1-15], it is seen that different feature extraction methods and neural architectures have generally emerged for hate speech detection. Indeed, it is evident that the classification performance of existing methods is around 90%. The learning capabilities of neural architectures that solve natural language processing problems can be strengthened by using learning strategies in addition to layering and feature extraction methods. For example, CLIP [16], an embedding architecture developed for image-text alignment, achieved very high performance by utilizing contrastive learning strategies. In a previous study, we developed an approach inspired by contrastive learning to summarize query-driven transcript text [17]. The performance contribution of the contrastive learning strategy to the developed approach was clearly demonstrated by the ablation evaluations. Therefore, the primary motivation for this study was to demonstrate the performance contribution of contrastive learning techniques for various NLP problems addressed in the literature. An examination of hate speech detection approaches reveals that embedding-based methods are frequently used for feature extraction. Given that contrastive learning strategies could be significantly useful for providing embedding architectures with a distinctive feature-based mission, we identified hate speech detection as the primary focus of our proposed approach, driven by this motivation.

2. Proposed Approach

In this study, a new approach based on the Siamese BERT encoder and contrastive learning is proposed for hate speech detection. The fine-tuned BERT encoder using the contrastive learning strategy will be used for feature extraction from the texts analyzed for hate speech detection. Figure 2 shows the block diagram representing the fine-tuning of the Siamese BERT encoder. Figure 2 shows that three BERT encoders with shared parameters encode one anchor and two samples from the same and different classes. For the problem under consideration, the anchors can consist of hate speech and normal text. To avoid data imbalances, an approach using equal numbers of anchors from both classes in the training triplets was employed. A loss function called the Triplett loss was used for the training process. Equation 1 provides the mathematical expression of the triplet loss function. Equation 1 shows that a denotes the anchor, p denotes the similar sample, and n denotes the non-similar sample.

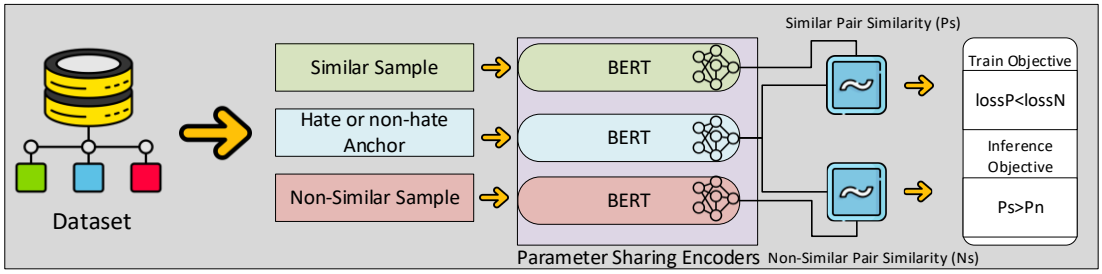


Figure 2. Block diagram of Siamese-BERT encoder fine-tuning.

$$L = \max(0,\ D(a,p)\ -\ D(a,n)\ +\ \text{margin}) \tag{1}$$

After fine-tuning the BERT encoder, the embedding features of both the default BERT and the Siamese-BERT model were visualized using the t-SNE technique. The t-SNE method enables the visualization of multidimensional feature vectors in two-dimensional space. In this context, the similarities of each data point with its neighbors are calculated. These similarities are calculated using a Gaussian normal distribution;

$$p_{j|i} = \frac{\exp(-\frac{\|x_i - x_j\|^2}{2\sigma_i^2})}{\sum_{k \neq i} \exp(-\frac{\|x_i - x_k\|^2}{2\sigma_i^2})} \tag{2}$$

When equation 2 is examined, xi and xj represent the data points in the high-dimensional space, σ represents the variance, and p represents the probability distribution.

Finally, the training of the logistic regression, support vector machine, random forest, multi-layer perceptron, naive Bayes, and decision tree classifiers that perform the detection process was performed. Figure 3 shows a block diagram illustrating the use of Siamese BERT embeddings for hate speech detection. When Figure 3 is examined, the embedding vectors of the incoming text input are first obtained with the fine-tuned BERT model. Then, the training of the seven different classifier models given in the diagram was performed. Of these models;

- Logistic Regression is a statistical classification method that models the probabilities between two classes by learning a linear decision boundary. It performs probability-based separation between classes using the sigmoid function.
- Support Vector Machines (SVM) are supervised learning algorithms that aim to find an optimal hyperplane that provides the maximum margin between classes. With a linear kernel, data is classified directly by a separable plane.
- An MLP is a fully connected artificial neural network with one or more hidden layers. It can generate nonlinear decision boundaries by learning abstract representations from inputs.
- Random Forest is an ensemble of decision trees generated by randomly sampling subsets. This structure provides high generalization performance by reducing overfitting.
- Gradient Boosting is an ensemble learning method that minimizes errors by repeatedly training weak classifiers. Each model focuses on the error patterns of the previous model.
- The Naive Bayes classifier estimates the probability distribution between classes using Bayes' Theorem. It assumes that features are independent of each other and is usually modeled according to a Gaussian distribution.
- Decision trees are hierarchical structures that perform decision-making by dividing data into branches based on attributes. Each node attempts to divide the data set into more homogeneous subgroups.

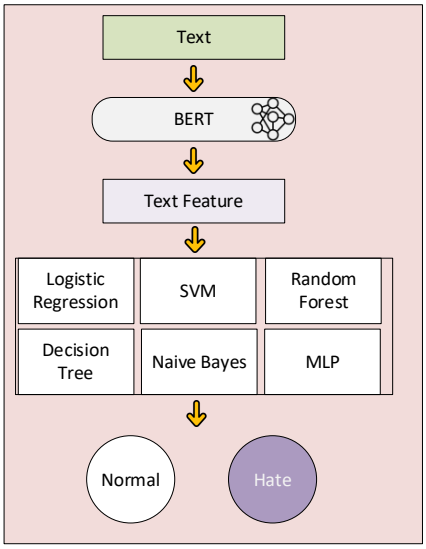


Figure 3. Block diagram of hate-speech detection method proposed in this study.

3. Experimental Results

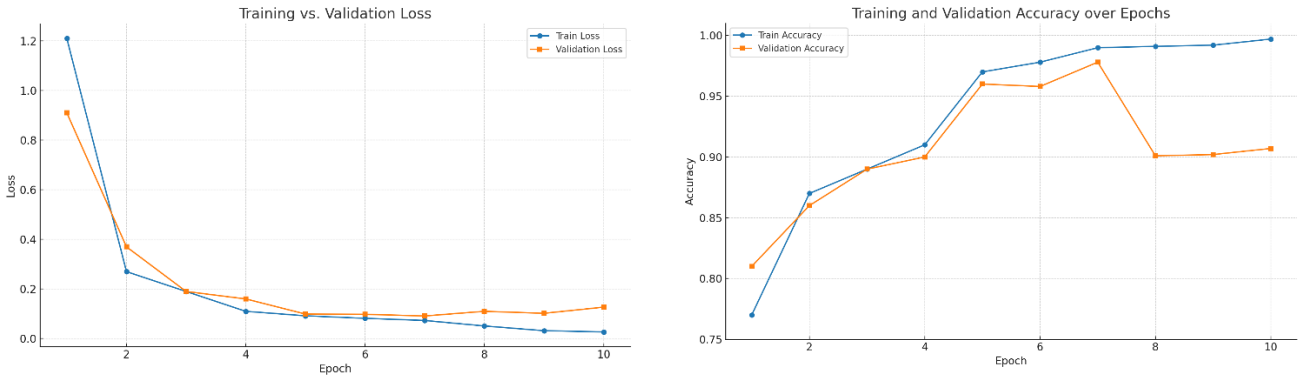
In this study, the Davidson Hate Speech dataset was used for all training, validation, and testing processes of the proposed method [18]. Quantitative information about the dataset and the training, validation, and test partitions are presented in Table 1. A downsampling preprocess was applied to reduce the intra-class imbalance of the dataset. Following this process, the Siamese-BERT fine-tuning process was performed. The model's train, validation, and test loss and accuracy evaluations are presented in Table 2. During model training, an early-stopping mechanism was used to prevent overfitting after the 7th epoch, ensuring the best epoch was recorded. The model's train, validation loss, and accuracy graphs across epochs are presented in Figure 4. Figure 5 shows the t-SNE visualization results of the default BERT and the fine-tuned Siamese-BERT model. Figure 5 shows that the Siamese-BERT model produces embeddings with much higher discrimination. Table 3 presents comparisons of recall, precision, F1-score, and accuracy metrics for different classification models. Comparative results of the classification models reveal that the proposed method significantly improves performance over all classification approaches. The comparative results reveal that the SVM model has the best hate speech detection performance.

Table 2. Train, test and validation splits of dataset.

Class	Total	Down-sample	Train	Validation	Test
Normal	4163	4163	3500	500	163
Hate	20120	4163	3500	500	163

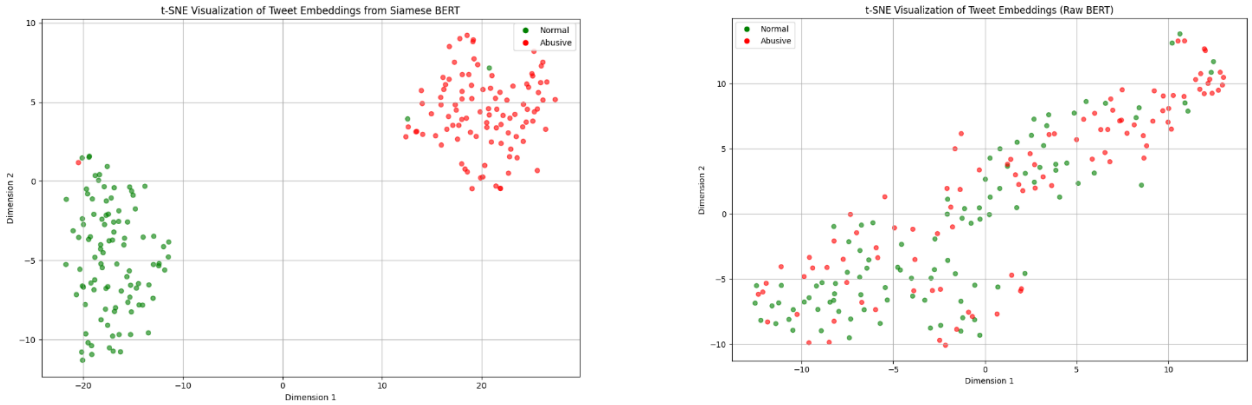
Table 2. Training and validation results and parameters of Siamese-BERT.

	Maximum	Early Stopping
Epoch	10	7
	Training	**Validation**
Triplet Count	35000	5000
Loss	0.0731	0.0914
Accuracy	98.99%	97.81%



(a) (b)

Figure 4. Loss (a) and accuracy (b) graph of Siamese-BERT fine-tuning.



(a) (b)

Figure 5. T-SNE visualization of Siamese-BERT (a) and default BERT (b) embeddings for 100 test sample.

Table 3. Ablation and test evaluation of different classifier methods using default and Siamese-BERT.

	Siamese-BERT				**Default BERT**			
Method	**Acc**	**Recall**	**Precision**	**F1-Score**	**Acc**	**Recall**	**Precision**	**F1-Score**
Logistic Reg.	98.38%	**98.5%**	98.39%	98.44%	84.3%	86.7%	82.56%	84.775
SVM	**98.44%**	98.38%	**98.61**%	**98.5%**	84.57%	82.56%	87.09%	84.77%
MLP	98.38%	98.16%	98.73%	98.45%	79.59%	66.40%	92.15%	77.18%
Random For.	**98.44%**	**98.5%**	98.5%	**98.5%**	77.01%	76.33%	78.78%	78.88%
Grad. Boost.	98.08%	98.15%	98.15%	98.15%	78.21%	78.29%	79.48%	78.88%
Naïve Bayes	97.78%	98.25%	97.46%	97.86%	58.58%	48.85%	63.13%	55.08%
Decision Tree	96.64%	95.81%	97.81%	96.80%	64.95%	66.74%	66.13%	66.44%

Finally, Table 4 presents the F1-score comparisons of the proposed method with studies in the literature. When the comparative results are examined, it is seen that the highest-performing result belongs to the Siamese-BERT + SVM model proposed in this study.

Table 4. Comparative results.

Study	Approach	F1-Score
[1]	CNN	92%
[2]	Autoencoder	95%
[3]	RNN	90%
[4]	Glove	80%
[5]	Transformer	96%
[6]	DNN	72%
[7]	DNN	66%
[8]	LR, RF, SVM, DT	54%
[9]	CNN, Bi-LSTM	93%
[10]	CNN+RNN	91.4%
[11]	Transformer	88%
[12]	CNN+RNN	88.1%
[13]	RNN	95.54%
[14]	Transformer	90.08%
[15]	BERT	90%
Ours	Siamese-BERT + SVM	**98.5%**

4. Conclusions

In this study, a new high-performance approach for detecting hate speech on social media is developed. The proposed method provides an effective embedding removal method by fine-tuning the Siamese-BERT encoder using a contrastive-learning strategy. The superiority of the embedding vectors of the fine-tuned BERT model over the embeddings of the default BERT model is demonstrated with t-SNE visualization. These embeddings are then classified using Logistic Regression, SVM, MLP, Random Forest, Gradient Boost, Naive Bayes, and Decision Trees. Comparative results indicate that the proposed embedding removal approach provides significant performance improvement over classification approaches performed with the default BERT encoder. Furthermore, the proposed method demonstrates the superiority of the Siamese-BERT + SVM approach with an F1-score of 98.5% compared to existing approaches. Future studies will utilize the proposed contrastive-learning strategies to solve various classification problems in the NLP field.

5. References

[1] Roy, P. K., Tripathy, A. K., Das, T. K., & Gao, X. Z. (2020). A framework for hate speech detection using deep convolutional neural network. IEEE Access, 8, 204951-204962.

[2] Kumar, G., Singh, J. P., & Singh, A. K. (2023). Autoencoder-based feature extraction for identifying hate speech spreaders in social media. IEEE Transactions on Computational Social Systems, 11(4), 4819-4827.

[3] Ahmed, U., & Lin, J. C. W. (2022). Deep explainable hate speech active learning on social-media data. IEEE Transactions on Computational Social Systems, 11(4), 4625-4635.

[4] Alatawi, H. S., Alhothali, A. M., & Moria, K. M. (2021). Detecting white supremacist hate speech using domain specific word embedding with deep learning and BERT. IEEE Access, 9, 106363-106374.

[5] Sreelakshmi, K., Premjith, B., Chakravarthi, B. R., & Soman, K. P. (2024). Detection of hate speech and offensive language codemix text in dravidian languages using cost-sensitive learning approach. IEEE Access, 12, 20064-20090.

[6] Rodriguez, A., Chen, Y. L., & Argueta, C. (2022). FADOHS: framework for detection and integration of unstructured data of hate speech on facebook using sentiment and emotion analysis. IEEE Access, 10, 22400-22419.

[7] Miran, A. Z., & Yahia, H. S. (2023). Hate speech detection in social media (Twitter) using neural network. Journal of Mobile Multimedia, 19(3), 765-798.

[8] Watanabe, H., Bouazizi, M., & Ohtsuki, T. (2018). Hate speech on twitter: A pragmatic approach to collect hateful and offensive expressions and perform hate speech detection. IEEE access, 6, 13825-13835.

[9] Khan, S., Kamal, A., Fazil, M., Alshara, M. A., Sejwal, V. K., Alotaibi, R. M., ... & Alqahtani, S. (2022). HCovBi-caps: hate speech detection using convolutional and Bi-directional gated recurrent unit with Capsule network. IEEE Access, 10, 7881-7894.
[10] Riyadi, S., Andriyani, A. D., & Sulaiman, S. N. (2024). Improving hate speech detection using Double-Layers hybrid CNN-RNN model on imbalanced dataset. IEEE Access.
[11] Sharmila, P., Anbananthen, K. S. M., Chelliah, D., Parthasarathy, S., & Kannan, S. (2022). PDHS: Pattern-based deep hate speech detection with improved tweet representation. IEEE Access, 10, 105366-105376.
[12] Ibrahim, M. A., Winarto, T. S. Y., & Sulistiya, Z. D. (2025, April). Dual-Class Prompt Generation: Enhancing Indonesian Gender-Based Hate Speech Detection through Data Augmentation. In 2025 8th World Conference on Computing and Communication Technologies (WCCCT) (pp. 468-473). IEEE.
[13] Goswami, P., & Daniel, A. (2024, December). Hate Speech Detection Using Multiple Channel Recurrent Neural Network. In 2024 1st International Conference on Advances in Computing, Communication and Networking (ICAC2N) (pp. 1150-1155). IEEE.
[14] Nabila, P., & Setiawan, E. B. (2024, July). Adam and adamw optimization algorithm application on bert model for hate speech detection on twitter. In 2024 International Conference on Data Science and Its Applications (ICoDSA) (pp. 346-351). IEEE.
[15] Arif, M. F., & Nayak, J. (2024, December). Fine-Tuned BERT Model for Accurate Hate Speech Detection in Social Media. In 2024 International Conference on Modeling, Simulation & Intelligent Computing (MoSICom) (pp. 406-410). IEEE.
[16] Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., ... & Sutskever, I. (2021, July). Learning transferable visual models from natural language supervision. In International conference on machine learning (pp. 8748-8763). PmLR.
[17] Altundogan, T. G., & Karakose, M. (2025). QUBVIS: query based multi-modal summarization system using CLIP based transformer and vision language models. SoftwareX, 31, 102303.
[18] Davidson, T., Warmsley, D., Macy, M., & Weber, I. (2017, May). Automated hate speech detection and the problem of offensive language. In Proceedings of the international AAAI conference on web and social media (Vol. 11, No. 1, pp. 512-515).

An Autonomous Aerial Surveillance System for Early Detection of Wildfire Anomalies Using Smart Sensor Integration

Ebru KARAKÖSE[1], Hazar Diyar SARİ[2], Emirhan GÜLKIRAN[3], Ertuğrul MERT[4]

[1,2,3,4] Aviation Electrical and Electronics Department, School of Civil Aviation, Fırat University, Elazığ, Türkiye
eozbay@firat.edu.tr, hazardiyar0202@gmail.com, emirhangulkiran54@hotmail.com, ertugrulm242@gmail.com
[1]0000-0003-1191-6375, [2]0009-0000-9706-2934, [3]0009-0009-7397-6294, [4]0009-0009-7397-6294

Abstract
Forest fires, as a manifestation of the global climate crisis, have shown a dramatic increase both in frequency and severity, evolving into a multidimensional environmental disaster that threatens ecosystem integrity, biodiversity, human settlements, and the security of natural resources. In this context, the early detection of pre-fire anomalies is of critical importance. This study introduces an advanced balloon-based surveillance platform, autonomously operated and integrated with sensor systems, developed for deployment in fire-prone regions. The platform is built upon a lightweight balloon capable of remaining airborne at low altitudes for extended periods, featuring an aerodynamic design and constructed from graphene-coated, recyclable polyurethane material. Equipped with high-resolution thermal cameras, smoke and gas sensors, temperature-humidity detectors, and multispectral imagers, the system enables real-time monitoring of environmental deviations indicative of fire risk. Its energy demand is met through a hybrid system comprising flexible solar panels and micro wind turbines, ensuring an uninterrupted and sustainable operational cycle. Collected data are processed via an ESP32 microcontroller unit and transmitted to cloud-based systems through low-power LoRa/NB-IoT communication modules. Artificial intelligence-powered data analysis algorithms facilitate map-based visualization and early warning mechanisms. The proposed system aims to contribute innovatively to sustainable fire management by enhancing effective response capacity in disaster management.

Keywords: forest fire, early warning, balloon platform, sensor technology, hybrid energy, disaster management.

1. Introduction

In recent years, it has been observed that temperature extremes, decreases in relative humidity, and fluctuations in wind dynamics resulting from climate change have significantly increased the frequency and intensity of forest fires worldwide. Especially in the Mediterranean climate zone and other fire-prone ecosystems, this situation has led to the disruption of ecosystem integrity, losses in agricultural production, damage to residential areas, and the irreversible destruction of biodiversity. In this context, technologies such as satellite-based detection systems and terrestrial sensor networks have been developed and have been implemented for the early detection and rapid response to fires. However, due to the spatial coverage limitations, high installation and operational costs, and delays in detection time of these systems, it has been determined that they are unable to provide sufficient effectiveness.
Considering the current situation, it has been concluded that an alternative system, which is low-cost, has a wide coverage area, and can operate for a long duration, needs to be developed to make the fire early warning process more efficient. In line with this, the design of a next-generation monitoring platform that will overcome the limitations of existing technologies has been planned. The system to be designed is envisaged to be configured to continuously monitor environmental parameters, support early warning mechanisms, and be integrated into existing monitoring infrastructures. Furthermore, it is aimed to provide that the system to be developed will be powered by sustainable energy sources, reduce maintenance and operational costs, and also ensure long-term and uninterrupted operability.

The work to be carried out in this context is aimed to offer a new solution approach in fire management strategies while complementing the deficiencies of existing satellite and land-based systems. Thus, significant contributions in strengthening ecosystem resilience, increasing rapid response capacity against fires, and reducing environmental damage is planned to be provided[1-6].

[1]Corresponding author

A novel solution, a low-altitude, sensor-integrated, balloon-based early warning system powered by hybrid renewable energy, is planned to be designed. Unlike conventional satellite or ground-based methods, the platform is aimed to combine multi-sensor integration, real-time AI-based anomaly detection, and energy autonomy into a single, low-cost system. It is foreseen that data collected from thermal cameras, smoke and gas detectors, and temperature-humidity sensors will be processed on an ESP32-based embedded system and transmitted to cloud servers via LoRa or NB-IoT communication protocols.

In contrast to prior studies, the integration of AI-driven anomaly detection algorithms is aimed to allow for immediate identification of subtle changes in environmental parameters, thereby significantly reducing the time between fire onset and alert transmission. The system's hybrid energy infrastructure, consisting of flexible photovoltaic panels and mini wind turbines, is intended to ensure uninterrupted, long-term operation, surpassing the limited endurance of battery-dependent alternatives. Additionally, the graphene-coated polyurethane carrier structure is expected to offer a unique combination of low density and high strength, while aerodynamic optimization is expected to enhance stability under variable wind conditions. With GPS-based positioning and directional control, the platform is planned to remain stationary at designated coordinates, ensuring continuous localized surveillance a feature often lacking in high-mobility UAV-based monitoring systems.

By combining these innovations, the proposed platform is aimed to differentiate itself from existing technologies in terms of mobility, sustainability, and real-time responsiveness. This system is intended to not only complement current monitoring infrastructures but also address their temporal and spatial limitations, thereby strengthening wildfire prevention strategies and improving environmental resilience.

2. Proposed System

In this study, an innovative airborne, sensor-integrated balloon (zeppelin) platform is proposed for the early detection of forest fires. A key feature of the design is planned to be its ability to provide continuous and effective surveillance even under challenging terrain conditions, with the capacity to remain stationary at low altitudes or to be steered when required. The system, which is planned to be powered by hybrid renewable energy sources, is conceived to support long-term missions by ensuring energy independence. In this way, it is aimed to contribute to minimizing fire risks through real-time data transmission. The balloon hull is planned to be aerodynamically optimized and manufactured from advanced materials such as graphene-coated polyurethane, which is foreseen to increase payload capacity and ensure long-term durability against environmental conditions such as UV radiation, rain, and wind. Continuous energy production and the assurance of system sustainability are aimed to be achieved through flexible photovoltaic panels to be integrated onto the balloon surface and mini wind turbines to be positioned at the bottom.Within the scope of sensor integration, thermal cameras, smoke detectors, temperature-humidity sensors, and multispectral imaging units are planned to be mounted on the underside of the balloon, with the aim of detecting fires at their early stages. The collected data is foreseen to be pre-processed on an ESP32 or similar embedded microcontroller and transmitted to central control units via low-power communication technologies such as LoRa or NB-IoT. This communication infrastructure is expected to provide fast and reliable data transfer for early warning applications thanks to its wide coverage and low latency [7–9]. The processing of the obtained data through cloud-based analysis software and its visualization on Geographic Information Systems (GIS) is planned, which is foreseen to present critical parameters such as temperature increase, humidity decrease, and smoke density via color-coded risk maps. Furthermore, the automatic transmission of this information to disaster management units is also conceived as an essential component of the design. Upon examining similar studies, it is seen that while satellite-based observation systems are able to cover wide areas, they remain limited due to detection delays, whereas terrestrial sensor networks are not always sufficient because of their high installation costs and restricted coverage [10–12]. In this context, the proposed balloon-based early warning system is conceived as a complementary and sustainable solution, with the capacity for long-term autonomous operation.

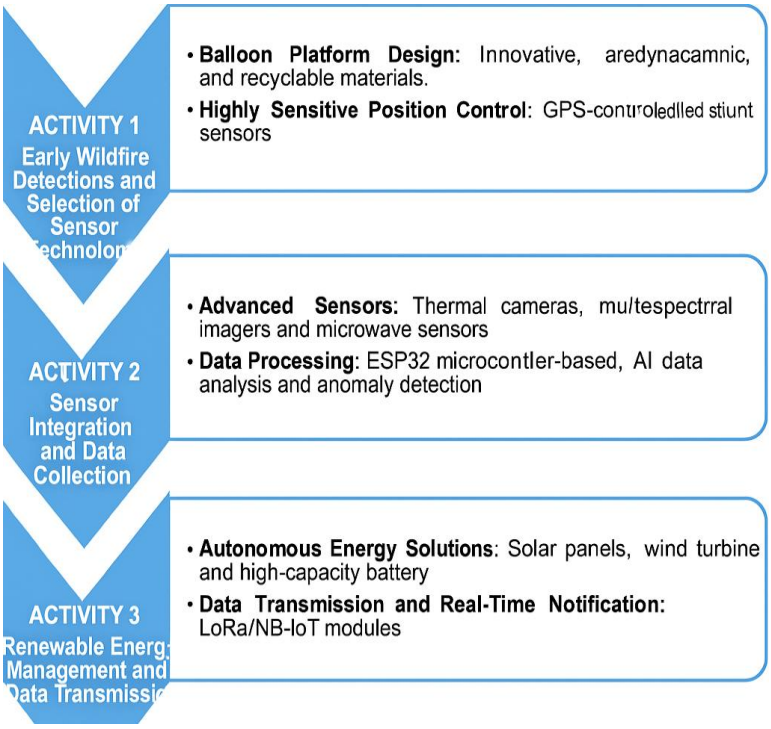


Figure 1. Smart Balloon: A Renewable Energy-Powered Autonomous Platform for Early Wildfire Detection

2.1. System Design and Balloon Platform

Innovative Balloon Design: In order to effectively monitor large forest areas, a specialized balloon system has been designed that can operate at low altitudes and remain airborne for extended periods. The main components of this system are as follows [13-15].

- **Balloon Platform:** Manufactured from lightweight, environmentally friendly, and recyclable materials (e.g., graphene-coated polyurethane). These materials enhance the balloon's durability and longevity, ensuring resistance under harsh weather conditions [16-17].
- **Aerodynamic Design:** The balloon possesses an aerodynamic form that minimizes wind resistance. This design aims to maintain stability in flight and improve energy efficiency [18-19].
- **High-Precision Stabilization Systems:** Utilizing GPS-based control mechanisms, the balloon can remain stationary over designated areas and, when necessary, navigate along pre-programmed routes [20-22].

2.2. Validation Stages of the Wildfire Detection Balloon System

The wildfire detection balloon system is planned to be evaluated through three key stages. First, it has been planned that laboratory calibration will be carried out to ensure the accuracy of smoke and temperature sensors by testing them under various fire scenarios and optimizing their performance. Next, software integration and functional testing are envisaged to be implemented in order to verify that all sensors and software components operate seamlessly together. Finally, pilot applications in forest areas are planned to be conducted to test the system in real environments, with the aim of confirming its capability to deliver accurate, real-time wildfire alerts through controlled fire simulations. Through this planned process, it is aimed that the system's reliability and readiness for future operational use will be ensured.

As illustrated in Figure 2, the sensor integration and data processing workflow of the balloon-based early wildfire detection system is planned to be designed in a holistic structure. Thermal cameras, multispectral imaging units, and environmental sensors measuring temperature, humidity, and wind speed are envisaged to be integrated into the

balloon platform, enabling detailed monitoring of microclimatic variations within the forest ecosystem. The collected data are planned to be pre-processed through an ESP32-based microcontroller, and anomalies are aimed to be identified with the assistance of artificial intelligence-supported algorithms. In addition, Figure 2 is planned to illustrate the process of transmitting the collected data to a cloud-based platform, where they are aimed to be analyzed using machine learning algorithms to generate real-time wildfire alerts. In this way, both pre-fire risk factors and early indicators of ignition are planned to be detected at an early stage and promptly communicated to the relevant authorities.

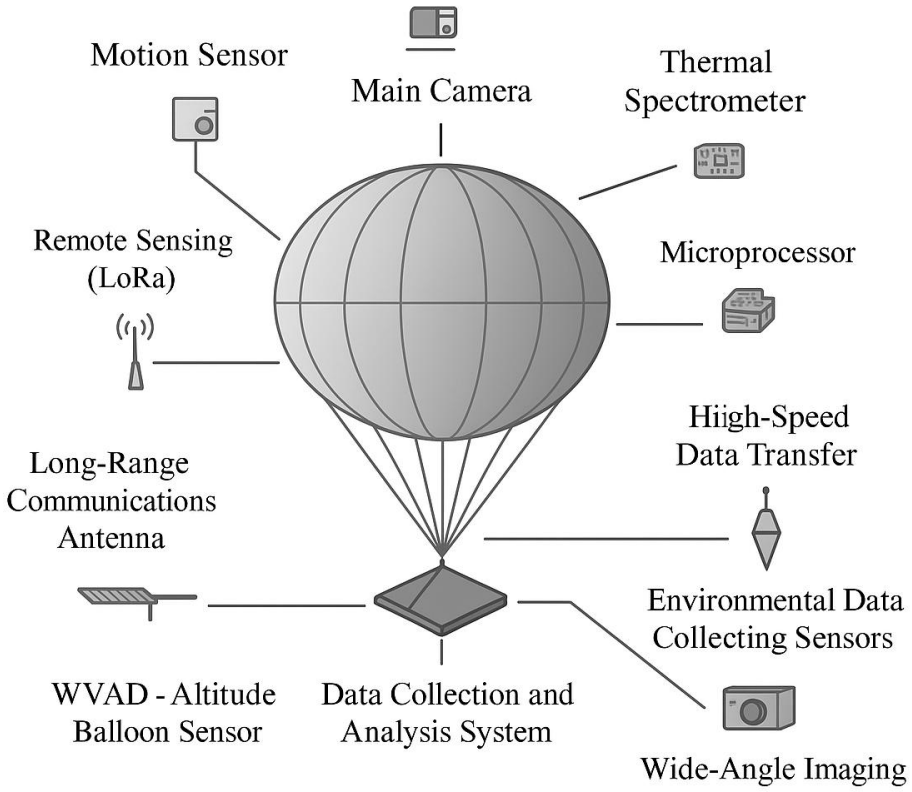


Figure 2. Sensor integration and data processing workflow of the balloon-based wildfire detection system

2.3. Sensor Integration and Data Acquisition

2.3.1. Advanced Sensor Technology and Data Processing

The balloon platform is equipped with advanced sensor technologies designed for early wildfire detection. Thermal cameras with infrared capability can detect temperature variations in forest areas within milliseconds, enabling identification of abnormal temperature rises that indicate the early stages of fire. Complementing this, multispectral imaging sensors monitor smoke density and air quality, allowing assessment of fire risk through the analysis of smoke composition. In addition, environmental monitoring sensors that measure microclimatic variables such as temperature, humidity, and wind speed provide essential data for predicting the potential spread and intensity of fires. The collected data are processed on ESP32-based microcontrollers integrated into the platform, where artificial intelligence algorithms detect anomalies that may signal a fire risk. These data are further analyzed on a cloud-based platform using machine learning methods capable of distinguishing normal from abnormal conditions. As a result, real-time alerts are generated whenever critical thresholds are exceeded, ensuring immediate response capability.

2.3.2. Renewable Energy Management and Autonomous Operation

To ensure energy independence and long-term autonomous operation, the system incorporates a hybrid renewable energy infrastructure. Ultra-lightweight flexible solar panels, integrated into the balloon surface, maximize solar energy capture during daylight hours. Meanwhile, mini wind turbines positioned at the lower section of the balloon provide additional power generation, particularly under cloudy or nighttime conditions. The harvested energy is stored in high-capacity lithium-ion batteries, and an intelligent charge management system optimizes power consumption to maintain efficiency. Through this integrated renewable energy management system, the balloon can sustain uninterrupted missions while minimizing environmental impact.

2.3.3. Data Transmission and Real-Time Notification

The communication infrastructure of the system is designed to ensure reliable operation over large geographical areas. Sensor data are transmitted from deep forest regions to the central control unit through long-range, low-power communication technologies such as LoRa and NB-IoT. Once received, the information is processed in a cloud-based analytics environment and visualized on GIS-supported maps. Critical parameters such as sudden temperature increases, humidity decreases, or smoke detection are represented with color-coded risk indicators, providing intuitive situational awareness.
An artificial intelligence–based decision support system further enhances the effectiveness of the platform by classifying fire risk levels, prioritizing intervention strategies, and automatically generating early warnings for disaster management units. In this way, the balloon system not only reduces wildfire response time but also strengthens proactive risk management capacity through continuous and precise monitoring.

2.4. Prototype Development and Testing Processes

2.4.1. Indoor Testing

Testing and Validation Process for Sensor-Equipped Balloon System

The balloon system underwent laboratory calibration, software integration tests, and forest pilot trials. Sensors were optimized for accuracy, software was verified for compatibility, and real-world simulations confirmed its real-time wildfire alert capability. As shown in Figure 3, the testing and validation process of the developed sensor-equipped balloon system was carried out in three stages. First, calibration experiments conducted in the laboratory optimized the sensitivity of the temperature, smoke, and environmental sensors, thereby enhancing measurement accuracy. In the second stage, software integration and functionality tests confirmed the seamless and compatible operation of the hardware and software components. In the final stage, pilot field trials conducted in forested areas evaluated the system's performance under real environmental conditions, demonstrating its capability to generate real-time alerts through controlled fire scenarios. In this context, Figure 3 visually summarizes the multi-stage validation process extending from laboratory experiments to field applications, and it highlights the operational reliability of the prototype at an academic level.

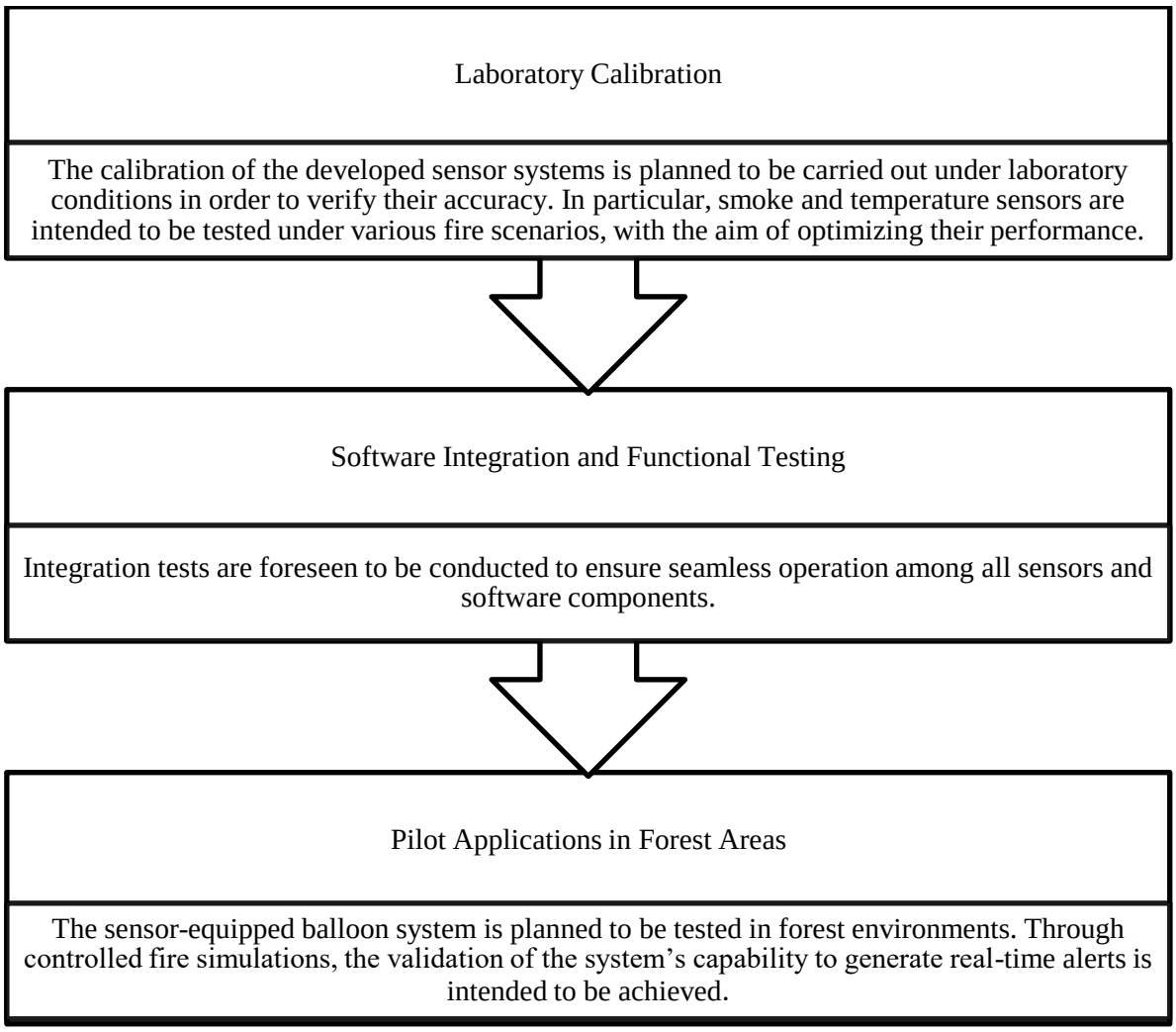


Figure 3. Process for validating a forest fire detection system

3. Experimental Results

3.1. Advanced Balloon Platform for Efficient Atmospheric Operations

This article discusses the conceptual design, development, and validation of a balloon-based, sensor-integrated aerial surveillance platform designed for the early detection of forest fires. By integrating advanced sensor technologies such as thermal imaging, and smoke and gas detection, the system is designed to be compatible with high-sensitivity detection of fire indicators in both controlled laboratory conditions and real-world forest environments.

The system aims to achieve uninterrupted autonomous operation with a hybrid renewable energy source that combines flexible photovoltaic panels and miniature wind turbines, thus planning to offer a solution to one of the major limitations of remote environmental monitoring systems: energy dependency.

The new balloon platform is designed using graphene-coated polyurethane for lightweight, durability, and recyclability, which may be in line with environmental sustainability goals. Its aerodynamic design is planned to minimize air resistance and wind drag, reduce energy consumption, and ensure long-term, stable operations. With a wide surface area, the platform can provide strong lift, wind resistance, and space for additional solar panels to increase energy collection efficiency. When filled with helium, it has been designed to provide sufficient buoyancy to carry the system's total weight, including sensors, communication systems, and power sources. This optimized

balance of weight, buoyancy, and efficiency suggests that the platform could be an ideal solution for environmental monitoring, communication networks, and renewable energy applications.

The operational parameters of the system design are given in Table 1 below.
Table 1 contains the main parameters determined during the design phase of the system, which detail the physical and operational characteristics of the balloon platform. These parameters include:

- **Dimensions and Material Properties:** The surface area of the balloon, the type of material used (graphene-coated polyurethane), and its properties (lightweight, durability, recyclability).
- **Energy Systems:** The configuration of the hybrid energy source (photovoltaic panels and wind turbines) and the expected energy generation capacity.
- **Operational Performance:** The impact of the designed aerodynamic features on energy consumption and wind resistance.

Table 1. System design and operational parameters

PARAMETRY	EXPLANTION
Material of the Balloon Platform	The platform is manufactured from advanced materials such as graphene-coated polyurethane. This material combination is distinguished by its light weight and environmental durability. Furthermore, its recyclable nature aligns with the platform's environmental sustainability objectives.
Aerodynamic Structure	The balloon features an aerodynamic profile designed to minimize air resistance. This design reduces the platform's wind drag, enabling it to remain stable in the air with low energy consumption. Aerodynamic efficiency allows for prolonged operations.
Surface Area	The balloon has a surface area of approximately 20 m², a value that supports its lift and wind resistance. The increased surface area also allows for the integration of more solar panels, strengthening its energy collection capacity..
Buoyant force	The buoyancy provided by the helium fill and optimized surface area has been calculated to be approximately 500 Newtons. This force is designed to carry the weight of the integrated sensors and support stable, long-duration flight.
Total Weight	The total weight of the balloon platform, including the integrated sensors, communication systems, and power sources, has been determined to be approximately 20 kg. This weight has been optimized for long-duration mission performance at low altitudes.
Resistance to wind	The drag force on the balloon was measured to be approximately 50 Newtons at a wind speed of 10 m/s. This drag force is balanced by the platform's center of gravity, which has been configured to maintain stability."
Stability and Balancing Systems	The balloon is equipped with GPS-based control and balancing systems, enabling it to maintain a stable position in designated areas. This stability is essential for autonomous monitoring systems and keeps energy consumption at a minimum.
Energy Efficiency	The flexible solar panels integrated into the balloon's surface allow the platform to collect energy throughout the day. Additionally, mini wind turbines located at the bottom of the balloon convert kinetic energy into electrical energy, providing a backup power source.
Durability	The balloon's surface, coated with UV-resistant and waterproof materials, provides protection against harsh environmental conditions. The material properties offer stable operational capability even in extreme temperatures and high humidity.
Potential Risks and Solutions	Interruptions in data transmission are overcome with buffer memory and reconnection strategies. In case of power shortages, backup batteries and mini wind turbines activate to ensure a continuous energy supply.

During the laboratory calibration phase, the accuracy and reliability of the sensor suite, particularly in detecting subtle thermal and particulate anomalies preceding ignition events, were demonstrated. In field trials, the platform's stability, communication range, and real-time alert capabilities were successfully validated, even under challenging meteorological conditions. The results provide evidence that the system can operate sustainably across diverse ecological and climatic contexts, thereby indicating its potential deployment in high-risk wildfire zones.

From an operational perspective, the low-altitude aerostatic design, aerodynamic optimization, and GPS-based position control offer spatial stability and targeted monitoring flexibility. Through the integration of a cloud-based data processing pipeline and AI-driven anomaly detection algorithms, rapid situational awareness and decision support are enabled, with the objective of shortening the time between fire onset and the initiation of mitigation measures. The study's findings suggest that the proposed platform not only enhances early warning capabilities but is also planned to be tested as a complementary tool to existing satellite- and ground-based detection infrastructures.

In terms of scalability, the modular architecture of both hardware and software components has been designed for adaptation to various environmental monitoring applications, such as detecting illegal logging, monitoring agricultural conditions, or assessing post-disaster damage. Furthermore, the selection of recyclable and durable materials has been structured in alignment with environmental sustainability principles, providing a practical solution for long-term operational use.

In conclusion, this study emerges as a viable, cost-effective, and sustainable technological solution for strengthening wildfire prevention and response strategies. By enabling precise, continuous, and autonomous monitoring of forested areas, the system is expected to contribute to reducing the ecological, economic, and societal impacts of wildfires. Future work will focus on large-scale field trials, the advancement of AI-based predictive models, and the integration of the system into national and international disaster response frameworks.

4. Conclusion and Discussion

In recent years, the increasing number of forest fires due to climate change and human interaction poses a serious threat to ecosystems and human communities. One of the most critical factors in mitigating the impact of these disasters is the early detection of fires before they spread. Traditional detection methods often face challenges such as limited coverage, high costs, or insufficient real-time data provision. This study presents a design that aims to overcome these challenges by proposing a balloon-based aerial monitoring platform with the potential to serve as an effective and sustainable early warning system.

In the subsequent phases of this project, it is planned to conduct laboratory calibration tests of the developed system and verify whether the thermal, smoke, and environmental sensors demonstrate high accuracy and responsiveness in various simulated fire scenarios. Following the completion of these tests, the system's operational stability, reliable data transmission, and capacity to generate timely alerts, even under challenging weather conditions such as high winds, extreme temperatures, and high humidity, will be tested through field trials. The hybrid renewable energy configuration, consisting of flexible photovoltaic panels and miniature wind turbines, will also be proven to be sufficient for uninterrupted autonomous operation, overcoming the common energy limitations of remote monitoring systems. The integration of GPS-assisted positioning, aerodynamic design, and AI-based anomaly detection will enable both spatial stability and intelligent situational awareness, allowing for the rapid identification of potential fire events.

Compared to traditional satellite or ground-based detection methods, it is expected that this system will strengthen real-time decision-making capabilities with a shorter latency in threat recognition and increased coverage flexibility. These results suggest that the platform has a strong potential for integration into broader wildfire prevention and disaster management strategies, and its modular and environmentally sustainable design is projected to support scalability to other ecological monitoring applications.

Acknowledgements

This study has been supported by the TÜBİTAK 2209-A project numbered is 1919B012453694.

5. References

[1] Bennert, T., Hanson, D. & Maher, A.. "Influence of pavement surface type on tire/pavement generated noise". Journal Yıldız, M. ve Kaya, M. "Orman Yangınlarının Erken Tespiti ve Müdahale Yöntemleri: Teknolojik Çözümler ve Uygulamalar". *Orman Mühendisliği Dergisi*, 32(1), 55-70(2021).
[2] Öztürk, M. & Doğan, E. "Orman Yangınlarıyla Mücadelede Yapay Zeka ve İleri Teknolojilerin Kullanımı". *Anadolu Bilim Dergisi*, 4(2), 90-105(2019).
[3] Chavez, F. & Sharma, N. "Thermal and Smoke Sensors in Forest Fire Detection: A Review". *Journal of Environmental Monitoring*, 9(4), 413-426(2020).
[4] Kara, Y. & Akay, G. "Türkiye'de Orman Yangınları ve Erken Uyarı Sistemlerinin Geliştirilmesi". *Gazi Üniversitesi Orman Fakültesi Dergisi*, 25(1), 88-101(2018).
[5] Hersbach, H. et al. "Seasonal Forest Fire Risk Prediction Using Satellite Data and Machine Learning Techniques". *Remote Sensing*, 13(10), 2038-2050(2021).
[6] Tuncer, E. & Kılıç, D. "Rüzgar Enerjisi ve Güneş Enerjisinin Orman Yangınları Erken Tespit Sistemlerinde Kullanımı". *Enerji Verimliliği Dergisi*, 7(3), 203-220. (2020)
[7] Barton, C. & Freeman, C. "IoT-Enabled Forest Fire Detection and Prevention Systems". *International Journal of Environmental Science*, 11(4), 1470-1484(2021).
[8] López, M. & Santos, R. "Development of Smart Sensors for Early Fire Detection in Forested Areas". *Sensors and Actuators A: Physical*, 295, 356-365(2019).
[9] Simsek, A. & Sarac, O. "Sustainable Technologies in Environmental Protection: A Case Study of Forest Fire Detection". *Sustainability*, 12(6), 1820-1835(2020).
[10] Yang, Y. & Wang, Z. . "Advanced Remote Sensing Technologies for Forest Fire Detection: A Review". *International Journal of Remote Sensing*, 40(1), 256-270(2019).
[11] Zhao, J. et al. "Smart Balloon-based IoT Sensors for Remote Environmental Monitoring". *Environmental Monitoring and Assessment*, 192, 309-317(2020).
[12] Wu, T. et al. "Design and Applications of Autonomous Airships for Environmental Surveillance and Fire Detection". *Environmental Technology & Innovation*, 14, 100312(2019).
[13] Nikolov, G. & Ivanov, P. "Wireless Sensor Networks for Real-Time Fire Detection and Monitoring: A Review of the Recent Trends". *Journal of Environmental Engineering*, 9(3), 134-145(2020).
[14] Ravi, R. & Sharma, S. "Energy-Efficient Smart Monitoring Systems for Forest Fires: A Comprehensive Review". *Journal of Renewable and Sustainable Energy*, 13(2), 160-171(2021).
[15] Fang, L. & Li, D. "Low-Cost UAV and Balloon-Based System for Monitoring Forest Fires in Remote Areas". *Sensors*, 20(19), 5642-5658(2020).
[16] Gong, H. & Liu, C. "Application of Thermal Imaging and IoT-Based Sensor Systems for Early Detection of Forest Fires". *IEEE Sensors Journal*, 21(7), 8121-8130(2020).
[17] Wang, H. & Zhang, S. "Real-Time Forest Fire Detection Using Drone and Satellite Data: A Comparative Analysis". *Remote Sensing Letters*, 11(5), 101-113. (2020).
[18] Zhou, Z. & Chen, Y. "Development of an Autonomous Zephyr Airship for Environmental Monitoring and Disaster Response". *Advanced Intelligent Systems*, 4(3), 291-305(2019).
[19] Aydin, A. & Akbulut, K.. "Orman Yangınlarının Erken Tespiti İçin Geliştirilen Dron ve Balon Tabanlı Sistemler: Kullanım Alanları ve Zorluklar". *Teknoloji ve Bilim Dergisi*, 5(2), 145-159(2018).
[20] Wong, S. & Whelan, D. "Data Fusion Techniques for Enhanced Fire Detection in Forests Using Remote Sensing and Ground-Based Sensors". *International Journal of Wildland Fire*, 29(6), 457-468(2020).
[21] Bektas, I. & Yildirim, M.. "Energy Harvesting Solutions for Wireless Sensor Networks in Remote Forest Fire Detection Applications". *Journal of Green Energy*, 14(5), 1150-1165(2020).
[22] Vazquez, A. & Munoz, M. "Smart Sensor Networks for Forest Fire Early Detection: A Case Study from Spain". *Sensors and Actuators B: Chemical*, 300, 126853-126866 (2020).

Design and Calculations of Rocket Launching Mechanism Using Centrifugal Technique

Tolga MERTYÜZ[1] , Oğuz YAKUT [2]

[1]Graduate School of natural and Applied Sciences, Firat University, Elazig, Turkey.
[2] Mechatronics Engineering, Faculty of Engineering, Firat University, Elazig, Turkey.
[1]tolga.mertyuz@gmail.com, [2]oyakut@firat.edu.tr
[1](ORCID: 0000-0002-8888-4298), [2](ORCID: 0000-0002-0986-1435)

Abstract
In this study, an alternative launch mechanism that enables the rocket to be launched without chemical fuel using centrifugal force has been designed and analyzed theoretically. In contrast to the disadvantages of conventional solid or liquid fuel engines, such as high cost, safety risks, and structural complexity, this system, which operates entirely on mechanical principles, offers a low-cost and environmentally friendly solution. In the designed system, the rocket is fixed to a horizontal arm; after the arm reaches a certain angular velocity with the help of an asynchronous motor, the rocket is released by a servo motor-controlled separation mechanism. In theoretical analyses, a rotation speed of 400 rpm was achieved with a 3 kW motor and 90% efficiency. With these values, it was calculated that a rocket with a mass of 5 kg rotating at a radius of 1 meter would reach an exit velocity of approximately 41.89 m/s. When the launch angle was set to 85, it was determined that the rocket would reach an altitude of approximately 71 meters and a horizontal range of 3 meters. In these calculations, the highest achievable altitude value was obtained by assuming that aerodynamic drag forces cause approximately 20% energy loss. The results obtained indicate that this system, which operates using centrifugal force, holds potential for short-range missions and low-cost applications for educational purposes.

Keywords: centrifugal force, rocket launch mechanism, rocket separation mechanism

1. Introduction

Today, rockets are widely used in various space research programs, particularly in the defense industry. Rockets can perform missions at different ranges and altitudes depending on their structure and power capacity. As the amount of energy required for propulsion increases, the cost of rocket operations also rises accordingly [1]. Solid fuel engines are commonly used in rockets. However, the use of solid-fuel engines is limited in long-range and high-altitude missions. In such missions, liquid-fuel or hybrid rocket engines are generally preferred [2,3]. However, these engines are structurally more complex and much more expensive than solid-fuel engines. A significant portion of the fuel is consumed during the launch phase when rockets ascend against gravity. Therefore, vertical takeoff aircraft consume much more fuel than those that take off horizontally. Some rockets can be mounted on a winged aircraft and launched from the air once sufficient altitude is reached. The ability of these carrier aircraft to carry dozens of rockets in a single flight provides a significant advantage. In this case, rockets launched from the air either reach the target vertically at a certain glide ratio under the influence of gravity or reach longer ranges by using a small rocket motor to propel themselves horizontally. This method is both economically advantageous and ensures that the rocket reaches its target more quickly [4].

Air-launched rockets require a guidance kit to accurately reach their target by gliding. For this purpose, fixed or active control surfaces are integrated into the rocket body. In Turkey, institutions such as SAGE, ROKETSAN, and ASELSAN are developing products with these features and leading related R&D efforts. These institutions have undertaken important and valuable projects for the needs of our defense industry. Some of these products include KUZGUN (Figure 1), MK-82/83, and M5 miniature bombs [5].

[1]Corresponding author



Figure 1. KUZGUN M5 winged rocket

The increasing use of unmanned aerial vehicles (UAVs) today has proven their effectiveness in the defense industry and ensured that they take their rightful place. The primary expectations for fixed-wing or rotary-wing UAVs are long endurance and high payload capacity [6–8]. Although their speeds are significantly lower compared to conventional combat aircraft, UAVs are highly effective for intelligence purposes, particularly in the processes of obtaining ground imagery and transmitting this data to ground control stations. Rather than high speed, they are expected to perform safe and stable flights at specific altitudes for extended periods. When necessary, they can also be armed in threatening situations. Therefore, high payload capacity is a significant advantage. In this regard, large-scale UAV systems can be produced specifically for mission requirements. However, this also leads to increased fuel consumption. Therefore, UAV systems with high energy efficiency are gaining prominence. The most critical factors directly affecting UAV performance are the selection of the engine type and the design of the aerodynamic structure of the body. The BAYRAKTAR TB2 (Figure 2), produced domestically in Turkey, is a prominent example in this field and has become one of the leading products of our defense industry [9].



Figure 2. BAYRAKTAR TB2 armed unmanned aerial vehicle (UCAV)

In Turkey, STM is actively conducting research and development work on micro UAVs, swarm UAVs, and loitering munitions (kamikaze drones) [10]. Among these systems, three micro UAVs that have entered the inventory of the Turkish Armed Forces and are currently in use are shown in Figure 1.3. One of these, KARGU, is also classified as a loitering munition. Equipped with an artificial intelligence-based system, KARGU has the ability to autonomously detect and neutralize targets. Its advanced algorithms, detection systems, and sensors are key features that distinguish KARGU from traditional drones [11]. ALPAGU shares similar features but is designed to be effective against faster-moving targets, unlike KARGU. TOGAN, on the other hand, is primarily used for surveillance and reconnaissance purposes in urban operations. These micro UAVs have a range of approximately 5 km and an operational altitude of 1–2 km. Their average flight duration is approximately 10 minutes [11].



Figure 3. a.KARGU, b.ALPAGU, c.TOGAN

Unmanned aerial vehicles (UAVs) are primarily designed for tasks such as airspace surveillance, intelligence gathering, and reconnaissance. Vertical take-off and landing (VTOL) UAVs are a special type of UAV that can take off and land

vertically. When comparing these two systems, VTOL UAVs offer significant advantages thanks to their vertical take-off and landing capabilities. They can operate in confined spaces and have high maneuverability. In addition, they can be deployed quickly and operations can be started in a short time, which is a significant advantage, especially in emergency situations. VToL UAVs do not require long runways or special infrastructure for takeoff and landing, which makes them more practical for field operations. Their ability to take off and land without runway requirements allows them to adapt more easily to different terrain conditions.

However, there are also some disadvantages. Compared to fixed-wing UAVs, VToL UAVs generally have shorter ranges, which can limit operational duration and surveillance/reconnaissance distances. Additionally, their speeds and payload capacities are typically lower, which can reduce the amount and variety of equipment that can be carried. The additional energy required for vertical takeoff and landing leads to higher energy consumption, which can shorten flight time or make energy management more complex. Furthermore, due to their more complex structures and specialized design requirements, VToL UAVs are generally more expensive than traditional fixed-wing UAVs [12,13].

In this study, an innovative launch system utilizing centrifugal force was designed to reduce fuel consumption and provide a low-cost, portable solution for small-scale space missions. With this design, the rocket's exit velocity, torque value, rocket launch force, maximum altitude, flight duration, and horizontal range were theoretically calculated based on the dimensions of the rocket to be launched.

1.1. Rocket Launch System

The use of solid fuel in rocket propulsion raises various issues in terms of safety, supply difficulties, and potential technical failures. Therefore, a safer and more mobile launch system is being targeted. Figure 4 shows a rocket launch system based on centrifugal technology.



Figure 4. Launch Pad model

Centrifugal rocket launch technology offers several notable advantages over traditional chemical rocket engines. The centrifugal launch method minimizes the use of chemical propellants to achieve high speeds. This significantly reduces the weight of the rocket and, consequently, the total launch cost. Additionally, the reduced fuel requirement lowers operational costs, as chemical fuels are quite expensive, and the centrifugal method can eliminate a significant portion of these costs.

Traditional rocket launches release large amounts of carbon dioxide and other harmful gases into the atmosphere. In contrast, the centrifugal technique offers an environmentally friendly alternative by greatly reducing these emissions. Furthermore, the environmentally harmful waste produced by the combustion process of chemical propulsion systems does not occur with the centrifugal method. Centrifugal systems also allow for the reuse of rocket components. This contributes to reducing operational costs and environmental waste. Furthermore, this method allows for more precise control of launch angle and speed, resulting in safer and more efficient launches. Chemical rocket fuels pose an

explosion risk. The centrifugal method significantly reduces this risk because there is no need to use explosive materials. Additionally, ground-based centrifugal launch systems are simpler and safer compared to traditional rocket launch pads. Lower fuel and launch costs make it economically feasible to launch more satellites and space vehicles. This enables space agencies and private companies to carry out more missions and accelerates space research. The centrifugal system can accelerate the rocket to very high initial speeds. This allows the rocket to pass through the atmosphere more quickly and be exposed to less aerodynamic drag.

The centrifugal method is an extremely effective solution, especially for launching satellites into low Earth orbit (LEO). The ability to achieve higher speeds with less energy makes it easier to launch small satellites, both technically and economically. The advantages offered by rocket launches using centrifugal technology hold great potential for space research and commercial space transportation. Reduced fuel costs, minimized environmental impact, and increased safety levels could contribute to the widespread adoption of this method in the future.

2. Rocket Launch System Design

The purpose of this system is to design a propulsion ramp that utilizes centrifugal force to launch a low-altitude rocket. The ramp to be designed will feature one or more rotating arms that generate thrust. The object to be launched will be mechanically secured to the rotating arm using a special locking mechanism. When the arm reaches the specified angular velocity, the locking mechanism will be disengaged in a controlled manner, allowing the object to be released and launched.

In the preliminary design, it is anticipated that the rotating arm will have a radius of approximately 1 meter and will be driven by a motor with a minimum power of 3 kW. The motor will be connected to the rotating arm via a belt-pulley system, and the system will be designed to reach a rotational speed of 400 rpm. According to theoretical calculations, objects with a mass of approximately 5 kg launched from the ramp are expected to reach a height of approximately 100–120 meters, depending on the aerodynamic drag coefficient. These calculations were made for cylindrical bodies with rocket-like geometries, 30–50 cm in length and 8–10 cm in diameter.

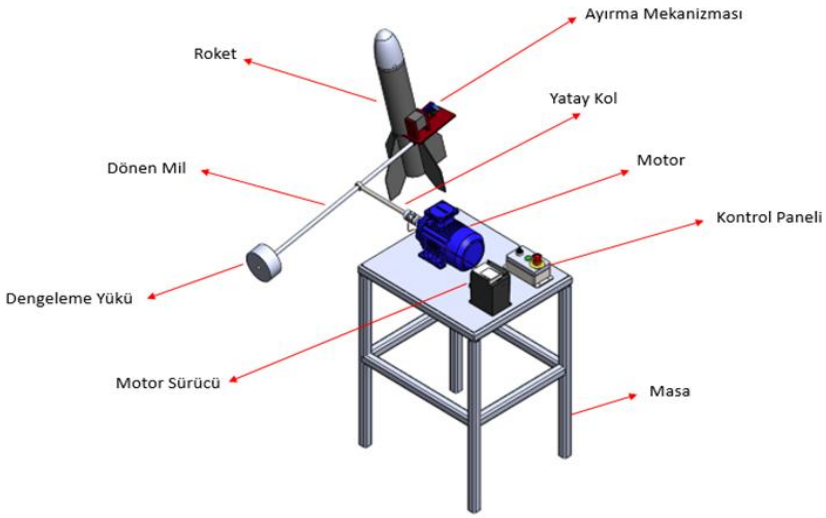


Figure 5. Launch system design using centrifugal force

The system shown in Figure 5 is mounted on a sturdy table so that it can operate without being affected by vibrations and imbalances. The rotational movement of the system is provided by an asynchronous motor with a power of 3 kW and a rated speed of 1500 rpm. The motor is connected to the drive system via a horizontal shaft, and the rotational movement is transmitted to the launch arm via this shaft. The arm that holds the rocket and performs the launch task is directly connected to this horizontal shaft. The rocket is positioned to be launched outward at a specific speed due to the centrifugal force generated by the rotation of this arm. As the motor runs, the arm begins to rotate, and the rocket at its tip accelerates outward with increasing centrifugal force. When the arm reaches the targeted angular velocity and

centrifugal force, the rocket is released in a controlled manner and separates from the system at a specific launch angle. This launch process can be performed either through an automatic lock-release mechanism or a manual trigger mechanism. A special control panel has been designed to ensure the safe and controlled operation of the system. This panel includes a start button to activate the motor, an emergency stop button for critical situations, and a potentiometer to adjust the motor speed. A VFD (variable frequency drive) is incorporated into the system to ensure the motor operates at the desired RPM and in the correct direction.

2.1. Rocket Seperation Mechnanism

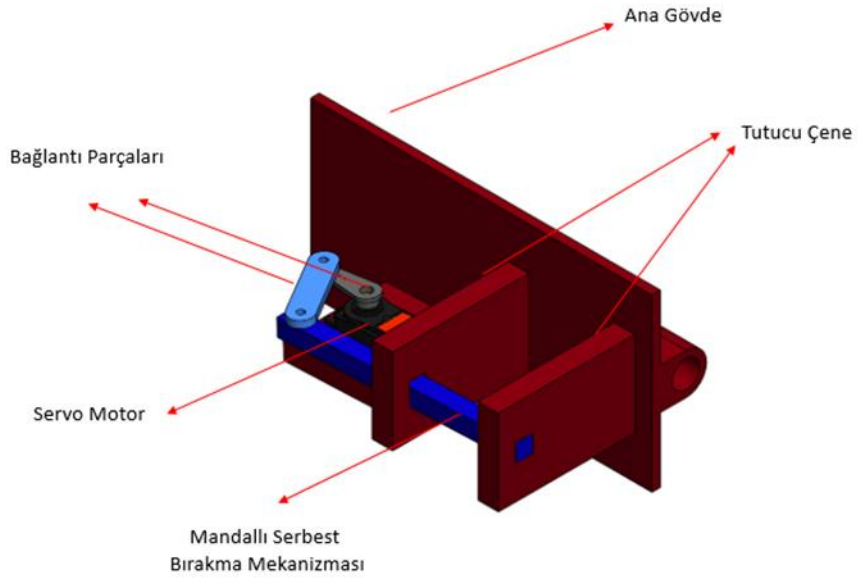


Figure 6. Seperation mechanism design

Figure 6 shows the rocket release mechanism used in the centrifugal force-driven rocket launch system. This mechanism is designed to ensure that the rocket is safely separated from the system when it reaches the desired rotational speed and angular position. The blue section indicates the latch release mechanism. The section shown in red consists of the main body and the retaining jaw that secures the rocket. The main body is mounted directly on the rotating arm, and the retaining jaw ensures that the rocket remains securely in place.

To initiate the launch, the latching mechanism is designed to work in conjunction with a servo motor. When the servo motor is activated, the retaining jaw opens, allowing the rocket to be released. At this point, the rocket separates from the rotating arm and is launched forward under the influence of centrifugal force. The launch direction of the rocket is directly related to the direction of rotation of the arm and the moment the rocket is released.

After the rocket is launched, the separation mechanism returns to its previous position, and the system becomes ready for a new rocket to be placed. This allows the mechanism to be reused for repeatable tasks.

3. Rocket Launch Dynamics and Calculations

In launch systems that operate with centrifugal force, the angular velocity, linear velocity, torque, and aerodynamic cross-sectional area, which are fundamental physical quantities, must be calculated to determine the rocket's exit velocity and the forces it is exposed to. Below are the theoretical calculations made according to the motor and geometry used in the system.

The most decisive force in rocket launch is centrifugal force. Defined in rotating systems, this force is the effect that causes the rocket to be pushed outward during its rotational movement. Centrifugal force (Fc) is proportional to the rocket's mass (m), angular velocity (ω), and radius of rotation (r) and is expressed as follows:

$$F_c = m.w^2.r \quad (1)$$

According to Equation (1), when the radius of rotation is constant, the centrifugal force exerted on the rocket will increase quadratically as the angular velocity increases. Therefore, angular velocity control of the system is critical for a safe launch.

Angular velocity (ω) refers to the amount of angular displacement of an object per unit time and is usually calculated in radians per second. If the rotational speed of the system is known in revolutions per minute (RPM), the angular velocity is calculated using the following formula:

$$\omega = \frac{2\pi N}{60} \quad (2)$$

Using equation (2), the angular velocity for a system rotating at 400 rpm is approximately:

$$\omega = \frac{2\pi \cdot 400}{60} \approx \frac{2513.27}{60} \approx 41.89 \ \square\square\square\square\square/\square\square\square\square\square\square \quad (3)$$

Linear velocity (v) is the linear velocity of the rocket on the rotating arm and determines the exit velocity when the rocket is released. It is calculated using the angular velocity and radius of rotation as follows:

$$V = w.r \quad (4)$$

If r = 1 m is assumed here, the rocket's exit velocity according to Equation (4) is:

$$v = 41.89 \cdot 1 \approx 41.89 \, m/s \quad (5)$$

The torque (T) produced by the motor used to drive the system is calculated based on the relationship between the motor's output power (Pcikis) and angular velocity. The output power is obtained by multiplying the nominal power by the efficiency (η). This relationship is as follows:

$$T = \frac{P_{\text{ç}\square\square\square\text{ş}}}{\square} \quad \text{ve} \quad P_{\text{ç}\square\square\square\text{ş}} = \eta \cdot P \quad (6)$$

If the nominal power of the motor is 3000 W and the efficiency is η=0.9:

$$P_{\text{çıkış}} = \eta \cdot P = 0.9 \cdot 3000 = 2700 \, W$$
$$T = \frac{2700}{41.89} \approx 64.47 Nm \quad (7)$$

This torque value is sufficient for the system to rotate the arm to the desired angular velocity. The frontal area (A) of the rocket to be used in aerodynamic analyses can be calculated as follows based on the diameter (d) for a cylindrical structure:

$$A = \pi \left(\frac{d}{2}\right)^2 \quad (8)$$

When the rocket diameter is assumed to be 0.1 meters, the corresponding solution becomes available.

$$A = \pi \left(\frac{0.1}{2}\right)^2 \approx \pi \cdot 0.005^2 \approx 7.85 \times 10^{-3} \, m^2 \quad (9)$$

3.1. Calculation of the Rocket's Flight Performance

The trajectory that the rocket will follow after launch, integrated into the centrifugal system, can be defined by parameters such as the net force it is subjected to, launch velocity, maximum altitude, and range. This section covers

the fundamental physical quantities and theoretical performance calculations that are effective throughout the rocket's flight process.

One of the most important external factors encountered during the launch of the rocket is aerodynamic drag force. This force is calculated using equation (10), taking into account air density (ρ), the square of the rocket's exit velocity (v^2), the drag coefficient (C_d), and the frontal area (A).

$$F_d = \frac{1}{2}\rho v^2 C_d \tag{10}$$

$$F_d = \frac{1}{2} \cdot 1.225 \cdot 41.89^2 \cdot 0.75 \cdot 7.85 \times 10^{-3} \tag{11}$$

Here, ρ is the air density (1.225 kg/m³), v is the linear velocity, C_d is the drag coefficient (0.75), and A is the frontal area of the rocket. The calculations resulted in Fd ≈ 6.31 N. The aerodynamic drag force calculated here results in a load moment of M_d ≈ 6.31 Nm on a 1 m radius arm.

The net torque value obtained by subtracting the aerodynamic drag moment M_d from the previously defined motor torque (Equation 7) serves as the thrust moment that enables the rocket to launch. Accordingly, the net moment value is:

$$M_{\square\square\square} = T - M_d$$
$$M_{\square\square\square} = 58.16\ Nm \tag{12}$$

This net moment acts until the rocket reaches the launch velocity required for its initial kinetic energy. The linear velocity of the rocket at launch was calculated using Equation 4. Using this value and assuming a launch angle of 85 degrees, the horizontal (v_X) and vertical (v_y) components of the velocity can be determined as follows:

$$v_y = v_{firlatma} \cdot sin(85°)\ \ v_y \approx 41.89 \cdot 0.9962\ \ v_y \approx 41.73\ m/s\ \ v_x = v_{firlatma} \cdot cos(85°)\ \ v_x \approx 41.89 \cdot 0.0872\ \ v_x \approx 3.65\ m/s \tag{13}$$

The theoretical maximum altitude that a launched rocket can reach can be calculated by equating the rocket's kinetic energy to its potential energy. However, part of the rocket's kinetic energy will be lost due to aerodynamic drag. Assuming that this energy loss is approximately 20%, the energy equation can be written as follows.

$$0{,}8.(E_{\square\square\square\square\square\square\square\square}) = (E_{\square\square\square\square\square\square\square\square\square\square\square\square})$$
$$0{,}8.(\frac{1}{2}mv_y^2) = mgh \tag{14}$$
$$0{,}8.(\frac{1}{2}.5.41{,}73^2) = 5.9{,}81.h$$

The altitude that the rocket can reach is calculated using this expression.

$$h_{max} = \frac{0{,}4.v_y^2}{g} \approx 71.00\ metre \tag{15}$$

4. Results

In this study, a rocket launch system operating on the principle of centrifugal force was designed, and the system's performance was evaluated using basic engineering calculations. The developed mechanism aims to launch the rocket by accelerating it to a specific angular velocity using a rotating arm and a servo-controlled separation mechanism. Based on the applied system parameters and simulation-based calculations, it was observed that the system demonstrates suitable performance for low-altitude missions.

According to the data given in Table 1, the system is propelled by a 3 kW motor with 90% efficiency. As a result of reaching a rotation speed of 400 rpm at the end of a rotating arm with a radius of 1 meter, the angular velocity was calculated to be approximately 41.89 rad/s, and the linear exit velocity was approximately 41.89 m/s. With these speeds and system parameters, a net thrust torque of 58.16 Nm was achieved, demonstrating that this torque value is sufficient to provide the necessary thrust for launch.

Table 1. Technical parameters related to the rocket launch system

Parametre	**Değer**
Motor Power (P)	3 kW = 3000 W
Rotational Speed (N)	400 rpm
Rotational Radius (r)	1 meter
Roket massi (m)	5 kg
Roket length	0.5 meter
Roket diameter	0.1 meter
Motor efficiency (η)	%90 = 0.9
Air density (ρ)	1.225 kg/m³ (at sea level)
Aerodynamic drag coefficient (Cd)	0.75 (estimated value for rocket)
Launch angle	85^0

The effect of aerodynamic drag force on the system varies depending on the rocket diameter, and this value can be kept small to minimize energy loss. With a speed control unit used in the system, launches can be performed at different speed values. With this designed system, it has been concluded that rockets of different sizes can be launched at different launch speeds within the limits of engine power. These results demonstrate that the system designed at the theoretical level is physically feasible and could serve as an alternative launch method, particularly for low-altitude missions.

In conclusion, under the system parameters presented in Table 1, this centrifugal-based launch system, supported by the flight data obtained, can be considered a suitable solution for both academic and defense applications with its low-cost, compact, and reusable structure. Future studies can be expanded to include experimental prototyping of the system, structural analyses, and integration with different rocket designs.

Acknowledgments

This study was supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK) under the 1001-Scientific and Technological Research Projects Support Program, titled "Design, Production and Control of a Throwable Rocket-Like Unmanned Aerial Vehicle (RIHA) with a Centrile Launch Ramp" Project No: 223M001, under the Scientific and Technological Research Projects Support Program (1001).

Additionally, this work was derived from the first author's doctoral thesis titled "Design and Control of a Solar-Powered Unmanned Aerial Vehicle Capable of Long-Term Endurance Mission Under Stratospheric Conditions."

5. References

[1] Heister, S. D., Anderson, W. E., Pourpoint, T. L., Cassady, J., & Cassady, R. J. (2019). Rocket propulsion (Vol. 47). Cambridge University Press.

[2] Pérez-Roca, S., Marzat, J., Piet-Lahanier, H., Langlois, N., Farago, F., Galeotta, M., & Le Gonidec, S. (2019). A survey of automatic control methods for liquid-propellant rocket engines. Progress in Aerospace Sciences, 107, 63-84.

[3] Mahjub, A., Mazlan, N. M., Abdullah, M. Z., & Azam, Q. (2020). Design optimization of solid rocket propulsion: a survey of recent advancements. Journal of Spacecraft and Rockets, 57(1), 3-11.
[4] Misra, A., Jayachandran, S., Kenche, S., Katoch., A., Suresh, A., Gundabattini, E., ... & Legesse, A. A. (2022). A Review on Vertical Take-Off and Landing (VTOL) Tilt-Rotor and Tilt Wing Unmanned Aerial Vehicles (UAVs). Journal of Engineering, 2022.
[5] Füze Ve Roketlermühimmatlar, Tübitak Sage Kuzgun Mühimmatı, https://www.millisavunma.com/tubitak-sage-kuzgun-muhimmati/ (Erişim tarihi: 19.12.2023)
[6] Ucgun, H., Yuzgec, U., & Bayilmis, C. (2021). A review on applications of rotary-wing unmanned aerial vehicle charging stations. International Journal of Advanced Robotic Systems, 18(3), 17298814211015863.
[7] Cai, G., Lum, K. Y., Chen, B. M., & Lee, T. H. (2010). A brief overview on miniature fixed-wing unmanned aerial vehicles. IEEE ICCA 2010, 285-290.
[8] Sujit, P. B., Saripalli, S., & Sousa, J. B. (2014). Unmanned aerial vehicle path following: A survey and analysis of algorithms for fixed-wing unmanned aerial vehicless. IEEE Control Systems Magazine, 34(1), 42-59.
[9] Bayraktar TB2, https://baykartech.com/tr/uav/bayraktar-tb2/ (Erişim tarihi: 19.12.2023)
[10] Kamikaze dronlar sürü harekâtına hazırlanıyor, https://www.iha.com.tr/haber-kamikaze-dronlar-suru-harekatina-hazirlaniyor-777594 (Erişim tarihi: 19.12.2023)
[11] ÇETİNKAYA, S. G. (2023). Türkiye'nin İnsansız Hava Araçları Serüveni. Anadolu Strateji Dergisi, 5(1), 1-27.
[12] Novák, M., & Hulek, D. (2017). Types of unmanned aerial vehicles for package transport. In New Trends in Process Control and Production management (pp. 377-381). CRC Press.
[13] Smith, J. K., & Johnson, L. M. (2020). "Comparative Analysis of VTOL RIHAs and Fixed-Wing UAVs for Airspace Monitoring and Intelligence Gathering." Journal of Unmanned Systems Technology, 15(3), 45-62.
[14] Michael Sheetz, Alternative rocket builder SpinLaunch completes first test flight,2

Effect of Boron on the Microstructure and Wear Behavior in Laser Surface Coating of SAE 4140 Steel

Emir KARACAM[1], Mehmet EROĞLU[2], Amir MOTALLEBZADEH[3]

[1,3]Şenmak Makine, Akçaburgaz Mah.Hadımköy Yolu Cad. No: 94 Esenyurt, İstanbul, Türkiye.
[2]Metallurgical and Materials Engineering, Engineering Faculty, Fırat University, Elazığ, Turkey
[1]emirkaracam@senmak.com, [2]meroglu@firat.edu.tr, [3]amir@senmak.com
[1](ORCID: 0009-0000-3809-0476), [2](ORCID: 0000-0002-5097-1943), [3](ORCID: 0000-0001-6753-9316)

Abstract
Boronizing is a thermochemical treatment applied to increase the surface hardness and wear resistance of steels. However, the traditional thermochemical boronizing process has certain disadvantages, such as the need for high processing temperatures and long treatment durations. As an alternative to conventional boronizing, laser-assisted surface boronizing has emerged as a significant innovation. In this study, the surface of SAE 4140 steel was boronized using laser processing with powders containing 4%, 6%, and 8% boron. After the coating process, hardness measurements were performed, microstructures were examined, and abrasive wear tests were conducted. As a result, it was observed that an increase in the boron content of the coating led to a higher formation of the Fe_2B phase, which in turn enhanced both hardness and wear resistance.

Keywords: laser coating, microstructure, hardness, wear

1. Introduction

Today, materials used in engineering applications are evaluated not only by their basic mechanical properties but also by the resistance of their surfaces to environmental effects. Surface damages such as abrasion, corrosion and thermal oxidation cause significant economic losses, especially in sectors such as automotive, defense, mold and machinery industries. In this context, surface engineering methods have become indispensable for extending material life and improving performance [1, 2].

Laser surface coating method is a technology that can create high-adherence, homogeneous and dense structures with the substrate material by providing high temperature and rapid cooling rate in a short time. Thanks to this method, only the surface properties are modified while the general mechanical structure of the substrate is preserved. Laser coating is a process that allows alloys in powder, wire or paste form to be melted with laser beam and coated on the surface [3-6].

The use of boron in surface coating greatly increases wear resistance through the formation of boride phases (e.g. FeB, Fe_2B) that provide high hardness. These borides exhibit superior properties such as high hardness (1500–2000 HV), temperature resistance and chemical stability. However, pure boron powders are limiting in terms of both cost and application difficulties. Therefore, iron-based alloys containing boron, such as ferrobor, stand out as a suitable and economical alternative for laser surface coating processes. Since Ferrobor contains boron phases dispersed in the iron matrix, it contributes directly to boride formation during the laser melting process. At the same time, during this process, parameters such as melting depth, laser scanning speed, laser power and powder feeding rate directly affect the quality of the coating. With the right parameter combinations, crack-free, porous and high-hardness coating layers can be obtained.

In the process of laser surface coating with ferrobor, boride phases are formed in the coating area and these phases harden the microstructure of the material and significantly increase the wear and corrosion resistance. At the same time, it is possible to obtain fine-grained and homogeneous coating structures with the rapid cooling effect of laser energy [7-9].

In this study, the coating characteristics formed by ferrobor powder on steel surfaces under different ferroboron contents were examined and the obtained microstructure, phase formation and abrasive properties were evaluated.

[1]Corresponding author

2. Material and Method

In the experimental part of the study, SAE 4140 steel was used for surface coating purposes, and the chemical composition of this steel is given in Table 1. The dimensions of the steel samples used are 60 mm x 20 mm x 10 mm. For laser boronizing of the surface, boron-containing powder mixtures were first prepared. For these powders, ferroboron powder (70 microns) was added to iron powder, and the analysis of the ferroboron powder used is presented in Table 2

Table 1. Chemical composition of SAE 4140 Steel (wt.%)

C	Mn	Si	P	S	Cr	Mo	Fe
0.42	0.80	0.26	0.019	0.023	1.10	0.19	Rest

Table 2. Chemical composition of ferroboron (wt.%)

B	C	Si	Al	P	Fe
15-20	Max. 0.2	Max. 1.5	Max. 2.5	Max. 0.03	Rest

The amount of ferroboron in the powder mixture was adjusted to be 4 wt.%, 6 wt.%, and 8 wt.% and the powders were subjected to a mixing process. After the preparation of the powder mixtures, they were applied to the substrate surface with an organic binder at a height of 3 mm and left to dry. In this way, the samples were made ready for laser melting and then subjected to the melting process. During the melting process, a power of 1.4 kW and a scanning speed of 5 mm/s were used. The laser track is schematically shown in Figure 1. The tracks were applied side by side with 50% overlap.

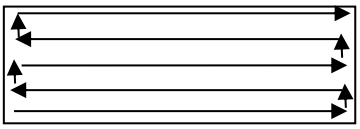


Figure 1. Schematic representation of the tracks followed during laser coating.

After the laser coating process, samples were extracted from the coated region for microstructural examination, microhardness measurements, XRD analysis, and abrasive wear tests, and were subjected to the necessary preparation procedures. Microhardness measurements were carried out starting from the coated surface toward the substrate material under a load of 300 g. In the abrasive wear test, the pin-on-disc method was used, and wear losses were determined under 20N load with a wear path of 150 m.

3. Results

3.1. Microstructural Analysis Results

The microstructures obtained from the samples are shown in Figure 2. When analyzing the microstructure formation for the sample containing 4 wt.%B, it is observed that the boron content is slightly above the eutectic boron concentration. According to the Fe-B binary phase diagram, the eutectic composition corresponds to approximately 3.8 wt.% boron. Therefore, the alloy with 4 wt.% boron lies just above the eutectic point, and thus contains some primary Fe_2B phases with eutectic in the microstructure. When the boron content increases from 4% to 6%, the amount of primary Fe_2B phases increases in the microstructure, and when increased to 8%, the microstructure becomes predominantly composed of primary Fe_2B phases [7-9]. The phases formed are also supported by the XRD analysis results given in Figure 3. According to the Fe-B equilibrium phase diagram, the observed phases are consistent with those expected based on equilibrium conditions.

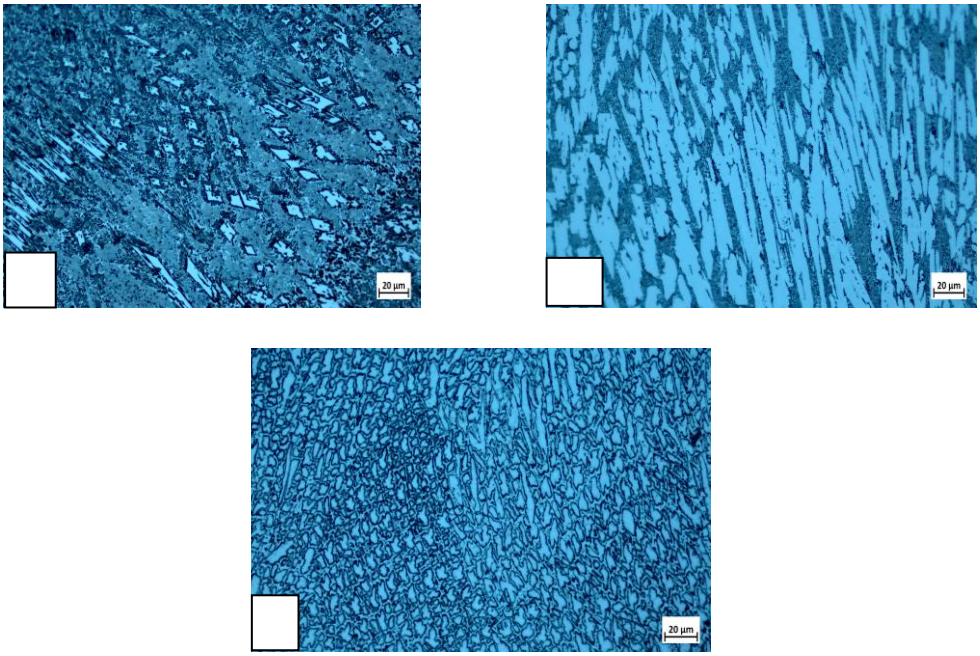


Figure 2. Optical microscope microstructure images taken from the coatings (a: 4 wt.%B, b: 6 wt.% B, and c: 8 wt.% B)

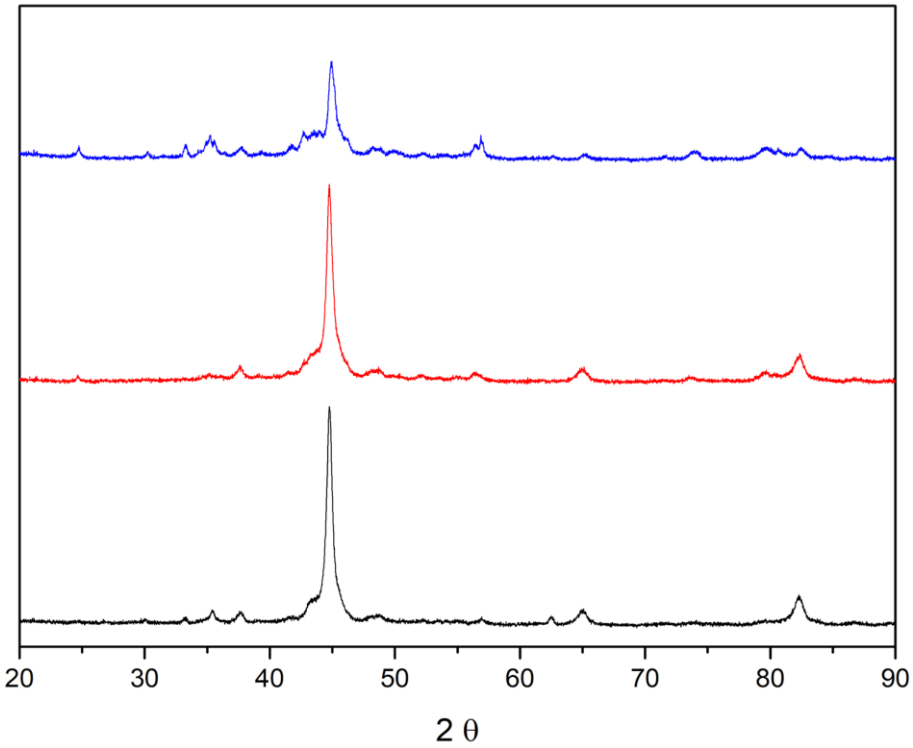


Figure 3. XRD results obtained from the coatings.

3.2. Hardness Test Results

The hardness results obtained from the coatings are given in Table 3. As can be seen from these values, the hardness of the coatings increases with the increasing boron content. This increase in hardness is attributed to the rise in the amount of the Fe_2B phase formed in the coating. Considering that the hardness of this phase can reach up to 1800 HV, these results are consistent with those reported in the literature [7-9].

Table 3. Hardness values obtained from the coatings

Distance(mm)	4 wt.% B	6 wt.% B	8 wt.% B
0.2	1091	1212	1351
0.4	1072	1191	1365
0.6	1091	1205	1384
0.8	1060	1195	1353
1.1	1060	1204	1375
1.3	1053	1209	1399
1.6	983	1213	1375
1.8	941	1204	1353
2.0	1000	1241	1409
Interface	603	712	581
2.2	365	303	405
2.4	327	326	326

3.3. Wear Test Results

The results of the weight loss measurements from the samples are presented in Figure 4. As can be seen from this figure, the wear losses decrease as the boron content in the coating increases. This behavior is attributed to the presence of Fe_2B phases in the coating. The amount of Fe_2B increases with the boron content, and the resulting increase in surface hardness protects the surface against abrasive particles [7].

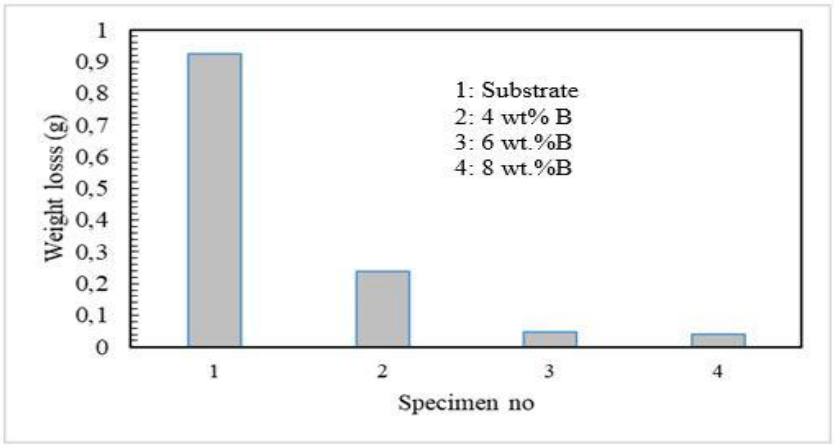


Figure 4. Weight loss values

4. Conclusions

In this study, SAE 4140 steel was coated using a laser beam with boron contents of 4 wt.%, 6wt.%, and 8 wt.%. The microstructure, hardness, and wear properties of the coatings were examined. The general conclusions drawn from the study are summarized below:

- The coating microstructure changed from a primary Fe_2B + eutectic structure to a predominantly primary Fe_2B structure.
- As the boron content in the coating increased, the hardness also increased.
- As the boron content in the coating increased, the wear resistance improved.
- Based on the obtained results, this study appears to have high applicability for components used in the wear industry.

5. Acknowledgment

The laser device used in the experimental part of this study was obtained within the scope of project ADEP 25.49, supported by the Scientific Research Projects Coordination Unit of Fırat University. The authors would like to express their gratitude to the Scientific Research Projects Coordination Unit of Fırat University for their support.

6. References

[1] Wang C, Li R, Bi X, Yuan W, Gu J, Chen, J, Yan M , Zhang Z. "Microstructure and wear resistance property of laser cladded CrCoNi coatings assisted by ultrasonic impact treatment", *Journal of Materials Research and Technology*, 22, 853–864, 2023.

[2] Zhu Z, Liu X, Y. Liu Y, Zhang S, Meng, Y, Zhou H, Zhang S. "Effects of Cu/Si on the microstructure and tribological properties of FeCoCrNi high entropy alloy coating by laser cladding", *Wear*, 512-513, 204533, 2023.

[3] Farahmand P, Liu S, Zhang, Z, Kovacevic R. "Laser cladding assisted by induction heating of Ni–WC composite enhanced by nano-WC and La_2O_3", *Ceramics International*, 40, 15421–15438, 2014.

[4] Bartkowski D. Młynarczak A, Piasecki A, Dudziak B, Goscianski M, Bartkowska A. "Microstructure, microhardness and corrosion resistance of Stellite-6 coatings reinforced with WC particles using laser cladding", *Optics and Laser Technology*, 68, 191–201, 2015.

[5] Davydova A, Domashenkov A, Sova A, Movtchan I, Bertrand P, Desplanques N, Peillon N, Saunier S, Desrayaud C, Bucher S. "Selective laser melting of boron carbide particles coated by a cobalt-based metal layer", Journal of Materials Processing Technology, 229, 361–366, 2016.

[6] Saravanan M, Senthilkumar P, Channankaiah D. "Wear resistance improvement of a laser cladded B_4C–NiCr on mild steel", *IJIRST* 2, 157–160, 2016.

[7] Aneta Bartkowska A, Bartkowski D, Radosław Swadźba R, Damian Przestacki D, Miklaszewski A. "Microstructure, chemical composition, wear, and corrosion resistance of FeB–Fe_2B–Fe_3B surface layers produced on Vanadis-6 steel using CO2 laser", *International Journal of Advanced Manufacturing Technology*, 95, 1763–1776, 2018.

[8] Bartkowska A, Swadźba R, Popławski M, Bartkowski D, "Microstructure, microhardness, phase analysis and chemical composition of laser remelted FeB-Fe_2B surface layers produced on Vanadis-6 steel". *Optics and Laser Technology* 86, 115–125, 2016.

[9] Bartkowska A, Bartkowski D, Przestacki D, Talarczyk M, "Microstructure and selected properties of boronized layers produced on C45 and CT90 steels after modification by diode laser". *Archives of Mechanical Technology and Materials*, 36, 51–58, 2016.

Use of *Lavandula angustifolia* M. Plant in Plant Protection

Aslıhan HAROĞLU ORHAN[1], İnanç ÖZGEN[2]

[1,2] Department of Biotechnology, Graduate School of Natural and Applied Sciences, Firat University, Elazig, Turkey.
[1] 242111103@firat.edu.tr, [2] iozgen@firat.edu.tr
[1] (ORCID: 0009-0000-1840-0641), [2] (ORCID: 0000-0003-1742-9324)

Abstracts
Lavandula angustifolia M., commonly known as English lavender, has been used for centuries in the Mediterranean region for its pleasant scent, calming properties, and ability to extend the shelf life of food. Its essential oil has demonstrated strong antibacterial, antifungal, insecticidal, repellent, and herbicidal activities, making it a promising natural alternative in plant protection. Scientific studies have shown that this oil is effective against various plant pathogens, storage pests, disease-carrying insects such as mosquitoes and ticks, and even certain weed species. Its main active components—such as linalool, linalyl acetate, and borneol—contribute to its broad-spectrum bioactivity. As chemical pesticides increasingly threaten environmental and human health, Lavandula angustifolia M. offers an eco-friendly, sustainable, and efficient solution for controlling plant diseases, pests, and weeds in agricultural systems.

Keywords *:* *lavandula angustifolia* m., herbal insecticide, herbal fungicide, herbal bactericide, herbal herbici

1. Introduction

Lavender (Lavandula); mostly of Mediterranean origin *L. dentata* , *L.* It is a genus belonging to the Lamiaceae family , which includes many species such as *intermedia* , *L. latifolia* , *L. stoechas* and *L. angustifolia* (*officinalis*). The most cultivated species in this genus are; *L. dentata*, known as French lavender and spreading in the Mediterranean, Yemen and the Arabian Peninsula; *L. latifolia* , known as Portuguese Lavender, grown in the Western Mediterranean ; L. stoechas, known as Spanish Lavender, grown in North Africa and around the Mediterranean; *and L. angustifolia*, known as English Lavender .

Essential oils are generally easy to extract, rapidly decompose in the soil, are broad-spectrum, and environmentally friendly. Essential oils derived from certain plants in *the Lamiaceae*, *Myrtaceae*, *Asteraceae, Rutaceae, Apiaceae,* and *Laureaceae* families not only have aromatic and public health effects, but also possess insecticidal, antiviral, antibacterial, and antifungal properties. Various studies have shown that essential oils derived from different plant species have ovicidal, larvicidal, antifeedant, insecticidal, and repellent properties against various insect species. They alter insects' behavior during feeding, egg-laying, and mating, inhibiting their growth and development, and possessing neurotoxic mechanisms.

For example, limonene and linalool are two main components of some citrus oils. Limonene has been shown to be effective against adult bean beetles (*Callosobruchus phasecoli*) and the Formosan subterranean termite (*Coptotermes formosanus), while linalool* has been found to be toxic to the eggs and larvae of the Caribbean fruit fly (*Anastrepha suspense). Essential oils from Mentha* species have been shown to be effective against the common stored grain pests the cowpea beetle (*Callosobruchus maculatus*) and the red flower beetle (*Tribolium castanum*). Lemongrass essential oils have long been sold in the form of candles as an insect and mosquito repellent. Tasmanian blue gum (*Eucalyptus globulus) oil* has been found to have significant activity against larvae and pupae of the house fly (*Musca domestica L.). The essential oils of some eucalyptus species (Eucalyptus dunnii, E. saligna, E. benthamii, E. globulus and E. viminalis)* have been found to have effective insecticidal and repellent activity in the control of the corn beetle (*Sitophilus zeamais). The chemical constituents of the essential oils of some thyme species (Satureja thymbra* L. and *Origanum onites* L.) and myrtle (*Myrtus communis* L.) have been tested for their insecticidal activity against adult stages of the stored product pests Mediterranean flour moth (*Ephestia kuehniella*), Indian flour moth (*Plodia interpunctella*) and bean beetle (*Acanthoscelides obtectus*). Myrtle oil has been found to have effective insecticidal and repellent activity in the control of the corn beetle (*Sitophilus zeamais) . There are numerous studies indicating that* the survival rate of adults and larvae *of cabbage root flies (Phorbia brassicae*) is reduced [1].when treated with essential oils obtained from the aerial parts of dill (*Anethum graveolens L.*

[1]Corresponding author

Numerous studies have investigated the effects of essential oils derived from various plants. This review will evaluate the potential use of *Lavandula angustifolia M., also known as English lavender, as a pesticide.*

2. Description of *Lavandula angustifolia* M. Plant

Lavandula angustifolia M. plant means narrow -leaved lavender . Its old name was *Lavandula officinalis* , meaning medicinal lavender (Figure 1.). As can be seen from this, *L. angustifolia* M. has been widely used in folk medicine, aromatherapy, and perfumery for thousands of years. Oils obtained from this species are known worldwide as lavander. Lavandula oil, obtained from *Lavandula intermedia* , a hybrid of *Lavandula angustifolia* x *Lavandula latifolia* , and called 'lavandin', is the closest competitor to the essential oil obtained from *L. angustifolia and is frequently confused with the other. In this study, Lavandula angustifolia,* also known as lavander, is considered Studies conducted with the essential oil of M. have been investigated.



Figure 1. Image of lavandula angustifolia m. plant

2.1. History of Lavandula angustifolia M.

In a study that reviewed the literature on the history of lavender, it was stated that Lavandula was first mentioned as Stoechas by Dioscorides before 50 AD in his review of lavender published by Baron Frederic de Gingins-Lassaraz in 1826; Dioscorides wrote, "Stoechas grows on the Stoechades Islands near Gaul and takes its name from there. It is a plant with a thin stem, its flowers are clustered in the head like in *Thymus . Its leaves are long and have a bitter taste."* In 76 AD, Pliny, in his description of a plant called Nard, points to a plant that may have belonged to Stoechas of Dioscorides: "Nard is always falsified with the False Nard, which grows everywhere, is whitish in color and has thicker leaves than the Nard of Crete, and has small leaves in clusters on top[2].”

2.2. Use of the Lavender Plant in the World

Lavender is used worldwide in the food industry, perfumery, and cosmetics, in the production of many products such as soaps and scented candles, and in the pharmaceutical industry as a carminative, relaxant, and sedative in aromatherapy. The countries that cultivate and produce the most lavender oil are Bulgaria, France, China, Russia, and some Eastern European countries. Bulgaria ranks first in the production of essential oil from [3]*L. angustifolia* .

In a study conducted to protect food safety, the antimicrobial activity and antioxidant capacity of the oil obtained from Lavandula angustifolia M. (Denizli/Turkey) on some Gram-negative/positive pathogenic microorganisms and lactic acid bacteria were investigated. The highest antimicrobial activity of L. angustifolia oil was found on Bacillus subtilus and Methicillin-resistant Staph. aureus with zone diameters of 24-22 mm, respectively, followed by Staph. aureus, E. coli, Campliobacter jejuni, Salmonella poona, Lactobacillus reuteri and Lactobacillus plantarum. However, a significant antioxidant effect was determined by DPPH, FRAP and metal chelation analyses [4].

Another study found that lavender has significant antibiosis effects against both C. albicans and E. coli at different concentrations , demonstrating that *Lavandula angustifolia* M. possesses both antifungal and antibacterial activity [5]. Therefore, its antibacterial and antifungal properties are also used in the global food chain to extend shelf life.

2.3. Use of the lavender plant in Anatolia

Although [6]*Lavandula angustifolia M. has been recognized as a plant in recent years, some research indicates that its parts have been used for various purposes in traditional Turkish medicine over the years. For example, the flowers, inflorescences, seeds, leaves, and branches, whether extracted from essential oils, infused or boiled, have been used for psychological disorders, hepatitis, insomnia, diuretic, stomach ailments, dyspepsia, hair and skin care, rheumatism, migraine, acne treatment, depression, and as a sedative*. It has also been used in small pouches in chests, wardrobes, and drawers, both for its pleasant scent and as an alternative to mothballs, believed to have mothball-repelling properties.

Components Found in the Essential Oil of Lavandula Angustifolia M.

Lavandula angustifolia Numerous studies exist on the essential oil content of Lavandula angustifolia *M* .. A comprehensive study on the counterfeiting of commercial *Lavandula angustifolia M. essential oils and a study to assess industrial standards* revealed 122 identified compounds from 55 *L. angustifolia essential oils from Poland,* the proportions of which were included in the study [7]. The average essential oil compositions obtained by averaging the data are shown in Table 1.

Table 1. Components and average amounts found in the essential oil of *lavandula angustifolia m.* grown in Poland

Serial Number	Component Name	Percentage in whole fat (%)
1	linalyl acetate	32.79
2	linalool	32.77
3	(Z)-β-ocimene	4.28
4	terpinen−4-ol	4.05
5	α-terpineol	4.05
6	(E)-caryophyllene	3.43
7	lavandulyl acetate	3.02
8	(E)-β-farnesene	2.46
9	(E)-β-ocimene	2.42
10	3-octanone	1.10
11	borneol	0.96
12	1,8-cineole	0.91
13	1-octen−3-yl acetate	0.88
14	lavandulol	0.82
15	Mirsen	0.69
16	geranyl acetate	0.51
17	hexyl acetate	0.50
18	neryl acetate	0.46
19	α-pinene	0.46
20	α-santalene	0.43
21	limonene	0.37
22	hexyl butanoate	0.34
23	germakren D*	0.33
24	p-cymene	0.31
25	caryophyllene oxide	0.31

Table 1. (Continue) Components and average amounts found in the essential oil of *lavandula angustifolia m.* grown in Poland

26	3-octanol	0.30
27	1-octen−3-ol	0.30
28	camphor	0.29
29	β-phellandrene	0.24
30	3-Karen	0.23
31	allo-ocimene	0.22
32	camphene	0.18
33	krypton	0.17
34	cis -linalool oxide (furanoid)	0.17
35	trans -linalool oxide (furanoid)*	0.17
36	β-pinene	0.16
37	Nerol	0.16
38	γ-terpinene	0.15
39	p-cymene−8-ol	0.14
40	terpinolene	0.14
41	α-humulene	0.13
42	trans-α-bergamoten	0.13
43	α-thujene	0.11
44	sesquithujene	0.10
45	1-methoxyhexane	0.10
46	γ-cadinene	0.09
47	bornyl acetate	0.09
48	(Z)−3-hexen−1-ol	0.09
49	1-hexanol	0.09
50	hexyl hexanoate	0.08
51	cis -sabinene hydrate	0.08
52	seskisabinen	0.07
53	1,3,5-undecatriene	0.07
54	furfuryl alcohol	0.07
55	geraniol	0.07
56	hexyl isobutyrate	0.07
57	hexyl 3-methylbutanoate	0.07
58	τ-cadinol	0.06
59	m-cyme	0.06
60	Sabinen	0.06
61	hexyl tiglate	0.06
62	cumin aldehyde	0.06
63	citronellol	0.06
64	α-terpinene	0.06
65	α-phellandrene	0.05
66	trans-linalool oxide (pyranoid)	0.05
67	γ-terpineol	0.05
68	thyretienol	0.05
69	hexyl 2-methyl butanoate	0.05

Table 1. (Continue) Components and average amounts found in the essential oil of *lavandula angustifolia m.* grown in Poland

70	bornil format	0.05
71	myroxide	0.05
72	alloaromadendren	0.04
73	p -cymene−7-ol	0.04
74	phellandral	0.04
75	coumarin	0.04
76	β-bisabolene	0.04
77	lavandulil isovalerate	0.04
78	dausen	0.04
79	butyl acetate	0.04
80	lavender lactone	0.04
81	β-bourbonene	0.04
82	trtcyclene	0.03
83	nerol oxide	0.03

In a study conducted in Diyarbakır Province, when the components of *L. angustifolia were examined,* A total of 46 components were identified, constituting 98.92% of the total oil. The main components of lavender oil are linalool (22.46%), 2-isopropenyl-5-methyl-4-hexenyl acetate (15.42%) and bergamiol (14.19%) (Table 2) [8].

Table 2. Components and average amounts found in the essential oil of *L. angustifolia plant* grown in Diyarbakır Province [8]

Serial Number	Component Name	Percentage in Whole Fat
1	Linalool	22.46
2	2-Isopropenyl-5-methyl	15.42
3	Bergamiol	14.19
4	Caryophyllene oxide	5.15
5	Isothymol	4.43
6	Tau.-Cadinol	3.64
7	Borneol	2.48
8	Geraniol acetate	2.12
9	Cis-Ocimene	2.03

The findings of a study investigating the essential oil components of L. angustifolia grown in Algeria are also included in Table 3 [9].

Table 3. Components and average amounts found in the essential oil of *L. angustifolia* grown in algeria [9]

Serial Number	Component Name	Percentage in Whole Fat
1	Linalool	31.27
2	Camphor	16.21
3	Linalool oxide	11.98
4	Linalyl acetate	11.93
5	2-Furanmethanol	7.49
6	Eucalyptol; 1,8-cineole	6.76
7	Borneol	5.34
8	L-octen-3-yl acetate	1.56
9	L-Hexyl butyrate	1.25
10	L-Cyclohexene-l-methanol	1.21
11	Trimethylbicyclo(2.2.1) heptan-2-ol	0.85
12	3-Octanone	0.79
12	Hexyl acetate	0.73
14	Cyclohexane	0.66
15	Lavandulyl acetate	0.63
16	Caryophyllene epoxide	0.39
17	Lavandulol	0.38
18	β-Ocimene	0.21
19	Tetrahydrofuran	0.20
20	β-Myrcene	0.18

The essential oil components found in another study on the L. angustifolia plant growing in Bulgaria are given in Table 4 [10].

Table 4. Components and average amounts found in the essential oil of *l. angustifolia* grown in Bulgaria [10]

Serial Number	Component Name	Percentage in whole fat (%)
1	Linalyl acetate	27.53
2	Linalool	24.11
3	E-β-ocimene	7.01
4	Terpinen-4-ol	5.11
5	Caryophyllene	4.46
6	Carvacrol	4.42
7	Lavandulyl acetate	3.52
8	Z-β-farnesene	3.33
9	Z-β-ocimene	3.15
10	lavandulol	1.47
11	α-terpineol	1.36
12	Eucalyptol	1.09

Table 4. (Continue**)** Components and average amounts found in the essential oil of *l. angustifolia* grown in Bulgaria [10]

13	O-cymene	0.84
14	γ-terpinene	0.82
15	1-octen-3-yl-acetate	0.78
16	β-myrcene	0.69
17	1-octen-3-ol 0.222 3-octanone	0.67
18	Geranyl acetate	0.62
19	Limonene	0.53
20	α-santalene	0.53
21	Germacrene D	0.52
22	Acetic acid, hexyl ester	0.51
23	Butanoic acid, hexyl ester	0.36
24	Nerol acetate	0.36
25	α-pinene	0.32
26	Camphor	0.26
27	3-octanol	0.24
28	Caryophyllene oxide	0.24
29	β-thujene	0.23
30	Bornyl acetate	0.18
31	Camphene	0.17
32	E-α-bergamotene	0.16
33	Z-geraniol	0.16
34	Humulene	0.14
35	Crypton	0.13
36	Terpinolene	0.13
37	3-octanol,acetate 0.05 allo-ocimene	0.12
38	β-sesquiphellandrene	0.12
39	3-carene	0.11
40	α-phellandrene	0.06
41	β-pinene	0.06
42	β-terpinene	0.04
43	Tricyclene	0.02
44	Caren-4-ol	0.01

Based on the findings from these four studies , the most abundant components in *L. angustifolia* essential oil are linalool, linalyl acetate, borneol, caryophyllene oxide, and geraniol acetate. These are generally pleasant-smelling and are terpenoids.

4. Use of *Lavandula Angustifolia* M. in Agriculture

Lavandula angustifolia M. is widely used in areas such as human health, nutrition, and cosmetics. In recent years, due to the high cost of pesticides and the damage they cause to human health and the environment, they have begun to be used as a homemade herbal pesticide by agricultural producers as an alternative to pesticides. This has attracted the attention of researchers, who have investigated the utility *of L. angustifolia* M. against plant diseases, pests, and weeds. It has been determined that both *L. angustifolia* M. and other *Lavandula* species can be used for this purpose. This study compiles studies on *L. angustifolia M.*

4.1. Phytopathological effect

There are many studies on the use of *Lavandula angustifolia M. as a pesticide*. In a study evaluating the effects of organic essential oils *(Thymus vulgaris , L. angustifolia M. , Rosmarinus officinalis)* against apple pathogens, both in vitro and in vivo as an integrated management tool against postharvest diseases such as *Fusarium avenaceum, Botrytis cinerea , Penicillium expansum* and *Neofabraea vagabunda* which cause significant losses in apple, it was found that 1.4 ml dose of *L. angustifolia* M. essential oil showed a reduction of 52.8%, 15.2%, 51.4% and 61.4% for *F. avenaceum, B. cinerea, P. expansum* and *N. vagabunda* , respectively [11].

In a study investigating the antifungal effects of different types of lavender oils against fungi of both medical and agricultural importance, it was found that *L. angustifolia* essential oil had a very good antifungal effect in general, especially against the medically important fungi *A. nidulans and T. mentagrophytes* [12].

4.2. Insecticidal Effect

In a study investigating the fumigant toxicity and biochemical effects of selected essential oils on the red flour beetle, *Tribolium castaneum (Coleoptera: Tenebrionidae*), it was observed that essential oil, 1,8-cineole (71.9%), camphor (10.0%), limonene (2.7%) and borneol (6.5%) obtained from *L. angustifolia* showed good fumigation activity (LC50 34.1 µL/L air) and increased the insect mortality rate in *Tribolium castaneum* beetle alone and in combination with β-caryophyllene (24.4%), limonene (15.3%), sabinene (13.7%), β-pinene (8.3%) and α-pinene (5.1%) as the main compounds in the essential oil obtained from *P. nigrum* [13].

A study on *Culex pipiens* mosquitoes, the most common primary vector of West Nile virus worldwide, found that lavender oil caused 87.20% larval mortality and 95.55% adult mortality [14].

In a study conducted to investigate the bioacaricidal activity of *Lavandula angustifolia* M. essential oil against *Rhipicephalus (Boophilus) annulatus (Acari; Ixodidae)* females, various concentrations of *L. angustifolia* M. essential oil were applied to *R. annulatus*. As shown in Table 5, although there was no difference between 6.0% and 8.0% for all measured criteria, a significant mortality rate was observed at concentrations greater than 4.0%. At 24 hours after inoculation, mortality rates were 73.26% and 100% in the groups treated with 4.0% and 8.0% essential oil, respectively. Lavender essential oil also reduced tick egg weight in a concentration-dependent manner, as shown in Table 6, and when the data in Table 7 were examined, it was seen that it caused 100% failure in oviposition. Therefore, a positive correlation was observed between *L. angustifolia* M. essential oil concentration and tick control, evaluated with relative mortality rate and egg-laying weight. When calculated using the Probit test, values were found to be 2.76 times higher than the control. [15].

Table 5. Mortality rate in *R. annulatus* females exposed to *l. angustifolia essential oil* 24 hours after inoculation [15].

Essential Oil Concentrations (%)	Mortality Rates (%)
Control	0 ± 0
0.5	0 ± 0
1.0	6.6 ± 6.6
2.0	33.3 ± 6.6
4.0	73.3 ± 17.6*
6.0	100 ± 0*
8.0	100 ± 0*

* Indicates a significant difference between the application doses and the control (P = 0.05)

Table 6. Response of *R. annulatus to L. angustifolia* essential oil 24 hours after inoculation the mass of eggs produced by exposed females [15].

Essential Oil Concentrations (%)	Mass of Eggs Produced (g) *
Control	0.1220 ± 5.033E-03
0.5	0.1233 ± 3.528E-03
1.0	0.1040 ± 7.211E-03
2.0	0.1028 ± 4338E-03
4.0	2.777E-02 ± 2.295E-0.2*
6.0	TE (zero)*
8.0	TE (zero)*
TE = Not Detected	

* Indicates a significant difference between the application doses and the control (P = 0.05)

Table 7. Egg-laying failure rate of *R. annulatus* females exposed to *L. angustifolia essential oil* 24 hours after inoculation [15].

Essential Oil Concentrations (%)	Ovulation Failure Rates (%)
Control	0 ± 0
0.5	0 ± 0
1.0	6.6 ± 6.6
2.0	33.3 ± 6.6
4.0	80 ± 11.7*
6.0	100 ± 0*
8.0	100 ± 0*

* Indicates a significant difference between the application doses and the control (P = 0.05)

Dengue fever is a disease caused by a virus in the *Aedes aegypti* mosquito. One way to prevent the spread of dengue is to interrupt the mosquito's life cycle. One of the most easily broken cycles in the mosquito life cycle is the larval cycle. In a study to determine the larvicidal effectiveness of *L. angustifolia* M. essential oil on *Aedes aegypti* larvae, 60 larvae were exposed to *L. angustifolia* M. essential oil. They found that lavender extract killed the larvae. The mortality rate of *Aedes aegypti* larvae was significantly affected by the number of drops of the extract and the number of larvae per container [16].

Another study was conducted on *Aedes aegypti* adults. As can be seen in Table 8, they found that various concentrations of essential oil obtained from the leaves of *Lavandula angustifolia* M. caused adult mortality of *A. aegypti* at different rates, from 26.25% to 100%, after 24 hours of exposure [17].

Table 8. *A. aegypti* after 24 hours of exposure to L. angustifolia leaf juice adult mortality rate[17]

Concentrations	Mosquito Amount	Number of Deaths in Replay				Total Deaths	Average Mortality Rate	Mosquito Mortality Rate (%)
		1	2	3	4			
25%	20	5	4	7	5	21	5.25	26.25
50%	20	8	10	12	9	39	9.75	48.75
75%	20	15	14	17	16	62	15.5	77.5
Control (-)	20			0		0	0	0
Control (+)	20			20		20	20	100

In a study evaluating the effects of the essential oil content obtained from *L. angustifolia* on *Rhyzopertha dominica (Coleoptera: Bostrichidae*) adults, GC-MS analysis showed that this oil contained 56 compounds, the main components of which were linalool (20.42%) and linalyl acetate (13.24%). This essential oil was found to exhibit insecticidal activity against *R. dominica* adults, depending on the concentration and exposure time, as seen in Figure 4.2.5. [18].

A study on the *Ephestia cautella* moth found that L. angustifolia essential oil was an effective insecticide against *E. cautella*, with an LC50 value of 0.39 µl/L and a toxicity index of 46.15% after 7 days of exposure.

The results showed that the essential oils tested had fumigant toxicity against *E. cautella*, and that Peppermint and Lavender oils were more effective against E. cautella than star anise and coconut oils after 7 days of exposure [19]

In another study investigating the effect of L. angustifolia essential oil on Sitophilus oryzae (rice beetle) adults, it was found that adult mortality rates increased with increasing essential oil concentration and exposure time, reaching 98.2% after 24 hours and 100% after 48 hours [20].

In a study investigating the fumigant insecticide activities of essential oils obtained from thyme (*Origanum syriacum* L.), lavender (*Lavandula angustifolia* M.), sage (*Salvia officinalis* L.), fennel (*Foeniculum vulgare* Mill.) and laurel (*Laurus nobilis* L.) plants against adults of the cowpea beetle *Callosobruchus maculatus* (*Coleoptera: Bruchidae*), the LD50 value of Lavender was found to be 11.64 µg ml-1 and it was revealed that it can be used as a fumigant against the cowpea beetle, which is a storage pest. [21].

In another study, L. angustifolia showed extremely potent insecticidal effects against *Pyrrhocoris apterus* (flame bug) at various concentrations. Additionally, *L. angustifolia* also showed potent insecticidal activity against *Oulema melanopus* (cereal leaf bug); in the control group, all tested survived [22] .

The chemical compositions of essential oils obtained from three plants, *Salvia officinalis*, *Lavandula angustifolia* M. and *Mentha pulegium*, were investigated for their insecticidal activities and feeding deterrent properties against *Plodia interpunctella*. As a result, lavender essential oil was found to be an effective contact insecticide [23].

In a study on the toxic, repellent and anti-feeding activities of essential oil of *Lavandula angustifolia* M. against adults of *Sitophilus granarius*, adult mortality upon contact with the essential oil reached 91.7% and 100% after 24 and 48 hours of exposure, respectively, and the LD50 and LD90 values were 83.8 and 379.7 µg/adult after 24 hours, which decreased to 58.3 and 208.3 µg/adult after 48 hours, respectively. When the essential oil was applied as a fumigant, 100% mortality was achieved at volumetric doses of 11.9 and 47.5 mg/L in the absence and presence of wheat grains, respectively. LC50 and LC90 values were found to be 1.6 and 4.1 mg/L volume, respectively, in the absence of wheat grains, and 10.9 and 47.6 mg/L volume, respectively, in the presence of wheat grains, indicating that *Lavandula angustifolia* M. essential oils were effective as insecticides against *Sitophilus granarius* [24].

4.3. Repellent effect

In a study conducted on Empoasca flavescens, it was found that different doses of thymol and 1-methoxy-4-methyl-2-(1-methylethyl)-benzene obtained from the plant Lavandula angustifolia M. significantly repelled male aphids, and their combinations/mixtures significantly repelled both male and female aphids [25].

In another study, the chemical composition of *Lavandula angustifolia* M. was determined by gas chromatography-mass spectrometry. Then, the fumigant toxicity, repellent and anti-baiting properties of the essential oil obtained from L. angustifolia and some of its effects on *Rhyzopertha dominica* F. *(Coleoptera: Bostrichidae)* adults were evaluated and an increase in repellency percentage was found depending on the concentration of the essential oil and the exposure time, the values shown in Table 9. Repellency rates; The mean values were found by classifying them as Class 0 (RP < 0.1%), class I (RP = 0.1%–20.0%), class II (RP = 20.1%–40.0%), class III (RP = 40.1%–60.0%), class IV (RP = 60.1%–80.0%) and class V (RP = 80.1%–100%) [26].

Table 9. Repellent effect of *l. angustifolia* essential oil against *r. dominica* adults at different exposure times [26].

Concentrations	Post-Application Time	Repellency Rate (%)	Class
1 µ/ml	30 minutes	26.66	II
	3 hours	40.00	II
	6 hours	43.33	III
	24 hours	50.00	III
2 µ/ml	30 minutes	32.66	II
	3 hours	50.00	III
	6 hours	53.33	III
	24 hours	60.00	III
4 µ/ml	30 minutes	50.00	III
	3 hours	60.00	III
	6 hours	63.33	IV
	24 hours	73.33	IV
8 µ/ml	30 minutes	65.00	IV
	3 hours	73.33	IV
	6 hours	83.33	V
	24 hours	86.96	V

Again, *L. angustifolia*, In a study investigating its effect on *Hyalomma marginatum rufipes* , high repellency (70-100%) was recorded against adults, and this repellency increased up to 120 minutes at high concentrations and decreased to 40 minutes at low concentrations.

In general, the repellent power of *L. angustifolia* essential oil against *H. m. rufipes* adults was found to be similar to that of N,N-diethyl-m-toluamide (DEET), a synthetic acaricide against the same tick species [27].

A study on adults of the rice beetle *Sitophilus oryzae found that the repellency of L. angustifolia* essential oil increased with increasing essential oil concentrations [20].

Lavandula angustifolia M. against adults of *Sitophilus granarius*, the repellent activity of lavender essential oil increased significantly with increasing dose, but was not significantly affected by increasing exposure time, as seen in Table 10. In filter paper biological assays, mean PR values were over 80% (repellent class V) from the dose of 0.441 mg/cm2 and significantly higher than those recorded at the lowest doses 60 minutes after the start of the experiment. In behavioral biological experiments in the field, it was observed that increasing essential oil concentrations led to significant reductions in insect numbers both in the absence and presence of wheat grain odors [24].

Table 10. Statistical analysis of repellency rates of *S. granarius adults* exposed to different doses of *l. angustifolia essential oil for different durations* using Turkey HSDD test.[24]

Dose (mg/ ^{cm2})	Exposure Time (Min)			
	30	60	90	120
3,531	95.0 ± 5.0 a	100 ± 0.0 a	100 ± 0.0 a	95.0 ± 5.0 a
1,765	95.0 ± 5.0 a	100 ± 0.0 a	95.0 ± 5.0 a	100 ± 0.0 a
0.883	100 ± 0.0 a	95.0 ± 5.0 a	100 ± 0.0 a	95.0 ± 5.0 a
0.441	85.0 ± 9.6 ab	100 ± 0.0 a	90.0 ± 5.8 a	90.0 ± 5.8 a
0.221	55.0 ± 12.6 b	60.0 ± 0.0 b	55.0 ± 12.6 b	45.0 ± 17.1 c
0.110	60.0 ± 8.2 b	25.0 ± 9.6 c	15.0 ± 12.6 c	5.0 ± 9.6 c
0.055	20.0 ± 8.2 c	25.0 ± 9.6 c	25.0 ± 9.6 bc	20.0 ± 11.6 c
F	15.65	41.68	25.73	18.81
Df	6	6	6	6
P	<0.001	<0.001	<0.001	<0.001

4.4. Nutritional inhibition and developmental retardation effects

In a study investigating the effects of Lavandula angustifolia M. on adults of Rhyzopertha dominica F. (Coleoptera: Bostrichidae), enzymatic measurements showed neurotoxic activity, as evidenced by inhibition of acetylcholinesterase (AChE). Furthermore, stimulation of the detoxification system was observed, as indicated by an increase in glutathione-S-transferase (GST) activity and a decrease in glutathione (GSH) levels. The nutritional indices of the essential oil were investigated. The results showed a decrease in the tested digestive enzymes, α-amylase and protease, together with an increase in the relative growth rate, relative consumption rate, conversion efficiency of digested food, and feeding deterrence index in the treated series compared to the control group.

The same study also examined its effects on feeding, and the data in Table 11 were found. Accordingly, when L. angustifolia essential oil was consumed as food by R. dominica adults, an increase in feeding deterrence was observed at 1% and 4% concentrations, but a decrease in food conversion was found at the 4% concentration. They also found that relative growth rate was inversely proportional to essential oil concentrations, with the highest concentration reducing R. dominica growth and feeding, but not affecting the relative consumption rate [26].

Table 11. Effect of *l. angustifolia* essential oil on nutritional indexes of *r. dominica* adults [26]

Concentrations	Relative Growth Rate (mg/mg/hour)	Relative Consumption Rate (mg/mg/hour)	Conversion Efficiency of Food Purchased	Nutrition Deterrence Index (%)
0%	0.17 ± 0.01 a	223 ± 0.46 a	8.39 ± 2.09 a	-
1%	0.07 ± 0.01 b	1.73 ± 0.31 a	4.51 ± 1.41 a	34.11 ± 10.08 b
4%	0.02 ± 0.00 c	1.60 ± 0.05 a	1.55 ± 0.14 b	54.39 ± 5.80 a
P	0.0002	0.4120	0.0438	0.0035
F	45.79	1,032	5,513	16.76
df	2	2	2	1

In a study conducted on *Glyphodes pyloalis , various doses* of essential oils from *Lavandula angustifolia M. were applied and it was* observed that they affected the nutritional indices of 4th instar larvae of *Glyphodes pyloalis ; compared to the control, the essential oil reduced the total protein amount of G. pyloalis* larvae as seen in Figure 2., and the total lipid amount as seen in Figure 3.. All concentrations of the essential oil (0.25%, 0.5%, 1% and 2%) affected the main macromolecules and the activities of the key enzyme in *G. pyloalis* [28].

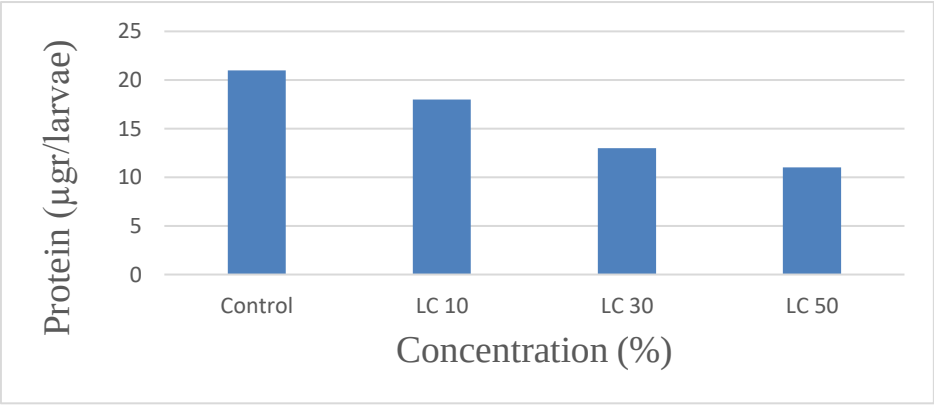


Figure 2. Total protein amounts in 4th instar larvae of *g. pyloalis* after application of *l. angustifolia essential oil* [28].

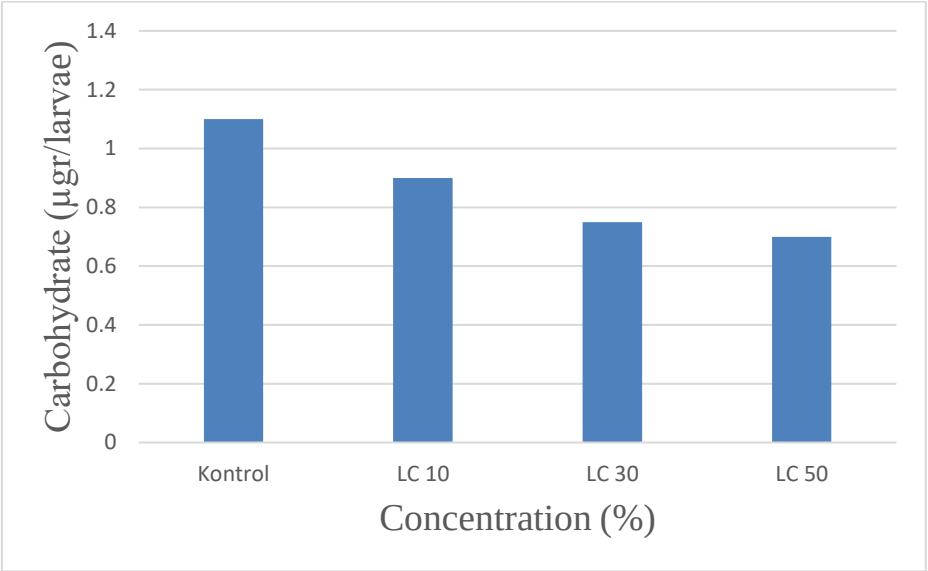


Figure 3. Total lipid amounts in 4th instar larvae of *g. pyloalis* after application of *l. angustifolia* essential oil [28].

A study on the feeding deterrent properties of the chemical compositions of essential oils obtained from three plants, *Salvia officinalis, L. angustifolia* M., and *Mentha pulegium*, against *Plodia interpunctella* found that *L. angustifolia* M. alone acted as a potent environmentally friendly inhibitor of the feeding behavior of *Plodia interpunctella* larvae, as shown in Table 12. Applications of *L. angustifolia* essential oil resulted in a decrease in relative growth rate, relative consumption rate, and conversion efficiency of ingested food, while a significant increase in the feeding deterrent index was observed [23].

Table 12. Feeding indices of *P. Interpunctella 3rd instar larvae* exposed to different concentrations *of salvia officinalis, mentha pulegiun, l. angustifolia essential oils* [23].

	Concentration (µL/pellet)	Relative Growth Rate (mg/mg/day)	Relative Consumption Rate (mg/mg/day)	Conversion Efficiency of Food Purchased (%)	Nutrition Deterrence Index (%)
S. officinalis	0	0.89 ± 0.005 a	29.1 ± 0.13 a	7.19 ± 0.99 a	0.00 ± 0.00 a
	8	0.35 ± 0.05 b	5.95 ± 0.38 a	0.06 ± 0.01 b	89.36 ± 0.003 b
	16	0.48 ± 0.18 b	3.49 ± 0.55 a	0.11 ± 0.04 b	89.52 ± 0.05 b
	32	0.08 ± 0.01 b	4.03 ± 0.95 a	0.04 ± 0.00 b	89.44 ± 0.06 b
L. angustifolia	0	0.89 ± 0.05 a	29.09 ± 0.13 a	7.19 ± 0.99 a	0.00 ± 0.00 a
	8	-0.13 ± 0.14 a	4.41 ± 0.77 a	-0.04 ± 0.03 a	89.40 ± 0.05 b
	16	-0.08 ± 0.02 b	3.73 ± 0.55 b	-0.03 ± 0.01 b	89.49 ± 0.05 b
	32	-0.04 ± 0.12 b	2.98 ± 0.49 b	-0.04 ± 0.05 b	89.52 ± 0.05 b
M. pulegium	0	0.8 ± 0.05 a	29.09 ± 0.13 a	7.19 ± 0.99 a	0.00 ± 0.00 a
	8	-0.16 ± 0.11 a	14.76 ± 0.88 a	-1.19 ± 0.69 a	8.49 ± 1.64 a
	16	-0.06 ± 0.08 b	10.09 ± 0.46 a	-0.75 ± 0.31 b	35.48 ± 3.03 b
	32	-0.31 ± 0.042 b	4.53 ± 0.36 b	-1.19 ± 1.37 c	68.18 ± 1.89 c

4.5. Herbiside effect

In a study conducted to evaluate the phytotoxic effects of commercially widely used essential oils of *Eucalyptus citriodora, Lavandula angustifolia* M., and *Pinus sylvestris* on the weeds *Portulaca oleracea, Lolium multiflorum, Echinochloa crus-galli*, and *Nicotiana glauca*, the essential oils were tested at different concentrations (0.125–1 µL/mL). Germination index and seedling length data were recorded for 14 days. The essential oil of *L. angustifolia* M., containing highly oxygenated monoterpenes linalool (38.7 ± 0.1%), 1,8-cineole (26.5 ± 0.1%), and camphor (14.2 ± 0.1%), affected the seed germination and development of *L. multiflorum, E. crus-galli*, and *N. glauca*, but did not show

any significant phytotoxic effect on *Portulaca oleracea* seed germination. *L. multiflorum* was the most sensitive weed, especially to lavender essential oil, and reduced hypocotyl and radicle growth by 87.8% and 76.7%, respectively, at a dose of 1 µL/mL [29].

5. Conclusions

L. angustifolia M. has been frequently used in folk medicine, particularly in aromatherapy, for centuries. Its teas and essential oils have been used to soothe neurological disorders, as a painkiller for various ailments, as a diuretic for some urinary system disorders, to relieve insomnia, and to induce and maintain deep sleep. Its antibacterial and fungicidal properties are known, and it is particularly utilized in food storage. The cosmetics industry and perfumery are also among its most common uses. It is also widely used to repel insects and create a pleasant fragrance in clothing and fabric storage. However, the use of *L. angustifolia* M. in phytopathology, for controlling plant pests and weeds, is still a relatively new field.

In this study, in comparison with previous studies, it was determined that *L. angustifolia* M. has lethal properties on some warehouse pests such as *Tribolium castaneum*, *Culex pipiens*, *Rhipicephalus (Boophilus) annulatus*, *Aedes aegypti*, *Rhyzopertha dominica*, *Ephestia cautella*, *Sitophilus oryzae*, *Callosobruchus maculatus*, *Pyrrhocoris apterus*, *Plodia interpunctella*, *Sitophilus granarius*, tick and mosquito species, and adults and larvae of some leaf-eating pests; repellent properties on pests such as *Empoasca flavescens*, *Rhyzopertha dominica*, *Hyalomma marginatum rufipes*, *Sitophilus oryzae*, *Sitophilus granarius*; feeding and growth inhibitory properties on *Rhyzopertha dominica*, *Glyphodes pyloalis*, *Plodia interpunctella*; and also has anti-inflammatory properties on some plant diseases such as *Fusarium avenaceum*, Botrytis It has been shown that it can be used in the control of cinerea, *Penicillium expansum*, and *Neofabraea vagabunda*, and even in weed control on plants such as *Portulaca oleracea*, *Lolium multiflorum*, *Echinochloa crus-galli*, and *Nicotiana glauca*. In today's world where pesticides threaten human health and nature is increasingly polluted, it is hoped that *L. angustifolia* M. will be a natural and healthy alternative.

Acknowledgements

We would like to thank Fırat University Biotechnology Department for providing all kinds of support in our study.

6. References

[1] Grdiša M, Grsic K. “Botanical Insecticides in Plant Protection”. *Agriculture Conpectus Scientificus* , 78, 2, 85-93, 2013.
[2] Audette FL. An Historical and Horticultural Perspective. Master's Thesis, Ohio University, Ohio, USA, 1997.
[3] Republic of Turkey Ministry of Industry and Technology. “Lavender Agriculture and Industry Feasibility Report 2020”. Republic of Turkey Ministry of Industry and Technology, Ankara, Turkey, 2020.
[4] Özdemir O, Yılmaz N, Gök M, and Kaya MO. “Determination of Antimicrobial and Antioxidant Activities of Lavandula angustifolia Volatile Oil”. *Turkish Journal of Agricultural Research* , 9, 3, 265-273, 2022.
[5] Khalif DS, Ozbek A, Gedik TN, and Taner Z. “Investigation of Antifungal and Antibacterial Effect of Lavandula angustifolia Against Candida albicans and Escherichia coli”. *Aurum Journal of Health Sciens* , 5, 3, 121-128, 2023
[6] Bozyel ME, Bozyel EM. “Ethnomedicinal Uses of Genus Lavandula (Lamiaceae) in Turkish Traditional Medicine”. *International Journal of Academic and Applied Research* , 4,2,5-16,2020.
[7] Pokajewicz K, Lamaka D, Wieczorek PP, Bortkiewicz O, and Paluch E. “Lavender essential oil market integrity: A comprehensive study of commercial Lavandula angustifolia essential oils adulteration and assessment of industrial standards”. *Industrial Crops & Products* , 222, 2024.
[8] Kizil S, Haşimi N, Kizil S, Kilinç E, and Tolan V. “Essential oil components, microelements content and biological activity of Lavandula angustifolia miller”. First Mediterranean Symposium on Medicinal and Aromatic Plants, Diyarbakır, Türkiye, 17-20 April 2013.
[9] Moussa AHH, Benaliouche F, Sbartai I, and Sbartai H. “Biological activities and chemical characterization of Lavandula angustifolia essential oil from Seraïdi, Northeastern Algeria”. *Biodiversitas* , 24, 8, 4535-4542, 2023.
[10] Kozuharova E, Simeonov V, Batovska D, Stoycheva C, Valchev H and Benbassat N. “Chemical composition and comparative analysis of lavender essential oil samples from Bulgaria in relation to the pharmacological effects”. *Pharmacia* , 70, 2, 395-403, 2023.

[11] Di Francesco A, Aprea E, Gasperi F, Parenti A, Placi N, Rigosi F, Baraldi E. "Apple pathogens: Organic essential oils as an alternative solution", *Sci Hortic* , 300, 2022.
[12] Moon T, Cavanagh HMA, and Wilkinson JM. "Activity of Australian grown Lavandula spp. essential oils against Aspergillus nidulans, Trichophyton mentagrophytes, Leptosphaeria maculans and Sclerotinia sclerotiorum". *Journal of Essential Oil Research* , 19, 171-175, 2007.
[13] Oviedo-Sarmiento JS, Cortes JJB , Avila WAD, Suarez LEC, Daza EH, Patino-Ladino OJ, Prieto-Rodriguez JA. "Fumigant toxicity and biochemical effects of selected essential oils towards the red flour beetle, Tribolium castaneum (Coleoptera: Tenebrionidae)". *Pesticide Biochemistry and Physiology* , 179, 2021.
[14] Al-Kasem Bosly HA. "Larvicidal and adulticidal activity of essential oils from plants of the Lamiaceae family against the West Nile virus vector, Culex pipiens (Diptera: Culicidae)". *Saudi Journal of Biological Sciences* , 29, 8, 2022.
[15] Pirali-Kheirabadi K, Teixeira da Silva JA. "Lavandula angustifolia essential oil as a novel and promising natural candidate for tick (Rhipicephalus (Boophilus) annulatus) control", *Experimental Parasitology* , 126, 2, 184-186, 2010.
[16] Farah K, Harahap N. "Analysis of Lavender Flower Extract (Lavandula Angustifolia) as a Natural Larvicide for the Control of Aedes Aegypti Larvae". *Journal of Society Innovation and Development* , 5, 2, 202-211, 2024.
[17] Pakpahan ELE, Pane DY. "Bay Leaf Juice (Syzygium Polyanthum) And Lavender Leaf Juice (Lavandula Angustifolia) As Ae. Aegypti Adult Mosquito Control". *Jurnal Keperawatan Priority* , 7, 2, 2024.
[18] Sayada N, Tine S, Soltani N. "Evaluation of a Botanical Insecticide, Lavender (Lavandula Angustıfolıa (M.)) Essential Oıl As Toxicant, Repellent And Antifeedant Against Lesser Graın Borer (Rhyzopertha Dominica (F.))". *Applied Ecology and Environmental Research* , 20, 2, 1301-1324, 2022.
[19] AL-Nujiban AA. "Efficacy of Certain Selective Plant Essential Oils against Ephestia cautella (Walker) (Lepidoptera: Pyralidae)". *Egyptian Scientific Journal of Pesticides* , 6, 4 25-32, 2020.
[20] Al-Harbi NA, Al Attar NM, Hikal DM, Mohamed SE, Latef AAHA, Ibrahim AA, Abdein MA. "Evaluation of insecticidal effects of plants essential oils extracted from basil, black seeds and lavender against sitophilus oryzae". *Plants* , 10, 5, 2021.
[21] Kaya K, Sertkaya E, Türemiş İ, Soylu S. "Determination of chemical composition and fumigant insecticidal activities of essential oils of some medicinal plants against the adults of cowpea weevil, Callosobruchus maculatus", *Kahramanmaraş Sütçü İmam University Journal of Agriculture and Nature* , 21, 5, 708-714, 2018.
[22] Perry T, Batterham,P. Daborn P.J. "The biology of insecticidal activity and resistance". *Insect Biochemistry and Molecular Biology* , 7, 411-422, 2011.
[23] Moullamri M, Rharrabe K, Annaz H, Laglaoui A, Alibrando F, Cacciola F, Mondello L, Bouayad N, Zantar S, Bakkali M. "Salvia officinalis, Lavandula angustifolia, and Mentha pulegium essential oils: insecticidal activities and feeding deterrence against Plodia interpunctella (Lepidoptera: Pyralidae)". *Journal of Essential Oil-Bearing Plants* , 27, 1, 16-33, 2024.
[24] Di Stefano MG. Toxic, Repellent And Antifeedant Activities Of Lavandula Angustifolia Müller (Lamiaceae) Essential Oil Against Sitophilus Granarius (L.) (Coleoptera, Curculionidae) Adults. Doctoral Thesis, Universita Degli Studi Del Molise, Campobasso, Italy, 2015.
[25] Han SJ, Wang MX, Wang YS, Wang YG, Cui L, Han BY. "Exploiting push-pull strategy to combat the tea green leafhopper based on volatiles of Lavandula angustifolia and Flemingia macrophylla". *Journal Integrative Agriculture* , 19, 1, 193-203, 2020.
[26] Mkolo MN, Magano SR. "Repellent effects of the essential oil of Lavendula angustifolia against adults of Hyalomma marginatum rufipes". Journal of the South African Veterinary Association, 3, 78, 149-152, 2007.
[27] Yazdani E, Sendi JJ, Aliakbar A, Senthil-Nathan S, "Effect of Lavandula angustifolia essential oil against lesser mulberry pyralid Glyphodes pyloalis Walker (Lep: Pyralidae) and identification of its major derivatives". Pestic Biochem Physiol, 107, 2, 250-257, 2013.
[29] Ibáñez MD, Blázquez MA. "Phytotoxic Effects of Commercial Eucalyptus citriodora, Lavandula angustifolia, and Pinus sylvestris Essential Oils on Weeds, Crops, and Invasive Species." *Molecules* , 24, 15, 2019.

Current Development of Soft Grippers Used in Robotic Applications

Gökhan GÜNEŞ[1] , Murat ŞEN[2]

[1] Department of Mechanical Engineering, Engineering Faculty, Firat University, Elazığ, Türkiye.
[2]Department of Mechanical Engineering, Engineering Faculty, Firat University, Elazığ, Türkiye.
[1]gokhan_gns_@hotmail.com, [2]msen@firat.edu.tr
[1](ORCID: 0009-0001-7667-9705), [2](ORCID: 0000-0002-3063-5635)

Abstract

This study examines the design principles, actuation mechanisms, material selection criteria, and application domains of soft robotic grippers. Compared to rigid grippers, soft robotic grippers enable the safe manipulation of delicate, irregular, and fragile objects. The roles of under-actuation and morphological computation are highlighted as critical for achieving adaptive grasping with minimal control complexity. Pneumatic, hydraulic, tendon-driven, and shape memory-based actuation technologies are systematically analyzed, and their limitations are comparatively evaluated. The functional characteristics of elastomers, shape memory polymers, and magnetorheological materials, particularly in terms of flexibility, durability, and operational performance, are explored. Bio-inspired designs, sensor integration, and hybrid actuation approaches are discussed alongside recent advances and application-specific implementations. From agricultural tasks to medical interventions, a wide range of use cases is examined. Owing to their intrinsic compliance and morphological computation capabilities, soft robotic grippers demonstrate superior performance in such domains. Existing challenges such as limited gripping force, material fatigue, and energy efficiency are critically assessed. Furthermore, the future directions of soft grippers concerning aspects such as smart materials, AI-assisted control systems, 4D printing technologies, and miniaturization are also addressed. This research reinforces the transformative potential of soft robotics in human-robot interaction and in navigating unstructured environments.

Keywords: soft robotic grippers, adaptive grasping, pneumatic actuation, smart materials, bio-inspired design

1. Introduction

Soft grippers belong to the class of robotic end-effectors and are generally manufactured from flexible materials. Unlike traditional rigid grippers, these grippers do not rely on fixed kinematic pairs and precise mechanical movements; instead, they achieve this through deformation of their structural components with environmental interaction. This approach allows them to mimic the dexterity and adaptability of biological systems. The operation of Soft grippers is based on the principle of morphological computation. In this principle, the robot's physical shape, material properties, and natural dynamics significantly contribute to its behavior and control, reducing the need for complex open control algorithms. The natural softness and mechanical compliance of materials allow Soft grippers to adapt their shape to contact surfaces, distribute stress over a wider area, and absorb impacts, establishing interaction with objects with less damage [1-5].

Soft grippers go beyond the limitations of precise control, utilizing the natural compliance of materials to enable robots to safely interact with unpredictable environments and delicate objects. This marks a deviation from the traditional robotics focus on high precision and rigidity toward a more biologically inspired, flexible approach. Traditional robots rely on rigid connections and precise control of discrete joints. This naturally limits their ability to adapt to irregular shapes or grasp fragile objects without damage. In contrast, soft robot grippers achieve movement and interaction through deformation of their compliant bodies [2, 3, 6, 7].

Soft grippers have a series of important advantages compared to traditional rigid grippers. These grippers can grasp irregular, fragile, and delicate objects without damage thanks to their adaptability and versatility; with their safety features, they offer ideal use in human-robot interactive environments. Thanks to morphological computation and under-actuated designs, they enable complex adaptive grasping with simplified control systems, reducing the need for meticulous force control and comprehensive sensing. Additionally, while their lightweight and compact structures provide energy efficiency and maneuverability, soft contact interfaces demonstrate superiority in damage reduction, minimizing the risk of irreversible damage on the surfaces of grasped objects [1, 2, 8].

[1]Corresponding author

This study examines the fundamental principles of soft robotic grippers, focusing on their ability to handle delicate and irregularly shaped objects. It highlights the importance of under-actuation and morphological computation for simple and adaptive grasping. The research analyzes various actuation methods (pneumatic, tendon-driven, etc.) and materials (elastomers, shape memory alloys), while also discussing bio-inspired designs and sensor integration. It concludes by addressing current challenges like material fatigue and a look at future trends such as AI-assisted control and 4D printing.

2. Design Principles and Mechanisms

With the rise of the concept of soft robotic grippers, a new class of grippers has emerged that overcomes the limitations of traditional, rigid-based robots. These types of robots offer important advantages such as flexibility and adaptability thanks to being manufactured from inherently soft and compliant materials. Soft grippers have the capability to interact safely and gently with objects in human-machine interaction and in irregular, unknown environments. Unlike their rigid counterparts, soft robots have the ability to gently grasp and maneuver objects with open-loop kinematic control. This capability is largely based on the morphological computation principle provided by the gripper's physical structure. Morphological computation allows some of the control tasks supervised by control units in traditional robotic systems to be undertaken by the mechanical properties of the robot's hardware. This simplifies control strategies while helping to achieve better compliance with objects having complex shapes. The main focus point in the design of soft grippers is shape design that enables effective handling of objects with low rigidity and various forms. These designs determine the gripper's fundamental performance characteristics such as compliance, reliability, and flexibility. During design, material selection, actuation mechanisms, and optimization techniques play critical roles. For example, rod-based models are used to analyze how soft robot fingers provide the gripper and how performance can be affected by changing system parameters. Generally, soft grippers gain movement capability through pneumatic, hydraulic, or cable-driven approaches. These different approaches enable soft grippers to adapt to various applications. Additionally, the need to develop complex-shaped devices to produce complex movements can increase design and manufacturing costs. Therefore, systematic design methodologies are of critical importance for optimizing soft robots to suit the needs of targeted applications. In general, soft gripper design aims to simplify interaction with complex objects while increasing safety and adaptability [1, 2, 3].

2.1. Under-actuation and Morphological Computation

In the field of soft robotic grippers, under-actuation and morphological computation principles offer significant advantages compared to traditional rigid robots. Under-actuation means controlling a robot with fewer actuators than the number of degrees of freedom (DOF) it possesses. This situation simplifies control systems and reduces hardware costs while allowing the robot to adapt to complex interactions with the environment. Soft grippers, thanks to their structural compliance, can generally gently grasp and maneuver objects with open-loop kinematic control. This takes advantage of soft robots' ability to spontaneously adapt shape, whereas rigid robots require complex feedback mechanisms for precise control. Morphological computation enables the robot's physical form and material properties to passively perform part of the tasks such as sensing and control. That is, rather than relying on complex control algorithms, the mechanical compliance of the robot's body enables it to spontaneously adapt to the object's shape and surface. This is particularly important when working with uncertain or irregularly shaped objects, as it allows the gripper to gently wrap around the object and provide stable grip [9].

As an example, the "Fin Ray" effect can be given; this effect mimics the principle of fish fins bending in an S-shape when exposed to external load to form a wrap around the object and provides adaptable grasping. This type of morphological compliance is achieved through the natural flexibility of elastomeric materials and asymmetric geometries or reinforcing fibers incorporated into the design [10].

Soft grippers can also benefit from under-actuation through techniques such as granular jamming. In this method, the gripper wraps around the object in its soft configuration and then stabilizes the object by applying vacuum or hardening with another mechanism. This allows obtaining high holding forces by providing shape changes in multiple

degrees of freedom with a single actuation input. However, this approach may not be suitable for deformable objects, as difficulties may be experienced in adapting to object deformation when the gripper structure hardens [11].

In summary, under-actuation and morphological computation are fundamental design principles in enhancing soft robotic grippers' simplified control, high adaptation capability, safe interaction, and ability to gently grasp various objects. These approaches pave the way for soft robots to be used in a wide range from industrial to home applications [12].

2.2. Basic Movement and Structure

Soft robotic grippers are built upon two fundamental movement forms of bending and bending-twisting thanks to their flexible material and continuous deformation capability. Bending-type grippers encompass designs where gripping surfaces are controllably bent to wrap around objects. This approach is divided into two subcategories. Non-continuous bending grippers combine elastic finger joints with rigid components; motors and cable-driven manipulators fall into this group. Micro-scale four-fingered designs and shape memory alloys also provide a balance of flexibility and speed. Continuous bending grippers offer passive compliance to complex geometries using monolithic structures manufactured entirely or largely from flexible materials. Pneuflex and origami-inspired actuators, starfish-like modules, and pressure-operated silicon cells are typical examples of this type. Foam structures produced with the "Lost-salt" method or 3D printing techniques simplify manufacturing processes, while fiber-reinforced bending actuators provide theoretical modeling and design guidance possibilities. Twisting-type grippers exhibit twisting behavior around objects in addition to bending movement. Helical inflatable designs along with multi-segment continuum robots like OctArm generally offer complex movement capability with hydraulic or pneumatic actuation [13]. This structure is suitable for applications requiring additional degrees of freedom and three-dimensional grasping strategies. Tentacle-like grippers grasp by surrounding and wrapping around objects with structures resembling an elephant trunk. McKibben muscles; control both bending and extension movements thanks to their fibrous structures that contract or extend when pressure is applied [14]. Nitrile rubber tubes and multiple FEA modules are examples specific to this category. Finally, multi-fingered hands mimic human anatomy. Generally manufactured using durable silicones like Dragon Skin; passive compliance to object shape is provided through under-actuated arrangements and hybrid rigid-soft designs reinforce safe grasping. Three or four-fingered soft grippers offer the ability to successfully grasp a wide variety of objects with a limited number of actuators [3, 15, 16].

The most distinctive feature of soft robotic grippers is their inherent flexibility and adaptation capabilities. These systems with high degrees of freedom can perform complex and compliant actions with their simple mechanical structures thanks to their almost unlimited movement capabilities; they do not need numerous joints or complex control algorithms as in rigid robots. The passive adaptation feature called internal compliance or morphological computation automatically adapts to complex surfaces of objects with the natural flexibility of material and structural design. Some designs guarantee both gentle and safe grasping by adjusting their stiffness with the variable compliance principle; as in the example of opposing McKibben muscles, it is possible to increase bending stiffness without changes in finger position. Open-loop kinematic control supports the gripper's elastic structure to adapt to objects without the need for feedback mechanisms and significantly reduces the complexity of control systems. Additionally, the ability to successfully grasp objects of varying size, shape, and fragility plays a critical role especially in sensitive applications like food processing. Shape-compliant grasping achieved through the granular jamming method increases grip reliability by providing high-level adaptation to object contours. On the other hand, some PneuNet-based soft grippers having limited degrees of freedom reveals the necessity of additional mechanisms for more complex tasks. Some materials offering self-sealing or puncture resistance increases durability and reliability in challenging working conditions, while in-hand manipulation capability enables complex in-hand movements through synchronized operation of multiple actuators. This classification reveals the fundamental design philosophies and operational capabilities of soft grippers, showing that future developments will aim to further optimize the balance between these mobility and constraint levels [17, 18].

The structural design of soft robotic grippers directly determines their adaptation capabilities and force application methods. Unlike traditional rigid robotic systems, these grippers adopt organic approaches such as continuous

structures and modular designs. Rod-based models enable precise prediction of grasping performance by analyzing finger deformation through internal curvature profiles. Continuous structures offer passive compliance to complex-shaped objects by using completely flexible materials instead of discrete joints; thus reducing the need for control and planning algorithms. Multi-segment continuum manipulators provide wide movement flexibility through multiple flexible sections. Modular designs increase the scalability of the gripper by showing rapid adaptation to different grasping tasks; for example, the ability to easily change the number of multiple linear volume actuators according to application requirements provides access to different force levels. Under-actuated arrangements offer passive adaptation to object shapes with fewer actuators than degrees of freedom while also minimizing control complexity. Multi-fingered hands gain superior manipulation capability in complex manipulation tasks with their anthropomorphic structures and integrated sensors. Origami-inspired foldable structures enable complex shape changes with simple actuation, while hollow injection molding and foam porous structures also simplify manufacturing processes and reduce costs. This classification emphasizes that each structure offers optimized features for specific application needs [19, 20].

2.3. Grasping Method

Soft robotic grippers adopt different methods according to their object grasping shapes, and these methods determine the gripper's physical interaction with objects and versatility. In the method called enveloping or compliant grasping, the gripper's fingers or surfaces fully conform to the object's shape and achieve grasping by wrapping around it like a cage. In typical examples known as "universal soft grippers," the hardening of a bag filled with particles like ground coffee beans through vacuum is the basis. This way, variable-shaped objects such as small flashlight bulbs, plastic bags, bottle caps, and various office materials are successfully grasped, but limitations arise in complete wrapping of flat-surfaced objects and flexibility after solidification. Vacuum mode (suction) enables holding objects through pressure difference via a suction cup or similar apparatus and is particularly used in packaging difficult-to-absorb substances like wrapped rice or sauce. This method, suitable for smooth, flat-surfaced objects, has gained widespread use in packaging and conveying robot systems. The clamping or pinching method is based on mechanical fingers grasping objects through friction. In two, three, or more fingered designs, the number of fingers and geometry are determined according to the shape of the workpiece to be grasped. While traditional industrial grippers can be equipped with rigid jaws, compliant mechanisms and fingers developed through topology optimization for irregular objects offer superior performance [21, 22].

Adhesion methods are divided into two subtypes: gecko-inspired dry adhesion and electroadhesion. Dry adhesion provides high gripping force by creating strong van der Waals forces with microfiber surfaces and can lift both regular and irregular objects on both smooth and rough surfaces. In electroadhesion, flexible material fingers adhere to object surfaces through electrical charging; for example, it shows success in processing metallized PET film. The hooking method is seen in multimodal systems based on the combination of layer jamming and tendon-driven mechanisms. This approach offers secure grasping by catching objects like hooks and holding them from support points. The actuation principles of soft grippers differ in terms of performance, size, and complexity. Pneumatic actuators provide movement through pressure-induced inflation and deformation of chambers filled with air or liquid. Subtypes such as PneuNets, fast PneuNet, fiber-reinforced SPA, variable-slope air cushions, and bladder motors exist; however, they require large compressors for high force and speed. Tendon-driven actuators offer control through motor-pulled tendons embedded in the structure [23]. These systems enable precise control of position and force, can pass through complex paths, and can be hybridized with other methods like granular jamming. Electroactive polymers (EAPs) are materials that change shape with electrical stimulation. Dielectric elastomer actuators (DEAs) and ionic polymer-metal composites (IPMCs) experience slow response and hysteresis problems despite offering self-sensing capability. Shape memory alloys (SMAs) generate force by transforming to predetermined shapes at specific temperatures. Used in thin wire or spring form and can increase stiffness up to 120 times; however, response times and control precision are limited. Magnetorheological fluids (MR) allow the gripper to gain variable stiffness by changing viscosity with magnetic fields. They have been used in processing various objects like food items and plastic containers. The granular jamming method enables hardening of gripper structure containing particles through vacuum pressure. It offers both compliance and sufficient stiffness together; however, it requires large vacuum pumps and experiences difficulty in adapting to object deformation after hardening. Dexterity can be increased without increasing complexity by combining with pneumatic systems [3, 15].

In conclusion, each grasping and actuation method diversifies the adaptation capability and performance of soft robotic grippers. Designers should evaluate the advantages and limitations of these methods according to application requirements to select the optimum combination.

3. Actuation Mechanisms

Fluidic Elastomer Actuators (FEAs) are one of the fundamental actuation methods of soft robotic grippers and enable deformation of elastic structures with liquid or gas pressure. Generally using chambers with asymmetric geometry or anisotropic reinforcement fibers, inflation movement is converted to bending or extension in a specific direction. FEAs can produce 80 N at 300 kPa pressure, 112 N block force at 200 kPa; can exhibit bending angles up to 300° and 0.05-1.0 s response time. Main examples are PneuNets, McKibben muscles, and tentacle-like grippers. PneuNets create bending movement with multiple chambers embedded in elastomer; McKibben muscles work by contraction of braided corded tubes with pressure; tentacle designs use a series of FEA elements to wrap around objects. High force and lifting capacity is obtained by combining positive and negative pressures, and FEAs have also been used in applications like octopus arms or planar manipulators.

Pneumatic actuators are widespread with advantages of fast response, high output force, and easy manufacturing. Gas pressure provides deformation by changing the internal curvature profile in flexible structures [24, 25].. Manufactured by injection molding or 3D printing; PneuNets and fiber-reinforced FRSAs are basic design examples. Fiber winding creates higher bending stiffness by restricting radial expansion. McKibben muscles offer antagonistic control with braided tubes. Pneumatic actuators can produce complex movements like rotation and bending, can increase fingertip force with rigid components, and provide grasping sizes between 84-141 mm in four-fingered grippers. The disadvantage is that the system gains volume and weight due to pump and valve requirements [26]. Hydraulic actuators offer higher force density and rigid control with liquid pressure. Deformation is obtained by pressurizing channels in elastomer with liquid. They have innovative applications like printable hydraulics and self-healing electrostatic actuators. Viscosity change like MR fluid can be used for controlled stiffness. However, pump-valve complexity, leakage risk, and prolonged response time due to high density of liquids are limiting [23, 27].

Tendon-driven mechanisms create bending and extension movement by pulling cables through motors or rigid actuators. FDM printing, molding, and hybrid structures are used. Variable stiffness is obtained with antagonistic tendon configuration. Used in systems like SDM hand, SIMBA arm, and "Farm Hand," it has grasped objects up to 1 kg with topology optimization. Tendon systems with advantages of low inertia and flexible compliance carry disadvantages of wear and control complexity [28, 29]. Shape memory alloys (SMAs) are wires or strips that return to programmed shape when heated. Used in micro grippers, curved actuators, and 3D printed hybrid systems. Offers 0.2-20 N force, 0.67-11 s response time. Can provide variable stiffness; however, low speed and high energy consumption are limiting [30]. Electroactive polymers (EAPs) are divided into dielectric elastomer actuators (DEA) and ionic polymer-metal composites (IPMC). DEAs show fast deformation with high voltage, 1-8 N force and 0.1-1 s response while having fatigue problems. IPMCs respond between 1-10 s with 0.1-4 N force; they have disadvantages of hysteresis and low efficiency. Other EAPs like liquid crystal elastomers (LCE) are also being researched [31].

Magnetic actuators offer variable stiffness with MR fluids. Pouches at fingertips provide compliance and grip by hardening and softening with magnetic field. Magnetic actuation is also applied in micro-robots. While having contactless control advantage, low force and magnetic interference are disadvantages [32].

Shape memory jamming combines granular or fiber jamming with SMPs. Granular jamming provides membrane hardening with vacuum; SMP joints provide stiffness control with heat. Offers high adaptation and safe grasping with a single system [33]. Other mechanisms include moisture-sensitive materials, hydrogels, LCEs, electro-adhesion, gecko-adhesion, explosive and super-coiled polymer (SCP) actuators. These "stimuli-responsive" technologies offer unique advantages for specific applications. The diversity of shape-changing mechanisms expands the flexibility and application range of soft grippers [34].

4. Material Selection and Properties

The performance of soft robotic grippers largely depends on the properties of the materials used. These materials determine the gripper's flexibility, durability, actuation mechanism, and adaptation to environmental conditions. The functionality of soft robotic grippers largely depends on the mechanical and chemical properties of structural materials. Proper material selection determines the gripper's flexibility, durability, actuation efficiency, and ability to adapt to environmental conditions. Below, the main material categories and their basic characteristics are summarized.

Elastomers constitute the most widely used material class in soft gripper designs. Thanks to their high elongation capability and rapid return properties, elastomeric fingers passively adapt to the object's geometry. The Shore degree defining elastomer hardness is critically important for optimizing the flexibility-stiffness balance required by the application. The hyperelastic behavior of elastomers means the ability to return to original shape despite large deformations. Mooney-Rivlin and Ogden models are widely used in defining hyperelasticity by adapting to experimental stress-strain data of rubber-like materials. In finite element analysis, these models form the basis for predicting the nonlinear deformations occurring during gripper grasping. The viscoelastic properties of elastomers describe time-dependent deformation and stress relaxation. Hysteresis and deformation rate dependency affect the dynamic response of soft fingers; therefore dynamic modeling becomes complex and quasi-static assumptions are preferred to analyze slow deformation processes. During the manufacturing phase, elastomeric grippers are generally manufactured by casting/molding or 3D printing methods; adjusting crosslinking density in PDMS mixtures provides adhesive strength and elasticity control. Silicon-based elastomers stand out with their low cost, biocompatibility, and processability advantages. For example, Ecoflex and PDMS (polydimethylsiloxane) have been successfully used in starfish-inspired gripper structures and flexible sensor skins with embedded liquid metal microchannels. Dragon Skin is preferred in three and four-fingered manipulators with its high tear strength. Nitrile rubber is a suitable material especially for tentacle-like pneumatic actuators; while thermoplastic elastomers like NinjaFlex (Shore 85A) are preferred in producing 3D printed high force capacity grippers. However, NinjaFlex's relatively high hardness can be limiting in applications requiring low pressure and sensitive force [35, 36].

Shape memory alloys (SMAs) and polymers (SMPs) belonging to the smart material class offer stiffness control and active actuation with heat or electrical stimulation. Shape memory alloys are flexible wires that gain plasticity at specific temperatures and return to their initial shape when heated. Curved SMA actuators provide controllable bending movements in soft grippers. Thin SMA wires can operate in the range of 0.2-20 N output force and 0.67-11 s response time; however, low actuation speed can be limiting in industrial applications. Shape memory polymers are smart materials that allow elastic deformation up to 200% and respond to heat or light. SMP-based actuators gain the ability to grasp randomly shaped objects by increasing their stiffness up to 120 times with external stimulation. SMP joints produced with 3D printing can return fingers to their initial position with rapid shape recovery stress. However, low actuation speed in both SMPs and SMAs can create difficulties in meeting industrial speed requirements [30, 37].

Magneto-rheological (MR) materials offer variable stiffness by changing viscosity when a magnetic field is applied. MR fluid-filled bags behave like soft molds during the grasping phase while hardening with magnetic field to provide strong grip. This transition occurs on the order of milliseconds, offering sensorless adaptation and low system complexity. MR grippers have produced effective results in grasping delicate food products like tomatoes, strawberries, and grapes without damage; they have safely held objects like bulbs and oranges. MR-based joint stiffness control has also been applied in wearable supernumerary limbs. Dielectric elastomer actuators (DEAs) are thin elastomer membranes placed between two flexible electrodes, where thickness decreases and surface increases with Maxwell stress when voltage is applied. They offer low weight, fast response (0.1-1 s), and high deformation rate (1-8 N force). High operating voltage, fatigue, and limited stress generation are the main disadvantages of DEAs. Variable stiffness actuators are obtained with rigid fiber integration and LMPA compounds; manufacturing is simplified by creating thermoplastic frames with 3D printing [38].

The mechanical behaviors of soft gripper elastomers are defined by hyperelasticity and viscoelasticity models. Hyperelastic materials provide adaptive grasping opportunity by maintaining elasticity at large deformations.

Experimental data is integrated into FEA analyses with Mooney-Rivlin and Ogden models. Viscoelasticity explains that deformations show relaxation and hysteresis over time depending on the applied load and loading speed. When quasi-static approach is adopted in dynamic deformations, modeling is simplified; however, viscoelastic properties remain critical for functions such as energy dissipation, shock absorption, and vibration damping [39].

Each of these material classes must be carefully evaluated for the suitability of soft grippers to different application requirements. Designers can maximize the flexibility, durability, and control efficiency of grippers by selecting optimal material combinations considering mechanical properties, manufacturing methods, and environmental factors [2, 40].

5. Application Areas

Soft robotic grippers undertake critical roles in many areas where traditional rigid grippers are inadequate thanks to their flexibility, compliance, and gentle interaction capabilities.

Holding fragile objects in unknown environments increases the complexity of control algorithms while making the integration of sensory feedback and computer vision methods mandatory. Magnetorheological grippers show success in agricultural automation by offering adaptive stiffness [8, 41]. Agriculture and food industry constitute priority application fields for these grippers. "Fin Ray"-like bending mechanisms are successfully used in greenhouses for fruit and vegetable harvesting; gecko-inspired "Farm Hand" designs can lift heavy products like pumpkins with high contact surface and evenly distributed stress. Four-fold pneumatic models with force feedback and grippers integrated with vacuum suction can collect products between 5-150 g like tomatoes, carrots, and strawberries without damage. Various configurations—tentacle, finger, hand, or envelope grasping—are optimized according to product diversity, balancing reliability, accuracy, flexibility, and stability requirements [42-44].

Soft grippers also offer unique advantages in food processing and packaging processes. In Japanese bento packaging tests, grippers worked in suction mode for wrapped rice and sauce, grasping modes for fried chicken and boiled eggs, safely carrying loads up to 75.2 g. Various components like spaghetti, ohitashi, bread, and salmon could be successfully placed; additionally, MR fluid-based models were tested in handling variable-shaped products like onions and artichokes [45].

In healthcare, soft grippers enhance tissue compliance in minimally invasive surgery, "Stiff-FLOP"-like manipulators reduce risks in interaction with internal organs. Pneumatic gloves and tendon-driven support devices strengthen movement capability in rehabilitation of paralyzed patients [46,47].

In industry, universal granular jamming grippers offer adaptive grasping in "pick-and-place" tasks of regular or irregular parts, while contact sensing is integrated through 3D printed conductive circuits.

Miniaturization is critical in microsurgery and nano-manipulation fields, and reducing micro-grippers while maintaining their performance is a fundamental engineering problem in healthcare [26]. Material limitations; slow response times, low force output, and fatigue tendencies of SMA, SMP, DEA, IPMC, and hydrogels are restrictive in industrial grasping. New materials—sustainable biogels, hybrid composites, electronic skin—are aimed at increasing durability and performance. Energy efficiency is on the agenda with McKibben muscles replacing bulky pneumatic and vacuum systems, voltage requirements in DEA/IPMCs, and the need for internal power sources for untethered applications [48, 49].

Additionally; underwater explorations, acoustically controlled soft robot fish, amphibious arms and adaptation to challenging pressure conditions with hydrostatic systems; studies continue on inflatable "POPUP" manipulators and cargo handling designs in space applications. This diversity shows that soft robotic grippers have become a universal tool in providing safe, compliant, and effective manipulation in industrial, agricultural, medical, underwater, and space conditions [50-52].

Universal gripper development is still a significant challenge: It is necessary to design systems that can grasp objects of different shapes, materials, and fragility, and can harmoniously use actuation, controlled stiffness, and adhesion methods. While morphological computation, where the mechanical properties of the body undertake some control tasks, replaces traditional rigid robotic control paradigms, insufficient output force and the penetrability of materials by sharp objects limits durability. In dual-mode applications like food packaging, the integration of suction and grasping modes is supported by advanced sensing and path planning software (Faster R-CNN, MoveIt). In the future, hybrid and multi-material designs, 3D/4D printing technologies, and bio-inspired mechanisms (octopus, Fin-Ray, gecko adhesion) will increase the adaptability and functionality of soft grippers; AI-supported control and sensor integration will optimize real-time grasping strategies.

6. Conclusion

Soft robotic grippers represent a paradigm shift from traditional, rigid manipulators, offering an exceptional capability to handle delicate, irregularly shaped, and fragile objects without causing damage. This study delves into the fundamental principles that govern their design, function, and future potential. mAt the heart of this technology lies the synergy between under-actuation and morphological computation. Unlike their rigid counterparts that require complex control systems for every degree of freedom, soft grippers achieve adaptive grasping through their inherent mechanical design and material properties. This allows them to conform to an object's shape with minimal control input, a concept known as morphological computation, significantly reducing the complexity of the control system and enhancing their adaptability in unstructured environments.

The research systematically analyzes a wide array of actuation mechanisms that power these grippers. It provides a detailed comparative evaluation of pneumatic and hydraulic systems, which rely on fluid pressure to induce deformation, as well as tendon-driven mechanisms that use cables to manipulate gripper fingers. Furthermore, it explores the transformative potential of shape memory-based systems, which utilize smart materials to generate force and motion through thermal or electrical stimuli. The analysis not only highlights the unique advantages of each method but also critically assesses their respective limitations in terms of speed, force output, and energy efficiency.

In addition to actuation, the choice of materials is paramount. The study meticulously examines the functional characteristics of key materials such as elastomers, known for their flexibility and stretchability, and shape memory polymers, which can undergo significant shape changes. It also explores the use of magnetorheological fluids, which enable variable stiffness control, allowing the gripper to be both compliant and rigid as needed. These material selections are crucial for ensuring the gripper’s durability, flexibility, and operational performance in diverse applications.

Looking beyond current capabilities, the study discusses cutting-edge developments, including bio-inspired designs that mimic natural biological systems for enhanced grasping efficiency. It also addresses the integration of sensor technology for haptic feedback and explores hybrid actuation approaches that combine multiple methods to overcome individual limitations. The broad range of applications, from intricate tasks in medicine to large-scale operations in agriculture, underscores the versatility of this technology. While the field is rapidly advancing, significant challenges remain, such as achieving higher gripping forces, mitigating material fatigue, and improving energy efficiency. The study concludes by envisioning the future of soft robotics, driven by innovations in smart materials, AI-assisted control systems, and 4D printing technologies, all of which promise to pave the way for more sophisticated, autonomous, and safer human-robot interactions.

Acknowledgements

The corresponding author dedicates this symposium paper to his mother Immihan GÜNEŞ.

7. References

[1] Zongxing, L., Wanxin, L., & Liping, Z. (2020). Research development of soft manipulator: A review. Advances in Mechanical Engineering, 12(8), 1687814020950094.

[2] AboZaid, Y. A., Aboelrayat, M. T., Fahim, I. S., & Radwan, A. G. (2024). Soft robotic grippers: A review on technologies, materials, and applications. *Sensors and Actuators A: Physical*, 115380.

[3] Samadikhoshkho, Z., Zareinia, K., & Janabi-Sharifi, F. (2019, May). A brief review on robotic grippers classifications. In *2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE)* (pp. 1-4). IEEE.

[4] Long, Z., Jiang, Q., Shuai, T., Wen, F., & Liang, C. (2020, March). A systematic review and meta-analysis of robotic gripper. In *IOP Conference Series: Materials Science and Engineering* (Vol. 782, No. 4, p. 042055). IOP Publishing.

[5] Blanco, K., Navas, E., Emmi, L., & Fernández, R. (2024). Manufacturing of 3D printed soft grippers: a review. IEEE Access.

[6] Zhang, H., Kumar, A. S., Fuh, J. Y., & Wang, M. Y. (2018, April). Topology optimized design, fabrication and evaluation of a multimaterial soft gripper. In *2018 IEEE International Conference on Soft Robotics (RoboSoft)* (pp. 424-430). IEEE.

[7] Liu, C. H., Chung, F. M., Chen, Y., Chiu, C. H., & Chen, T. L. (2020). Optimal design of a motor-driven three-finger soft robotic gripper. *IEEE/ASME Transactions On Mechatronics*, *25*(4), 1830-1840.

[8] Elfferich, J. F., Dodou, D., & Della Santina, C. (2022). Soft robotic grippers for crop handling or harvesting: A review. *Ieee Access*, *10*, 75428-75443.

[9] Hughes, J., Culha, U., Giardina, F., Guenther, F., Rosendo, A., & Iida, F. (2016). Soft manipulators and grippers: A review. *Frontiers in Robotics and AI*, *3*, 69.

[10] Crooks, W., Vukasin, G., O'Sullivan, M., Messner, W., Rogers, C. Fin ray® effect inspired soft robotic gripper: From the robosoft grand challenge toward Optimization, Frontiers in Robotics and AI, 3(70), 1-9.

[11] Götz, H., Santarossa, A., Sack, A., Pöschel, T., Müller P. Soft particles reinforce robotic grippers: robotic grippers based on granular jamming of soft particles, Granular Matter (2022) 24:31, https://doi.org/10.1007/s10035-021-01193-4.

[12] Hassan, T., Manti, M., Passetti, G., d'Elia, N., Cianchetti, M., & Laschi, C. (2015, August). Design and development of a bio-inspired, under-actuated soft gripper. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 3619-3622). IEEE.

[13] McMahan, W., Chitrakaran, V., Csencsits, M., Dawson, D. Walker, I.D., Jones, B.A., Pritts, M., Dienno, D., Grissom, M., Rahn, C.D. Field Trials and Testing of the OctArm Continuum Manipulator, Proceedings of the 2006 IEEE International Conference on Robotics and Automation, Orlando, Florida - May 2006, 2336-2341.

[14] Ai, W., Hou, K., Wu, J., Long, Y., Song, K. Miniaturized and untethered McKibben muscles based on photothermal-induced gas-liquid transformation, Nature Communications, (2024) 15:1329.

[15] Zhou, L., Ren, L., Chen, Y., Niu, S., Han, Z., & Ren, L. (2021). Bio-inspired soft grippers based on impactive gripping. *Advanced Science*, *8*(9), 2002017.

[16] Zhou, X., Majidi, C., & O'Reilly, O. M. (2015). Soft hands: An analysis of some gripping mechanisms in soft robot design. *International Journal of Solids and Structures*, *64*, 155-165.

[17] Giannaccini, M. E., Georgilas, I., Horsfield, I., Peiris, B. H. P. M., Lenz, A., Pipe, A. G., & Dogramadzi, S. (2014). A variable compliance, soft gripper. *Autonomous Robots*, *36*, 93-107.

[18] Al Abeach, L. A., Nefti-Meziani, S., & Davis, S. (2017). Design of a variable stiffness soft dexterous gripper. *Soft robotics*, *4*(3), 274-284.

[19] Park, W., Seo, S., & Bae, J. (2018). A hybrid gripper with soft material and rigid structures. *IEEE Robotics and Automation Letters*, *4*(1), 65-72.

[20] Subramaniam, V., Jain, S., Agarwal, J., & Valdivia y Alvarado, P. (2020). Design and characterization of a hybrid soft gripper with active palm pose control. *The International Journal of Robotics Research*, *39*(14), 1668-1685.

[21] Zhou, J., Chen, S., & Wang, Z. (2017). A soft-robotic gripper with enhanced object adaptation and grasping reliability. *IEEE Robotics and automation letters*, *2*(4), 2287-2293.

[22] Liu, C. H., Chen, T. L., Chiu, C. H., Hsu, M. C., Chen, Y., Pai, T. Y., ... & Chiang, Y. P. (2018). Optimal design of a soft robotic gripper for grasping unknown objects. *Soft robotics*, *5*(4), 452-465.

[23] Mannsa, M., Moralesa, J., Frohn P. Additive manufacturing of silicon based PneuNets as soft robotic actuators, Procedia CIRP 72 (2018) 328–333.
[24] Nguyen, A.K., Russell, A., Naclerio, N., Vuong, V., Huang, H., Chui, K., Hawkes, E.W. A tri-stable soft robotic finger capable of pinch and wrap grasps, Proc - IEEE Int Conf Robot Autom 2020;9028–9034.
[25] Ariyanto, M., Munadi, M., Setiawan, J.D., Mulyanto, D., Nugroho, T. Three-fingered soft robotic gripper based on pneumatic network actuator. 2019 6th Int Conf Inf Technol Comput Electr Eng ICITACEE 2019. https://doi.org/10.1109/ICITACEE.2019.
[26] Ye, Y., Cheng, P., Yan, B., Lu, Y., & Wu, C. (2022). Design of a novel soft pneumatic gripper with variable gripping size and mode. *Journal of Intelligent & Robotic Systems*, *106*(1), 5.
[27] Shi, H.; Tan, K.; Zhang, B.; Liu, W. Review on Research Progress of Hydraulic Powered Soft Actuators. Energies 2022, 15, 9048.
[28] Saputro, D.J., Yang, E. Investigation of Soft Robotic Hand Using Tendon-DrivenMechanism of Robotic Manipulator, Engineering Physics International Conference 2021 – EPIC 2021, AIP Conf. Proc. 2580, 040011-1–040011-10; https://doi.org/10.1063/5.0123384.
[29] Chen, F., Xu, W., Zhang, H., Wang, Y., Cao, J., Wang, M. Y., ... & Zhang, Y. F. (2018). Topology optimized design, fabrication, and characterization of a soft cable-driven gripper. *IEEE Robotics and Automation Letters*, *3*(3), 2463-2470.
[30] Liu M, Hao L, Zhang W, Zhao Z. A novel design of shape-memory alloy-based soft robotic gripper with variable stiffness. Int J Adv Robot Syst. 2020;17:1–12.
[31] Xu L, Gu G. Bioinspired Venus flytrap: a dielectric elastomer actuated soft gripper. 24th Int Conf Mechatronics Mach Vis Pract M2VIP, 2017-December:1–3.
[32] Wang, H., Zhu, Z., Jin, H., Wei, R., Bi, L., Zhang, W. Magnetic soft robots: Design, actuation, and function, Journal of Alloys and Compounds, Volume 922, 2022, 166219.
[33] Dezaki, M.L., Bodaghi, M. Shape memory meta-laminar jamming actuators fabricated by 4D printing, Soft Matter, 2023, 19, 2186–2203.
[34] Tsabedze, T., Zhang, J. Design, characterization, modeling, and comparison of helically wrapped super-coiled polymer artificial muscles, Sensors and Actuators A 331 (2021) 113018.
[35] Ranjbar, S., Lakhi, M., Bodaghi, M., Irani, M.S., Zolfagharian, A. Silicone elastomer soft robots via 4D printing, Smart Materials in Additive Manufacturing, Volume 3. DOI: https://doi.org/10.1016/B978-0-443-13673-3.00007-9.
[36] Wang, X., Kang, H. Soft robotic finger with variable effective length enabled by an antagonistic constraint mechanism, Smart Mater. Struct. 32 (2023) 055001.
[37] Scalet, G. Two-Way and Multiple-Way Shape Memory Polymers for Soft Robotics: An Overview, Actuators 2020, 9(1), 10; https://doi.org/10.3390/act9010010.
[38] McDonald, K.J., Kinnicutt, L., Moran, A.M., Ranzani, T. Modulation of Magnetorheological Fluid Flow in Soft Robots Using Electropermanent Magnets, in IEEE Robotics and Automation Letters, vol. 7, no. 2, pp. 3914-3921, April 2022, doi: 10.1109/LRA.2022.3147873.
[39] Ambaye, G., Boldsaikhan, E., Krishnan, K. Soft Robot Workspace Estimation via Finite Element Analysis and Machine Learning, Actuators 2025, 14(3), 110; https://doi.org/10.3390/act14030110.
[40] Hu, Q., Dong, E., & Sun, D. (2021). Soft gripper design based on the integration of flat dry adhesive, soft actuator, and microspine. *IEEE Transactions on Robotics*, *37*(4), 1065-1080.
[41] Navas, E., Fernández, R., Sepúlveda, D., Armada, M., & Gonzalez-de-Santos, P. (2021). Soft grippers for automatic crop harvesting: A review. *Sensors*, *21*(8), 2689.
[42] Liu, Y., Hou, J., Li, C., & Wang, X. (2023). Intelligent soft robotic grippers for agricultural and food product handling: A brief review with a focus on design and control. *Advanced Intelligent Systems*, *5*(12), 2300233.
[43] Wang, Z., Torigoe, Y., & Hirai, S. (2017). A prestressed soft gripper: design, modeling, fabrication, and tests for food handling. *IEEE Robotics and Automation Letters*, *2*(4), 1909-1916.
[44] Chen, K., Li, T., Yan, T., Xie, F., Feng, Q., Zhu, Q., & Zhao, C. (2022). A soft gripper design for apple harvesting with force feedback and fruit slip detection. *Agriculture*, *12*(11), 1802.
[45] Wang, Z., Or, K., & Hirai, S. (2020). A dual-mode soft gripper for food packaging. *Robotics and Autonomous Systems*, *125*, 103427.

[46] De FalcoIris, I., Cianchetti, M., Menciassi, A. STIFF-FLOP surgical manipulator: design and preliminary motion evaluation, Conference: 4th Joint WorkShop on Computer/Robot Assisted Surgery (CRAS), Genova, October 2014, DOI: 10.13140/RG.2.1.3972.9124.
[47] Rateni, G., Cianchetti, M., Ciuti, G., Menciassi, A., & Laschi, C. (2015). Design and development of a soft robotic gripper for manipulation in minimally invasive surgery: a proof of concept. *Meccanica*, *50*, 2855-2863.
[48] Tawk, C., Gillett, A., in het Panhuis, M., Spinks, G. M., & Alici, G. (2019). A 3D-printed omni-purpose soft gripper. *IEEE Transactions on Robotics*, *35*(5), 1268-1275.
[49] Trinh, H. X., Nguyen, H. H., Pham, T. D., & My, C. A. (2024). A novel rigid-soft gripper for safe and reliable object handling. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, *46*(4), 176.
[50] Zaidi, S., Maselli, M., Laschi, C., & Cianchetti, M. (2021). Actuation technologies for soft robot grippers and manipulators: A review. *Current Robotics Reports*, *2*(3), 355-369.
[51] Kim, D., Baek, S. R., Kim, M. S., Park, C. Y., & Lee, I. H. (2024). Design and manufacturing process of pneumatic soft gripper for additive manufacturing. *Sensors and Actuators A: Physical*, *370*, 115218.
[52] Ye, Y., Cheng, P., Yan, B., Lu, Y., & Wu, C. (2022). Design of a novel soft pneumatic gripper with variable gripping size and mode. *Journal of Intelligent & Robotic Systems*, *106*(1), 5.

A Hybrid Model Based on Image Processing for Violence Detection

Yunus Emre Seyyar[1]

[1]Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBITAK), Ankara, Türkiye.
[1]yunusemre.seyyar@gmail.com.tr
ORCID: 0000-0001-5767-8515

Abstract
Violence detection is crucial for maintaining security and social order in today's societies. Detecting and responding quickly to violent incidents, particularly in public spaces, plays a critical role in both individual safety and maintaining public order. Technological advancements have made it possible to perform automatic violence detection using artificial intelligence and deep learning methods. These methods can analyze video data to identify violent situations and quickly provide information to security forces and authorities. This study aims to develop an artificial intelligence model capable of violence detection. The study was conducted using LSTM (Long Short-Term Memory) and CNN models, which are effective in understanding long-term dependencies. By combining the strengths of CNN and LSTM models, a higher-performance violence detection model was created. Experimental results demonstrated that the proposed model detects violence with a test accuracy of 92.5%. The model's success in violence detection was demonstrated by the results obtained during the testing phase. The model's high accuracy demonstrates its effective ability to detect violence. The findings provide a significant contribution not only to the academic field but also to social security and preventive measures.

Keywords: violence detection, deep learning, lstm, cnn

1. Introduction

Detecting violent incidents is critical to maintaining security and social order in modern societies. Rapid and effective detection of violent incidents, especially those occurring in public areas, is essential for both individual safety and the preservation of public order. Artificial intelligence and deep learning methods are used to automatically detect violent incidents from video data. In addition to models such as CNN, RNN, and LSTM, hybrid models are also used in violence detection. Convolutional Neural Networks (CNN) are widely used deep learning models for feature extraction and classification in image data. CNN extracts features from input images with its layered structure and classifies these features. In violence detection, violent incidents are detected using features extracted from video frames. Long Short-Term Memory (LSTM) is a model belonging to the Recurrent Neural Network (RNN) family and was developed to learn long-term dependencies. LSTM makes future predictions by remembering past data thanks to its memory cells. In violence detection, LSTM models can be used to process video frames as a time series and detect violent incidents [1-6].

Negre et al. [7] presented a comprehensive literature review on deep learning-based violence detection methods. They reviewed violence detection studies using various deep learning models (CNN, RNN, LSTM) and evaluated the results of these studies. They discussed the challenges and future research directions in violence detection. For example, they reviewed studies on feature extraction from video frames and classification of these features using CNN models. Mumtaz et al. [8] provided an overview of violence detection techniques and discussed current challenges and future research directions. They provided examples of violence detection studies using image processing, machine learning, and deep learning methods. They emphasized the importance of violence detection for security and social order. Mumtaz et al. [9] used deep learning methods for violence detection in surveillance videos. They examined in detail how Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) models are used. They discussed violence detection studies conducted on data from surveillance videos. Zhou et al. [10] focused on detecting events and violent situations in videos using time-related reasoning. They presented detailed information on data processing in the time domain and discussed the effects of these methods on violence detection. They discussed deep learning techniques used to understand events in videos. Anomaly detection can be considered as a subset of violence detection. Suarez et al. [11] provided a comprehensive review of deep learning techniques for video anomaly detection. The article discussed various deep learning models and their effects on video anomaly detection. Rachna et al. [12] aimed to perform real-time violence detection using CNN and LSTM. The model architecture, dataset preparation, training process, and performance evaluations are explained in detail. They

[1]Corresponding author

discussed the advantages and challenges of the techniques used for real-time violence detection. Kayani et al. [13] aimed to perform real-time violence detection using deep learning methods. The system architecture is given in Figure 1. In the study, the deep learning models used, dataset preparation, training process, and performance evaluations are explained in detail. The advantages and challenges of techniques used for real-time violence detection are discussed. Examples of AI-based violence detection studies can be expanded [14-16]. A comparison of studies in the literature is provided in Table 1.

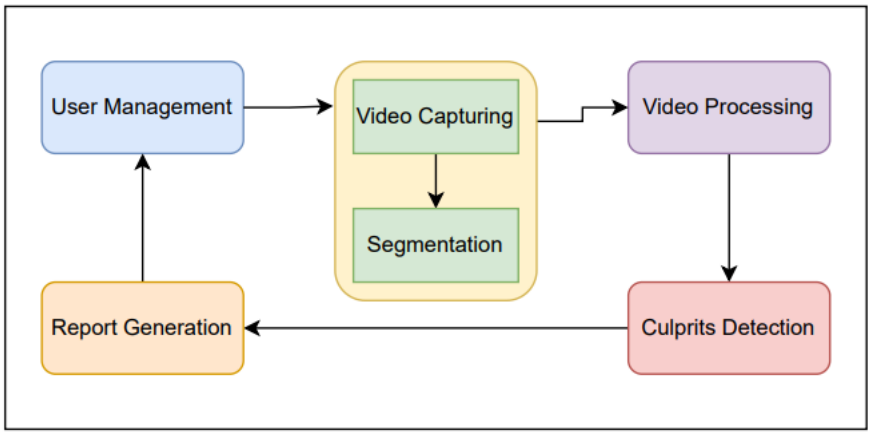


Figure 1. Violence detection method [13]

Table 1. Literature Comparison

Reference	Year of Publication	Methods Used
Literature Review of Deep-Learning-Based Detection of Violence in Video [7]	2024	CNN, RNN, LSTM
An Overview of Violence Detection Techniques: Current Challenges and Future Directions [8]	2022	CNN, RNN
Violence Detection in Surveillance Videos: A Deep Learning Approach [9]	2018	CNN, RNN, LSTM
Temporal Relational Reasoning in Videos [10]	2018	RNN
Survey on Deep Learning Techniques for Video Anomaly Detection [11]	2020	CNN
Real-Time Violence Detection Using Deep Neural Networks and DTW [12]	2023	CNN, RNN
Real-Time Violence Detection Using Deep Learning [13]	2022	RNN

This study employed data augmentation techniques not commonly used in literature. Various data augmentation methods were applied to increase the dataset size and improve the model's generalization ability. This allowed the model to perform better under different scenarios. Furthermore, while studies in the literature typically examine the performance of a single model, this study comparatively evaluated different deep learning models. Experiments using CNN, RNN, and LSTM models compared the effects of each model on violence detection, and the best-performing model was identified. The contributions of the AI-based violence detection method proposed in this study compared to other methods in the literature are listed below.

- Data augmentation techniques have increased the model's generalization ability.
- The hybrid CNN+LSTM model has achieved greater success in temporal correlation analysis.

- Optimization processes have been improved for real-time applications.

2. Proposed Approach

A block diagram containing the technical stages of the proposed method for developing an AI-based violence detection model is presented in Figure 2. The method developed within the scope of this study was optimized and tested to develop an AI model for violence detection. The model development process encompasses a series of stages, starting with dataset preprocessing, through model training, evaluation, and performance improvement.

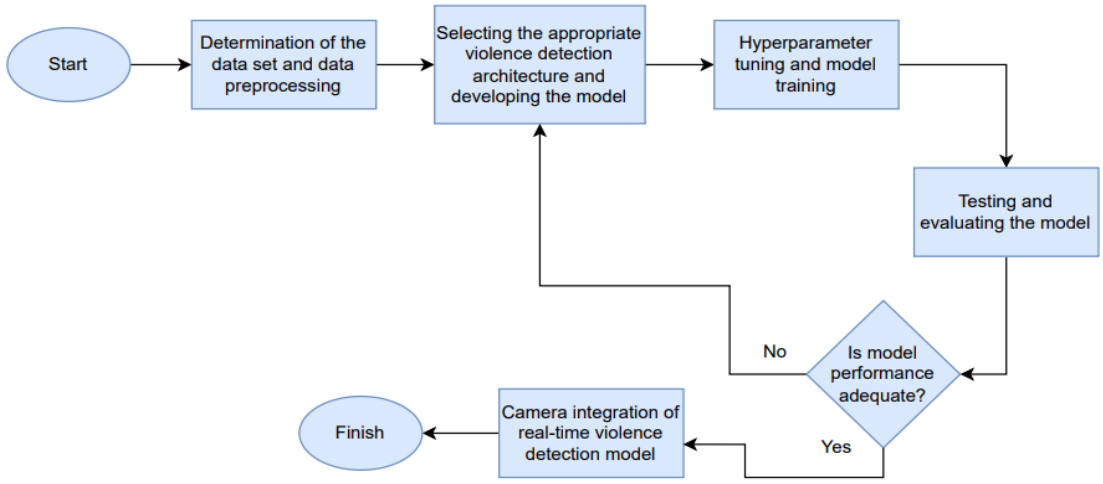


Figure 2. Block diagram of the proposed method

The diagram shows all the steps in detail, from dataset preparation to model testing and integration into a real-time system. If model performance is not sufficient throughout the process, the model development and optimization steps are repeated. Once performance expectations are met, the developed model is integrated into applied systems. The step-by-step steps of the proposed method are outlined below.

Data Collection and Preprocessing: Video data was obtained from sources such as Kaggle and processed in a specific format. The resulting videos were segmented into frames, and these frames were prepared for feature extraction.

- **Video Dataset Selection:** The Violence/NonViolence dataset was used, and the videos were converted to specific formats.
- **Frame Separation:** Videos were divided into frames at fixed time intervals, resulting in a specific number of images from each video.
- **Normalization:** The data was optimized for the deep learning model by normalizing the pixels between 0 and 1.

Feature Extraction: Each video frame was analyzed with CNN-based feature extraction layers and meaningful features were obtained.

- **CNN Usage:** A pre-trained CNN model (such as ResNet, VGG) was chosen to extract meaningful features from video frames.
- **Transfer Learning:** Feature extraction layers were optimized using pre-trained weights.

Time Relational Analysis: LSTM models have been used to understand the relationship of features over time and to detect violent events.

- **LSTM Model:** A specialized LSTM model was trained for time-series data. The model analyzed the behavior of violent events over time.
- **Hyperparameter Optimization:** Parameters such as batch size, epoch number, and learning rate were optimized to improve performance.

Classification and Decision Making: The extracted features are processed with a softmax classifier to make the final decision.

- **Softmax Layer:** LSTM outputs were separated into two classes: violent and non-violent.
- **Performance Measures:** Model performance was evaluated using accuracy, precision, recall, and F1 score.

LSTM models have a more complex structure than standard RNNs. An LSTM has three main gates: an input gate, a forget gate, and an output gate. These gates control the flow of information into the model and enable it to learn long-term dependencies.

- **Input Gate:** Controls how much new information the model will receive.
- **Forget Gate:** Controls how much past information the model will forget.
- **Output Gate:** Controls what information will be transferred to the model's output.

LSTM models are ideal for processing time-series video data. In violence detection, video frames are fed to the LSTM model as sequential data, and the model detects violent events by learning dependencies over time. For example, to detect a fight, dependencies between previous and subsequent frames must be learned. By learning these dependencies, the LSTM model can detect violent events with high accuracy.

Hybrid models are models created by combining different deep learning models, such as CNN and LSTM. These models combine the capabilities of both image feature extraction and time-series analysis. Hybrid models achieve higher accuracy and precision in violence detection. To achieve this, features are extracted from video frames using a CNN model, fed to the LSTM model to learn dependencies over time.

Datasets used for violence detection should be diversified to increase the model's generalization ability. The model's performance can be improved by using video data recorded in different scenarios. Data augmentation techniques can improve the model's generalization ability by increasing the size of the dataset. Examples of datasets frequently used for violence detection include the UCF Crime Dataset, the Hockey Fight Dataset, and the Violence in Movies Dataset.

Data Augmentation Techniques:

- **Rotation:** New images are added to the dataset by randomly rotating images within a specified angle range.
- **Flipping:** Images are flipped horizontally and/or vertically, creating new perspectives.
- **Shifting:** Images are shifted horizontally and/or vertically within a specified pixel range.
- **Scaling:** Images are resized using random scaling factors.
- **Cropping:** New sub-images are created by cutting off specific regions of the images.
- **Brightness Adjustment:** The brightness levels of images are randomly altered.
- **Adding Noise:** Variety is enhanced by adding random noise to images.
- **Channel Shifting:** Changes are made to color channels, such as RGB.
- **Shearing:** Images are cropped at a specified angle.

In this study, data augmentation techniques were used to improve model performance and increase the learning capacity of the model.

3. Experimental Results

This section presents detailed experimental findings from the application of the developed AI model for violence detection. The results are visually supported with graphs and tables, and model performance is compared across different scenarios. Furthermore, the computer and dataset specifications used are detailed, and a comparison is made with similar studies in the literature.

3.1 Summary and Structure of the Model

The developed model is based on a deep learning-based approach for violence detection in video data. The model's overall structure consists of two main components that aim to learn both spatial and temporal features:

- **Spatial Feature Extraction:** A VGG16-based neural network was used to extract visual features from video frames. This model was wrapped with a TimeDistributed layer to process each frame independently. Additionally, the spatial dimensions were reduced using a GlobalAveragePooling2D layer.
- **Temporal Information Processing:** LSTM (Long Short-Term Memory) layers were used to learn the temporal relationships of consecutive frames in the video. The first LSTM layer has 512 units, and the second LSTM layer has 256 units. These layers learn the correlations between consecutive frames and the long-term temporal relationships.
- **Classification Layers:** At the end of the model, fully connected (dense) layers were used to produce a two-class (violence/non-violence) output for violence detection. In this stage: 512 neurons were used in the first dense layer. A softmax activation function was used in the final layer to return the probabilities for the two classes.
- **Regularization and Regularization:** BatchNormalization layers were used to optimize the model's learning process and prevent overfitting. The dropout rate was set at 50% for each LSTM and dense layer. A summary of the generated deep learning model is provided in Figure 3..

Layer (type)	Output Shape	Param #
time_distributed (TimeDistributed)	(None, 5, 7, 7, 512)	14,714,688
time_distributed_1 (TimeDistributed)	(None, 5, 512)	0
lstm (LSTM)	(None, 5, 512)	2,099,200
batch_normalization (BatchNormalization)	(None, 5, 512)	2,048
dropout (Dropout)	(None, 5, 512)	0
lstm_1 (LSTM)	(None, 256)	787,456
batch_normalization_1 (BatchNormalization)	(None, 256)	1,024
dropout_1 (Dropout)	(None, 256)	0
dense (Dense)	(None, 512)	131,584
dropout_2 (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 2)	1,026

Total params: 17,737,026 (67.66 MB)
Trainable params: 3,020,802 (11.52 MB)
Non-trainable params: 14,716,224 (56.14 MB)

Figure 3. Model Summary

3.2. Application Environment

- **Computer Specifications:**
  - CPU: Intel Core i7-12700H
  - GPU: Intel IRIS Xe
  - Memory: 16 GB DDR4
  - Storage: 512 GB SSD
  - Operating System: Windows 11 Pro and Google Colab Pro
- **Software Tools:**
  - Python Version: 3.9
  - TensorFlow Version: 2.9
  - Keras: 2.9
  - Other Libraries: NumPy, Matplotlib, OpenCV, scikit-learn
- **Dataset:** The dataset used in this study consists of two classes: violent and nonviolent videos. Each class contains 1000 videos, and the total dataset consists of 2000 videos. The videos were diversified to represent various situations and environments, and each class was created with a balanced sample. The dataset characteristics can be summarized as follows:
- **Total Number of Videos:** 2000
  - Violent Videos (Violence): 1000
  - Nonviolent Videos (Nonviolence): 1000
- **Video Format:** MP4

- **Video Resolution:** 96x96 pixels
- **Video Length:** Maximum 6 seconds (180 frames)

The dataset was split into 70%, 20%, and 10% for the model's training, validation, and testing processes. Examples of images from the Violence class are shown in Figure 4, and examples of images from the NonViolence class are shown in Figure 5.

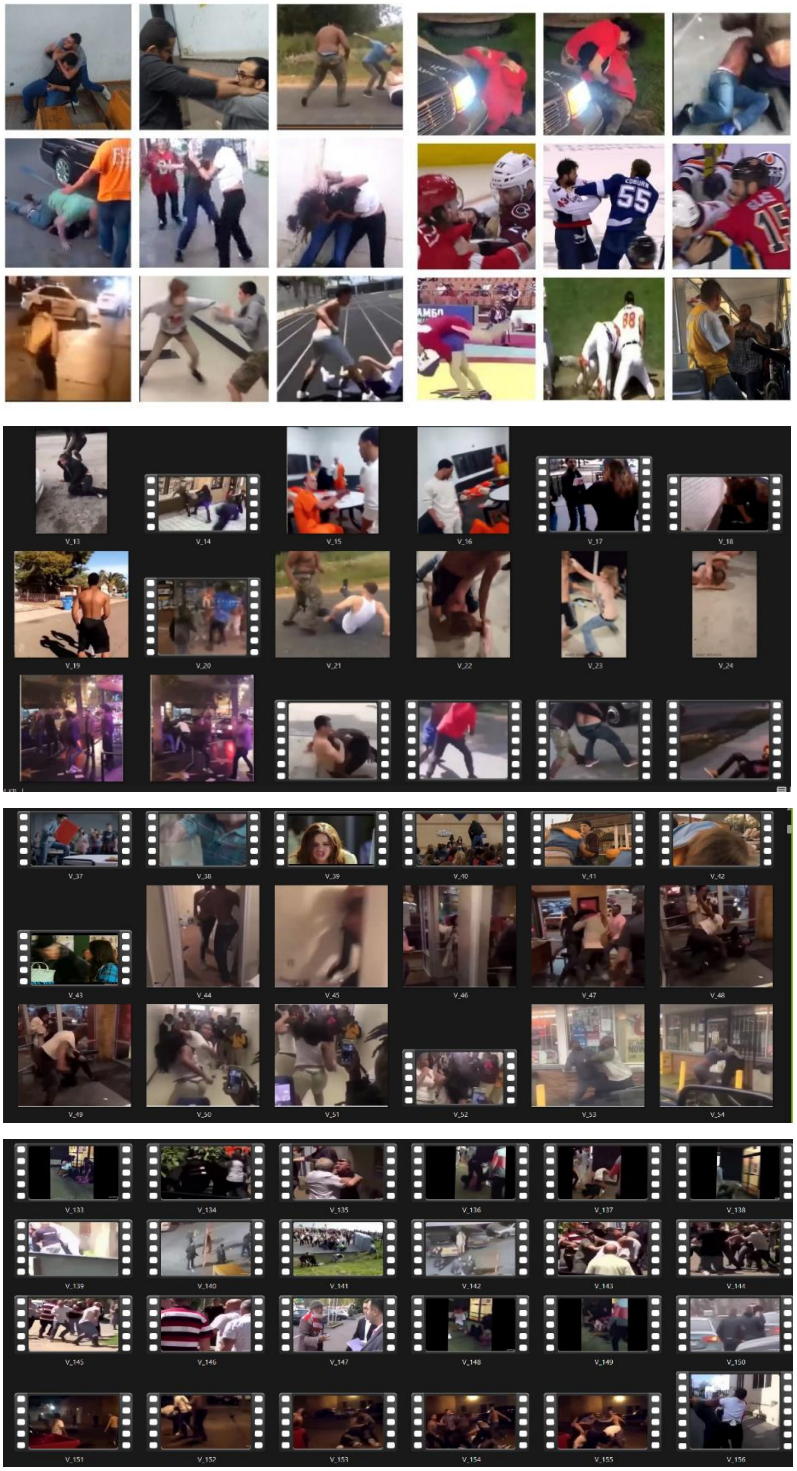


Figure 4. Sample images of the violence class

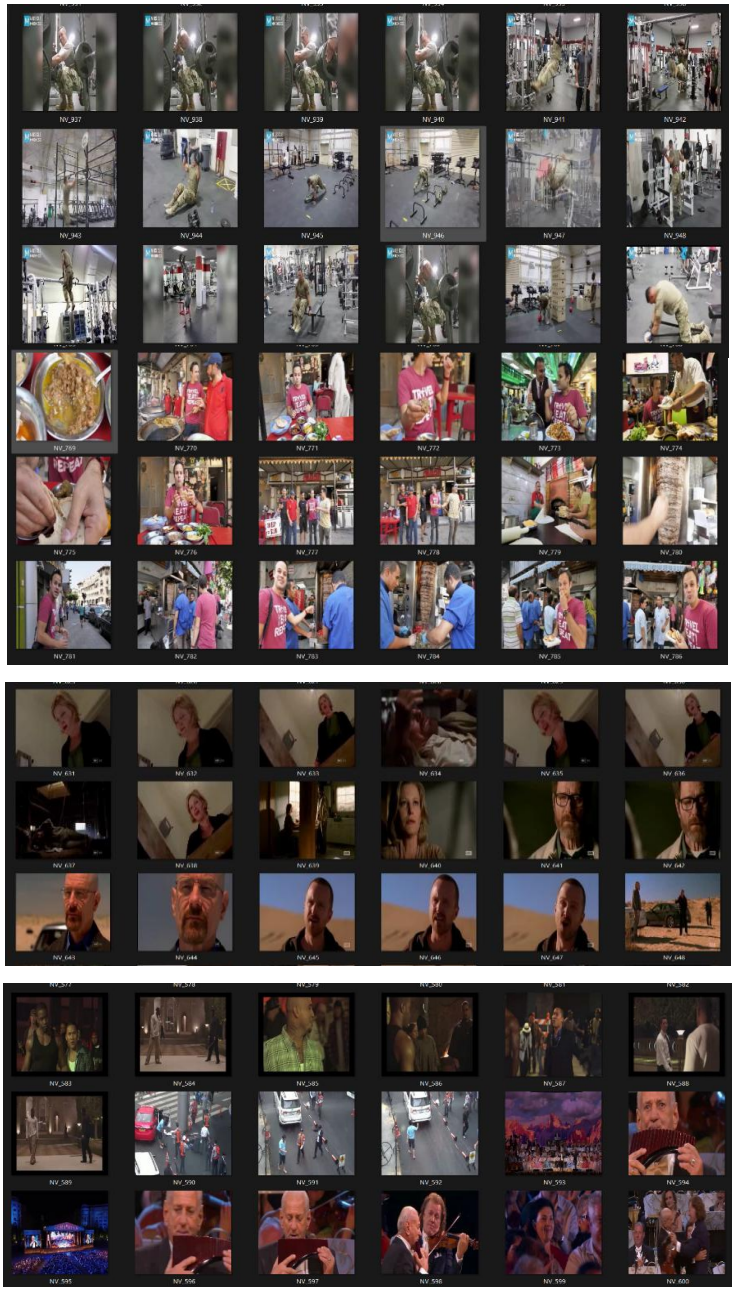


Figure 5. Sample images of the nonviolence class

3.3. Model Performance

The accuracy, precision, recall and F1 Score performance values obtained by the model in training, validation and test data sets are calculated according to Equation (1), Equation (2), Equation (3) and Equation (4) and the results obtained are given in Table 2.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

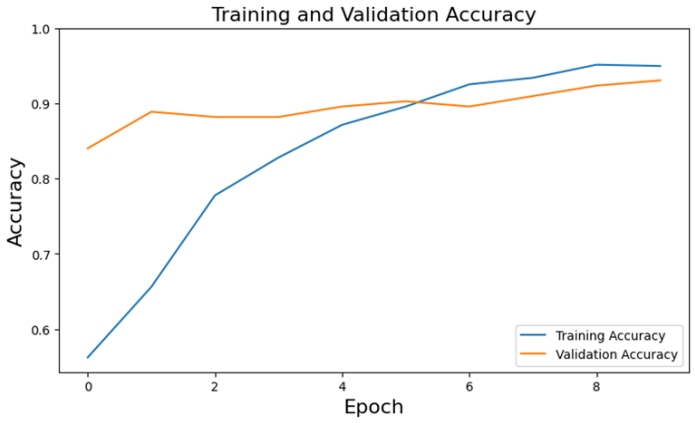
$$recall = \frac{TP}{TP+FN} \quad (3)$$

$$f1score = 2 \times \frac{recall \times precision}{recall + precision} \quad (4)$$

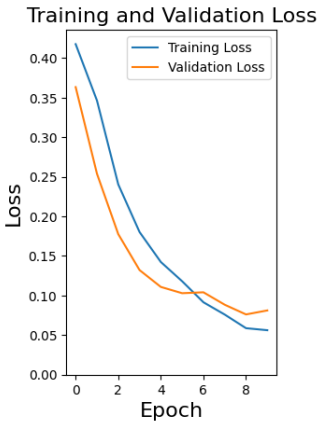
Table 2. Model evaluation results

Dataset	Accuracy (%)	Precision (%)	recall (%)	f1 score (%)
Train	98.2	97.6	98.9	98.2
Validation	94.5	92.7	95.4	94.0
Test	92.5	92.5	92.5	92.5

As seen in the graph in Figure 6(a), it was observed that the loss value of the model decreased and the validation loss stabilized during the training process. As seen in the graph in Figure 6(b), the training and validation accuracy increased. o The LSTM models used in the project provided high accuracy and sensitivity in detecting violent events in video data. Compared to CNN models commonly used in the literature, LSTM models learned time dependencies better and thus made more accurate predictions. By using hybrid models, the strengths of both CNN and LSTM models were combined, and higher performance was achieved. In particular, analyzing features extracted from CNN models with LSTM models provided a significant improvement in the detection of violent events. The data augmentation techniques used in the study increased the diversity of the dataset and improved the generalization ability of the model. Techniques such as rotation, translation, shift, and noise addition enabled the model to perform better under different scenarios. In addition, data augmentation techniques contributed to the creation of a more generalizable model by reducing the overfitting problem of the model.



(a) Training and Validation Accuracy Graph



(b) Training and Validation Loss Graph

Figure 6. Model training and loss graphs

The models used in this study were found to be suitable for real-time violence detection applications. By optimizing training time and computational costs, the models were enabled to process real-time video data. Real-time violence detection applications can play an important role in security cameras and surveillance systems used in public areas. This will enable rapid interventions in violent incidents. The developed model was also tested on different datasets. Table 3 presents the results obtained on these datasets. The developed method was compared with similar models in the literature. In this study, where CNN and LSTM models were used together, a significant advantage was achieved in time-relational analysis. Table 4 presents a comparison of the model proposed in this study with methods in the literature.

Table 3. Comparing model performance for different datasets

Dataset	Accuracy (%)	Precision (%)	recall (%)	f1 score (%)
UCF Crime Dataset	91.8	90.2	92.1	91.1
Hockey Fight Dataset	88.5	87.3	89.2	88.2
Custom Dataset	92.5	92.5	92.5	92.5

Table 4. Comparison of model evaluation with methods in the literatüre

Method	Accuracy (%)	Model Used
Real-Time Violence Detection Using CNN-LSTM [17]	92.3	CNN + LSTM (hybrid model)
An investigation of videos for abnormal behavior detection [18]	86.63	DenseNet-121
Proposed Method	92.5	CNN+LSTM

4. Conclusions

Artificial intelligence models that detect violence have great potential for maintaining security and social order. In the future, with the development of more advanced and sensitive models, violence detection applications will become more widespread. Furthermore, real-time violence detection applications will enable rapid interventions and provide significant support to security forces. This study has made significant contributions to the research on artificial intelligence models that detect violence. The use of LSTM models and hybrid models has been shown to have significant potential in detecting violent incidents. Data augmentation techniques have improved the generalization ability of the model, and high performance has been achieved under various scenarios. Models optimized for real-time violence detection applications can make a significant contribution to maintaining security and social order. One of the advantages of this study is that the model's training time is shorter than studies in the literature. Using hybrid models has reduced computational costs and created a more user-friendly system. This is a significant advantage for real-time violence detection applications. In conclusion, the work conducted within this study has laid an important foundation for the development of more advanced and sensitive models in the field of violence detection. Future research may further improve the performance of these models, enabling them to be used more widely in real-world applications.

5. References

[1] Ke, S. R., Thuc, H. L. U., Lee, Y. J., Hwang, J. N., Yoo, J. H., & Choi, K. H. (2013). A review on video-based human activity recognition. Computers, 2(2), 88-131.

[2] Ahmad, Z., Shahid Khan, A., Nisar, K., Haider, I., Hassan, R., Haque, M. R., ... & Rodrigues, J. J. (2021). Anomaly detection using deep neural network for IoT architecture. Applied Sciences, 11(15), 7050.

[3] Ramzan, M., Abid, A., Khan, H. U., Awan, S. M., Ismail, A., Ahmed, M., ... & Mahmood, A. (2019). A review on state-of-the-art violence detection techniques. IEEE Access, 7, 107560-107575.

[4] Jebur, S. A., Hussein, K. A., Hoomod, H. K., & Alzubaidi, L. (2023). Novel deep feature fusion framework for multi-scenario violence detection. Computers, 12(9), 175.

[5] Prabha, S. C., Kaviyadharshini, S., & Micheal, A. A. (2025, February). A Hybrid Deep Learning Model for Violence Detection. In 2025 3rd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT) (pp. 1604-1612). IEEE.

[6] Graves, A. (2012). Long short-term memory. Supervised sequence labelling with recurrent neural networks, 37-45.

[7] Negre, P., Alonso, R. S., González-Briones, A., Prieto, J., & Rodríguez-González, S. (2024). Literature Review of Deep-Learning-based detection of violence in video. Sensors, 24(12), 4016.

[8] Mumtaz, N., Ejaz, N., Habib, S., Mohsin, S. M., Tiwari, P., Band, S. S., & Kumar, N. (2023). An overview of violence detection techniques: current challenges and future directions. Artificial intelligence review, 56(5), 4641-4666.

[9] Mumtaz, A., Sargano, A. B., & Habib, Z. (2018, December). Violence detection in surveillance videos with deep network using transfer learning. In 2018 2nd European conference on electrical engineering and computer science (EECS) (pp. 558-563). IEEE.
[10] Zhou, B., Andonian, A., Oliva, A., & Torralba, A. (2018). Temporal relational reasoning in videos. In Proceedings of the European conference on computer vision (ECCV) (pp. 803-818).
[11] Suarez, J. J. P., & Naval Jr, P. C. (2020). A survey on deep learning techniques for video anomaly detection. arXiv preprint arXiv:2009.14146.
[12] Rachna, U., Guruprasad, V., Shindhe, S. D., & Omkar, S. N. (2022, November). Real-time violence detection using deep neural networks and DTW. In international conference on computer vision and image processing (pp. 316-327). Cham: Springer Nature Switzerland.
[13] Kayani, T. (2022, December). Real-time violence detection using deep learning techniques. In 2022 3rd International Conference on Innovations in Computer Science & Software Engineering (ICONICS) (pp. 1-8). IEEE.
[14] Bermejo Nievas, E., Deniz Suarez, O., Bueno García, G., & Sukthankar, R. (2011, August). Violence detection in video using computer vision techniques. In International conference on Computer analysis of images and patterns (pp. 332-339). Berlin, Heidelberg: Springer Berlin Heidelberg.
[15] Hsairi, L., Alosaimi, S. M., & Alharaz, G. A. (2024). Violence Detection Using Deep Learning. Arabian Journal for Science and Engineering, 1-11.
[16] Khan, H., Yuan, X., Qingge, L., & Roy, K. (2025). Violence detection from industrial surveillance videos using deep learning. IEEE Access.
[17] Patwal, A. Diwakar, M. Tripathi, V. Singh, P. (2023). An investigation of videos for abnormal behavior detection, Procedia Computer Science, vol. 218, pp. 2264-2272.

A New Record of the Masarine Wasp *Celonites rugiceps* (Hymenoptera: Vespidae, Masarinae) from Bingöl Province, Turkey

Ceyhan Sönmez[1], İnanç Özgen[2]

[1] Fırat University, Department of Bioengineering, Elazığ, Turkey
[1]csonmez@firat.edu.tr, [2]inancozgen@gmail.com
[1](ORCID: 0000-0001-5894-2445), [2](ORCID: 0000-0003-1742-9324)

Abstract
In this study, *Celonites rugiceps* (Hymenoptera: Vespidae: Masarinae) is reported for the first time from Bingöl Province, Turkey. The specimens were collected on August 18, 2025, in Gökçelik Village, Şenlik settlement, from *Heliotropium circinnatum* (Boraginaceae). A total of five individuals were obtained. This record expands the known distribution of the species within Turkey and contributes to a more comprehensive understanding of the distribution of the Masarinae subfamily in the country.

Keywords: celonites rugiceps, masarinae, bingöl, heliotropium circinnatum, Turkish fauna

1.Introduction

The Masarinae subfamily is the only group within the Vespidae family that feeds on pollen and nectar, thus distinguishing it from other Vespidae members, which exhibit a predatory or parasitoid lifestyle (Gess, 1996; Krenn et al., 2005). Masarines are a unique group of wasps displaying "bee-like" behaviors. Represented by approximately 300 species worldwide, this group primarily occurs in the arid and semi-arid regions of Africa, the Mediterranean, and Western Asia (Blüthgen, 1961; Gusenleitner, 1990).

Within the Turkish fauna, records of Masarinae members are relatively limited. The genus Celonites is one of the most significant representatives of this subfamily in Turkey (Özbek & Katılmış, 2009). The ecological role of this genus involves its function as a pollen collector in plant–insect interactions and its contribution to pollination in steppe regions. *C. rugiceps* (Figure:1) has been reported in Southern Europe, North Africa, and Western Asia (Blüthgen, 1961), with records in Turkey predominantly confined to Central and Eastern Anatolia. This study presents the first record from Bingöl, thereby providing a new contribution to the known distribution of the species in Turkey.



Figure 1. The habitus of *Celonites rugiceps* (Hymenoptera: Vespidae: Masarinae)

2.Material and Methods

The specimens were collected on August 18, 2025, using an aspirator from *Heliotropium circinnatum* (Boraginaceae), located in the village of Gökçelik, Şenlik settlement, Bingöl Province (Figure 2). A total of five individuals were obtained. Species identification was made with the help of the existing literature (Blüthgen, 1961; Gusenleitner, 1990).

[1]Corresponding author

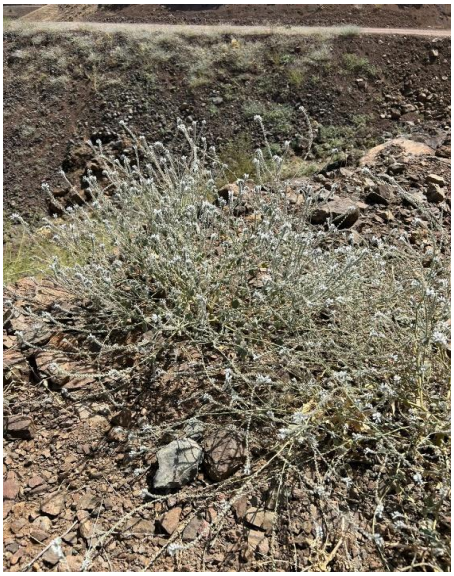


Figure 2. *Heliotropium circinnatum* (Family: Boraginaceae)

Findings
Species: *Celonites rugiceps*
Family: Vespidae
Subfamily: Masarinae
Host plant: *Heliotropium circinnatum* (Boraginaceae)
Collection location: Gökçelik Village, Şenlik Settlement, Bingöl Province
Date: August 18, 2025
Number of individuals collected: 5
This finding adds Bingöl Province to the known distribution of *C. rugiceps* in Turkey. Additionally, *Heliotropium circinnatum* is reported as the host plant of the species for the first time.

3.Discussion and Conclusion

Masarines occupy a unique position in ecosystems due to their non-predatory feeding habits (Krenn et al., 2005). As pollen collectors, they have developed close relationships with herbaceous plants, particularly in arid and steppe regions (Gess & Rozen, 1993). Records of Masarinae in Turkey are limited, which has resulted in insufficient ecological and biogeographical data about the group (Özbek & Katılmış, 2009).

The new record from Bingöl presented in this study reveals that the distribution range of *C. rugiceps* is broader than previously thought. Additionally, its presence on *Heliotropium circinnatum* provides new insight into the species' plant preferences. Similarly, Gess (1996) noted that masarines form close associations with certain plant families, particularly *Boraginaceae* and *Lamiaceae*. This finding is a significant contribution to understanding the ecological niche of *C. rugiceps*.

In conclusion, this new record from Bingöl is an important addition to the Turkish fauna. The identification of the species' host plants also provides valuable information for the study of pollination ecology. Future molecular studies and eco-faunistic research will allow for a more detailed understanding of the biogeographical history of *Celonites rugiceps* and its relationships with the Turkish flora.

Acknowledgements

We would like to express our sincere thanks to Assoc. Prof. Dr. Maruf Balos (T.C. Şanlıurfa Provincial Directorate of National Education) for the plant identification and to Milo van Loon (Naturalis Biodiversity Center, Leiden, Netherlands) for the species identification, for their valuable contributions.

4. References

[1] Gess, F. W. (1996). The pollen wasps: ecology and natural history of the Masarinae. Harvard University Press.
[2] Krenn, H. W., Mauss, V., & Plant, J. (2005). Evolution of the proboscis in pollen wasps (Masarinae). Biological Journal of the Linnean Society, 85: 43–52.
[3] Blüthgen, P. (1961). Die Faltenwespen Mitteleuropas (Hymenoptera, Diploptera). Abhandlungen der Deutschen Akademie der Wissenschaften zu Berlin.
[4] Gusenleitner, J. (1990). Studien über Faltenwespen der Unterfamilie Masarinae (Hymenoptera, Vespidae). Linzer Biologische Beiträge, 22(1): 139–174.
[5] Özbek, H., & Katılmış, Y. (2009). A contribution to the knowledge of the Masarinae (Hymenoptera: Vespidae) fauna of Turkey. Entomofauna, 30: 453–464.
[6] Gess, F. W., & Rozen, J. G. (1993). Biology of Masarinae (Hymenoptera: Vespidae). Journal of Hymenoptera Research, 2: 145–146.

The Role of Renewable Energy in the Global Energy Transition and Türkiye's Strategic Position

Ayca AYTAC [1]

[1] Department of Civil Engineering, Vocational School, Dogus University, Istanbul, Türkiye
[1]aaytac@dogus.edu.tr
[1](ORCID: 0000-0002-2108-6363)

Abstract
The energy sector, as a core component of economic development, is undergoing a major transformation shaped by climate change, energy security, and sustainability concerns. This study provides a multidimensional analysis of the global energy transformation and Türkiye's strategic role within this shift. The increasing share of renewable energy in global energy production is examined alongside its implications for electricity generation and consumption. The research further evaluates Türkiye's energy demand projections, installed capacity structure, and national energy policies. Findings emphasize the significance of this transformation for environmental and economic sustainability.

Keywords: renewable energy, energy transition, Türkiye's energy strategy, electricity market, installed capacity, climate change

1. Introduction

In recent years, the global energy sector has undergone a profound transformation driven by efforts to combat climate change, reduce carbon emissions, and ensure energy security. Rising population, economic growth, and industrialization worldwide continue to sustain energy demand, while the need to meet this demand through sustainable, environmentally friendly, and cost-effective sources is becoming increasingly critical. In this context, the share of renewable energy resources in electricity generation has risen rapidly, strengthening the global shift away from fossil fuels [1].

Türkiye holds a strategic position in this global transition. Over the past two decades, the country has increased its installed power generation capacity by approximately 2.5 times, with renewable resources now accounting for 46% of total capacity [2]. Furthermore, in line with the Türkiye National Energy Plan (2022), significant growth is targeted in solar- and wind-based electricity generation by 2035. This paper examines the drivers, development, and implications of the transformation in global and national energy markets and analyzes Türkiye's energy supply-demand structure, electricity generation profile, and renewable energy strategies. A descriptive analytical method is employed, drawing on structural trends in the energy sector, policy documents, and market data. The primary data sources include:

- International Energy Agency (IEA) 2024 World Energy Report,
- Turkish Electricity Transmission Corporation (TEIAS) 2024 statistics,
- Türkiye National Energy Plan (2022),
- Turkish Statistical Institute (TUIK) energy and population projections,
- Sectoral analysis reports from IRENA and the World Bank.

In addition, global electricity generation and consumption trends are examined in comparison with Türkiye's current figures and projections. The data are evaluated in terms of Compound Annual Growth Rate (CAGR), production and consumption distributions, and capacity increases by energy source.

2. Results

2.1 Trends in Global Energy Transition

Over the past two decades, global energy production and consumption have undergone a major structural transformation. According to data from the International Energy Agency (IEA), global electricity generation

[1]Corresponding author

increased by 2.5% in 2023, while the annual growth in installed capacity reached 7.6%. The primary driver of this expansion has been the rapid growth of renewable energy investments. In 2023, a significant portion of the 668 GW increase in global installed capacity was derived from sources such as wind and solar energy [1] (Figure 1).

Furthermore, under the IEA's Stated Policies Scenario, approximately 30% of global electricity generation was supplied by renewable sources as of 2023, and this share is projected to reach 73% by 2050 [1] (Figure 2).

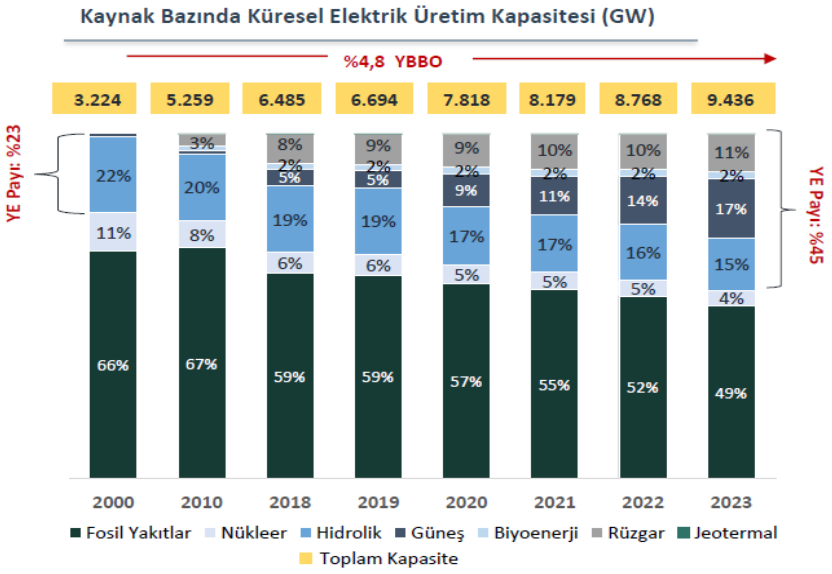


Figure 1. Global installed electricity generation capacity by source [1]

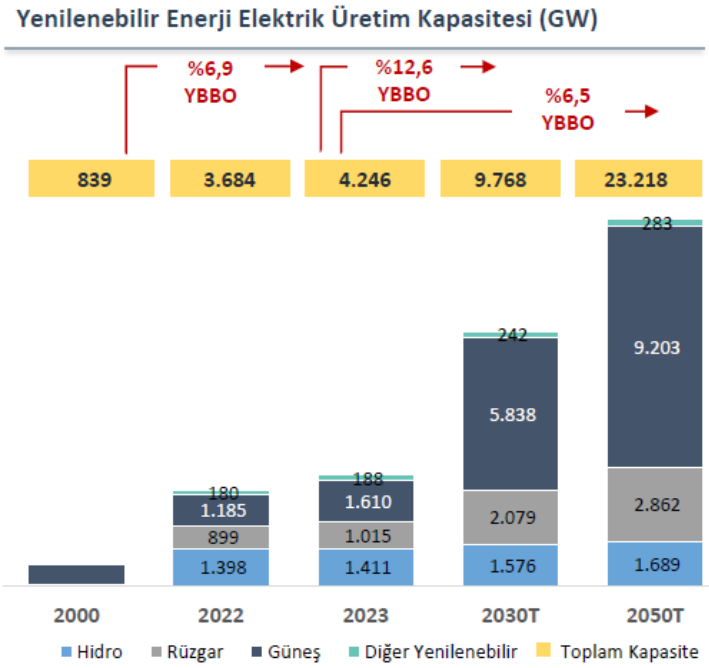


Figure 2. Global electricity generation capacity shares by source [1]

This figure illustrates that the total share of fossil fuels declined from 66% in 2000 to 49% in 2023, while renewable sources—particularly after 2010—showed a substantial increase.

As of 2023, total global energy investments reached USD 3.1 trillion, with approximately USD 2 trillion allocated to clean energy. This demonstrates a significant shift in favor of renewable energy technologies compared to fossil fuel resources [1].

2.2 Türkiye's Installed Power Generation Profile

As of 2024, Turkey's installed power generation capacity has reached 117.9 GW. An examination of its distribution by source shows that approximately 28% is derived from hydropower, 11% from wind, and 7% from solar energy [2] (Figure 3).

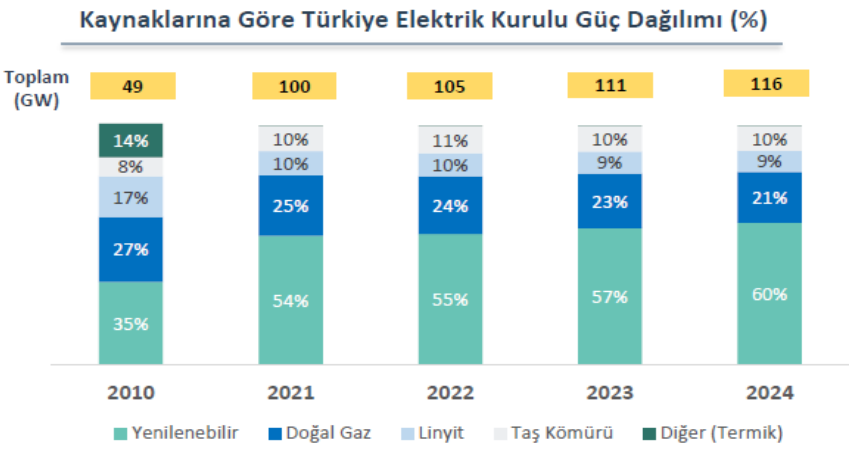


Figure 3. Distribution of Tükiye's installed electricity generation capacity [2].

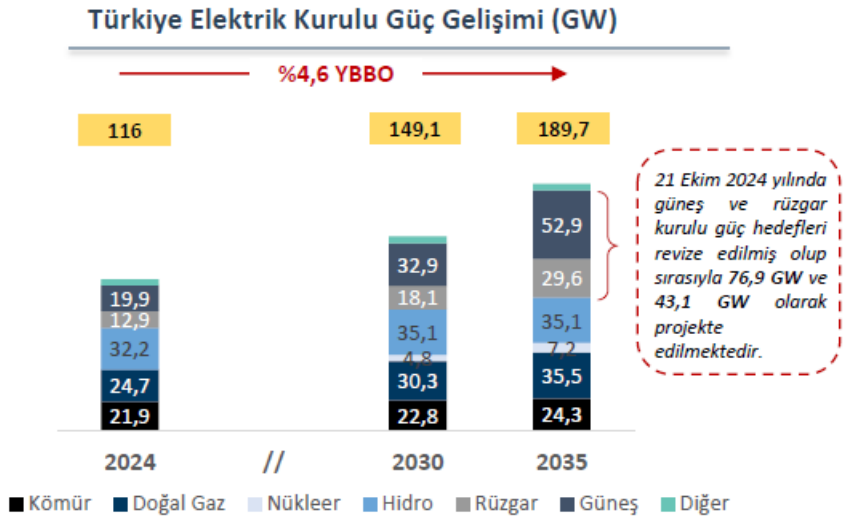


Figure 4. Development of Türkiye's installed electricity generation capacity [2].

The pie chart highlights that renewable resources now constitute a distinct 46% share of Turkey's installed capacity, with hydropower, solar, and wind playing a particularly critical role. Between 2000 and 2024, the Compound Annual Growth Rate (CAGR) of installed capacity was recorded at 4.6%. According to national targets, this capacity is planned to increase to 190 GW by 2035 [3] (Figure 4).

2.3 Electricity Generation Structure and Transformation Process

As of 2024, Türkiye's electricity generation stands at 349 TWh. The share of renewable energy sources in total electricity generation increased from 26% in 2010 to 46% in 2024 [2]. Figure 3 highlights a notable decline in dependence on natural gas and imported coal after 2010, alongside a growing share of domestic renewable sources in total generation. Similarly, an analysis of electricity generation license types shows that 82% of electricity generation is carried out by the private sector, indicating the strength of privatization and liberalization trends in Türkiye's energy market [4]

2.4 Electricity Consumption and Demand Projections in Türkiye

Between 2000 and 2024, Türkiye's electricity consumption grew at an average annual rate of 5.4%. In 2024, per capita electricity consumption reached 4,061 kWh, and it is projected to rise to 5,676 kWh by 2035 [5] (Figure 5).

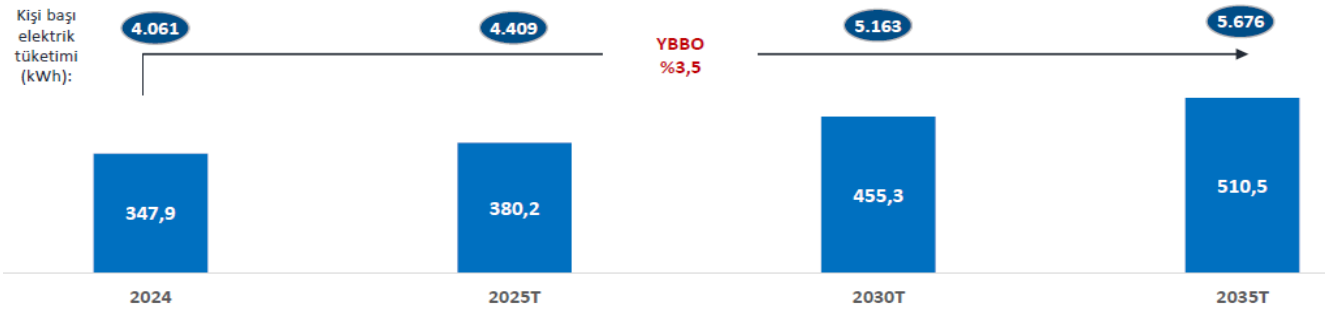


Figure 5. Electricity demand projection according to the Türkiye national energy plan [2], [4]

These data demonstrate that Türkiye's growing population, industrial output, and urbanization will lead to a significant increase in electricity demand. Therefore, implementing sustainable solutions on the supply side is of critical importance.

3. Discussion and Conclusions

The global energy transition affects not only supply–demand dynamics but also national security, economic growth, social development, and environmental sustainability in a multidimensional manner. Investments in renewable energy reduce energy dependency while minimizing the impact of fossil fuel price volatility. In this sense, energy investments have evolved beyond being merely environmentally friendly to become a cornerstone of economic stability [6].

In the case of Türkiye, this transformation has become both an opportunity—due to its geopolitical position—and a necessity, given its high energy import dependency. In 2023, Türkiye's energy imports exceeded USD 40 billion, representing a significant component of the current account deficit [5]. Therefore, increasing the share of domestic and renewable sources in electricity generation is not only an environmental imperative but also an economic one.

The potential of renewable resources, particularly solar and wind, is notably high. Türkiye enjoys an average of 2.740 hours of annual sunshine, which exceeds the European average. Likewise, the Aegean and Marmara regions offer high efficiency for wind energy investments [3]. However, fully leveraging these resources requires several structural reforms:

- Strengthening grid infrastructure: Current transmission and distribution lines lack sufficient capacity and flexibility to accommodate distributed renewable generation.
- Simplifying permitting procedures: Bureaucratic hurdles, particularly for solar and biomass projects, can discourage investment.

- Expanding energy storage technologies: Due to the intermittent nature of renewable sources, widespread deployment of storage systems has become essential.

Furthermore, the liberalization of Türkiye's energy market began with the enactment of the Electricity Market Law in 2001, evolving into a more competitive structure during the 2020s. As of 2024, 82% of electricity generation in Türkiye is carried out by the private sector [4]. This share is higher than that of many European countries and serves as a significant indicator of market development.

However, despite the rising role of the private sector in generation, there is still room for improvement in areas such as price stability, transparent market structures, and investor confidence. In particular, lowering the eligible consumer threshold and auditing distribution companies based on performance criteria are among the measures that could enhance competition in the market [7].

This study demonstrates that the energy sector will remain at the core of both economic and environmental policies in the coming years. Given its high renewable energy potential and growing energy demand, Türkiye must utilize this potential effectively and in a well-planned manner.

Global trends indicate that the energy sector will undergo profound transformations in decarbonization, distributed generation, digitalization, smart grids, and energy storage. For Turkey to keep pace with this transition, the following strategic measures are essential:

- Post-2024 energy planning should focus not only on increasing installed capacity but also on enhancing energy efficiency.
- Sectoral incentives should extend beyond production to include R&D, storage technologies, grid digitalization, and demand-side management.
- The domestic content ratio in renewable energy investments should be increased, and local equipment manufacturers should be supported.
- Transparency in capacity allocation should be ensured, and a long-term, secure investment environment should be established for investors.

In conclusion, building a sustainable energy future for Turkey depends not only on harnessing its natural resources but also on implementing integrated, scientifically grounded, and participatory energy policies.

4. References

[1] International Energy Agency (IEA). "World Energy Outlook 2024 and Net Zero Emissions by 2050 Scenario". https://www.iea.org (Access date: 02.09.2025).
[2] Turkish Electricity Transmission Corporation (TEIAS). "Electricity Statistics Report". https://www.teias.gov.tr (Access date: 02.09.2025).
[3] Ministry of Energy and Natural Resources of Turkey (ETKB). "Türkiye Ulusal Enerji Planı (2022–2035)". Ankara, Türkiye, 2022.
[4] Energy Market Regulatory Authority (EPDK). "Electricity Market Development Report". https://www.epdk.gov.tr (Access date: 02.09.2025).
[5] Turkish Statistical Institute (TUIK). "Energy and Consumption Statistics". https://data.tuik.gov.tr (Access date: 02.09.2025).
[6] International Renewable Energy Agency (IRENA). "World Energy Transitions Outlook: 1.5°C Pathway". https://www.irena.org (Access date: 02.09.2025).
[7] World Bank. "Unlocking Renewable Energy Investment in Emerging Markets". https://www.worldbank.org (Access date: 02.09.2025).

From Debris to Production: The Effect of Aggregates Obtained from Earthquake Debris on Concrete Performance

Elifnur AĞAOĞLU[1], Esra TUĞRUL TUNÇ[2]

[1, 2]Civil Engineering Department, Engineering Faculty, Fırat University, Elazığ, Turkiye.
[1]elifnuragaogluu@gmail.com [2]esratugrul@firat.edu.tr
[1](ORCID: 0009-0008-2317-3613), [2](0000-0001-9071-774X)

Abstract
Due to its geological location, Turkey has witnessed destructive earthquakes throughout history. In particular, the earthquakes centered in Kahramanmaraş on February 6, 2023 resulted in the generation of a substantial amount of construction debris. The recycling of this debris is of critical importance in terms of environmental sustainability and the reduction of economic losses. In this study, the usability of rubble aggregates obtained from earthquake debris in the production of structural concrete, in compliance with relevant standards, is investigated through an experimental approach. Concrete cube specimens with dimensions of 15×15×15 cm were produced using these aggregates at different water-to-cement ratios. Compressive strength and splitting tensile strength tests were conducted on the specimens, and all experiments were carried out at FÜYAMLAB in accordance with international standards. Furthermore, different mix proportions and mineral admixtures were evaluated to determine the optimum strength values. Throughout the process, aggregate selection, crushing, separation, curing, and experimental testing stages were systematically implemented. As a result, the recycling of earthquake debris for use in concrete production contributes to the conservation of natural resources, the mitigation of waste management problems, and the achievement of economic benefits.

Keywords: aggregate, recycling, strength, earthquake

1. Introduction

The earthquakes that struck our country on February 6, 2023 not only caused massive destruction but also led to severe economic losses. According to World Bank reports, the financial cost of these earthquakes is estimated at approximately 34.2 billion USD, corresponding to about 4% of the national gross domestic product [1]. In order to minimize such a substantial economic burden, it is of vital importance to recycle the debris generated from demolished or to-be-demolished structures. Data from the Turkish Statistical Institute (TÜİK) support this necessity; construction costs have increased by 106.87% annually and 3.47% monthly [2]. These cost increases clearly demonstrate that recovery practices are no longer an option but have become a necessity. Therefore, it is essential to enhance support for such researchs. The primary aim is to economically and effectively recycle construction rubble to ensure sustainability. The notion that industrial waste with recyclable potential should be reprocessed through various methods and reintroduced into the economy as raw materials forms the basis of this study [3,4].
Recycling practices, when properly managed, provide undeniable economic and environmental benefits. Today, the importance of recycling is increasing, and numerous studies are being conducted in this field. Particularly in the aftermath of the Kahramanmaraş-centered earthquakes in Türkiye on February 6, 2023—described as the “Disaster of the Century”—many structures in the affected provinces collapsed, and the demolition of severely damaged buildings is still ongoing. Utilizing the debris generated by this disaster as recycled aggregates in concrete production aims both to achieve structural concrete in compliance with relevant standards and to obtain significant economic and environmental benefits within the scope of recycling.

The difficulty of storing debris from collapsed structures in nature for long periods has made the recycling of earthquake debris for concrete production increasingly important. In this study, the usability of such debris as recycled aggregates in concrete production is examined in terms of structural and economic benefits in accordance with the standards. Although the literature contains significant research on the effects of recycled aggregates on structural concrete, this study specifically aims to investigate the suitability and performance of aggregates obtained from earthquake-induced debris for structural concrete production. The distinguishing feature of this research is its experimental evaluation of the usability of construction waste obtained from buildings demolished after the

[1]Corresponding author

Kahramanmaraş earthquakes of February 6 in the production of structural concrete in compliance with current standards.

Within the scope of this study, aggregates suitable for concrete production were selected from debris collected from different cities and designated sites in the earthquake-affected region. Using these aggregates, 15×15×15 cm concrete cube specimens were produced to observe their performance in structural concrete production. Compressive strength and splitting tensile strength tests were conducted on the specimens at Fırat University Construction Laboratory (FÜYAMLAB) in accordance with international standards. The primary objective of this study is to develop new concrete mixtures with different water-to-cement ratios and rubble aggregates (recycled concrete aggregates) without resorting to costly chemical admixtures, to identify optimum ratios that achieve strength performance comparable to conventional concrete, and thereby to contribute to recycling in the most efficient manner. For this purpose, Portland cement, sulfate-resistant cement, composite cement, and pozzolanic cement were employed. Previous studies have shown that increasing admixture ratios does not always linearly enhance the strength in every mixture [5,6].

The ultimate goal of this research is to achieve structural concrete strength with maximum efficiency. The primary reasons for selecting rubble aggregates for recycling are that concrete is not only advantageous in terms of workability but also the most widely used and economically feasible material in the construction industry. For example, concrete production requires approximately 20 times less energy compared to steel production [7,8]. Another advantage of concrete is that its production process requires less expertise and fewer skilled workers. In order for a concrete to achieve the required strength in compliance with standards, certain critical factors must be considered. This study aims to maximize economic, environmental, and structural benefits by taking these factors into account. Achieving these objectives is expected to make a significant contribution to the reconstruction processes in earthquake-affected regions

2. Materials and Method

This study examines the potential of producing structural concrete with rubble aggregates obtained from earthquake debris in five main stages.

In the first stage of the study, concrete rubble was collected from earthquake-affected regions. For this purpose, appropriate rubble procurement sites were identified in Elazığ, Malatya, and surrounding areas, where sampling and material collection were carried out on site.

In the second stage, the collected rubble samples were processed into finer particles using a Los Angeles abrasion testing machine. At the industrial scale, recycling processes generally include loading with a bucket, crushing (by jaw crushers), magnetic separation (for Fe particles), and sieving (e.g., 130 mm and 50 mm particle sizes). The separated gravel, sand, and other waste materials in these facilities are used in various fields such as road construction, drainage, and filling [9]. However, since this study was conducted at a laboratory scale, finer particles were obtained through the abrasion machine.

In the third stage, concrete cube specimens were prepared using the processed rubble aggregates with different cement ratios. Although the specimen dimensions and shapes used for compressive strength determination vary among countries, cube and cylindrical specimens are the most common. According to Turkish standards, cylindrical specimens are produced in 15×30 cm dimensions (height-to-diameter ratio 1.0), and cube specimens are produced in 15×15×15 cm dimensions [10]. In this study, 15×15×15 cm cube specimens were produced in accordance with the Turkish standards.

In the fourth stage, the prepared specimens were cured in a water tank to ensure maximum strength development. The curing period enhances compressive strength due to the formation of C-S-H gels during the hydration process. Testing the specimens after a 28-day curing period reveals their actual strength development [11-13].

In the fifth stage, the cured specimens were subjected to compressive strength and splitting tensile strength tests using a compression testing machine. The recorded strength values were obtained during the failure of the specimens. This

meticulously designed methodology aims to comprehensively evaluate the structural performance of concrete produced with recycled earthquake debris and to demonstrate its economic and environmental benefits.

After compressive strength testing, the compressive strength (f_c) is calculated in MPa using Equation (1):

$$f_c = P/A \quad (1)$$

f_c: Compressive strength, MPa (N/mm²)
P: Maximum applied load at failure, N
A: Cross-sectional area subjected to compressive loading, mm²

Within the scope of this proposed study, rubble obtained from earthquake-affected regions was processed into recycled aggregates and incorporated into concrete production. Accordingly, a series of 15×15×15 cm concrete cube specimens were produced. The compressive strength of the concrete specimens produced with recycled rubble aggregates and varying cement ratios was tested at the Fırat University Building Materials Laboratory (FÜYAMLAB). The aim of this study is to achieve economic benefits and minimize environmental damage through the recycling of earthquake debris generated after the February 6 earthquakes in Türkiye. The distinguishing feature of this work is its attempt to produce structural concrete with comparable strength to conventional concrete without the use of costly chemical admixtures, but instead by varying the proportions of water, cement, and rubble aggregate. Beyond preventing environmental damage, making such waste reusable is expected not only to reduce material losses but also to support psychological recovery for citizens affected by the earthquake by demonstrating that debris can be transformed into constructive materials, symbolizing resilience and continuity. Figure 1 presents images of the entire process described above.



Figure 1. The produced concrete specimens were subjected to fresh and hardened concrete tests.

3. Results

As shown in Figure 2, the graph presents a bar chart illustrating the compressive strength (fc) values of concrete. The vertical axis (Y-axis) is labeled “Concrete Compressive Strength, fc (MPa),” while the horizontal axis (X-axis) is

labeled “Specimens.” Accordingly, the compressive strength values of 12 different concrete specimens, designated as N1 to N12, are represented. These values correspond to the average results of three control specimens for each sample.

The compressive strength values range between 17.7 MPa (N1) and 34.1 MPa (N10). Among the tested specimens, N1 exhibits the lowest compressive strength at 17.7 MPa, whereas N10 demonstrates the highest strength at 34.1 MPa, indicating a substantial difference compared to the other samples. Specimens N8 (30.6 MPa), N10 (34.1 MPa), N11 (29.7 MPa), and N12 (29.0 MPa) display above-average compressive strength values, suggesting more durable concrete mixtures or potentially improved curing conditions. In contrast, specimens such as N2, N3, N4, N5, N6, N7, and N9 exhibit relatively close strength values, ranging between 22.5 MPa and 25.7 MPa.

No distinct linear increasing or decreasing trend is observed across the specimens; instead, the values fluctuate. For instance, a significant increase is observed from N1 to N2, followed by a slight decrease from N2 to N3. Similarly, a sharp rise occurs from N9 to N10, while a gradual decline is noted from N10 to N11 and N12. Since the curing duration was identical for all specimens, the differences in compressive strength can be attributed to variations in recycled aggregete properties, mix proportions, and casting conditions.

In conclusion, the graph clearly demonstrates the compressive strength data and highlights the performance differences among the specimens. The outstanding strength of specimen N10 is particularly noteworthy. These results provide valuable insights for concrete mix optimization, quality control, and further experimental investigations.

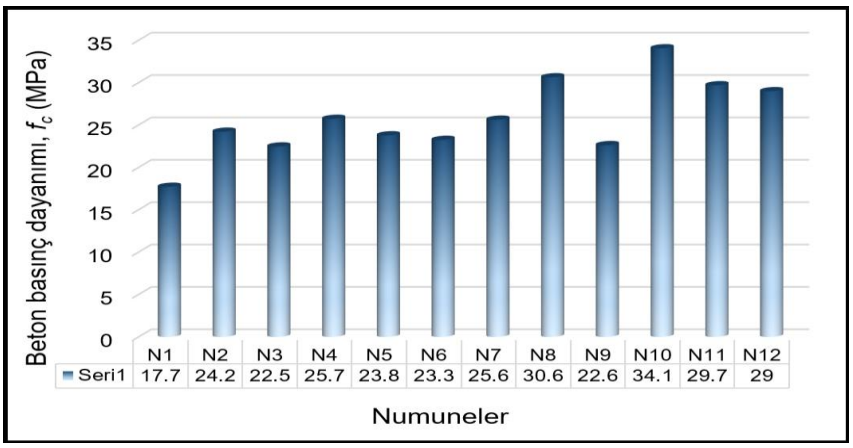


Figure 2. The compressive strength (f_c) values of the tested concrete specimens

Figure 3 presents a bar chart that displays the tensile strength values (f_t) in concrete splitting. The vertical axis (Y-axis) is labeled "Concrete Splitting Tensile Strength, f_t (MPa)," and the horizontal axis (X-axis) is labeled "Samples." Thus, the tensile strength values of 12 different concrete samples, ranging from N1 to N12, are represented. The splitting tensile strengths of different concrete samples, obtained by using construction waste resulting from the earthquake as recycled aggregates in the concrete mix, have been compared. The chart shows the splitting tensile strength values of the 12 different concrete samples (from N1 to N12), with each sample's value being determined by averaging the values of three control specimens.

The concrete splitting tensile strength values range from 1.71 MPa (N1) to 3.38 MPa (N10). This range indicates significant differences in strength among the samples. The N1 sample has the lowest splitting tensile strength at 1.71 MPa, while the N10 sample has the highest at 3.38 MPa. This represents a significantly higher value compared to other samples, indicating that this sample is the most resistant to tensile stresses. Samples N8 (2.98 MPa), N10 (3.38 MPa), N11 (2.91 MPa), and N12 (2.78 MPa) have splitting tensile strength values above the overall average, indicating that they perform relatively better under tensile stress. Samples such as N2, N3, N4, N5, N6, N7, and N9 show strength values between 2.15 MPa and 2.69 MPa, which are relatively close to each other. There is no

discernible linear increase or decrease trend among the samples. Instead, the values show a fluctuating pattern. For example, a notable increase is observed from N1 to N2, followed by a slight decrease from N2 to N3. Similarly, there is a sharp increase from N9 to N10, followed by slight decreases from N10 to N11 and N12. These variations align with the pattern observed in compressive strength values.

Sample N10 exhibits the highest values in both compressive strength (34.1 MPa) and splitting tensile strength (3.38 MPa). This suggests that concrete samples represented by N10 generally exhibit superior mechanical properties. On the other hand, sample N1 shows the lowest values in both strength types (fc = 17.7 MPa, ft = 1.71 MPa), indicating that the concrete represented by N1 has generally lower mechanical properties. Concrete's compressive strength is typically much higher than its tensile strength, a fact that can be observed from the differing scales on the graphs (Figures 4 and 5). While the compressive strength of concrete reaches up to 35 MPa, the splitting tensile strength reaches only up to 3.5 MPa. This is expected, as concrete is known for its resistance to compression but is relatively weak in tension.

In conclusion, this chart clearly demonstrates the splitting tensile strength performance of different concrete samples. Sample N10 shows a significantly higher tensile strength compared to the others. These data provide valuable insights, especially for concrete mix design, material selection, and structural applications where concrete will be subjected to tensile stresses.

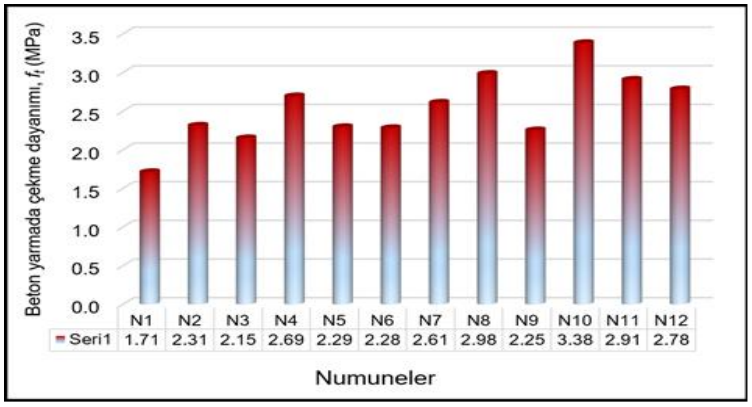


Figure 3. The splitting tensile strength (ft) values of the tested concrete specimens

Figure 4 presents a scatter plot that illustrates the relationship between concrete compressive strength (f_c) and concrete splitting tensile strength (f_t). The vertical axis (Y-axis) is labeled "Concrete Splitting Tensile Strength, f_t (MPa)," representing the dependent variable, which examines how the tensile strength changes in relation to compressive strength. The horizontal axis (X-axis) is labeled "Concrete Compressive Strength, f_c (MPa)," representing the independent variable. The correlation between the compressive strength and the splitting tensile strength of various concrete samples, obtained by using construction waste from the earthquake as recycled aggregates in the concrete mix, is analyzed.

In the plot, data points represented by circles correspond to each concrete sample. These points represent the combination of the fc and ft values of the samples from N1 to N12, as shown in the previous two figures (Figures 2 and 3). A dashed line (trend line or regression line) is plotted among the data points, visualizing the relationship between the two variables. An equation [Equation (2)] and the R^2 value are provided in the upper-left corner of the graph.

$$f_t = 0.1002 \times f_c - 0.0469 \quad (2)$$

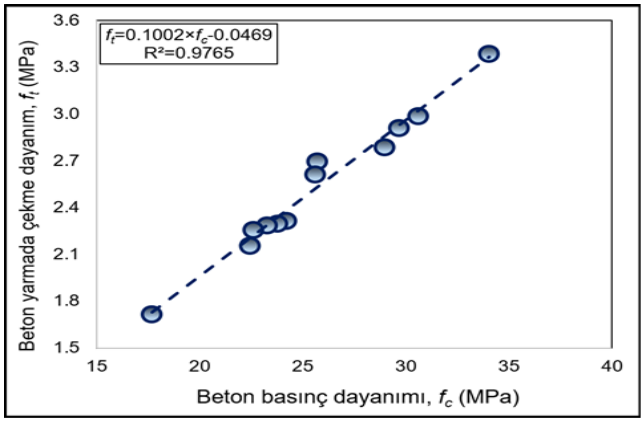


Figure 4. The correlation between the f_c and f_t values of the tested concrete specimens

Equation (2) represents a linear model used to predict the concrete splitting tensile strength (ft) based on the compressive strength (fc). The slope (0.1002) indicates that for every 1 MPa increase in compressive strength, the splitting tensile strength will increase by an average of 0.1002 MPa. The Y-intercept (-0.0469) represents the theoretical splitting tensile strength when the compressive strength is zero (this value may not have direct physical significance in the context of real concrete properties but is a mathematical component of the model). The coefficient of determination, $R^2 = 0.9765$, shows that approximately 97.65% of the variation in the dependent variable (tensile strength) is explained by the independent variable (compressive strength). This is a very high R^2 value, indicating that the linear model fits the data very well and that there is a strong positive linear relationship between compressive strength and splitting tensile strength.

The proximity of the data points to the trend line and the high R^2 value clearly demonstrate that as the compressive strength of concrete increases, the splitting tensile strength also increases. This is an expected relationship, as concrete typically exhibits proportional resistance to both compression and tension. The upward slope of the trend line confirms the positive correlation, meaning that as one variable increases, the other also increases. The obtained regression equation can be used to estimate the approximate splitting tensile strength of concrete with a known compressive strength. This can be a practical tool in material design and quality control processes. For instance, if the compressive strength of a concrete sample is known, the splitting tensile strength can be predicted without directly testing it, with a certain margin of error. Visually, most of the data points are very close to the line, and no significant outliers are observed, further supporting that the data align well with the model.

In conclusion, this graph (Figure 4) clearly highlights a very important relationship in concrete engineering—the strong linear and positive correlation between concrete compressive strength and splitting tensile strength. The high R^2 value indicates that this relationship is highly reliable and that the obtained regression equation can be practically used for prediction purposes. Such analyses play a critical role in concrete design and performance evaluation.

4. Conclusion

This study successfully demonstrates the feasibility of using debris waste resulting from the Kahramanmaraş-centered earthquakes of February 6, 2023, as recycled aggregates in concrete production. Based on experimental findings, it is shown that the recycled aggregates derived from earthquake rubble can replace traditional aggregates used in concrete production, providing a carrier concrete that complies with relevant regulations. The main conclusions of the study can be summarized as follows:

- According to the results of the conducted experiments, the compressive strength (f_c) values ranged from 17.7 to 34.1 MPa, while the splitting tensile strength (ft) values ranged from 1.71 to 3.38 MPa. This finding implies that when compressive strength is tested, the tensile strength can also be approximately predicted.
- The N10 sample exhibited the highest values in both compressive and tensile strength, while the N1 sample had the lowest values.

- It has been determined that concrete with strengths compliant with regulations can be produced using recycled rubble aggregates obtained from earthquake debris, without the need for costly chemical additives.
- A high level of positive correlation was obtained between compressive strength and splitting tensile strength of the samples ($R^2 = 0.9765$).
- The recycling approach mentioned in this study not only offers economic and environmental benefits but also has the potential to contribute to the construction process in earthquake-affected regions.

In conclusion, this study supports with scientific data that the reuse of earthquake debris waste in concrete production is both technically feasible and beneficial in terms of environmental and economic impact. These findings represent significant potential for future reconstruction efforts.

Acknowledgements

This work was supported by the TÜBİTAK 2209-A Undergraduate Research Project Support Program under grant number 1919B012326292.

5. References

[1] Öner, F. (2023, Şubat). Dünya Bankası'Ndan Deprem Açıklaması: 34 Milyar Doları Aşacak. https://m.haber7.com/ekonomi/haber/3305561-dunya-bankasindan-deprem-aciklamasi-34-milyar-dolari-asacak adresinden erişildi.
[2] https://data.tuik.gov.tr/Bulten/Index?p=Insaat-Maliyet-Endeksi-Haziran-2022-45627
[3] Gezen, A. F. (2015). İnşaat sektöründe sürdürülebilir kent yaşamı ve karayollarında uygulama alanları (Master's thesis, Fen Bilimleri Enstitüsü).
[4] Çorbacıoğlu, C. U. (2008). *Beton karışım tasarımının geçirimliliğe ve mekanik özelliklere etkisi* (Doctoral dissertation, Fen Bilimleri Enstitüsü).
[5] Büyüktaş, K., & Alagöz, T. (2004). Farklı çimento çeşitleri ve karışımlar kullanılarak kaplanan sulama kanalında sızma kayıplarının belirlenmesi. Akdeniz Üniversitesi Ziraat Fakültesi Dergisi, 17(1), 49-58.
[6] Tüdeş E., "Zeminlerin Kireç ve Çimento Katkısı ile Çözümü", Karadeniz Teknik Üniversitesi, Yüksek Lisans Tezi, 161s, 1996.
[7] TS 3068, "Laboratuvarda Beton Deney Numunelerinin Hazırlanması ve Bakımı" Türk Standartları Enstitüsü, Mart 1978
[8] Gönen, T., Onur, O. N. A. T., Cemalgil, S., Yilmazer, B., & Altuncu, Y. T. (2012). Beton Teknolojisi İçin Yeni Atık Malzemeler Üzerine Bir İnceleme. *Yapı Teknolojileri Elektronik Dergisi*, *8*(1), 36-43.
[9] Ölmez, E., & Yıldız, Ş. (2008). İnşaat ve yıkıntı atıklarının yönetimi ve planlanan İstanbul modeli. *Kent Yönetimi, İnsan ve Çevre Sorunları*, *8*, 02-06.
[10] Felekoğlu, B., Türkel, S., "Effects of Specimen Type And Dimensions on Compressive Strength of Concrete". Gazi Üniversitesi Fen Bilimleri Dergisi, 18 (4), S.639–645, 2005
[11] Tunç, E. T. (2019). An experimental study based on the strength properties of concrete containing chemical admixture. Avrupa Bilim ve Teknoloji Dergisi, (17), 901-908.
[12] Tugrul Tunc, E. (2025). Strength and durability of superplasticizer concrete based on different component parameters: An experimental and statistical study. Arabian Journal for Science and Engineering, 50(3), 1649-1664.
[13] Tunc, E. T. (2019). Recycling of marble waste: A review based on strength of concrete containing marble waste. Journal of environmental management, 231, 86-97.
[14] https://izmir.csb.gov.tr/yapi-malzemeleri-ve-pgd.sb.-md.-calismalari-i-1955
[15] https://earthquake.usgs.gov/storymap/index-turkey2023.html
[16] https://www.sciencedirect.com/science/article/pii/S0038080624000416

A Preliminary Experimental Investigation on the Compressive Strength of Fiber-Reinforced Polylmer (FRP) Reinforced Concrete Specimens

Ömer Faruk DEMİRCİ[1], Esra TUĞRUL TUNÇ[2]

[1,2]Civil Engineering Department, Engineering Faculty, Fırat University, Elazığ, Turkiye.
[1]omerfarukdemirci783@gmail.com, [2]esratugrul@firat.edu.tr
[1](ORCID: 0009-0003-5531-4453), [2](0000-0001-9071-774X)

Abstract
In this study, the benefits of Fiber Reinforced Polymer (FRP) materials which are products of advancing science and technology and come in various types have been investigated in terms of their potential to prevent loss of life and property. In particular, the study addresses major problems such as debris removal and storage following structural failures. To this end, structural elements were collected from the city of Malatya.In the first stage, the current strength of these elements was determined through concrete compressive strength tests (press test). Then, the elements were wrapped with FRP materials and retested under the same conditions. The results were thoroughly compared and presented with supporting graphs and tables.Findings from the study indicate that FRP application significantly increases the existing strength of structural elements. This increase ranges between 15% and 30%, and it is expected that future developments will further enhance this improvement. If the targeted increase is achieved, it is anticipated that many material and emotional losses in the event of structural failures—such as during earthquakes can be substantially mitigated.

Keywords: structural strengthening, evaluation of existing structures, damage detection, fiber reinforced polymer (FRP), concrete compressive strength test

1. Introduction

Due to its location on active fault lines, our country frequently faces destructive earthquakes. One of the most striking examples of this was the earthquake centered in Kahramanmaraş on February 6, 2023, which revealed some painful realities. Affecting 11 provinces both materially and emotionally, this earthquake once again highlighted the vital importance of structural strengthening. Among the most prominent strengthening methods is the use of Fiber Reinforced Polymer (FRP). Its application offers significant advantages over conventional methods such as column jacketing or shear walls.

To demonstrate these advantages, a series of 15x15x15 cm concrete cube specimens were tested at the Construction Materials Laboratory (FÜYAMLAB) of Fırat University. Some of the specimens were wrapped with FRP material, while others served as unwrapped control samples. Uniaxial compressive strength tests were carried out on both the FRP-wrapped and control specimens to determine the compressive strength of the concrete. The tests were conducted on a 2500 kN capacity hydraulic compression testing machine available at FÜYAMLAB, in accordance with ASTM or TS EN standards.

During testing, load–deformation data were recorded, and the maximum load-bearing capacity of each specimen was determined. The experimental data were analyzed to quantitatively assess the contribution of FRP wrapping to concrete compressive strength. This contribution was calculated by comparing the strength increase in the wrapped specimens with that of the control samples. The numerical results were further evaluated using statistical analysis methods.

FRP composites are recognized as an effective method to enhance the strength of damaged reinforced concrete members or structurally deficient load-bearing elements. The efficiency of FRP lies in its high tensile strength combined with its lightweight nature [2]. Composite materials have expanded their applications in the modern construction industry by addressing various structural needs. In particular, materials like FRP are widely used in structural rehabilitation, repair, and strengthening applications [3].

[1]Corresponding author

Although numerous studies have been conducted on the structural strengthening potential of FRP, the originality of this project lies in its aim to approximately quantify the strengthening capacity of FRP on load-bearing elements. The experimental results clearly present, with numerical data, the extent to which FRP improves the compressive strength of concrete.

2. Materials and Method

The project management consisted of four main phases. In the first phase, the tested concrete specimens were delivered to the laboratory and prepared for testing. In the second phase, Fiber Reinforced Polymer (FRP) materials with suitable properties were supplied and applied to half of the specimens. In the third phase, compressive strength tests were conducted on both FRP-wrapped and unwrapped (control group) specimens. In the fourth and final phase, the experimental data were analyzed, and the numerical contribution of FRP to the compressive strength of concrete was calculated. All project activities were carried out using the existing infrastructure of FÜYAMLAB (Construction Materials Laboratory of Fırat University), located within the Faculty of Engineering at Fırat University.

As part of the broader impact of this project, the findings clearly demonstrate that the FRP strengthening method significantly increases structural strength. Therefore, it is planned to share these results with the engineering community through seminars and technical meetings. This project also aims to serve as a foundational study for future large-scale research and graduate-level thesis projects.

The methodology of this project was designed to experimentally evaluate the effect of strengthening tested concrete specimens with Fiber Reinforced Polymer (FRP) on their compressive strength. All prepared specimens, both FRP-wrapped and unwrapped, were subjected to compressive strength testing using the compression testing machine in the laboratory. The maximum load-bearing capacities of the specimens were measured. The collected data were then analyzed to quantitatively determine the extent to which FRP improved the compressive strength of the concrete, and comparative evaluations were carried out.



Figure 1. Appearance of tested FRP wrapped concrete specimens

Three unwrapped (non-strengthened) specimens were used as a reference group to evaluate the effect of FRP, while the other three specimens were wrapped with FRP material. During the FRP application process, maximizing the bond between the FRP and the concrete surface was of critical importance. Therefore, the surface of the concrete was smoothed to create a uniform and even texture, ensuring optimal adhesion of the FRP to the surface. The joints of the FRP wrap were processed to be flat and aligned, reducing stress concentrations and enhancing the continuity of the wrap. Additionally, while applying FRP fibers to the corners of the structural elements, the edges were rounded with a radius of at least 2–3 cm to prevent stress concentrations at the corners. This application helped prevent early peeling of the FRP and improved the overall performance of the wrap.

3. Results

The graph presented in Figure 2 compares the variation in axial compressive strength of a series of concrete cube specimens (labeled from C1 to C10) between the “control” condition and the “FRP (Fiber Reinforced Polymer) wrapped” condition. The strength values are given in megapascals (MPa). The most prominent observation from the graph is that all FRP-wrapped specimens exhibit significantly higher compressive strength compared to the

unwrapped control specimens. This clearly demonstrates the effectiveness of FRP wrapping in enhancing the axial load-carrying capacity of concrete elements.

It is observed that the FRP wrapping increases the compressive strength of concrete by approximately 15% to 32% on average. The magnitude of this increase is a function of the FRP material, the wrapping technique, and the initial strength level of the concrete. The compressive strength of the control specimens gradually increases from C1 (12.06 MPa) to C10 (18.29 MPa). Correspondingly, the strength of the FRP-wrapped specimens follows a similar upward trend. This supports the hypothesis that FRP confinement is generally effective for both low- and medium-strength concretes, but the percentage increase in strength tends to be more pronounced in lower-strength concretes. In contrast, for higher-strength concretes, the effect of FRP confinement may be more associated with improvements in ductility rather than compressive strength alone.

The FRP wrapping acts as a lateral confining mechanism when applied around the concrete specimen. Under axial loading, concrete tends to expand laterally due to the Poisson effect. The FRP wrap resists this lateral expansion by generating a confining pressure, thereby delaying the onset of crushing and limiting the propagation of internal cracks. This passive confinement effect enhances the ultimate compressive strength of the concrete while also improving its deformation capacity and ductility. This increase in ductility is particularly critical for concrete, which is inherently a brittle material.

The presented graph reflects the results obtained for a specific type of FRP and wrapping technique. Factors such as the type of FRP fibers (e.g., carbon, glass, aramid), their modulus, fiber density, the number of wrap layers, adhesion to the concrete surface, and curing conditions directly influence the magnitude of the strength enhancement. In the experiments conducted within the scope of this project, a standardized FRP wrapping procedure was applied consistently to all specimens.

Figure 2 clearly illustrates that FRP wrapping is a highly effective strengthening technique for improving the axial compressive strength of concrete. The observed strength enhancements support the fundamental scientific rationale behind the widespread use of this method in structural strengthening and rehabilitation applications. Despite the variations in control strength among the specimens, the increase in strength observed in all FRP-wrapped samples suggests that FRP wrapping can provide consistent improvement across different concrete mixes or curing conditions.

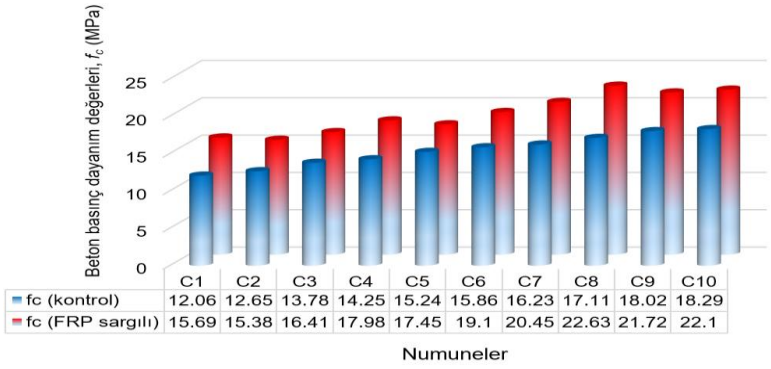


Figure 2. Comparison of concrete compressive strength values of FRP wrapped and FRP unwrapped samples

The graph presented in Figure 3 is a scatter plot illustrating the relationship between the compressive strength of control specimens (fc control) and FRP-wrapped specimens (fc FRP wrapped), based on the compressive strength results obtained from previous data. The horizontal axis represents the compressive strength values of the control specimens, fc control (MPa), while the vertical axis represents the compressive strength values of the FRP-wrapped specimens, fc FRP wrapped (MPa). Each data point corresponds to the strength values of a specimen pair (control and FRP wrapped). The plot demonstrates a positive correlation, indicating that as the control strength increases, the strength of the FRP-wrapped specimens also increases. A line of best fit has been applied to the data points, modeling the linear relationship between control strength and FRP-wrapped strength. The upward slope of this line confirms

that the FRP-wrapped strength increases with increasing control strength. The equation (1) and the R^2 value are presented in the upper left corner of the graph:

$$f_c(\text{FRP confined})=1.2\times f_c(\text{control})+0.45 \quad (1)$$

This equation can be used to estimate the expected strength of an FRP-wrapped concrete specimen based on the strength of the control specimen. The coefficient 1.2 indicates that for every 1 MPa increase in control strength, the FRP-wrapped strength increases by approximately 1.2 MPa. The constant term 0.45 represents an offset, theoretically indicating the FRP-wrapped strength when the control strength is zero (which is not a practical scenario). More practically, this constant can be interpreted as the initial confinement effect provided by the FRP wrap or a baseline strength contribution independent of the wrapping. The R^2 value of 0.91 is a statistical measure indicating how well the regression model fits the data points. The R^2 value ranges between 0 and 1; values close to 1 indicate an excellent fit. A value of 0.91 is quite high and demonstrates that control strength explains approximately 91% of the variation in FRP-wrapped strength. Therefore, it can be concluded that the regression equation is highly reliable for predicting the performance of FRP wrapping within this dataset.

The linear relationship shown by the graph and regression equation reveals that the strength enhancement provided by FRP wrapping is not independent of the initial strength level of the concrete but is proportional to it. In other words, concrete with higher control strength tends to exhibit greater absolute strength gains when wrapped with FRP. Equation (1) quantitatively expresses the confinement effect that FRP provides to concrete. By restricting lateral expansion, FRP enables concrete to carry higher axial loads under compression. This model can be practically used in engineering applications to predict the expected strength of an element strengthened with FRP based on its known original strength.

It is important to note that the regression equation is valid for the specific FRP material, wrapping technique, and concrete specimen properties used in this experiment. Different types of FRP, number of wrap layers, wrapping geometries, or concrete qualities may result in different coefficients. Additionally, the model applies only within the range of the existing data (approximately 12 MPa to 19 MPa control strength); strength values outside this range require further validation. The high R^2 value indicates that the model represents the experimental data well and that the results are consistent. This increases confidence in using the model for predictions in similar future applications.

In summary, this graph and equation scientifically support that FRP wrapping is an effective and predictable method to increase the compressive strength of concrete. For professionals involved in the field, it provides an important tool for design and performance estimation when strengthening concrete elements with FRP.

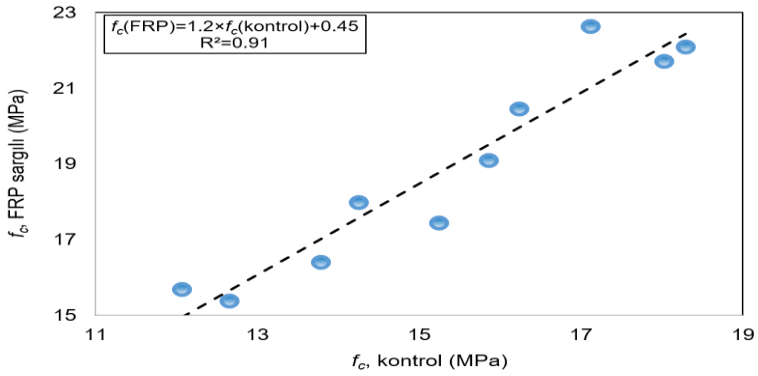


Figure 3. Correlation between measured fc-control values and fc-FRP wrapped values

he graph presented in Figure 4 is a scatter plot that relates the compressive strength of the control specimen (fc control) to the percentage change in compressive strength of the FRP-wrapped specimen (fc FRP wrapped) relative to the control specimen, based on previously obtained concrete compressive strength data. The horizontal axis represents the compressive strength values of the control specimens, while the vertical axis represents the percentage

change in strength of the FRP-wrapped specimens. The percentage increase in strength of the FRP-wrapped specimen is calculated relative to the control specimen strength. This value is computed using Equation (2).

$$Percentage\ change = \left(\frac{f_c(FRP\ confined - f_c(control))}{f_c(control)}\right) \times 100 \quad (2)$$

The data points represent the percentage change values obtained for each specimen pair (control and FRP-wrapped). The percentage changes range approximately between 14% and 32%, confirming the significant effect of FRP wrapping on increasing the compressive strength of concrete. When examining the graph, it is observed that as the control strength increases, the percentage change does not show a clear trend or exhibits a slight negative tendency. For example, one of the specimens with the lowest control strength, C1 (around 12 MPa), shows about a 30% increase, whereas C10 (around 18.29 MPa) shows about a 20.8% increase. The highest percentage change on the graph, approximately 32%, occurs for a specimen with about 17 MPa control strength, while the lowest percentage change, around 14%, is seen for a specimen with approximately 15.5 MPa control strength. This indicates some scatter among the data and that initial strength alone does not fully explain the percentage increase.

Some studies and engineering practices suggest that FRP wrapping tends to provide a relatively higher percentage increase in low-strength concrete. In this graph, a general trend could be interpreted where the percentage increase slightly decreases as control strength rises; however, the data scatter is not strong enough to conclusively confirm this, especially with some deviations in the medium-strength range. For high-strength concrete, the effect of wrapping is often more noticeable in improving ductility and deformation capacity rather than a large percentage increase in strength. The distribution of data points may reflect natural variability in experimental conditions or differences in the specimens used. Factors such as the quality of FRP application, concrete homogeneity, or minor differences in testing procedures are considered to influence this scatter. Such analysis is important to understand the relative effectiveness of FRP wrapping across different concrete strength levels.

Figure 4 clearly shows that FRP wrapping provides significant strength increases across various concrete strength levels. The examination of percentage changes offers insight into how effective FRP wrapping might be in strengthening designs relative to the existing concrete strength. This is especially important in the rehabilitation of existing structures or strengthening low-strength elements. It can be concluded that FRP wrapping not only increases absolute strength but also improves strength at varying rates depending on the initial properties of the concrete. The presented graph shows that FRP wrapping increases concrete compressive strength by between 14% and 32%, and although there is no clear linear relationship between control strength and percentage change, a general downward trend is observed. Findings from the current project provide valuable information to understand the potential and performance variability of FRP wrapping in engineering applications.

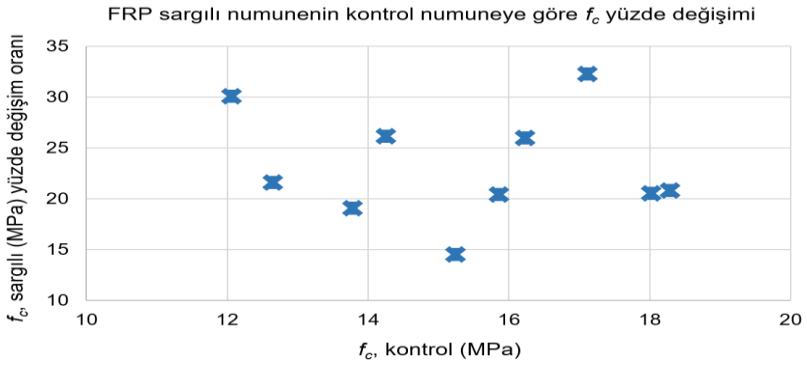


Figure 4. Percentage change of compressive strength (f_c) values of FRP-wrapped specimens relative to the compressive strength (f_c) values of control specimens.

4. Conclusions

- This study demonstrated that FRP materials are an effective method for enhancing the structural performance of reinforced concrete elements.
- Experimental processes showed that FRP wrapping has positive effects on concrete, particularly providing significant contributions in terms of strength.
- The application of FRP was evaluated as a faster, lighter, and more feasible alternative to traditional strengthening methods.
- It was understood that strengthening applications using FRP can play an important role in making existing structures more resistant to earthquakes.
- The findings indicate that FRP can be used not only for new constructions but also for the rehabilitation of existing and damaged structures.
- The study suggests that the wider and more systematic use of FRP materials in the construction industry would be an important step toward improving structural safety.

Acknowledgements

This work was supported by the TÜBİTAK 2209-A Undergraduate Research Project Support Program under grant number 1919B012326203.

5. References

[1] Utkucu, M., Durmuş, H., Uzunca, F., & Nalbant, S. (2023). 6 Şubat 2023 Gaziantep (Mw= 7.7) Ve Elbistan (Mw= 7.5) Depremleri Üzerine Bir Değerlendirme. Sakarya Üniversitesi Afet Yönetim ve Araştırma Merkezi, Sakarya.
[2] Parvin, A., & Jamwal, A. S. (2006). Performance of externally FRP reinforced columns for changes in angle and thickness of the wrap and concrete strength. Composite structures, 73(4), 451-457.
[3] Elsanadedy, H. M., Al-Salloum, Y. A., Alsayed, S. H., & Iqbal, R. A. (2012). Experimental and numerical investigation of size effects in FRP-wrapped concrete columns. Construction and Building Materials, 29, 56-72.
[4] https://cimsa.com.tr/formulhane/hazir-beton/beton-basinc-dayanimi-nedir-nasil-olculur/
[5] https://betonagrega.com/urun/beton-basinc-dayanimi-testi
[6] https://www.besoglu.com/betonlarin-guclendirilmesinde-karbon-fiber-takviyeli-polimer-yerine-cam-fiber-takviyeli-polimer-kullanilmasi-2/
[7] https://ctech-llc.com/tr/wikitech/frp-installation [5] Akbalık, H. H., & Sarıbıyık, A. (2019). Improving the Adhesion of BFRP Strips to the Concrete Surface. Academic Perspective Procedia, 2(2), 220-228.
[8] Akbalık, H. H., & Sarıbıyık, A. (2019). Improving the Adhesion of BFRP Strips to the Concrete Surface. Academic Perspective Procedia, 2(2), 220-228.
[9] https://www.t5yapilab.com.tr/beton-basinc-dayanim-testi.php.
[10] https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcSJjmyDsSQqgXbE0gmbhUBtY-atrCeLj5MRgw&usqp=CAU
[11] Parvin, A., & Jamwal, A. S. (2006). Performance of externally FRP reinforced columns for changes in angle and thickness of the wrap and concrete strength. Composite structures, 73(4), 451-457.
[12] Tunç, E. T. (2019). An experimental study based on the strength properties of concrete containing chemical admixture. Avrupa Bilim ve Teknoloji Dergisi, (17), 901-908.
[13] Tugrul Tunc, E. (2025). Strength and durability of superplasticizer concrete based on different component parameters: An experimental and statistical study. Arabian Journal for Science and Engineering, 50(3), 1649-1664.
[14] Tunc, E. T. (2019). Recycling of marble waste: A review based on strength of concrete containing marble waste. Journal of environmental management, 231, 86-97.

Experimental and Analytical Investigation of Flow Coefficient in Triangular Cross-Section Side Weirs

Mustafa TUNÇ [1], Muhammet Erdem URHAN [2]

[1,2]Civil Engineering Department, Engineering Faculty, Fırat University, Elazığ, Turkiye.
[1]mtunc@firat.edu.tr
[1](ORCID: 0000-0001-9756-8409), [2](ORCID: 0000-0000-0000-YYYY)

Abstract
Regulation of water level in open channel flows, flow measurement, and the discharge of excess flow are among the fundamental research topics of hydraulic engineering. Side weirs used for this purpose play a critical role in many engineering applications such as flood control, irrigation systems, and drainage networks. While the literature has generally focused on rectangular side weirs, and detailed studies have been conducted on their hydraulic performance, a lack of systematic analysis for triangular side weirs has been identified. This study aims to experimentally determine the flow coefficient (Cd) of triangular side weirs under different flow regimes using a specialized experimental setup constructed at the Fırat University Hydraulics Laboratory. The obtained data are compared with the results obtained from analytical and empirical methods widely used in the literature, such as the De Marchi, Schmidt, and Dominguez methods. As a result of experiments conducted over a wide flow range ranging from 5 L/s to 100 L/s, a new empirical equation specific to triangular side weirs was developed that provides the best fit to the experimental data. This developed equation predicts the flow coefficient with significantly higher accuracy than existing approaches, providing engineers with more reliable data for the design and sizing of triangular side weirs. This paper serves as a practical guide for the safer and more efficient design of hydraulic structures.

Keywords: flow coefficient, side weir, open channel, hydraulic engineering.

1. Introduction

Hydraulic engineering plays a vital role in the efficient and safe management of water resources. Controlling open channel flows is a fundamental engineering problem in areas such as flood management, irrigation canal regulation, and water distribution systems. Weirs are among the most common hydraulic structures used in these control mechanisms. Weirs are constructed to regulate the water level in a stream channel, measure discharge, or discharge excess water. Traditionally, weirs, placed perpendicular to the channel flow direction, were studied in detail by pioneering researchers such as Francis (1883), Rehbock (1929), and Henderson (1966). These studies contributed significantly to the hydraulic literature by developing fundamental empirical formulations on flow-load relationships. However, side weirs serve a vital function, especially in emergency situations such as flood control, by diverting excess water to a side discharge channel when the water level in the main channel exceeds a certain threshold.

Side weirs exhibit complex hydraulic behavior due to the gradual decrease in main channel flow along the weir. This flow structure makes it difficult to accurately determine the flow coefficient (Cd). While extensive research has been conducted on rectangular side weirs in the existing literature (Subramanya and Awasthy, 1972; Singh et al., 1994), studies on the hydraulic performance of side weirs with different geometries, such as triangular, are quite limited. It is known that triangular side weirs, in particular, exhibit a different flow regime due to the geometric narrowing of the flow, and that the flow coefficient cannot be fully explained with traditional empirical equations. This situation limits the potential of such structures in engineering applications.

The primary motivation of this study is to experimentally investigate the flow coefficient behavior in triangular side weirs, which cannot be fully explained by existing analytical approaches. This study aims to fill a significant gap in the literature by determining the flow coefficient of triangular side weirs. The unique value of the project lies in the development of an empirical model that does not rely solely on theoretical approaches but is supported by precise experimental data collected in the laboratory. The findings will provide engineers with more reliable data for the design and sizing of triangular side weirs, enabling more effective solutions in critical areas such as water management and flood control.

[1]Corresponding author

The study of side weirs accelerated in the early 20th century with developments in the theoretical and experimental hydraulics of open channel flows. The flow coefficient is the most important parameter in side weir design, and research in this area has generally focused on rectangular weirs.

De Marchi (1936): He developed the first systematic analytical approach using the principle of conservation of energy to calculate side weir flow rate. De Marchi's formula modeled the variation of flow rate depending on the water height above the side weir, but this approach required an empirical coefficient because it did not fully account for friction and energy losses.

Schmidt (1929) and Dominguez (1940): These researchers developed De Marchi's approach and presented different empirical formulations for the side weir flow coefficient. While Schmidt's approach included a correction factor based on experimental data, Dominguez developed a more detailed model that took into account the deceleration of the flow across the side weir.

Subramanya and Awasthy (1972): They experimentally investigated the flow characteristics in rectangular side weirs. Their work provided important information on the three-dimensional nature of side weir flow and the dependence of the discharge coefficient on the main channel discharge and water level. This work contributed significantly to the understanding of the concept of spatially varied flow in side weirs.

Singh, Manivannan, and Satyanarayana (1994): They conducted detailed experimental studies on the discharge coefficient of rectangular side weirs and proposed a new empirical equation valid for various hydraulic conditions. These studies, in particular, demonstrated the dependence of the discharge coefficient on dimensionless parameters such as the Froude number and weir height.

Swamee et al. (1994): They investigated the subcritical flow regime in rectangular side weirs and presented a formula for the discharge coefficient under these conditions. These studies have helped us understand the effects of different flow regimes on the discharge coefficient.

However, the existing literature is quite limited in the number of studies on the hydraulic performance of triangular side weirs. This study aims to fill this important gap and provide concrete and reliable data for the design of triangular side weirs.

2. Materials and Methods

Experimental studies were carried out on a specially designed open channel test setup in Fırat University Civil Engineering Hydraulics Laboratory to determine the triangular side weir flow characteristics.

2.1. Details and Components of the Experimental Setup

The experimental setup was equipped with high-precision equipment appropriate to the project's objectives. The main channel was a 10-meter-long, 40-cm-wide, and 50-cm-deep, acrylic and glass channel with a smooth bottom. This transparent structure allowed visual observation of the flow behavior over the side weir.

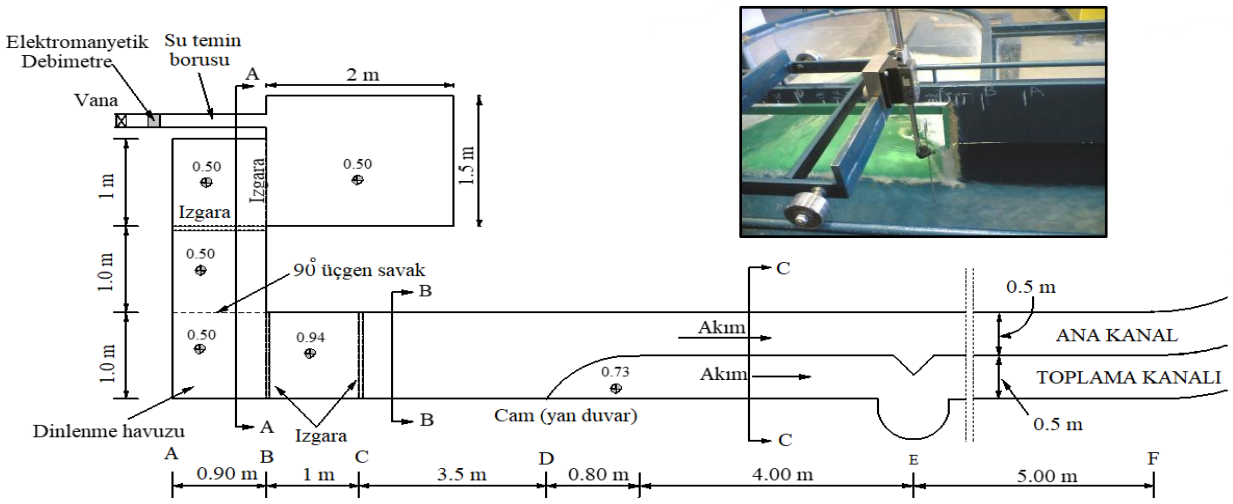


Figure 1. Schematic view of the experimental setup

The general layout of the experimental system is shown schematically in Figure 1. The system consists of a water supply tank, a pump, a magnetic flow meter for flow measurement, a test channel, a side weir, and water level measuring devices. The pump provides water supply and controls the flow rate. The electromagnetic flow meter, connected in series with the pump, allows for precise measurement of the flow rate entering the channel. A triangular side weir, located in the middle of the test channel, is positioned to discharge excess water from the main flow. Digital limnimeters located on the side weir and in the main channel measure water levels with millimeter precision and conduct data collection. This setup has been optimized to ensure repeatable and reliable data acquisition under different flow rates and flow regimes.

Experiments were conducted at 13 different flow rates, ranging from 5 L/s to 100 L/s, to cover practical applications and flow regimes. For each flow rate, the water level in the main channel and the height of the water flowing over the side weir were recorded using digital limnimeters with millimeter precision.

2.2. Data Analysis Methods and Theoretical Background

The experimental discharge and water level data obtained were used to calculate the discharge coefficient (Cd) of the side weirs. These calculations utilize analytical methods widely accepted in the literature and successfully applied to different weir types, such as those of De Marchi (1936), Schmidt (1929), and Dominguez (1940).

3. Results and Discussion

The data obtained as a result of experimental studies clearly revealed that the flow coefficient (Cd) in triangular side weirs varies according to the main channel flow rate and the water height (h) over the weir. It is clearly observed that both the main channel depth and the depth of water flowing over the side weir increase with increasing main channel discharge. Especially at low discharge values, the water flow can be assumed to be laminar or in the transition zone, causing the effects of surface tension and viscosity to become more pronounced. This explains the lower flow coefficient values at low flow rates. With increasing flow rate, the flow becomes more turbulent, which can be attributed to the greater efficiency of water flowing over the weir relative to the weir geometry, resulting in an increase in the flow coefficient (Cd). This correlation indicates an increase in the kinetic energy of the flow and a change in the hydraulic gradient over the side weir.

Comparison of the experimentally obtained Cd values with existing empirical equations in the literature revealed that these equations fail to fully model the complex hydraulic behavior of triangular side weirs. This finding highlights the need to develop a more reliable empirical equation specific to triangular side weirs.

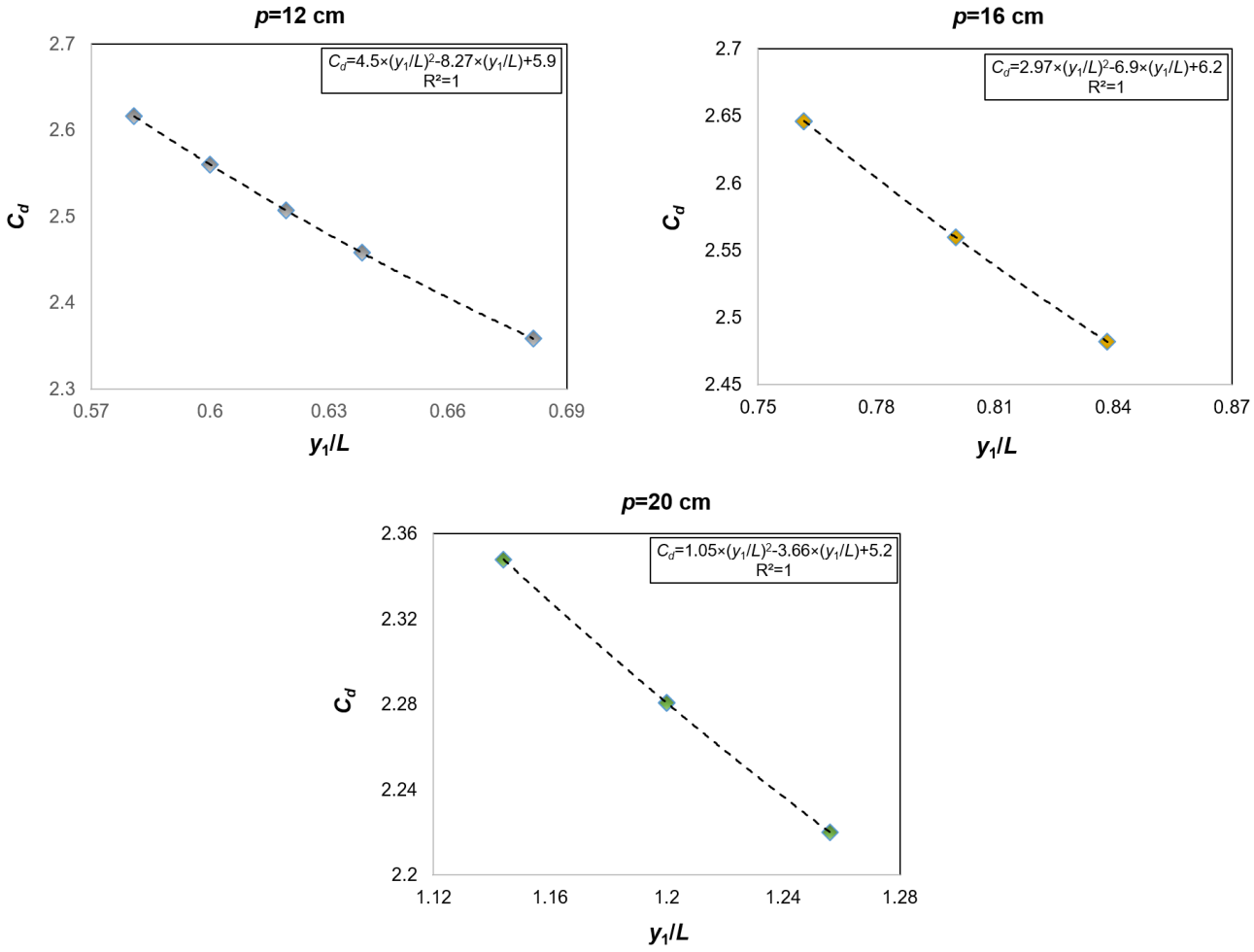


Figure 2. Comparison of experimental Cd values with the current and proposed equation results

Figure 2 summarizes the most important findings of this study. The graph compares experimental measurement points (blue dots) with existing approaches in the literature (De Marchi, Schmidt, Dominguez) and the new empirical equation developed within this study (red line). As is clearly evident from the graph, existing approaches in the literature (e.g., dashed and dotted lines) exhibit significant deviations from the experimental data, particularly in the low and high flow ranges. These deviations indicate that the existing equations do not adequately account for the specific flow geometry of triangular weirs and the resulting energy losses. The red line, represented by the proposed new empirical equation, follows a trend much closer to the experimental measurement points and perfectly models the behavior of the dataset. This high agreement demonstrates that the new equation will provide much more reliable results in the design and sizing of triangular side weirs compared to existing approaches in the literature.

This high agreement is achieved because the developed equation more accurately relates hydraulic parameters (e.g., the Froude number and the ratio of main channel depth to side weir height). The developed model can model the variable behavior of the flow coefficient more effectively by not only depending on the flow rate and water height but also using dimensionless parameters such as the Froude number of the fluid.

4. Conclusions and Recommendations

This study has made a significant contribution to the field of hydraulic engineering by experimentally determining the flow coefficient of triangular side weirs and developing a more reliable empirical equation based on these data. The findings provide practical guidance for the safer and more efficient design of such structures. The widespread

impact of the study is that it fills a significant gap in the existing literature and contributes to a better understanding of the hydraulic performance of side weirs.

4.1. Conclusions

- It has been experimentally confirmed that the flow coefficient of triangular side weirs varies depending on the main channel discharge and water level.
- It has been determined that existing analytical equations cannot fully model the complex flow mechanism of triangular side weirs.
- The new empirical equation developed using experimental data predicted the flow coefficient with significantly higher accuracy compared to existing approaches.

4.2. Recommendations

The findings of this study provide a strong basis for future research. The following future studies are recommended:

- The effects of side weirs with different triangle angles and different channel width/weir length ratios on the flow coefficient should be systematically investigated.
- Independent experiments should be conducted under real-world conditions or in different laboratory settings to increase the accuracy and general validity of the developed empirical equation.
- The application of the empirical equation can be expanded by investigating the effects of bottom slope and channel roughness on side weir flow.
- Comparative analyses can be conducted with experimental findings using numerical flow modeling (CFD) software, allowing for a more in-depth study of the three-dimensional behavior of the flow.

This study contributes to increasing scientific knowledge for the more effective and reliable design of water engineering structures.

Acknowledgements

This work was supported by the TÜBİTAK 2209-A Undergraduate Research Project Support Program under grant number 1919B012337844.

5. References

[1] De Marchi, G. (1936). Saggio di teoria del funzionamento degli stramazzi laterali. L'energia Elettrica, 13(12), 847-859.
[2] Dominguez, J. (1940). A Contribution to the Study of Side-Channel Weirs. Trans. of the Institution of Civil Engineers of Ireland, 66, 319-335.
[3] Francis, J. B. (1883). The Lowell Hydraulic Experiments. D. Van Nostrand.
[4] Gourley, H. J. F., & Crimp, W. H. (1915). Weir discharge coefficients: A new formula. Minutes of the Proceedings of the Institution of Civil Engineers, 199(2), 246-291.
[5] Henderson, F. M. (1966). Open Channel Flow. MacMillan.
[6] Jiwani, A. B., & Lucas, J. M. (2002). Discharge coefficient of sharp-crested weirs. Journal of Hydraulic Engineering, ASCE, 128(1), 1-8.
[7] Kindsvater, C. E., & Carter, R. W. (1957). Trapezoidal weirs. Journal of the Hydraulics Division, ASCE, 83(HY3), 1-32.
[8] Rehbock, T. (1929). Discussion of precise measurements, by K. B. Turner. Trans., ASCE, 93, 1143-1162.
[9] Singh, R., Manivannan, D., & Satyanarayana, T. (1994). Discharge coefficient of rectangular side weirs. Journal of Irrigation Drainage Engineering, ASCE, 120(4), 814-819.

[10] Şişman, H. Ç., Altan-Sakarya, A. B., & Aydın, İ. (2009). Dikdörtgen Kesitli Keskin Kenarlı Savaklarla Debi Ölçümü. IV. Ulusal Su Mühendisliği Sempozyumu, İstanbul, 27-35.
[11] Subramanya, K., & Awasthy, S. C. (1972). Spatially varied flow over side weirs. Journal of Hydraulics Division, ASCE, 98(HY1), 1-10.
[12] Swamee, P.K., Pathak, S.K., Mohan, M., Agraval, S.K., & Masoud, S.A. (1994). Subcritical flow over rectengular side weir. ASCE, 120(1), 212-217.
[13] Tokyay, N. D., & Turhan, Z. G. (2009). Küçük Savak Yükleri İçin Akım Katsayısı Değişimi. DSİ Teknik Bülteni, 89(125).
[14] Tullis, B. P., & Barfuss, S. L. (1995). Discharge coefficient for sharp-crested weirs. Journal of Irrigation and Drainage Engineering, ASCE, 121(2), 163-172.
[15] U.S. Bureau of Reclamation (USBR). (2001). Water Measurement Manual. Denver, Colorado.

International Conference on Advances and Innovations in Engineering

Bisphenol A (BPA): Properties, Environmental Impact, Health Risks, and Regulatory Actions

Meki EFE[1] , Sibel ASLAN [2]

[1,2] Department of Environmental Engineering, Faculty of Engineering, Firat University, Elazığ, Turkey.
[3]Department, Faculty, University, Province, Country.
[1]mekiefe@gmail.com, [2]sibela@firat.edu.tr
[1](ORCID: 0009-0002-1322-9421), [2](ORCID: 0000-0003-3454-798X)

Abstract
Bisphenol A (BPA) is a carbon-based chemical synthesized by the condensation of phenol and acetone in the presence of a strongly acidic ion-exchange resin acting as a catalyst. BPA is mainly utilized in the production of polycarbonates and epoxy resins, accounting for 65% and 30 % of its industrial use, respectively, as well as thermal paper, sulphides and pesticides, leather tanning agents, dye dispersants, fiber additives, medical devices, and electronic products. BPA is a widely used synthetic chemical that has raised concerns due to potential health risks and its widespread presence in the environment. It can enter the environment through various sources, such as industrial discharges, wastewater treatment plant (WWTP) effluents, landfill leachates, and the degradation of plastic materials over time. BPA has been detected in various environmental media, including air, soil, sediment, surface water, groundwater, biota, and food due to its high production, consumption, and subsequent release into the environment from various sources. BPA is not classified as a persistent organic pollutant (POP) because it has a short biological half-life. However, like many POPs, BPA is often grouped with them because it can accumulate biologically in human tissues and organs. This accumulation may contribute to the development of various disorders. BPA is often classified as an endocrine-disrupting compound (EDC) that can mimic or block natural hormones by binding with the estrogen receptors, thereby altering physiological and developmental processes in exposed organisms. BPA can bioaccumulate in aquatic and terrestrial organisms, leading to endocrine disorders, reproductive toxicity, neurotoxicity, and genotoxicity.

Keywords: bisphenol A (BPA), endocrine disruption, health concern, environmental risk.

1. Introduction

Bisphenol A (BPA) is a non-natural organic molecule with the chemical structure of two phenol rings connected by a central carbon atom [1, 2]. Due to its good chemical stability, BPA is widely used in the synthesis of polycarbonates, epoxy resins, and thermal paper [3, 4, 5]. However, the widespread use of BPA and its ubiquity in the environment are a cause for concern due to its potential adverse effects on human health and ecosystems [6, 7]. In particular, BPA is an EDC and is considered one of the most pervasive emerging pollutants in the environment [1, 8, 9]. Furthermore, it is potentially associated with numerous diseases such as breast cancer, infertility, headaches, dizziness, cognitive dysfunction, diabetes, cardiovascular diseases, and obesity [6, 10]. As a result, some countries have banned the use of BPA.

2. Characteristics of BPA

Definition and Chemical Characteristics of BPA

BPA is classified as a group of phenols [11] and is also known as 4,4'-isopropylidenediphenol and 2,2-bis(4-hydroxyphenyl)-propane. It is a chemical compound with a molecular weight of 228.29 Da and the chemical formula $(CH_3)_2C(C_6H_4OH)_2$ [3, 12]. The molecular structure of BPA is shown in Figure 1. BPA is colorless, odorless, and solid at room temperature [1, 13]. BPA exhibits low vapor pressure (5.3×10^{-6} Pa at 25 ◦C), moderate water solubility (0.3 g/L, at 25 ◦C), and low volatility [1, 2]. The octanol-water partition coefficient (log K_{ow}) of BPA is 3.32, indicating moderate hydrophobicity and good solubility in lipids. Due to its lipophilic nature, BPA has the ability to accumulate in various human and animal tissues [12].

[1]Corresponding author

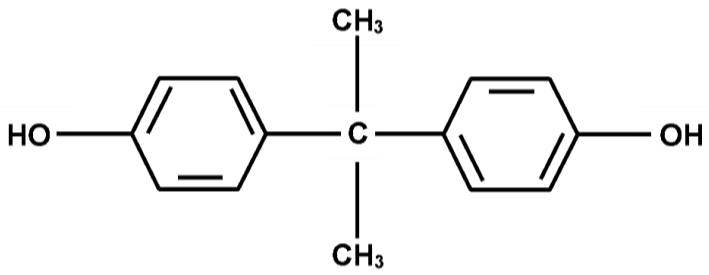


Figure 1. The molecular structure of BPA

Historical Context of BPA

BPA was initially synthesized in 1891 by the Russian chemist Aleksandr P. Dianin through the acid-catalyzed condensation of phenol and acetone [12]. However, the first mention of BPA appeared in a scientific paper in 1905 by Thomas Zincke from the University of Marburg, Germany [14]. In 1938, British chemist Charles Dodds uncovered the estrogenic activity of BPA because of the presence of hydroxyl groups in the para position, which is quite similar to estradiol. In 1953, Daniel Fox and Hermann Schell discovered BPA capacity for performing cross-linkages during the polymerization arising to polycarbonate. In 1957, the USA started the large-scale production of BPA [15].

3. Occurrence and Impacts of BPA in the Environment

Sources of BPA in the Environment

BPA is an important ingredient in the synthesis of in the synthesis of polycarbonates, epoxy resins, polyester resins, and thermal printer paper. Due to their superior mechanical strength across a broad temperature range (–40 to 145°C), hardness, acid resistance, and optical transparency, polycarbonates are extensively used in industry [11]. BPA is incorporated into numerous consumer products, incuding toys, plastic eating utensils, eyeglasses, reusable containers for food storage, food packaging materials, office products, and even in the thermal receipt paper to prevent contamination and extend shelf life [16]. The chemical composition of BPA comprises phenolic epoxy resins, which are utilized in the formulation of protective coatings for containers used in the food and beverage industry [7]. Furthermore, epoxy resins are used as coating for water pipes in houses and in the production of paints covering internal surfaces [11].

BPA is primarily utilized in the synthesis of polycarbonate and epoxy resins, which constitutes about 95% of its industrial applications. The residual 5% is allocated for utilizations including phenoplast resins, phenolic resins, unsaturated polyester resins, linings for cans, antioxidants, additives for PVC manufacturing and processing, ethoxylated BPA, dye developers for thermal paper, polyols, modified polyamide, compounding agents in automotive tires, flame retardants (notably tetrabromo BPA), equipment for automotive and transportation sectors, optical media (for instance, DVDs), electronic devices, construction materials, linings within potable water conduits, coatings for thermal and carbonless paper, and foundry casting [7].

Although BPA does not occur naturally, it has become pervasive in the environment as a result of high production, consumption, and subsequent release into the environment [17]. BPA can be introduced into the environment through various sources, including industrial discharges, effluents from wastewater treatment plants (WWTPs), landfill leachates, and the degradation of plastic materials over time [1, 5, 7]. Consequently, BPA has been detected in various environmental matrices such as soil, sediments, groundwater, surface water, and air, as well as in biota [1, 7]. The half-life of BPA in aquatic and terrestrial environments is approximately 4.5 days, whereas it is less than one day in the atmospheric context due to its low volatility. Nevertheless, the presence of BPA in the air can be ascribed to its adhesion to solid particulate matter suspended in the atmosphere [12].

BPA may be emitted into the atmosphere through the combustion of materials containing BPA, or it may be released into the environment in the form of gas. In the atmosphere, there are two predominant categories of BPA: BPA that is present in its form and BPA that is associated with particulate matter through adhesion [18].

Photodegradation and biodegradation are two mechanisms that contribute to the breakdown of this toxicant in aquatic ecosystems. Photodegradation serves as the main non-biological pathway for BPA degradation in these environments [19].

Environmental and Ecotoxicological Risks Associated with BPA

Although BPA is not classified as a persistent organic pollutant (POP) due to its relatively short half-life, it is often included with other POPs, because it can accumulate in human tissues and organs, potentially contributing to the pathogenesis of various disorders [12]. BPA has been widely used in the manufacture of polycarbonate plastics and epoxy resins. There are numerous sources of exposure, including dental sealants, canned foods and beverages, polycarbonate water dispensers, medical devices in neonatal intensive care units, vinyl products such as shower curtains and pillow protectors, cleaning products like dish and laundry detergents and tub and tile cleaners, personal care products such as sunscreen lotions, nail polish, body washes/lotions, bar soaps, shampoos, conditioners, shaving creams, and face lotions/cleansers, thermal paper, and toys [1]. Humans are exposed to varying concentrations of BPA (0.27-10.6 ng/ml) directly or indirectly through various routes, including oral ingestion, maternofetal transmission, inhalation, dermal absorption, and eye contact [12, 20-22].

BPA can cause adverse health effects in humans even at very low concentrations [21]. BPA has the ability to traverse the placental barrier and has been identified in maternal and fetal serum as well as in the human placenta. Consequently, BPA can infiltrate various tissues and fluids within the human uterus. Moreover, BPA may also be absorbed through inhalation or dermal exposure. Due to its structural similarity to hormones, BPA can stimulate cellular responses by binding with the estrogen receptors, thereby affecting body weight and tumorigenesis. BPA may also affect metabolism and cancer progression by interacting with GPR30 and may impair male reproductive function by binding to androgen receptors [12, 21]. Some health problems associated with BPA include the following:

- hormone-related cancers like breast cancer or prostate cancer,
- infertility,
- foetal growth restriction,
- feminization of male foetuses,
- early puberty,
- endometriosis,
- cardiovascular problems,
- obesity, diabetes, and overweight,
- auto-immune disease,
- brain disorders,
- hyper-insulinemia,
- glucose intolerance,
- insulin resistance,
- anxiety and depression,
- allergic contact dermatitis [2, 20, 21].

BPA poses significant ecological risks, particularly to aquatic ecosystems. It can cause disturbances in reproductive development, estrogenic effects, damage to the nervous system, and negative impacts on survival and growth [5]. Due to its lipophilic nature, BPA has the potential to bioaccumulate in aquatic organisms, raising concerns regarding biomagnification and long-term ecological consequences within aquatic ecosystems. Exposure to BPA in aquatic organisms has been associated with reproductive dysfunctions, developmental abnormalities, and behavioral patterns, thereby disrupting the integrity of aquatic ecosystems [2].

4. Regulatory Framework and Guidelines

Based on the health effects mentioned above, some countries have banned the use of BPA in certain consumer product categories such as food containers and baby bottles. The Canadian government was the first to declare BPA a 'dangerous substance' and prohibited the sale, import, and advertisement of baby bottles that contained BPA. In 2011, the European Commission decided to ban the use of BPA in polycarbonate baby bottles [23]. The European Chemicals Agency (ECHA) has incorporated BPA into the inventory of substances of significant concern owing to its toxicological properties, and numerous regulatory bodies have instituted specific limits on its utilization and migration. The United States Food and Drug Administration (FDA) published a report identifying potential health risks associated with BPA exposure [7]. In 2012, the FDA amended its regulations to no longer allow the use of BPA-based for polycarbonate resins in baby bottles and sippy cups, and the epoxy resins as a coating in packaging for infant formula in 2013 [15]. Most recently, on December 19, 2024, the European Commission adopted Regulation (EU) 2024/3190, which prohibits the use of BPA and other bisphenols in food contact materials-such as can coatings, reusable plastic bottles, and kitchenware-throughout the European Union [24, 25].

5. References

[1] Bhatnagar A, Anastopoulos I. "Adsorptive removal of bisphenol A (BPA) from aqueous solution: A review". *Chemosphere*, 168, 885-902, 2017.

[2] Matharage H, Jayaweera M, Bandara N, Manatunge J, Jayawardana D, Dissanayake J. "From waste to resource: King coconut biochar as a green adsorbent for bisphenol A removal". *Case Studies in Chemical and Environmental Engineering*, 12, 101261, 2025.

[3] Michałowicz J. "Bisphenol A–Sources, toxicity and biotransformation". *Environmental Toxicology and Pharmacology*, 37(2), 738-758, 2014.

[4] Kundu S, Das BK, Wodeyar A, Majumder P, Jana S, Biswas A, Das S, Besra R. "Clearing the path: Unraveling bisphenol A removal and degradation mechanisms for a cleaner future". *Journal of Environmental Management*, 373, 123558, 2025.

[5] Santos CR, Arcanjo GS, Araújo AA, Santos LV, Amaral MC. "Occurrence, environmental risks, and removal of bisphenol A and its analogues by membrane bioreactors". *Chemical Engineering Journal*, 494, 153278, 2024.

[6] Catenza CJ, Farooq A, Shubear NS, Donkor KK. "A targeted review on fate, occurrence, risk and health implications of bisphenol analogues". *Chemosphere*, 268, 129273, 2021.

[7] Datta S, Chauhan A, Ranjan A, Sardar AH, Tuli HS, Ramniwas S, Shahwan M, Sharma U, Jindal T. "A comparative review on bisphenol A sources, environmental levels, migration, and health impacts in India and global context". *Nature Environment and Pollution Technology*, 23(2), 2024.

[8] Cheng F, Wang J. "Removal of bisphenol A from wastewater by adsorption and membrane separation: Performances and mechanisms". *Chemical Engineering Journal*, 484, 149414, 2024.

[9] Halligudra, G, Sabbanahalli C, Singh S, Choudhary P, R M, Tigari G, Shantharaja, Lo SL. "Catalytic behaviour of iron-based nanomaterials for the remediation of hazardous chemicals from wastewater: A Review". *Journal of Physics and Chemistry of Solids*, 203, 112735, 2025.

[10] Tokula BE, Dada AO, Inyinbor AA, Obayomi KS, Bello OS, Pal U. "Agro-waste based adsorbents as sustainable materials for effective adsorption of bisphenol A from the environment: A review". *Journal of Cleaner Production*, 388, 135819, 2023.

[11] Mikołajewska K, Stragierowbi picz J, Gromadzińska J. "Bisphenol A–application, sources of exposure and potential risks in infants, children and pregnant women". *International Journal of Occupational Medicine and Environmental Health*, 28(2), 209-241, 2015.

[12] Cimmino I, Fiory F, Perruolo G, Miele C, Beguinot F, Formisano P, Oriente F. "Potential mechanisms of bisphenol A (BPA) contributing to human disease". *International Journal of Molecular Sciences*, 21, 5761, 2020.

[13] Besaratinia A. "The state of research and weight of evidence on the epigenetic effects of bisphenol A". *International Journal of Molecular Sciences*, 24, 7951, 2023.

[14] Almeida S, Raposo A, Almeida-González M, Carrascosa C. "Bisphenol A: Food exposure and ımpact on human health". *Comprehensive Reviews in Food Science and Food Safety*, 17(6), 1503-1517, 2018.

[15] Pereira, JPJ, Moura CS, Ayres C, Stávale LM. "An innovative and accessible chemical approach to bisphenol identification on plastic surfaces". *Revista Virtual de Quimica*, 13(1), 1-8, 2021.
[16] Ramakrishna MG, Girigoswami A, Chakraboty S, Girigoswami K. "Bisphenol A-An overview on its effect on health and environment". *Biointerface Research in Applied Chemistry*, 12(1), 105-119, 2022.
[17] Corrales J, Kristofco LA, Steele WB, Yates BS, Breed CS, Williams S, Brooks BW. "Global assessment of bisphenol A in the environment: Review and analysis of its occurrence and bioaccumulation". *Dose-Response*, 13(3), 1559325815598308, 2015.
[18] Xing J, Zhang S, Zhang M, Hou J. "A critical review of presence, removal and potential impacts of endocrine disruptors bisphenol A". *Comparative Biochemistry and Physiology, Part C*, 254, 109275, 2022.
[19] Biswal A, Srivastava P, NS, Kumar M. "Plastic originated bisphenol A: A potential endocrine disruptor for fish reproduction". *International Journal of Chemical Studies*, 8(1), 2085-2089, 2020.
[20] Ohore OE, Zhang S. "Endocrine disrupting effects of bisphenol A exposure and recent advances on its removal by water treatment systems. A review". *Scientific African*, 5, e00135, 2019.
[21] Abraham A, Chakraborty P. "A review on sources and health impacts of bisphenol A". *Reviews on Environmental Health*, 35(2), 201-210, 2020.
[22] Akash MSH, Rasheed S, Rehman K, Imran M, Assiri MA. "Toxicological evaluation of bisphenol analogues: preventive measures and therapeutic interventions". *RSC Advances*, 13, 21613–21628, 2023.
[23] Sharma P, Sharma K, Sharma G, Chadha P. "A review on the occurrence, exposure, and health impacts of bisphenol A". *Toxicology International*, 28(4), 337-356, 2021.
[24] https://foodpackagingforum.org/news/european-commission-bans-bisphenol-a-in-food-contact-materials (15.08.2025)
[25] https://eur-lex.europa.eu/eli/reg/2024/3190/oj/eng (15.08.2025)

Chemical Aspects of the Vanadium (V)-Oxide Extraction Process From Alunite Ore

Arzu Osmanova[1]

[1]Complex processing of mineral raw material, Chemistry, Institute of Catalysis and Inorganic Chemistry named after academician M.Nagiev, Baku, Azerbaijan
arzuo84@mail.ru,
[1](ORCID: 0009-0002-5754-7122)

Abstract
The new concept envisaged in the processing of alunite ore is its direct processing with an alkaline solution without burning. Since the alkaline aluminate solutions obtained at this time are periodically returned to the process, the amount of vanadium in the solution can be gradually and significantly increased. The dependence of the vanadium yield into solution during the dissolution of raw alunite ore on the concentration of alkali, temperature and the ratio of solid and liquid phases was studied.
The separation of vanadium from concentrated aluminate solutions by precipitation and from liquid solutions by adsorption has been investigated. The optimal conditions for the precipitation of vanadium in the form of meta- and polyvanadate have been found. Additionally, the phase compositions of the precipitates have been studied by physicochemical analysis methods and the kinetic parameters of the process have been determined.
The adsorption conditions of vanadium from model and technological solutions using iron-activated carbon (Fe-AC) nanocomposite have been found and the adsorption process has been clarified based on the construction of sorption isotherms.

Keywords: alunite, vanadium, Fe-AC nanocomposite, alkali leaching.

1. Introduction

The application of vanadium is constantly increasing all over the world. Thus, vanadium is widely used mainly in metallurgy, chemical production, aerospace due to its high physical and chemical properties. It is used as a key strategic resource in various technological areas, in steelmaking, petrochemicals and machine building [1,2].
Vanadium is used in oxidation-reduction batteries, steel production, lithium-ion batteries, titanium alloys, and vanadium salts as precursor reagents for cathode production. 80% of vanadium is mainly sold as ferrovanadium or is used as a strengthening additive in steel production. Vanadium is used in the production of stainless steel tools and steel tools. Compounds containing the vanadyl ion are mainly used for medical and biological purposes, mainly oxosulfates. Among vanadium compounds, V2O5 oxide plays an important role mainly in ceramics and as a catalyst in the production of sulfuric acid and maleic anhydride. Due to its universal chemical properties, vanadium is used in the production of ion batteries, which contributes to the development of technologies aimed at reducing CO2 emissions. The use of vanadium in oxidation-reduction and ion batteries leads to energy savings and is considered an alternative energy source. Thus, vanadium protects the environment from 185 million tons of CO2 emissions emitted into the atmosphere during various processes per year [3].

Vanadium plays an important role in non-ferrous and ferrous metallurgy due to its physical properties such as hardness, tensile strength, high temperature resistance and toughness. The use of vanadium in ferrous metallurgy accounts for 85% of the world's vanadium production. Vanadium used in construction gives materials high hardness. Alloys of vanadium with iron, titanium, nickel, aluminum, chromium and other metals are used in train tracks, in steel production, and as catalysts in the aerospace field, and many of its alloys are mainly biomedically safe materials and play an important role in production [4]. Vanadium is considered an indispensable raw material in the production of titanium alloys for aerospace applications. Globally, the main sources of vanadium supply include ores, slags, concentrates, metallurgical slags, and oil residues.

In industry, vanadium is primarily used in alloys in the form of solvan (25–30% V), ferrovanadium (35–80% V), korvan (83–86% V), and nitrovan (78–80% V) [5]. FeV alloys play an important role in steel production. The most widely used alloy is FeV_{80}, which contains 80% vanadium. Other existing alloys include FeV_{50}, FeV_{40}, and FeV_{60}.

[1]Corresponding author

2.Materials and Methods

Alunite ore is extracted from the Zəylik deposit in Azerbaijan. The chemical composition of alunite is represented by the following formulas: $KAl_3(SO_4)_2(OH)_6$, $K_2SO_4 \cdot Al_2(SO_4)_3 \cdot 4Al(OH)_3$, and $K_2O \cdot 3Al_2O_3 \cdot 4SO_3 \cdot H_2O$.
According to the results of X-ray fluorescence (XRF) analysis of an alunite ore sample obtained from the Ganja Clay and Soil Plant, the composition by mass is as follows (%): Al_2O_3 – 26.44; SiO_2 – 33.37; Fe_2O_3 – 3.975; TiO_2 – 0.482; Cr_2O_3 – 0.41; K_2O – 5.17; Na_2O – 1.78; SO_3 – 27.08; V – 0.0487; Ga_2O_3 – 0.003; MnO – 0.006; CaO – 0.0657; As_2O_3 – 0.0045; SrO – 0.1212; P_2O_5 – 0.215; F – 0.29.

Based on the results of X-ray diffraction (XRD) phase analysis, the main minerals in the alunite ore sample are: alunite (35.5%), kaolinite (dickite) (9.8%), quartz (53.4%), and hematite (1.5%) (Figure 1.).

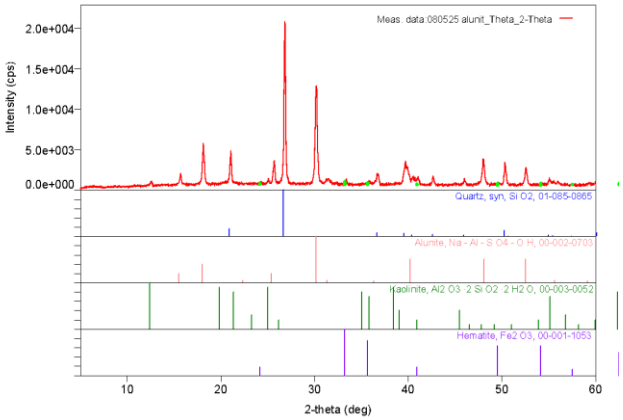


Figure 1. X-ray Diffraction (XRD) Analysis of Raw Alunite Ore

The composition of the Zaglik alunite sample was investigated using Scanning Electron Microscopy (SEM) and Energy Dispersive Spectroscopy (EDS) analyses. (Figure 2.).

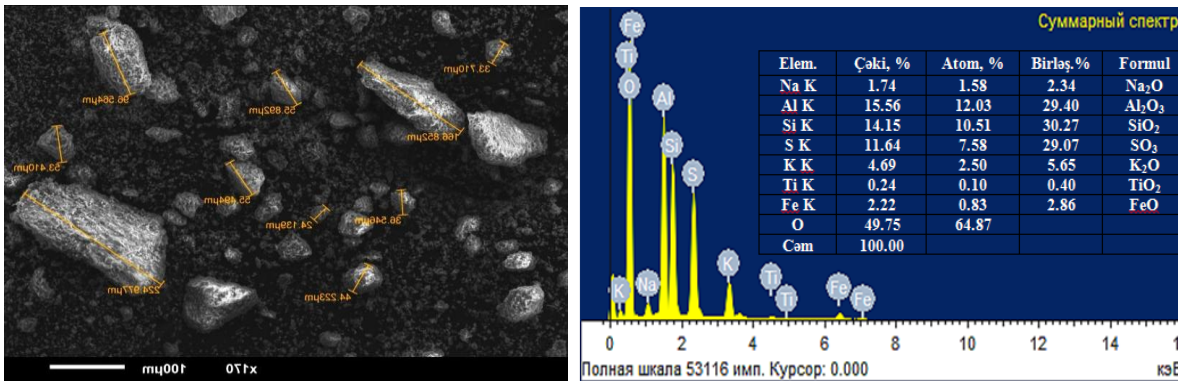


Figure 2. SEM and EDS Analysis of Raw Alunite Ore

The physical and chemical properties of alunite ore and the substances obtained during its processing, including vanadium-containing compounds, are investigated using a range of analytical methods. The determination and separation of vanadium are studied by the following methods:

Particle size determination: The degree of fineness of particles in the samples is determined using a SEISHNLMS-30 laser particle size analyzer (Japan).

Phase identification and mineralogical composition: The mineralogical composition and dispersion of the material are determined using X-ray diffraction (XRD) analysis. For this, the XRD analysis of the sample is performed using the AXS 200 diffractometer (Bruker-AS 200, Germany) with CuKα radiation.

Elemental analysis in solution: The concentrations of elements in the solution are measured using a Bruker S2-Picofox X-ray fluorescence spectrometer (Germany) and for comparison, a FEK-2 instrument.

Sample morphology: The morphology of the sample is examined using a scanning electron microscope (SEM/EDS, Hitachi-4800, Germany).

3.Results and Discussion

The Dissolution of Alunite Ore in Alkali Solutions. The dissolution of alunite ore samples in alkaline medium was carried out by varying a wide range of parameters: NaOH concentrations of 2.5%, 5%, 10%, and 20%; temperatures of 50°C, 60°C, 70°C, 80°C, and 90°C; solid-to-liquid ratios (S\:L) of 1:2, 1:4, 1:6, and 1:8 mL/g; and stirring durations of 0.5, 1, 1.5, 2, and 3 hours. The separation of vanadium from alunite was performed in a mechanical stirrer with continuous mixing.

An increase in NaOH concentration from 2.5% to 20% led to an increase in the vanadium extraction from 9.9% to 57.2%. Based on the experimental studies, it was found that the highest vanadium yield during 1-hour leaching at 60°C was achieved in a 5% NaOH solution. Additionally, when using 10% NaOH solution, increasing the temperature from 50°C to 90°C enhanced vanadium extraction into the solution. At 70°C, the vanadium extraction reached 67.52%, and raising the temperature to 90°C further increased this to 71.30%.

The optimal conditions for vanadium extraction from raw alunite—aligned with conditions that ensure complete dissolution of aluminum—are as follows: solid-to-liquid ratio: 1:6, Stirring speed: 700 rpm, NaOH concentration: 10%, temperature: 90°C, Leaching time: 120 minutes.

Under these conditions, up to 98% of the alunite mineral dissolves, and 71.30% of vanadium is extracted into the solution, with the vanadium concentration in the solution reaching 49.875 mg/L.

While the precipitation of aluminum hydroxide occurs at pH = 10.6 during the neutralization of obtained alkaline aluminate solutions with 1N H_2SO_4 solution, the precipitation of gallium(III) hydroxide occurs at pH = 9.4–9.7. The chemical reactions occurring during acid neutralization can be represented as follows:

pH=12.8-9.5 $Na[Al(OH)_4] + H^+ \rightarrow Al(OH)_3 + H_2O + Na^+$
pH=9.5-7 $Na[Ga(OH)_4] + H^+ \rightarrow Ga(OH)_3 + H_2O + Na^+$

After the precipitation of aluminum and gallium hydroxides from the aluminate solution, the mother liquor at pH = 7 consists of sodium vanadate, sodium sulfate, and potassium sulfate salts. At vanadium concentrations up to 0.6 g/L, no precipitation occurs in the solution. After the aluminum hydroxide precipitate is separated, the solution is heated to boiling temperature (25–130°C), evaporated, and then cooled to a temperature of 30–40°C. During this process, the alkaline metal sulfate salts crystallize and are separated from the sodium vanadate salt.During the complex processing of alunite ore, the concentration of vanadium in the recycled alkaline aluminate solutions increases and reaches up to 0.38–5 g/L. This concentration has been confirmed by the results of our experiments. After desiliconization of the solution, aluminum hydroxide is precipitated. At these stages, the vanadium concentration in the solution does not change and reaches 0.678 g/L. The vanadium-containing solution is evaporated at 125–130°C. In the solutions obtained after evaporation, vanadium is present in the form of $NaVO_3$ and KVO_3 salts, and its concentration reaches 1.5–2 g/L. After crystallization and the precipitation of alkaline metal sulfate salts from these solutions, the vanadium concentration increases up to 6.95 g/L. As can be seen, during the processing stages of alunite ore, the concentration of vanadium varies between 0.678 g/L and 6.95 g/L.

The precipitation of vanadium from a vanadium-containing concentrate composed of vanadates (sodium, potassium) in the form of ammonium metavanadate by ammonium sulfate has been studied. The experiments were conducted in solutions where the initial concentration of $NaVO_3$ was 6.9 g/L, at pH = 8, over a period of 2 hours [6]. Initially, no precipitation is observed; however, after a certain period (10 minutes), a white precipitate of vanadium begins to form. Based on the low values of the activation energy for the precipitation reaction (Ea = 16.39 kJ/mol, 8.77 kJ/mol), it was determined that the process proceeds in the diffusion-controlled region. A suspension of ammonium metavanadate in water is prepared and heated to 70°C. When mixed rapidly (200 rpm) in a ratio of B\:M = 1.6:5, ammonium polyvanadates are formed. The resulting slurry is then heated to 90–95°C and neutralized with sulfuric acid solution to a pH of 2. During this process, ammonium metavanadate reacts with sulfuric acid and is converted into hexavanadate:

$$6NH_4VO_3 + 2H_2SO_4 \rightarrow (NH_4)_2V_6O_{16} + 2(NH_4)_2SO_4 + 2H_2O$$

The phase and dispersive composition of both precipitates were examined using X-ray phase analysis, electron microscopy, and IQ spectroscopy methods. Figure 1 shows the diffractograms of the precipitates obtained at pH = 8 (a) and pH = 2 (b).

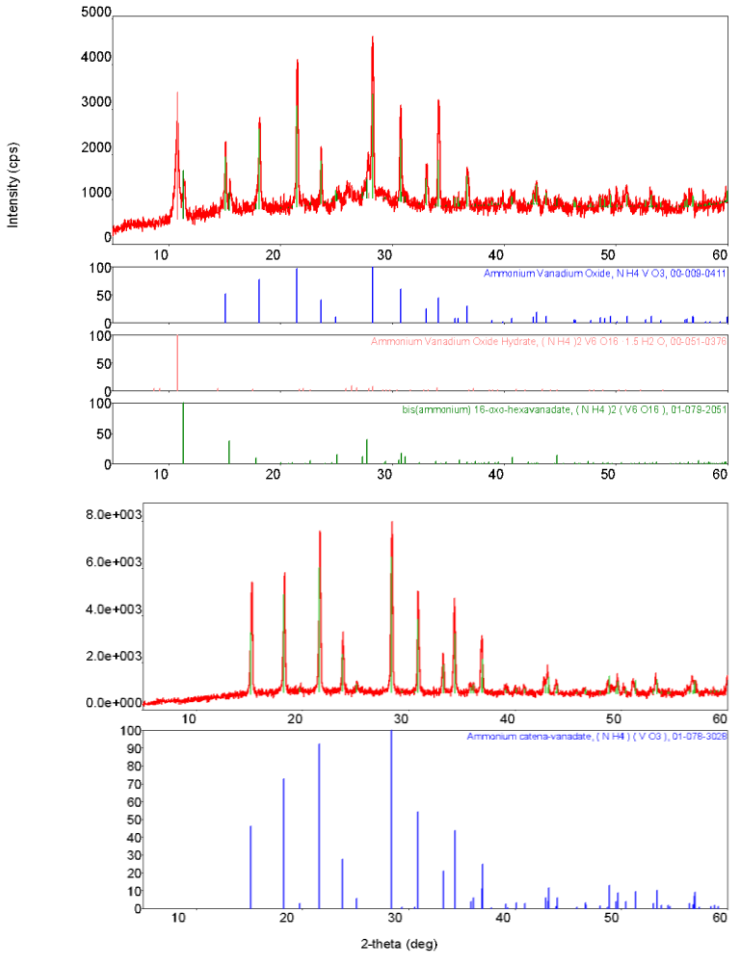


Figure 1. X-ray phase analysis of the precipitate obtained at different pH values: pH = 8 (a), pH = 2 (b).

The washed polyvanadate ammonium is dried and thermally decomposed in a muffle furnace at 550°C to obtain V_2O_5:

$$(NH_4)_2V_6O_{16} \xrightarrow{t} 2NH_3+3V_2O_5+H_2O$$

The separation of vanadium from clear solutions is primarily carried out using an adsorption method based on Fe-AC nanocomposite (iron-activated carbon nanocomposite). The optimum pH values corresponding to the maximum adsorption capacity are determined to be pH = 4 for activated carbon and pH = 4.5 for the iron-activated carbon

nanocomposite [7]. During vanadium adsorption with activated carbon, the adsorption capacity is 37 mg/g, whereas due to the presence of FeOOH groups in the nanocomposite (Fe-AC), the adsorption capacity increases up to 110 mg/g.

4. Conclusion

During the alkaline processing of alunite ore, the chemical aspects of vanadium (V) oxide extraction were studied, leading to the following conclusions: For the first time, the theoretical and experimental aspects of vanadium extraction through the complex processing of Zeylik alunite ore were systematically investigated. It was determined that at 90°C, over 2 hours, with a solid-to-liquid phase ratio of 1:6 and a stirring speed of 700 rpm, 10% alkaline solution dissolves up to 98% of the alunite mineral into the solution. Under these optimal conditions, 90.6% of gallium and 71.30% of vanadium pass into the aluminate solution, increasing the gallium concentration in the solution to 4.73 mg/L and vanadium concentration to 49.875 mg/L.

By adding ammonium sulfate to the sodium vanadate solution, the optimal conditions for precipitation of vanadium as NH_4VO_3 and $(NH_4)_2V_6O_{16}$ salts were determined. The individuality of the precipitates was studied by physicochemical analysis methods, and the kinetic parameters of the precipitation were identified. It was found that when ammonium sulfate is added in double the vanadium concentration, at pH = 8 and 25°C, ammonium metavanadate NH_4VO_3 precipitates up to 91% within 3.5 hours. When a suspension of ammonium metavanadate in water is mixed with sulfuric acid at 90–95°C and pH = 2, red-brick colored ammonium polyvanadate precipitates form. According to the kinetic parameters of precipitation (n = 0.46–0.64, Ea = 16 kJ/mol, 8.77 kJ/mol), the process likely occurs in the diffusion-controlled region.

The optimal conditions for vanadium adsorption using iron-activated carbon (Fe-AC) nanocomposite as an adsorbent from model and technological solutions were established. It was found that vanadium ion adsorption depends both on the pH of the solution and the concentration of vanadium ions. At pH = 4.5, vanadium in concentrations of 50–100 mg/L is adsorbed up to 88% from the solution within 3 hours. Adsorption results were analyzed according to Langmuir and Freundlich isotherms, revealing that while the adsorption capacity with activated carbon is 37 mg/g, with Fe-AC nanocomposite it reaches 110 mg/g. Finally, the chemistry of the adsorption process was elucidated.

5. References

[1] Summerfield, D. Bibliographical reference: Australian Resource Reviews: Vanadium. / D.Summerfield // Geoscience Australia – 2019. Canberra. http://dx.doi.org/10.11636/9781925848274

[2] M.Mohammadi, M.M.Salarirad. Kinetics of direct leaching of natural alunite in KOH / // Ind. Eng. Chem. Res. – 2013, 52 (40), – p. 14359-14365.

[3] Santos, D.A. Assessing the role of vanadium technologies in decarbonizing hard-to-abate sectors and enabling the energy transition / D.A.Santos, K.M.Dixit, P.P.Kumar [et al.] // İScience – 2021. V.24, I.11. – p. 103277.

[4] Y.L.Hao, S.J.Li, R.Yang Biomedical titanium alloys and their additive manufacturing. / // Rare Metals – 2016. – 35 (9), – p. 661-671.

[5] L.Snead, T.David, D.T.Hoelzer L.L. Chapter 13 - Refractory Alloys: Vanadium, Niobium, Molybdenum, Tungsten // Structural Alloys for Nuclear Energy Applications, – Elsevier: – 2019, – p. 585-640.

[6] Osmanova, A.Kh. Optimization of conditions for ammonium metavanadate and polyvanadate precipitation from alunite processing solutions // Chemistry of Sustainable and Development – 2023. 31, – p.1–8.

[7] Osmanova, A.Kh. Vanadium (V) adsorption from aluminate solutions by an Fe-AC nanocomposite / A.Kh.Osmanova, A.A.Geidarov, G.I.Alyshanly [et al.] // Russian metallurgy – 2022. № 3. – p. 211-216.

Dye Economy in Endemic and Medicinal Plants: An Analysis Based on *Isatis, Reseda,* and *Rubia*

Nevzat DOGAN[1] , Inanc OZGEN [2]

[1,2] Biotechnology Department, Faculty of Bioengineering, Firat University, Elazig, Turkey.
[1]nevzatdogan2323@gmail.com, [2]inancozgen@gmail.com,
[1](ORCID: 0009-0009-1862-5125), [2](ORCID: 0000-0003-1742-9324)

Abstract
This study aims to compare the cultivation and dye extraction processes of three natural dye plants grown in Elazig Province *Isatis constricta* (Hazar Indigo), *Reseda lutea* (Weld), and *Rubia tinctorum* (Madder) from an economic perspective. The analysis is based on demonstration trials conducted by the Elazig Provincial Directorate of Agriculture and Forestry between 2015 and 2023, as well as data from contract farming practices. The study analyzes parameters such as yield per decare, production costs, and powder dye output for each plant species. A cost-benefit analysis was performed accordingly. The findings indicate that *Isatis constricta* provides the highest dye income and profit under maximum yield scenarios, though it carries a risk of loss under minimum yield conditions. Reseda lutea has shown consistent and positive profitability in both low- and high-yield conditions, making it a low-risk option. Rubia tinctorum yields satisfactory profits under high-yield scenarios but requires long-term investment due to its three-year cultivation cycle. According to 2025 data, unit costs decrease as yield increases, while lower yields lead to higher unit costs. Comprehensive cost analyses show that, with proper planning, natural dye plants can be integrated into competitive and sustainable production models. Moreover, natural dyes have significant potential to be transformed into value-added products in sectors such as textiles, cosmetics, pharmaceuticals, and food.

Keywords: natural dyes, *Isatis constricta*, cost of dyes, *Rubia tinctorum, Reseda lutea*, pigment,

1. Introduction

Plant-based production costs were calculated based on the specific cultivation criteria of each dye plant. Since the growth conditions and harvesting periods vary for each species, especially Rubia tinctorum (madder), which can only be harvested at the end of the third year when grown from seed, represents a significant cost-increasing factor. By defining suitable production parameters, the cultivation costs per unit area were calculated. It was determined that factors such as the size of the cultivation area, use of equipment, availability of harvesting machinery, and labor input directly affect production costs. Additionally, the yield per unit area has been identified as a critical determinant of cost. Cost-benefit analyses, commonly used in crop production, revealed that as harvest yield increases, unit production costs decrease. The ratio of crop-specific cultivation costs to product quantity determines the unit production cost, emphasizing the importance of competitive pricing strategies. Reducing unit production costs enhances global market competitiveness. Natural dyes have strong potential for use in various sectors such as textiles, cosmetics, food, and pharmaceuticals, supporting the development of value-added products. The extraction of dye from plants represents a crucial stage in the production process. As the amount of dye extracted increases, the resulting products become more competitive and sustainable.

Cost, in economic terms, is defined as the sum of measurable sacrifices made to acquire a good or service with market value [1]. A well-structured income and cost database not only helps measure the economic performance of agricultural enterprises but also points to the root causes of problems and supports the development of solutions. The most effective method for accurately calculating production costs in agricultural businesses is the use of cost accounting techniques [2].

Agricultural cost accounting also plays a vital role in credit and financing processes. It can inform government policies and assist in determining income in sharecropping arrangements [3]. This branch of accounting records the financial transactions of agricultural enterprises, summarizes and analyzes them appropriately, and enables the calculation of unit product costs [4; 5].

In the context of target costing, accurate forecasting of market share and sales volume is crucial. Agricultural production should not be evaluated solely within the input-production-output cycle, as it is significantly affected by

[1]Corresponding author

external conditions such as climate, pollution, and global warming (Çetin & Bahşi, 2019). In conclusion, accurate cost calculations in dye production contribute to determining competitive price thresholds. Reducing the unit costs of dye plants supports the efficient and sustainable operation of agricultural enterprises.

2. Materials and Methods

This study is based on the research and demonstration activities conducted by the Elazig Provincial Directorate of Agriculture and Forestry between 2015 and 2023. Within the scope of the study, demonstration trials were carried out in suitable areas across Elazig to evaluate the cultivation potential of various dye plants. Data on implementation processes and yield records were collected. Additionally, prior to the COVID-19 pandemic, a contract farming model was established with a company operating in the natural dye and textile sector. Updated cost and contract data from this collaboration were also included in the study. The collected data cover various parameters such as the cultivation cost components of natural dye plants, current market prices, global pigment prices, importer and exporter countries, and the demands of companies that influence the market dynamics in these regions.

3. Results

This study presents data on the yield, production costs, harvested quantity, cost per unit area, pigment content, and dye yield per unit area of natural dye plants cultivated in Elazig Province. Most notably, the endemic species of the region, *Isatis constricta* PH Davis (Hazar Indigo), along with other dye plants, were cultivated under real farmer conditions, and their dye yields were determined. Detailed cost tables were prepared for each plant. For *Isatis constricta*, which produces a high amount of blue dye, it was concluded that new techniques are needed to improve dye extraction efficiency. The average yields recorded under field conditions are as follows:

***Isatis constricta* (Hazar Indigo):** Minimum 650 kg and maximum 1250 kg of fresh leaves per decare,

***Reseda lutea* (Weld):** Minimum 160 kg and maximum 447 kg of dried and ground product per decare,

***Rubia tinctorum* (Madder):** Minimum 300 kg and maximum 850 kg of washed and dried roots per decare in the third year of cultivation.

According to 2025 data, the cost of producing one kilogram of dye plant decreases as yield per unit area increases, while lower yields lead to higher unit production costs. The cost, dye yield, and production analysis of *Isatis constricta*, *Reseda lutea*, and *Rubia tinctorum* are presented in detail in Tables 1, 2, and 3, respectively.

Table 1. *Isatis constricta* (hazar indigo): dye cost, yield analysis, and production costs

Cost criterias for cultivation and dye production	Minimum	Maximum
Cultivation and Dye Extraction yield from 1 kg of fresh leaves (Blue) (gr)	3,5 gr	6 gr
Total cultivation fresh leaves of per decare (kg)	650 kg	1250 kg
Amount of powder dye extraction obtained per decare (kg)	2,27 kg	7,5 kg
Total cultivation cost per decare (Leaves) ($)	224,5 $	285,1 $
Cost per kg of powder dye (leaf + without processing) ($)	0,345 $	0,228 $
The powder of 1 kg dye production cost (leaves + processing) $	99 $	45 $
Sales price of 1 kg dye powder extract ($)	150 $	250 $
Amount of fabric dyeable with 1 kg of powder dye (kg)	10 kg	15 kg
Income dye powder extract obtained from per decare (kg)	340,5 $	1875 $

The production scenario of Hazar Indigo (*Isatis constricta*) is presented in Table 1.

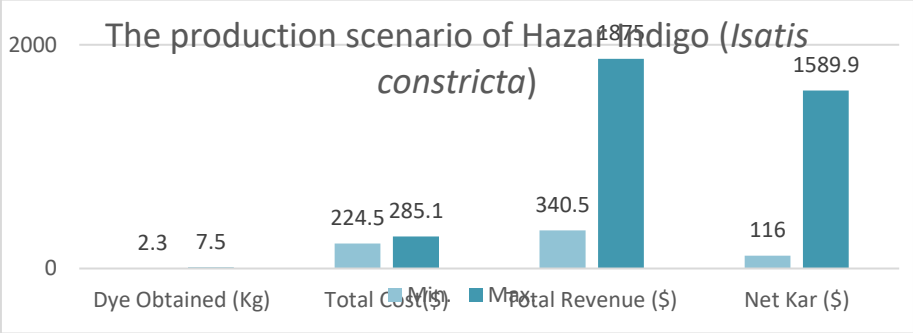


Figure 1. Hazar indigo (*Isatis constricta*) production and yield scenario

In Figure 1, the minimum and maximum yield scenarios for *Isatis constricta* (Hazar Indigo) are compared. In the maximum yield scenario, the amount of dye obtained and the resulting income increase significantly, while costs remain relatively controlled leading to a high net profit. In contrast, in the minimum yield scenario, the costs exceed the income, resulting in a loss.

Table 2. *Reseda lutea* (weld): dye cost, yield analysis, and production cost

Cost criterias for cultivation and dye production	Minimum	Maximum
Cultivation and Dye Extraction Costs Yield per decare (dried and ground product)	160 kg	447 kg
Cost per kg of dried and grounded product ($)	269,2 $	351,5 $
Cost per kg of powder dye ($)	1,68 $	0,79 $
Sales price per kg of powder dye ($)	90 $	45 $
The powder of 1 kg dye production cost $	150 $	220 $
Amount of fabric dyeable with 1 kg of powder dye (kg)	8 kg	12 kg
Yield per decare dried and ground product	960 $	1963,5 $

The production scenario of *Reseda lutea* (Weld) is presented in Figure 2.

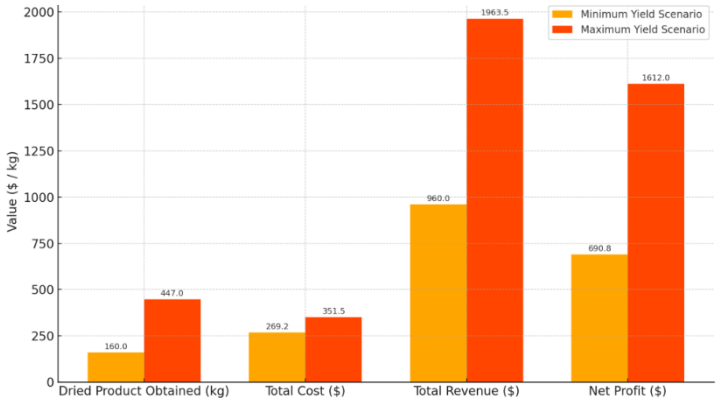


Figure 2. Production and yield scenario of *Reseda lutea* (weld)

In the maximum scenario, product yield and dye income are high, resulting in a highly satisfactory net profit. Even in the minimum scenario, the income exceeds the cost, providing a positive net profit. This indicates that *Reseda lutea* is a low-risk and economically sustainable natural dye plant.

Table 3. *Rubia tinctorum* (madder): dye cost, yield analysis, and production costs

Cost criterias for cultivation and dye production	Minimum	Maximum
1st year cultivation cost per decare	156,93 $	186,91 $
2nd year cultivation cost per decare	86,76 $	107,17 $
3rd year cultivation cost per decare	200,30 $	272,39 $
Total 3-year cost (per decare)	443,99 $	566,46 $
Product obtained per decare (madder roots washed and dried) (duplicate – may be omitted) (kg)	300 kg	850 kg
Cost per kg of dye product (washed and dried of madder roots) ($)	1,48 $	0,54 $
Cost per kg of powder dye ($)	85 $	40 $
Sales price per kg of powder dye extract ($)	160 $	240 $
Amount of fabric dyeable with 1 kg dye extract (kg)	6 kg	10 kg
Dye income per decare ($)	720 $	2040 $

The production scenario of *Rubia tinctorum* (Madder) is presented in Table 3.

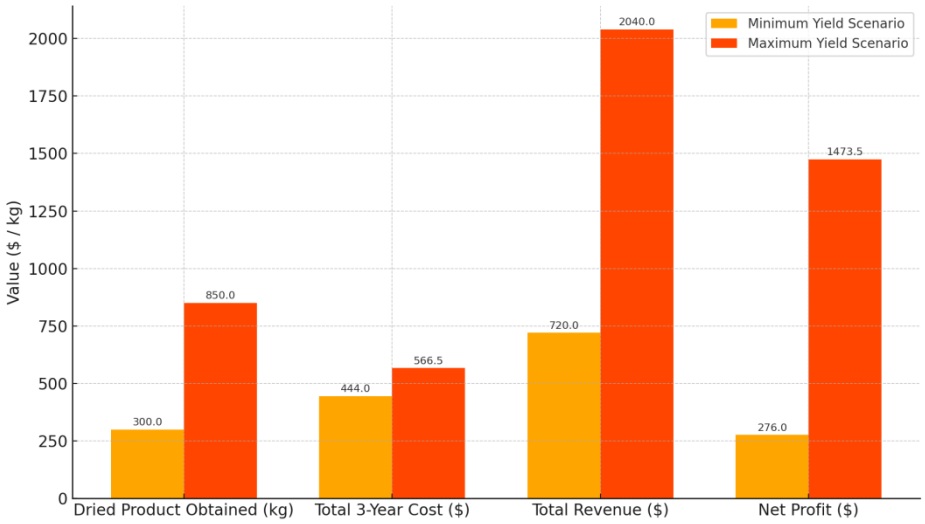


Figure 3. Production and yield scenario of *Rubia tinctorum*

The graph compares the minimum and maximum production scenarios for *Rubia tinctorum* (Madder). In the maximum yield scenario, both the product quantity and income are high, resulting in a highly satisfactory net profit. In the minimum yield scenario, due to lower product and income levels, the net profit remains limited. Within this context, from 1 kg of fresh *Isatis constricta* (Hazar Indigo), a minimum of 3.5 g and a maximum of 6 g of dye can be extracted. The cultivation cost per kilogram of dye plant has been calculated as follows:

For *Isatis constricta*: minimum 0.345/kg, maximum 0.228/kg

For *Reseda lutea* (Weld): minimum 1.68/kg, maximum 0.79/kg.

For *Rubia tinctorum* (Madder): minimum 1.48/kg, maximum 0.54/kg

3.1. Comparison of the Profitability of Three Different Dye Plants

The comparison of the profitability of the three different plants is presented in Figure 4.

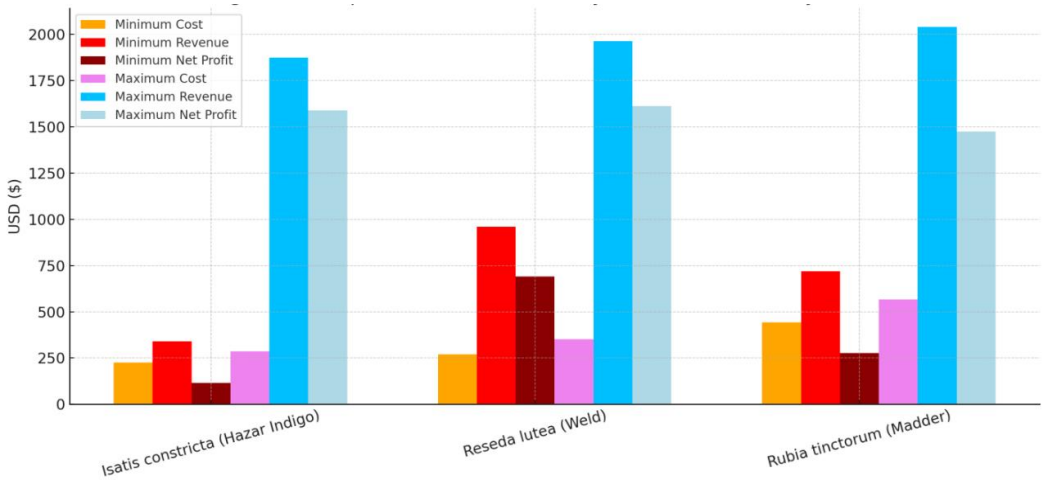


Figure 4. Profitability comparison of three different dye plants

In terms of profitability, *Isatis constricta* (Hazar Indigo) provides very high income and profit in the maximum scenario, while carrying a risk of loss in the minimum scenario. In *Reseda lutea* (Weld), there is positive net profit in both minimum and maximum scenarios, making it a low-risk and sustainable production alternative. For *Rubia tinctorum* (Madder), the maximum scenario is highly profitable, while in the minimum scenario the profit is very limited, yet not negative.

In conclusion, the highest potential profit is achieved with Isatis constricta under maximum yield conditions. The most stable and low-risk plant is Reseda lutea. The dye plant requiring long-term investment is *Rubia tinctorum* (Madder).

4. Discussion and Recommendations

This study examined in detail the economic performance of three different natural dye plants cultivated in Elazig Province. The analyses were carried out in terms of productivity, cost, sales potential, and profitability, taking into account the cultivation conditions of each species.

The most notable finding of the study is that yield levels directly affect costs and represent the key factor in determining profitability.

Isatis constricta (Hazar Indigo) emerged as the most economically advantageous species under maximum yield conditions, providing high income. However, under low-yield conditions, it carries a risk of loss, indicating that this plant requires yield-sensitive and technically informed cultivation practices.

Reseda lutea (Weld) demonstrated positive profitability under both low- and high-yield conditions, proving to be a low-risk and stable production alternative. This suggests that it may serve as a sustainable option for small- and medium-sized farming enterprises.

Rubia tinctorum (Madder), due to its three-year cultivation cycle, requires long-term investment. While it generates satisfactory profit under maximum yield conditions, its profitability remains limited under minimum yields. Thus, the cultivation of madder necessitates patience, technical capacity, and a long-term investment perspective.

5. Recommendations

- Yield-enhancing techniques should be promoted, especially for Isatis constricta and *Rubia tinctorum*; practices such as fertilization, irrigation, and harvest timing should be optimized.
- To reduce production costs, energy efficiency and shared equipment use in processing stages should be encouraged.
- The contract farming model should be expanded in natural dye production, as it provides producers with market security and increases profitability.
- Industrial integration of natural dyes (e.g., with textile, cosmetics, and pharmaceutical sectors) should be supported, along with R&D and product development activities.
- Government subsidies are crucial for the conservation of endemic species and the promotion of sustainable agricultural production. Special incentives should be provided particularly for regionally significant species such as Isatis constricta.
- International market research should be conducted to evaluate the export potential of natural dyes.

6. References

[1] Hatipoğlu, Z., & Gürsoy, C. T. Temel Maliyet-Kâr Planlaması ve Kontrolü. Beta Basım Yayım ve Dağıtım, İstanbul, 1979.
[2] Hopkins, J. A., & Heady, E. O. Farm Records and Accounting. Iowa State University Press.377 p, 1964.
[3] Aras, A. . Tarım Muhasebesi. Ege Üniversitesi Ziraat Fakültesi Yayınları No: 486, Ders Kitabı, İzmir, 1988.
[4] Anonymous, 1988. Tarım Muhasebesi: (Ders Kitabı) , Ege Üniversitesi Ziraat Fakültesi Yayınları No: 486, İzmir. Kitap 323 p, 1988.
[5] Çetin, E., & Bahşi, N. (2019). Tarımsal üretim gerçekleştiren işletmelerin faaliyetlerinin hedef maliyet yöntemine göre değ erlendirilmesi. OPUS Uluslararası Toplum Araştırmaları Dergisi, 13(19), 2365–2380. https://doi.org/10.26466/opus.563520

The Use of Algae in Agriculture

Siham MEÇİK[1] , İnanç OZGEN [2]

[1,2]. Fırat University, Faculty of Engineering, Department of Bioengineering, Elazığ, Turkey
[1]232132201@firat.du.tr ,[2]inancozgen@gmail.com,
[1](ORCID:0009-0007-8436-0829), [2](ORCID: 0000-0003-1742-9324)

Abstract
Algae have become critical in numerous sectors beyond traditional fields like pharmacology, playing a vital role in agriculture, environmental management, wastewater treatment, and biofuel production alongside industrialization. Microalgae, in particular, hold great potential for improving soil health, increasing carbon sequestration, supporting nutrient cycles, and strengthening sustainable agricultural practices. These organisms increase the bioavailability of nutrients by promoting beneficial microbial populations in the soil. They also secrete extracellular polymeric substances (EPS), which increase the carbon content of the soil and support its structural improvement. This effect increases both the water retention capacity of the soil and crop productivity. Microalgae species such as Chlorella contribute to healthier plant growth by directly affecting soil fertility. The use of algae as biofertilizers offers a natural and environmentally friendly alternative to synthetic fertilizers. This aims to reduce chemical pollution spreading into soil and water resources. Furthermore, thanks to the antifungal and antimicrobial properties of compounds derived from algae, disease control in plants can be achieved through natural means. This also contributes to reducing pesticide use. Algae also provide a dual benefit in wastewater treatment: they remove pollutants and the biomass produced during this process can be converted into agricultural products such as fertilizer or biofuel. Their CO_2 sequestration capacity, in particular, offers a significant advantage in combating climate change. In this respect, algae are a critical solution partner for a sustainable agriculture and environmental management model. This study examines in detail the multifaceted roles algae play in their integration into agriculture. Their potential contributions to utilizing microbial diversity, producing organic matter, improving soil, and mitigating climate change are addressed in both theoretical and practical dimensions. Furthermore, the environmental and economic advantages that can be achieved by integrating algae-based biofertilizers and biostimulants into existing agricultural systems are highlighted. In conclusion, algae are not only an environmentally friendly fertilizer alternative but also a strategic biological resource offering sustainable solutions to the fundamental challenges facing modern agriculture. In this context, the widespread adoption of algae-based applications is of great importance for future agricultural policies and food security.

Keywords: algae, biofertilizer, biofuel

1. Introduction

Algae are primary producers in water and play an important role in the food chain. They are vital resources used in food, agriculture, cosmetics, medicine, pharmacy, and industry. Algae possess secondary metabolites with potent biological activity, which are crucial for developing new drugs. (Aktar and Cebe, 2010). Phycocolloids used in pharmacy are obtained from seaweed. Research shows that algae have antimicrobial, cytotoxic, antimitogenic, anticancer, and antitumor properties. Algae are found globally, mostly in water, and can live in seas, lakes, rivers, and on land.

Along with animals and plants, they can live in ice-covered areas, hot springs, salt water, low light, and high-pressure environments. This study aims to gather information about algae, highlight their uses in various fields, and emphasize the importance of algae culture and industry. Algae are divided into eukaryotic and prokaryotic groups. (Cyanobacteria), euglenoids (Euglenozoa), diatoms (Bacillariophyta), dinoflagellates (Miozoa), green algae (Chlorophyta), cryptomonads (Cryptophyta), red algae (Rhodophyta), charophytes (Charophyta), Ochrophyta, and Haptophyta. Algae are studied in various phyla and are crucial in aquatic food chains as they are affected by environmental changes. Algae represent a sustainable resource with important applications in agriculture, environmental management, and waste treatment. This review highlights the multifaceted roles of algae, particularly microalgae, in improving soil health, facilitating carbon sequestration, treating wastewater, and producing biochar. Algae enhance soil fertility in agriculture by improving biological, chemical, and physical properties. They promote beneficial microbial populations that support the nutrient cycle while contributing to increased carbon content and soil aggregation by secreting extracellular polymeric substances (EPS).

[1]Corresponding author

Microalgae such as Chlorella have been shown to increase nutrient levels in the soil and improve water retention, ultimately promoting healthier crop growth. As global CO2 emissions increase, algae offer a promising solution for carbon sequestration. Through photosynthesis, they efficiently absorb atmospheric CO2 and convert it into biomass at significantly higher rates than traditional plant systems. This natural process positions microalgae as a suitable alternative for mitigating climate change due to their ability to sequester carbon and produce biomass for various applications. Algae play a crucial role in wastewater treatment by utilizing nutrient-rich wastewater for growth. Microalgae effectively remove pollutants while producing suitable biomass for biofuels and fertilizers. Recent studies show that municipal wastewater is an effective growth medium.

Algae are a large group of microorganisms capable of photosynthesis and capturing energy from sunlight. They are important in agriculture as bio-fertilizers and soil stabilizers, and they lead to a reduction in nitrogen and phosphorus runoff compared to animal manure. This improves water quality in rivers and oceans. Algae are cultivated worldwide as a food supplement and can grow in abandoned and arid areas with little water. Seaweed is an important source of iodine, which can increase iodine levels in milk when fed to cows. Algae also increase egg production in chickens. This article discusses the basic aspects of algae and their agricultural benefits for practitioners.

The integration of algae into agriculture offers a sustainable approach to improving crop production and soil health. Algae, particularly macroalgae and microalgae, act as biofertilizers and biostimulants, increasing nutrient availability and plant resistance to stress factors. Their application can increase crop yield and quality while minimizing environmental impacts.

Agricultural cost accounting also plays a vital role in credit and financing processes. It can inform government policies and assist in determining income in sharecropping arrangements (Hopkins & Heady, 1964). This branch of accounting records the financial transactions of agricultural enterprises, summarizes and analyzes them appropriately, and enables the calculation of unit product costs (Aras, 1988; Anonymous, 1988).

In the context of target costing, accurate forecasting of market share and sales volume is crucial. Agricultural production should not be evaluated solely within the input-production-output cycle, as it is significantly affected by external conditions such as climate, pollution, and global warming (Çetin & Bahşi, 2019). In conclusion, accurate cost calculations in dye production contribute to determining competitive price thresholds. Reducing the unit costs of dye plants supports the efficient and sustainable operation of agricultural enterprises.

2. Classification and Characteristics of Algae

2.1. Eukaryotic Organisms

Algae are eukaryotic microorganisms with structures such as cell walls, cell membranes, cytoplasm, and organelles. They may have special parts such as contractile vacuoles and stigmas and can store nutrients. Algae contain chlorophyll, can perform photosynthesis, and can be unicellular or multicellular. They reproduce vegetatively, sexually, and asexually, with vegetative reproduction being the most common. Algae can live in water and moist soil.

2.1.2. Classification of Algae

Macroalgae and microalgae, collectively referred to as “algae,” form a large and diverse group of organisms found primarily in aquatic environments. With over 156,000 described species, eukaryotic algae (including both macroalgae and microalgae) represent a significant source of biological diversity (Hughes et al., 2021; Kholssi et al., 2022). There are approximately 40,000 different microalgae species worldwide, constituting approximately 20-30% of all globally unidentified algae (Song et al., 2022). Algal classification has evolved over time using various characteristics, including morphological, physiological, biochemical, and, more recently, molecular features. William Henry Harvey (1811-1866) was an early pioneer in algal taxonomy in Flora Hibernica (Mackay, 1836).

In 1836, he divided algae into four groups based on the color of their thallus: Chlorospermae (Green), Melanospermae (Brown), Rhodospermae (Red), and Diatomacea (Diatoms). Later, A.W. Eichler (1883) (Eichler, 1883) introduced the division Thallophyta and divided algae into five groups based on their color: Cyanophyceae (Blue-green), Diatomeae (Diatoms), Chlorophyceae (Green), Phaeophyceae (Brown), and Rhodophyceae (Red). In 1916, West (West, 1916) classified algae according to their reproductive structures and the presence or absence of flagella. He divided algae into four groups: Isokontae (flagellates of equal size), Akontae (flagellates without flagella), Stephanokontae (flagellates with crown-shaped flagella), and Heterokontae (flagellates of unequal size). Similarly, A. Pascher (Pascher, 1931) developed a phylogenetic classification system focusing on evolutionary relationships between algal groups. His system placed algae at the top level of divisions and divided them into eight divisions, then further into classes.

J.E. Tilden (1935) (Tilden, 1935) divided algae into five groups based on characteristics such as pigmentation, food reserves, and flagellation, emphasizing the importance of pigments in the evolutionary development of algae. F.E. Fritsch (1935), considered the “Father of Phycology,” proposed one of the most comprehensive and widely accepted classifications. In his two-volume work, “The Structure and Reproduction of Algae,” Fritsch used characteristics such as pigmentation, reserve nutrients, flagellar structure, nuclear organization, and reproductive modes. He grouped algae into 11 classes. G.M. Smith (1938) supported Pascher's classification (1914, 1931) but made some changes by dividing algae into divisions and then classes. Later, Prescott (1969) based his classification on the presence or absence of a true nucleus, pigmentation, cell wall composition, reserve nutrients, life cycles, and reproductive strategies. His system divided algae into nine phyla (Sybil Parker, 1982).

He introduced a classification based on the presence or absence of membrane-bound organelles, distinguishing between Prokaryota and Eukaryota. Similarly, Lee (2008) classified algae as Prokaryota and Eukaryota; Prokaryota consists of a single division, Cyanophyta, while Eukaryota is further subdivided according to chloroplast membrane structures.

2.1.3. Chromista (Algae Containing Chlorophyll C)

Fucoxanthin, the internal structure of chloroplasts, and the presence of mastigonemes have led to the temporary inclusion of some small groups of organisms in the Chromista kingdom. Gene sequence analyses have shown that they share a common ancestor.

2.1.4. Chrysophyta (Yellowish-Green Algae, Brown Algae, and Diatoms)

The classification of algae is based on color, which depends on cell pigments. Subsequent studies have shown that characteristics such as flagella structure, energy storage substances, and cell walls also reveal similarities among algae. Most algae classified under Chrysophyta are brown or yellow due to carotenoids and contain chlorophyll a and c. In addition, their walls may contain silicon or calcium. Some species are unicellular, while others have a colonial structure, and asexual reproduction is common. Diatoms occupy an important place in freshwater and saltwater ecosystems, and planktonic organisms are divided into two main groups: phytoplankton and zooplankton.

2.1.5. Phaeophyta (Brown Algae)

The Phaeophyceae class encompasses brown algae ranging from small filamentous species to giant seaweeds, with approximately 1,836 species in 285 genera; less than 1% of these are freshwater species. These algae are mostly found in the sea, particularly in cold ocean regions, and are common on rocky shores in places like New England. Generally multicellular, these algae can reach lengths of up to 45 meters. Their cells contain cellulose, alginic acid, chlorophyll a, and fucoxanthin. Asexual reproduction occurs via flagellated zoospores, while sexual reproduction takes place in gametangia organs. Heterogamous species have two different gametes, while isogamous species have similar gametes. For example, the reproductive cycles of species such as Ectocarpus and Laminaria show the development of zygotes and the formation of new plants. Laminaria is notable for its large sporophyte structure and exhibits characteristics similar to terrestrial plants.

2.1.6. General Characteristics of Cyanobacteria

Cyanobacteria, recognized as the first photosynthetic organisms on Earth and present on the planet for approximately 3.5 billion years, are the only group among other algal groups that have a prokaryotic structure. Due to the absence of a nucleus and plastid membrane in their cells, as well as the lack of mitochondria, Golgi apparatus, and vacuoles, and the absence of sexuality, they are classified by bacteriologists in the Cyanobacteria division. Phycologists, however, study them in the Cyanophyta division due to their characteristics, as they possess Photosystem I and Photosystem II (Cirik & Cirik, 2011; Inman, 1940, Meinesz, 2010). Depending on the species, their cells, which are of a single color, exhibit blue-green and purple color changes. Phycocyanin and phycoerythrin phycobiliproteins are found between lamellae, while chlorophyll is a pigment found in double-lamellar thylakoids where photosynthesis occurs. In addition to photosynthesis, thylakoids also play a role in respiration and nitrogen fixation from the environment (Cirik & Cirik, 2011; Eriksen, 2008). Cellulose and pectin form cell walls, while cyanocyanin and volitin function as reserve nutrients (Cirik and Gökpınar, 2008; Cirik and Cirik, 2011). Although blue-green algae are primarily found in aquatic environments (seas and inland waters), they can also be found in unusual habitats such as deserts, glacial lakes, and hot springs (Stal, 2002). Furthermore, all types of microalgae can withstand higher levels of UV radiation than other photosynthetic organisms because they synthesize different extracellular polymeric compounds at different stages of their growth (Christensen et al., 1985; Schüßler et al., 1997). The glycans found in Nostoc commune maintain the integrity of cell membranes and prevent the freezing or drying of cell contents (Hill et al., 1997; Tamaru et al., 2005). When environmental conditions change, some blue-green algae species use sulfate compounds as electron donors to perform anoxygenic photosynthesis (Cohen et al., 1986).

Heterocysts are homogeneous, thick-walled structures with apical pores and fix nitrogen from the atmosphere in some genera, including Anabaena (Cirik & Cirik, 2011). Prokaryotic and eukaryotic algal communities, along with bacterial, fungal, and protozoan communities, thrive in areas where mineral-rich hot springs emerge from the earth. According to Seckbach (1994; 1996), the world's oldest prokaryotic cells may have originated in areas with hot water sources. Archaea (up to 114°C) and Eubacteria (up to 95°C) are two examples of organisms that can withstand extremely high water temperatures (Seckbach and Oren, 2007). In contrast, the maximum temperatures for photosynthetic organisms range from 74°C (prokaryotic algae) to 57°C (eukaryotic algae). Therefore, Cyanophyceae (blue-green algae) are the dominant algal class in geothermal sources (Brock, 1978; Seckbach and Oren, 2007; Ward et al., 2012).

As one moves away from the water source, the temperature decreases from the highest point where thermal waters emerge to the surface. Consequently, the distribution of algal species increases with distance from the source. Single-celled genera such as Chroococcus and Synechococcus are more common in areas with high water temperatures, while filamentous algae such as Oscillatoria are more common in areas with low water temperatures (Castenholz, 1969; Seckbach and Oren, 2007). The concentration of dissolved hydrogen ions in thermal waters determines the limits of algal life. Generally, the pH values of geothermal waters range from 3 to 10. While eukaryotic algae such as Cyanidium caldarium are found in waters with low pH values, species belonging to the blue-green algae class become dominant as the pH value increases (Brock, 1978). According to Meeks and Castenholz (1978), species of the genus Synechococcus are not found in Icelandic geothermal sources. On the other hand, the global species Mastigocladus laminosus Cohn ex Kirchner has been found in numerous geothermal areas (Brock, 1978; Castenholz, 1969; Thomas and Gonzalves, 1965).

3. Nutritional and Environmental Benefits

Algae are rich in macronutrients, essential amino acids, and fatty acids, making them a valuable food source for both humans and animals. They provide high protein content without requiring arable land, offering a sustainable alternative to traditional agricultural products (Akgül and Akgül, 2024; Diaz et al., 2023).

The ability of algae to capture CO2 and utilize non-potable water sources increases their sustainability and reduces the carbon footprint of agricultural practices (Diaz et al., 2023).

Nutrient Enrichment: Algae improve soil properties by providing essential nutrients such as nitrogen, phosphorus, and potassium, which are crucial for plant growth (Patel et al., 2024).

Stress Reduction: Algal extracts increase the plant's resistance to biotic and abiotic stresses, promoting better growth under adverse conditions (Sariñana-Aldaco et al., 2025). Improved Soil Health: The use of algae contributes to the structural, chemical, and biological improvement of the soil, promoting a healthier ecosystem for crops (Prisa et al., 2024).

Algae are rich in essential nutrients such as nitrogen, phosphorus, and potassium, which are vital for plant growth ("Traditional Applications of Algae in Cultivated Plants," 2023).

Incorporating seaweed into soil can increase nutrient availability and improve soil structure, leading to better water retention and aeration ("Traditional Applications of Algae in Cultivated Plants," 2023) ("The Role of Algae in Agriculture," 2022). Algae can reduce the environmental impact of fertilizers by reducing nutrient runoff into water bodies (Potharaju & Aruna, 2024).

Due to the excessive use of chemical-based products to increase crop production, soil health is deteriorating worldwide. The overuse of chemical fertilizers and pesticides damages the quality of soil and groundwater, disrupts ecosystems, and ultimately poses health risks to humans and wildlife. Therefore, sustainable agricultural practices are crucial for preserving soil quality. Algae and fungi play an important role in maintaining soil fertility and health by positively affecting the physical, chemical, and biological properties of the soil. Both algae and fungi help reduce heavy metal toxicity and detoxification and serve as indicators of soil health.

When used as bio-fertilizers, algae and fungi increase the availability of nutrients and crop yields. Some species also function as biological pesticides that help reduce pathogens in soil and crops. Beyond this, they pose no threat to environmental health. With this in mind, this study highlights the role of algae and fungi in maintaining good soil quality.

4. The Use of Algae in Agriculture

Algae fertilizers are valuable, especially in potato cultivation on potassium-deficient soils, due to their aeration, moisture retention, nutrient richness, and trace elements.

4.1. Biological Fertilizer Applications

Algae in many coastal regions around the world are used as fertilizers due to their phosphorus, potassium, and trace element content.

The excessive use of synthetic agricultural chemicals has caused significant global ecological damage, leading to ocean dead zones, eutrophication, soil infertility, and biodiversity loss (Garcia-Gonzalez and Sommerfeld, 2016; Li et al., 2020). In this context, bio-fertilizers are considered an environmentally friendly, cost-effective, and sustainable alternative to synthetic fertilizers. They increase agricultural productivity and reduce environmental pollution (Garcia-Gonzalez and Sommerfeld, 2016; Kawalekar, 2013). Similarly, algae, primarily aquatic and autotrophic organisms, lack true roots, stems, leaves, and different multicellular structures. These autotrophic organisms are typically classified as microalgae or macroalgae. Blue-green algae (BGA) such as Nostoc, Anabaena, Tolypothrix, and Aulosira in rice fields can fix atmospheric nitrogen. The main categories of biofertilizers include Azolla, Anabaena, and Rhizobium (Ammar et al., 2022; Choudhary et al., 2024). Algae, which are widely found throughout the world, are photosynthetic and naturally integrate into the environment by producing biomass. They include both single-celled organisms (microalgae) such as Chlorella, diatoms, and Prototheca, as well as multicellular organisms (macroalgae) such as green and brown seaweeds (Bhatt et al., 2022).

4.2. Biostimulants

Algal products stimulate seed germination and biomass accumulation, leading to higher productivity (Sariñana-Aldaco et al., 2025).

Pest Control: Compounds derived from algae exhibit fungicidal and antimicrobial properties, helping to protect plants against pathogens (Prisa et al., 2024).

Sustainable Practices: Algal biofertilizers promote environmentally friendly agricultural practices by reducing dependence on chemical fertilizers (Patel et al., 2024).

Algae contribute to sustainable agriculture by increasing soil fertility, improving nutrient availability, and reducing greenhouse gas emissions. Their cultivation requires fewer resources, making them a viable alternative to traditional farming practices while supporting environmental sustainability and food production. Brown algae serve as sustainable agricultural biostimulants that increase crop yield and quality by improving seed germination, increasing biomass, activating antioxidant systems, and aiding nutrient absorption. They are compatible with synthetic fertilizers and pesticides and promote balanced agricultural ecosystems. Marine macroalgae improve soil properties and serve as biofertilizers, pest control agents, and plant growth conditioners. They contain beneficial compounds such as fatty acids and polyphenols and show significant potential for sustainable crop production and organic farming practices.

Microalgae and cyanobacteria act as biofertilizers and biostimulants in agriculture, promoting plant growth and offering a sustainable alternative to chemical fertilizers by improving soil health through the production of N-fixing enzymes, phytohormones, and other beneficial biomolecules. Algae are used in agriculture as biofertilizers and soil stabilizers, reducing nitrogen and phosphorus discharge compared to animal manure. They require minimal water, can grow on marginal lands, and enhance livestock products such as increasing iodine in milk. Algae can improve agriculture by serving as a sustainable biomass source for animal feed, absorbing salinity and drought, capturing carbon, and reducing agricultural pollution through nitrogen and phosphorus utilization, thereby contributing to food security and environmental sustainability. Algae improve soil health by enhancing the physical, chemical, and biological properties of the soil. They act as bio-fertilizers, increasing nutrient availability and crop yields, while also promoting sustainable agricultural practices by acting as biological pesticides to reduce soil and crop pathogens.

5. Integrated Biorefinery Systems: Converting Algal Biomass into Agricultural Solutions

Many products derived from industrial biorefineries, including growth stimulants for plants, bio-fertilizers, oils for biofuels, and electricity produced from hydrogen, can be applied in agriculture. This wide range of products can be very important for large farms seeking independence. In line with the closed-loop economy, algal biorefinery products can play a very important role in industrial food production. These products enable more efficient use of residues and completion of the nutrient cycle (Galanopoulos et al., 2019). Integrating multiple systems ensures the most efficient use of equipment, energy, raw materials, and products. A combined biorefining system concept for agricultural companies is shown below.

5.1. The Role of Algae in Sustainable Agriculture

Agricultural lands are facing productivity losses due to climate change and intensive farming practices. This situation is exacerbated by the increased use of synthetic fertilizers, which contribute to environmental pollution and health problems. Interest in microorganisms, particularly algae, as environmentally friendly alternatives to chemical fertilizers is growing. These microorganisms function as biopesticides, biostimulants, and biofertilizers, increasing crop yields and soil quality. Algae are considered a promising source for bio-based organic fertilizers.

Technological advances can increase the commercial production of algae-based biofertilizers in a sustainable and cost-effective manner by integrating algae production with wastewater treatment. More research using genomic and proteomic tools is needed to understand the mechanisms behind algae-based biofertilizers. Algae are used globally

in agriculture to reduce nitrogen and phosphorus emissions, provide minimal water requirements, and improve livestock nutrition.

The use of algae in agriculture has grown significantly over the decades as the benefits they provide for sustainable practices have been recognized. This development highlights the potential of algae to address food security and environmental challenges. The integration of algae into agriculture has progressed from traditional food uses to sophisticated applications such as biofertilizers, contributing to the circular bioeconomy.

5.2. Algae in Wastewater Treatment and Biofertilizer Production

Water, along with air, is one of the most vital resources for human survival. However, rapid social development and increasing dependence on freshwater resources in recent years have led to the formation of significant amounts of wastewater from various sources, including agricultural runoff, industrial discharges, and domestic sewage.

Wastewater is often loaded with pollutants, including essential nutrients such as phosphorus and nitrogen, as well as heavy metals such as lead and zinc, which brings new environmental concerns (Gururani et al., 2022). In 2020 alone, total phosphorus and nitrogen emissions reached alarming levels of 336,700 tons and 3,223,400 tons, respectively.

Traditional wastewater treatment methods encompass a range of physical, chemical, and biological techniques (Mohsenpour et al., 2021). While these methods effectively remove pollutants, they also present various limitations, such as inconsistent removal efficiencies, high installation and operating costs, and the potential for secondary pollution due to residual pollutants. Microalgae have emerged as a promising alternative that can effectively remove nutrients and pollutants while also producing biomass that can be used for various applications such as biofuel and fertilizer (Rashid et al., 2018). By leveraging the natural processes of microalgae, wastewater treatment systems can not only increase treatment efficiency but also contribute to resource recovery, thereby promoting a healthier wastewater treatment system.

Algae are increasingly used in wastewater treatment as they help reduce nutrients such as nitrogen and phosphorus, thereby mitigating environmental impacts. This process also supports biofertilizer production, which improves soil health and promotes sustainable agriculture (Sartori et al., 2023) (Zou et al., 2021). Despite ongoing challenges such as developing technologies, high costs, and potential environmental threats, the focus remains on integrating algae-based wastewater remediation with biofertilizer production (Zou et al., 2021).

6. Algal Case Studies in Soil Remediation and Carbon Sequestration

Algae, including microalgae and cyanobacteria, are recognized as effective and sustainable tools for improving soil health and supporting carbon sequestration efforts to mitigate climate change. These photosynthetic organisms provide various benefits, such as increasing soil fertility, stabilizing degraded land, and capturing pollutants, making them valuable for tackling agricultural and environmental challenges. The versatility of algae has led to their application in various real-world contexts, including promoting soil remediation and sustainable agricultural practices. A study conducted in India evaluated the effect of marine microalgae as fertilizer on the growth and yield of maize (Zea mays L.). In this case, maize plants were grown for 75 days in greenhouse conditions in soil enriched with two marine microalgae species (Chlorella vulgaris and Spirulina platensis) and cow manure. Treatment with marine microalgae increased early-stage growth, improved seed germination, and provided better yield characteristics, highlighting the potential of marine microalgae as effective fertilizers for corn cultivation (Dineshkumar et al., 2019). In another study conducted by Dineshkumar et al. (2018), the use of a vertical semi-closed air-lift photobioreactor (PBR) for microalgae cultivation and the evaluation of the effectiveness of microalgae fertilizers on rice plant growth were investigated. The system was produced using polyculture of N2-fixing cyanobacterial strains (Anabaena sp. UTEX 2576, Nostoc muscorum UTEX 2209S) and C. vulgaris (UTEX 2714) and Scenedesmus dimorphus (UTEX 1237).

Under greenhouse conditions, rice plants treated with microalgal polyculture biomass showed similar or superior growth compared to urea treatments. However, inoculation with N2-fixing cyanobacteria had no significant effect on seedling growth. The PBR cultivation system developed in this study has proven to be a simple and effective method for microalgae cultivation (Dineshkumar et al., 2018). A study by Tsai et al. (2017) found that microalgae exhibited high CO2 uptake (159 mg L−1 day−1) and consumption efficiency (93%), which is similar to the performance of a continuously stirred tank reactor (178 mg L−1 day−1 and 96%). This research emphasized that algae, unlike other plants, have the ability to rapidly accumulate carbon and play an important role in the carbon cycle. Additionally, they found that forest plantations in tropical regions exhibited higher growth rates and CO_2 uptake compared to other terrestrial plant species (Tsai et al., 2017).

Therefore, algae have been able to play a critical role in the carbon cycle as biological sinks by accumulating large carbon stores in a short period of time. Similarly, a study conducted in an apple orchard showed that the application of microalgae biofertilizers increased fruit yield by up to 29.6%.

Notably, this did not lead to increased greenhouse gas emissions and suggested that algae could increase carbon sequestration and organic carbon accumulation in soils over time, while avoiding contributing to atmospheric CO2 levels (Ma et al., 2024). Additionally, a study by Tu et al. (2019) focused on the effect of CO_2 from power plant tail gas on the growth of Chlorella pyrenoidosa in wastewater. The results showed a significant increase in dry weight (84.92%) and lipid productivity (74.44%) in the presence of power plant tail gas. The microalgae exhibited optimal carbon fixation at a rate of 1,121 g/L, which was 134.22% higher than the control group.

The study also emphasized the importance of pre-treatment of power plant tail gas through desulfurization and denitrification for its effective use in algae cultivation (Tu et al., 2019). Finally, (Gupta et al., 2017) investigated two microalgal species, Chlorella sorokiniana and Scenedesmus obliquus, cultivated in different types of flocculated wastewater (e.g., BlueGreen, polymer, chitosan, gypsum). The results showed that these species were more effective in treating wastewater treated with chitosan. Finally, (Gupta et al., 2017) investigated two microalgae species, Chlorella sorokiniana and Scenedesmus obliquus, cultivated in different types of flocculated wastewater (e.g., BlueGreen, polymer, chitosan, gypsum). The results showed that these species grew most effectively in chitosan-treated wastewater and that their biomass yields were significantly higher than those cultivated in other types of wastewater. Additionally, lipid and carbohydrate yields were 16.0-19.8% and 19.98-20.43% (weight/weight) of dry weight, respectively. The study concluded that wastewater treated with chitosan is a suitable medium for microalgae cultivation, and experimental modeling provided reliable estimates of biomass and metabolic product yields.

7. Antifungal Activity of Macroalgae Extracts

Macroalgae are ubiquitous and serve as renewable resources used in food, feed, medicine, wastewater treatment, and phytocolloid production. Their biostimulant properties have been tested in agriculture. In this study, the antifungal activity of various algae was evaluated against different fungi. The results showed that most extracts possessed antifungal effects. This research is the first study on the antifungal activity of algal extracts in Romania. Macroalgae may assist in the management of fungal diseases in sustainable agriculture, but further research is needed for applications in product management.

Red and brown algae are important healthy food sources for humans due to their high levels of polysaccharides, minerals, fatty acids, and vitamins. They contain biologically active compounds with various biological activities, including antibacterial and antiviral effects. This study focuses on the antifungal properties of seven green and brown algae species, including Spirulina platensis, Fucus vesiculosus, Alaria esculenta, and Ecklonia maxima.

7.1. Materials and Methods

Algae extracts obtained from five species were tested for their antifungal activity against nine fungal isolates. While dried samples of two species were cut and ground, the others were used in powdered or liquid form. Each powdered

sample was subjected to lipid-soluble extraction with ethanol overnight. The extracts were filtered and prepared at a 10% concentration.

Nine strains of different genera were used for the fungal isolates, all of which were purified and stored at 4°C in malt agar medium. The effect of algal extracts on mycelial growth was tested in vitro after 7 days of incubation at 24°C. Agar discs impregnated with extracts at various concentrations were placed in malt agar. The percentage of growth inhibition was calculated and the results were expressed as EC50/EC90 values.

7.2. Results and Discussion

The antifungal properties of five algal extracts were tested against important fungal species, including eight plant pathogens and a known antagonist. Different responses were observed with Alaria esculenta extracts. For example, Aspergillus niger (An 205) was the least sensitive, while A. alternata (Aa 100) was the most sensitive. All tested Alternaria isolates were sensitive, with EC50 values ranging from 0.67% to 0.98%. Other isolates had EC50 values ranging from 1.27% to 1.54%. Fungal growth was significantly inhibited at concentrations ranging from 1.67% to 2.46%. Fucus vesiculosus extracts also exhibited fungitoxic effects, particularly against Alternaria isolates with average EC50 values ranging from 0.81% to 0.85%. The least sensitive isolate was Penicillium expansum (Pe 2712), which had higher EC50 values. The effects of Bioalgua® on fungal growth were also discussed, showing varying sensitivity among the tested isolates.

8. Development of Algae Use in Agriculture

The evolution of algae use in agriculture to date has been marked by a significant shift towards sustainable practices. Algae-based biofertilizers and biostimulants are attracting interest by reducing dependence on chemical fertilizers and pesticides while providing essential nutrients and promoting plant growth. Furthermore, algae-based plant protection methods, which offer environmentally friendly alternatives to traditional pesticides, are also expected to increase. The integration of microalgae into animal feed is also expected to increase, contributing to a more sustainable livestock industry by minimizing the ecological footprint of animal agriculture. Furthermore, Spirulina cultivation is predicted to expand as a nutrient-rich source for both agriculture and human nutrition. Collectively, these developments highlight a transformative approach to agriculture that leverages the benefits of algae, promotes sustainability, and reduces environmental impacts across various sectors.

This study demonstrated that ethanol extracts obtained from macroalgae possess antifungal effects against various plant fungi and significantly reduced growth by less than 2%. This is the first report on the antifungal activity of algal extracts against important pathogens in Romania. Further research is needed to identify the active molecules responsible for this antifungal activity and to evaluate the protective role of seaweed extract on plants.

9. Results

Algae are a large and diverse group of microorganisms capable of photosynthesis. Algae play an important role in agriculture and are used as biofertilizers and soil stabilizers. Algae, especially seaweed, are used as natural fertilizers. This prevents the pollution of rivers and oceans compared to chemical and animal fertilizers. These organisms are cultivated around the world and are also used as human food supplements (Abdel-Raouf et al., 2012).

Current agricultural methods rely on synthetic chemicals that harm the environment for plant nutrition and pest control. Algae offer a promising alternative to reduce the use of these chemicals. This study shows that ethanol extracts from macroalgae can combat plant fungi and reduce growth by less than 2%. Further research is needed to identify the active compounds and evaluate the protective effects of seaweed extract on plants.

10. References

[1] Aktar, S., and Cebe, G. E. (2010). General characteristics of algae, their uses, and their importance in pharmacy.

[2] Kholssi, R., Lougraimzi, H., Grina, F., Lorentz, J. F., Silva, I., Castaño-Sánchez, O., & Marks, E. A. (2022). Green agriculture: a review of the application of micro-and macroalgae and their impact on crop production on soil quality. Journal of Soil Science and Plant Nutrition, 22(4), 4627-4641.

[3] Gurau, S., Imran, M., & Ray, R. L. (2024). Algae: A cutting-edge solution for enhancing soil health and accelerating carbon sequestration–A review. Environmental Technology & Innovation, 103980.

[4] Ak, İ., & Cirik, S. (2017). Blue-green algae (Cyanobacteria) and thermalism. Ege Journal of Fisheries and Aquatic Sciences, 34(2), 227-233.

[5] Ak, İ., & Cirik, S. (2017). Blue-green algae (Cyanobacteria) and thermalism. Ege Journal of Fisheries and Aquatic Sciences, 34(2), 227-233.

[6] Matt, G., Marinucci, A., Guainazzi, M., Brenneman, L. W., Elvis, M., Lohfink, A., ... & Zhang, W. W. (2014). The soft-X-ray emission of Ark 120. XMM–Newton, NuSTAR, and the importance of taking the broad view. Monthly Notices of the Royal Astronomical Society, 439(3), 3016-3021.

[7] Seckbach, J., Chapman, D. J., Garbary, D., Oren, A., & Reisser, W. (2007). Algae and cyanobacteria under environmental extremes: final comments (pp. 781-786). Springer Netherlands.

[8] Seckbach, J., & Oren, A. (2007). Oxygenic photosynthetic microorganisms in extreme environments. Algae and cyanobacteria in extreme environments, 3-25.

[9] Konopka, A., & Brock, T. D. (1978). Effect of temperature on blue-green algae (cyanobacteria) in Lake Mendota. Applied and environmental microbiology, 36(4), 572-576.

[10] Brock, T.D. (1978). Thermophilic Microorganisms and Life at High Temperatures. New York: Springer Verlag.

[11] Cirik, S., Ak,İ. & Cirik, Y. (2008). Microalgae of Allianoi (Bergama, Turkey) Thermal Waters. III. National Limnology Symposium, Izmir

[12] Cirik, Ş. & Cirik, S. (2011). Aquatic Plants I (Biology, Ecology, and Cultivation Techniques of Marine Plants). Ege University Faculty of Fisheries Publications No:58, Textbook Index No.26, Bornova, Izmir.

[13] Schüßler, A., Meyer, T., Gehrig, H. & Kluge, M. (1997). Variations of lectin binding sites in extracellular glycoconjugates during the life cycle of Nostoc punctiforme, a potentially endosymbiotic cyanobacterium. European Journal of Phycology, 32: 233–239.

[14] Cohen, Y., Jørgensen, B.B., Revsbech, N.P. & Poplawski, R. (1986). Adaptation to hydrogen sulfide of oxygenic and anoxygenic photosynthesis among cyanobacteria. Applied and Environmental Microbiology, 51: 398–407.

[15] Hill, D., Keenan, T., Helm, R., Potts, M., Crowe, L., & Crowe, J. (1997). The extracellular polysaccharide of Nostoc commune (cyanobacterium) inhibits the fusion of membrane vesicles during desiccation. Journal of Applied Phytology. 9: 237–248. doi: 10.1023/A:100796522956

[16] Stal L. (2002). Cyanobacterial Mats and Stromatolites. In: Whitton, B.A., Potts, M., eds.: Ecology of Cyanobacteria: Diversity in Time and Space. Springer, Netherlands, Dordrecht. doi: 10.1007/0-306-46855-7_4

[17] Seckbach, J. & Oren, A. (2007). Oxygen-producing, photosynthetic microorganisms in extreme environments. Ed: Seckbach, J., in: Algae and Cyanobacteria in Extreme Environments (pp. 3-25). Springer. Netherlands, Dordrecht. doi: 10.1007/978-1-4020-6112-7_1

[18] Tamaru, Y., Takani, Y., Yoshida, T. & Sakamoto, T. (2005). Important role of extracellular polysaccharides in desiccation and freezing tolerance in the terrestrial cyanobacterium Nostoc commune. Applied and Environmental Microbiology, 71: 7327–7333. doi: 10.1128/AEM.71.11.7327-7333.2005

[19] Ward, D.M., Castenholz, R.W. & Miller, S.R. (2012). Cyanobacteria in Geothermal Habitats. Eds: Whitton, B.A., In: Ecology of Cyanobacteria II. Diversity in Space and Time (pp. 37–59). Springer. Dordrecht, Netherlands.

[20] Arman, M., Riahi, H., Yousefzadi, M., Sonboli, A. (2014). A floristic study on Cyanophyta in three hot springs in Hormozgan Province, Iran. Iranian Journal of Botany, 20(2): 240–247.

[21] Aysel, V., Çelik, A., Yayıntaş, A., Gezerler, U.Ş. (1992). Algal Flora of Zonguldak Ilıksu Spa. Journal of the Faculty of Fisheries, Ege University. 9 (33–36): 72–82.

[22] Castenholz, R.W. (1969). Thermophilic Blue-Green Algae and the Thermal Environment. Bacteriological Reviews, 33(4): 476–504.
[23] Castenholz, R.W. (1976). Effect of sulfide on blue green algae of hot springs. II. Yellowstone Park. Microbial Ecology Series No:3.
[24] Davis, B.M. (1897). The vegetation of the hot springs of Yellowstone park. Science New Series, 6(135): 145–157.
[25] Güner, H. (1967). Preliminary observations on the algal vegetation of thermal waters in the Aegean Region. Proceedings of the 5th Turkish Biology Congress. July 10–15, 1967. pp: 69–78.
[26] Güner, H. (1970). Study on the Algal Vegetation of the Thermal and Mineral Waters of the Aegean Region. Ege University Faculty of Science Scientific Reports Series No:99. Ege University Press, Izmir.
[27] Thomas, J. & Gonzalves, E.A. (1965). Thermal Algae of Western India. V. Algae of the hot springs at Tuwa. Hydrobiologia, 26 (1): 41 -54. doi: 10.1007/BF00142252
[28] Ulcay, S.Ö. (2005). Blue-green algae species found in hot springs in Manisa province and its surroundings. Celal Bayar University Institute of Science Master's thesis. Muradiye, Manisa.
[29] Ünal, T. (1996). Microscopic and Macroscopic Algal Flora of İzmir Balçova (Agamemnon) Hot Spring. Ege University Institute of Natural Sciences Department of Biology. Master's Thesis. Bornova, İzmir.
[30] Ward, D.M., Castenholz, R.W. & Miller, S.R. (2012). Cyanobacteria in Geothermal Habitats. Eds: Whitton, B.A., In: Ecology of Cyanobacteria II. Diversity in Space and Time (pp. 37–59). Springer. Dordrecht, Netherlands.
[31] (Akgül & Akgül, 2024)("Microalgae as a Sustainable Source of Vegan Protein and Omega-3 Fatty Acids", 2022).
[32] Galanopoulos, C., Kenkel, P., Zondervan, E., 2019. Optimization of an integrated algal biorefinery superstructure. Comput. Chem. Eng. 130, 106530. https://doi.org/ 10.1016/J.COMPCHEMENG.2019.106530.
[33] Abdel-Raouf, N., Al-Homaidan, A. A., & Ibraheem, I. (2012). Microalgae and wastewater treatment. Saudi Journal of Biological Sciences, 19(3), 257-275.
[34] Awad N.E., 2000. Biologically active steroids from the green alga Ulva lactuca. Phytother. Res., 14 (pp. 641-643).
[35] Jayaraj J., A. Wan, M. Rahman, Z.K. Punja, 2008. Seaweed extract reduces leaf fungal diseases in carrots. Crop Protection 27 (pp. 1360-1366).

Development of a Mobile Application Supporting Sustainable Living Habits for Individual Carbon Footprint Tracking

Fatma Nur KULLAP[1], Hicran AKTAŞ[2], Canan Koç[3], Fatih ÖZYURT[4]

[1,2,3,4]Department of Software Engineering, Faculty of Engineering, Firat University, Elazig, Turkey
[1]fatmanurkullap03@gmail.com, [2]hicranaktass9@gmail.com, [3]canankoc@fırat.edu.tr, [4]fatihozyurt@firat.edu.tr
[1](ORCID: 0009-0001-4557-8629), [2](ORCID: 0009-0004-2699-8179), [3](ORCID:0000-0002-2651-9471), [4](ORCID: 0000-0002-8154-6691)

Abstract
Today, the environmental impacts of individual energy consumption, particularly its role in increasing global carbon emissions, are at the center of sustainable development policies. This study aims to develop a mobile application that analyzes individuals' daily energy consumption habits, calculates their carbon footprints, and encourages sustainable living practices. The application is based on users uploading their digital energy bills or transferring their energy consumption data to the system via smart device integration. The transferred data is analyzed using Optical Character Recognition (OCR) technology and internationally recognized carbon emission calculation algorithms to measure individual carbon footprints. The developed system aims to increase environmental awareness by providing personalized savings and sustainability recommendations based on users' energy usage habits. Supported by a user-friendly interface design and interactive visual reports, the application is expected to contribute to the transition to sustainable lifestyles on a societal scale through individual behavioral change.

Keywords: carbon footprint, sustainability, energy consumption, mobile application, ocr

1. Introduction

The global climate crisis is deeply affecting not only ecosystems but also individuals' daily lifestyles, consumption habits, and social structures [1,10]. It is emphasized that the increase in greenhouse gas emissions brings environmental and socio-economic risks and that monitoring and managing individual energy consumption is becoming increasingly critical in achieving sustainable development goals [2,3]. Indeed, the largest share of the global carbon footprint comes from sectors such as electricity production (40%), transportation (21%), and manufacturing (16%); this necessitates policymakers to clearly identify priority areas for emission reductions [4]. In particular, under the Paris Agreement and the United Nations Sustainable Development Goals, individuals are encouraged to take conscious steps to reduce their carbon emissions and increase energy efficiency [3,11]. However, with the widespread use of information technologies, it is estimated that the contribution of hardware production processes in areas such as data centers and mobile device manufacturing to the total carbon footprint has exceeded 80% and that this trend will continue to increase [5].

Recent scientific studies have shown that measuring and tracking individual carbon footprints plays an important role in understanding and reducing environmental impact [6,17]. However, existing digital applications and tracking systems often fall short in terms of practical use, real-time data processing, and providing user-specific recommendations [13,15]. The lack of data-driven, user-friendly, and daily life-integrated solutions is frequently highlighted as a gap in the literature [14,16].

Mobile technologies, on the other hand, offer innovative solutions for sustainability by enabling users to continuously monitor their energy consumption and carbon emissions [12,14]. Through developed mobile applications, energy usage becomes more transparent and easier to analyze, and users can receive real-time notifications and personalized recommendations [12,16]. However, it is observed that many applications highlighted in the literature still rely heavily on manual data entry and cannot fully meet advanced features such as automatic data transfer and integrated recommendation systems [13,17].

This study aims to develop an integrated mobile application that can address the shortcomings of existing approaches and help individuals easily share their energy consumption data in a digital environment, automatically calculate their carbon footprint, and develop sustainable lifestyle habits. Thus, the goal is to increase sustainability awareness at both the individual and societal levels and take more effective steps in combating climate change [16,20].

[1]Corresponding author

2. Literature Review

Climate change and environmental sustainability have become increasingly important research areas, especially with the rise of individual energy consumption and carbon emissions [1,10]. There are some internationally recognized standards for carbon footprint calculations. Standards such as the IPCC Guidelines, DEFRA, and PAS 2050 define different emission factors for various energy types, primarily electricity and natural gas, and ensure methodological consistency in calculations [3,8]. However, most existing applications in the literature have been developed for large-scale organizations or industrial analyses, and there are various shortcomings in terms of automatic and practical solutions that individual users can easily use in their daily lives [2,20].

Mobile technologies have been seen as an important tool in achieving sustainability goals in recent years. Mobile applications such as EcoBuddy, JouleBug, and Oroeco, in particular, enable users to track their daily energy consumption and carbon footprint [14,15]. Thanks to these applications, users can more easily see their environmental impact and develop sustainable behaviors. However, the majority of these applications rely on manual data entry and lack more advanced features such as real-time automatic data transfer or personalized recommendations [13,16]. Indeed, studies indicate that fully integrated and user-friendly individual carbon footprint applications are still insufficient in the literature [16]. Additionally, it is emphasized that sustainability policies are not limited to the energy sector; comprehensive strategies to reduce carbon footprints are necessary in areas such as agriculture, food production, and natural resource management [9].

The use of OCR (Optical Character Recognition) technology in energy data analysis stands out as an innovative approach [12]. Recent studies have shown that OCR-based systems provide high accuracy and speed in obtaining energy consumption information from digital or paper bills [18,19]. This allows text extraction and data verification processes to be automated, eliminating the need for users to manually enter invoice data [18]. However, a significant portion of current systems still have limitations in terms of full automation and integrated recommendation generation [19]. In other words, existing solutions typically rely on manual input and do not include components such as automatic data transfer, personal analysis, and sustainable recommendation systems [13].



Figure 1. Image-to-text conversion using ocr technology

However, sustainability recommendation systems and gamification strategies are increasingly being used on digital platforms to change user behavior [15]. It is stated that techniques such as allowing users to earn rewards within the application, track their progress, and be motivated can be effective in encouraging energy savings [16]. However, it is understood that many applications in this field still lack fully integrated automatic data transfer and advanced analysis capabilities and are limited to behavioral recommendations [16,17].

In conclusion, the current literature highlights the lack of individual carbon footprint applications that offer automatic data transfer [18], real-time analysis [12], and user-specific recommendation systems [16] in an integrated manner.

It is believed that the mobile application proposed in this study will offer an innovative solution to address these needs [20,21].

3. Material And Method

The objective of this study is to develop an innovative mobile application model that enables users to accurately, practically, and automatically calculate their carbon footprint based on their daily energy consumption. The application's database is planned to be created using digital electricity and natural gas bills, measurements from smart meters, and energy usage data obtained from mobile devices. In the design of the study, methodological consistency will be ensured by considering both national and international carbon footprint calculation standards (particularly IPCC Guidelines, DEFRA, and PAS 2050) [3,8].

The use of OCR (Optical Character Recognition) technology is targeted for obtaining information from digital invoices so that users can quickly and accurately transfer their energy data to the application [12]. The OCR-based module is planned to automatically detect and transfer basic information such as consumption amount, unit price, and date from invoices to the system [18,19]. This aims to provide a user-friendly and reliable data flow without the need for manual data entry [18]. Technological transformation not only reshapes production and consumption patterns but also directly affects the implementation of sustainability policies by reshaping the labor force structure and social organizations [7]. In addition, it is planned to add real-time energy consumption data from smart meters and mobile devices to the application through API integrations [21].

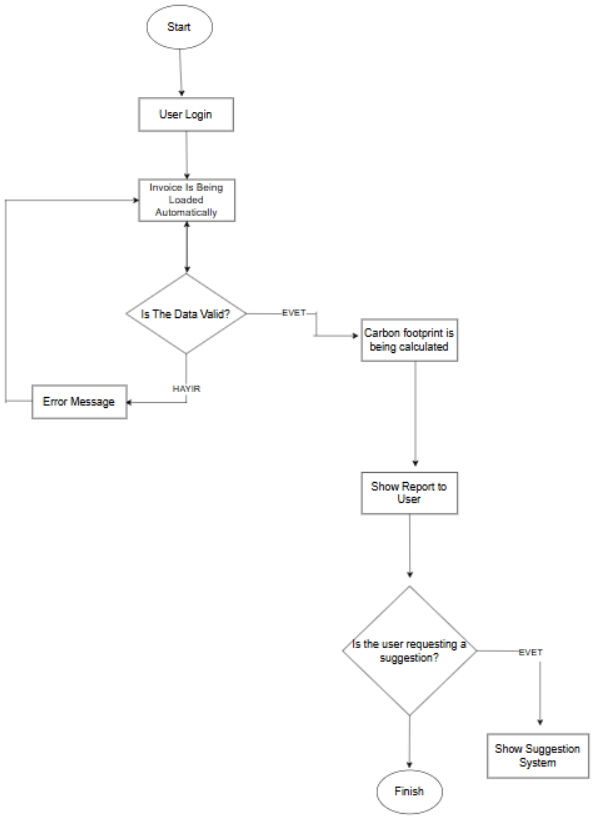


Figure 2. Application flow chart

All collected data is expected to be analyzed using a carbon footprint calculation algorithm developed in accordance with international standards such as the IPCC [20]. Considering specific emission factors for different energy types such as electricity and natural gas, detailed statistical reports such as total carbon footprint and per capita carbon emissions for the time period selected by the user are aimed to be presented [3]. This algorithmic approach is expected to enhance both the scientific validity and the meaningfulness of the results for the user.

One of the unique aspects of this study is the development of a system that can automatically provide personalized recommendations based on the user's energy consumption history. The designed system aims to analyze consumption habits and provide warnings, savings recommendations, and practical tips tailored to the user for periods of increased energy usage [14,15]. For example, personalized recommendations will be communicated to users regarding reducing natural gas consumption during winter months or developing alternative solutions for high electricity usage on weekends [16].

The application's interface design is planned to be simple, understandable, and interactive. Users will be able to easily review changes in their energy consumption over time through graphs, track improvements in their carbon footprint, and receive visual reports to increase their motivation [14]. Additionally, achievement badges and gamification elements are planned to be integrated into the system to reward users for their progress in energy savings [15]. This aims to increase both individual awareness and the widespread adoption of sustainable living habits at the societal level [16,17].

In conclusion, it is anticipated that the mobile application developed within the scope of this study will be an important tool in spreading sustainability awareness at the individual level by distinguishing itself from existing applications with its automatic data transfer [18], integrated recommendation system [16], and user-friendly interface [20,21].

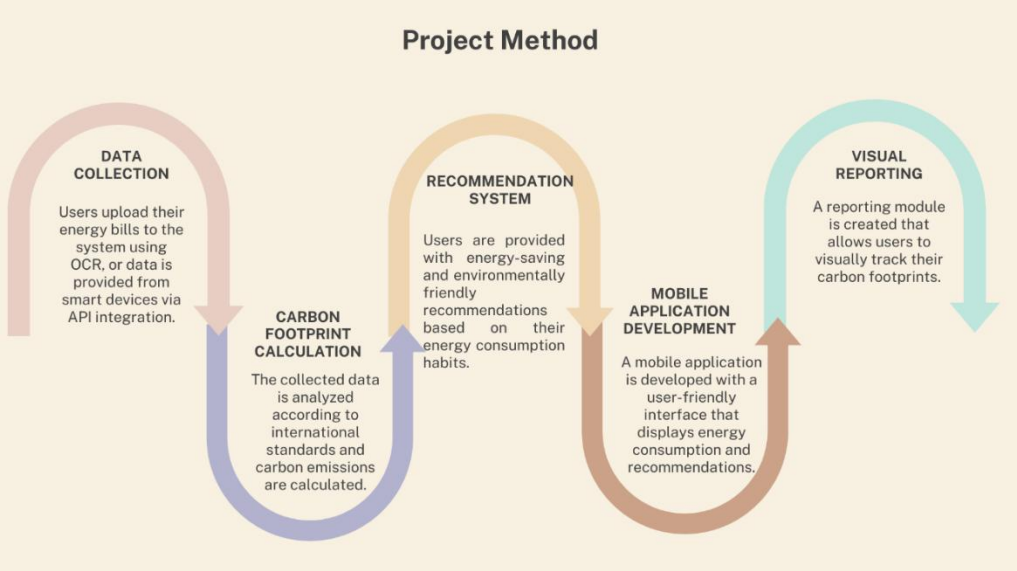


Figure 3. Step-by-step explanation of the project methodology

4. Expected Results And Contributions

Upon completion of the project, it is expected that users will be able to analyze their energy consumption habits more effectively, easily track their carbon footprints, and achieve energy savings through personalized recommendations. It is anticipated that this will increase awareness and a sense of responsibility at the individual level. Furthermore, it is believed that the mobile application to be developed will contribute to the spread of sustainability awareness in

society by reaching a wide audience. It is also aimed that the data obtained within the scope of the project will serve as a source for academic studies to be conducted in the future. The prototype of the application will be shared in scientific conferences and symposiums and published as articles. Furthermore, users will be encouraged to analyze their energy consumption habits and make them more efficient to save energy. It is expected that the reduction of carbon emissions across society will make significant contributions to environmental health [1]. This approach will contribute to both the setting of priority reduction targets in high-emission sectors [4] and the widespread adoption of sustainable practices in other sectors such as agriculture [9]. In the later stages of the project, it is also planned to present the application as a commercial product or prototype.

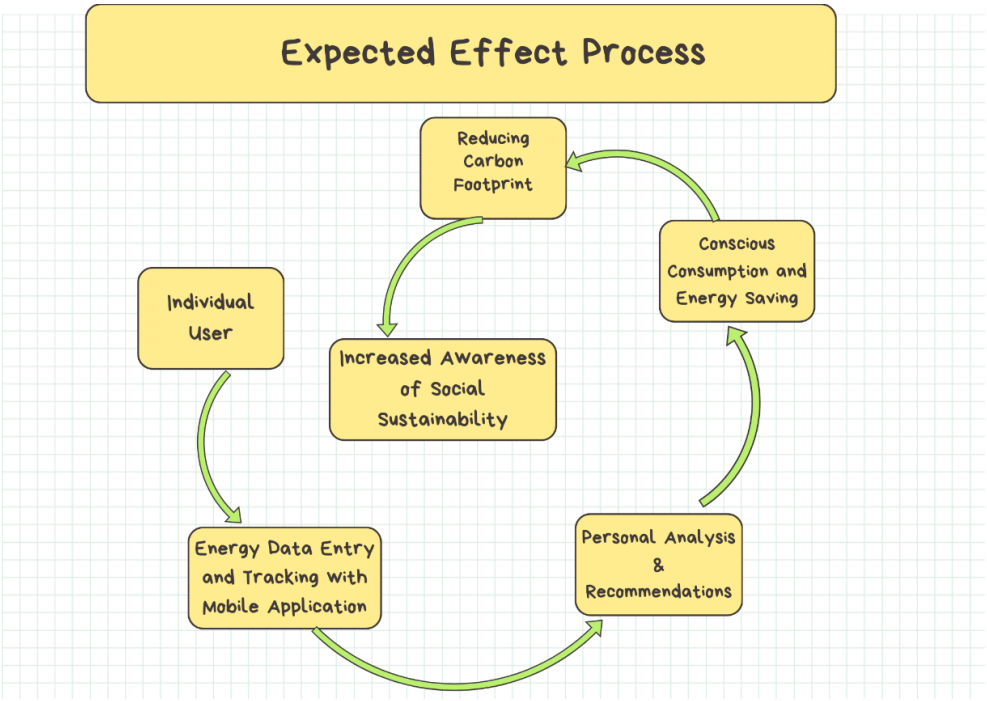


Figure 4. Expected impact process of the mobile application

5. Discussion And Conclusion

This study has demonstrated the potential for the effective use of mobile technologies in the process of tracking individual carbon footprints and fostering sustainable lifestyle habits [12]. In particular, the automatic collection of energy consumption data from digital environments reduces the risk of errors caused by manual input and increases data reliability [18]. Additionally, the aim is to increase individuals' motivation toward energy conservation through user-specific automatic recommendation systems [16]. In this regard, the system not only monitors but also guides users [15].

In future versions of the application, the scope of carbon footprint calculation modules will be expanded to include different sources such as fossil fuels used in transportation and water consumption [20]. Additionally, the development of social modules supported by reward/incentive mechanisms, enabling comparative analyses among users and providing community-level reporting functions, is planned [17]. Such community-focused approaches are expected to play a significant role in translating individual behavior changes into societal impact [16].

The mobile application approach proposed in this paper enables individuals to analyze their energy consumption habits, automatically calculate their carbon footprints, and develop sustainability-oriented behaviors [14]. Through a user-friendly interface, visual reporting tools, and a personalized recommendation system, both individual awareness is increased and concrete steps toward reducing environmental impacts are encouraged [16]. The project outputs are

expected to provide an innovative digital tool that will contribute to energy savings and carbon emissions reduction, especially in densely urbanized areas [21].

6. References

[1] M. E. MiRiCi and S. Berberoğlu, “Green Deal and Carbon Footprint from a Turkish Perspective: Threat or Opportunity?,” Journal of Natural Disasters and Environment, 2022.
[2] G. Çolak and B. A. Türkmen, “Corporate carbon footprint analysis: example calculation for a chemical factory,” Journal of Natural Disasters and Environment, 2023.
[3] V. Sebestyén, T. Czvetkó, and J. Abonyi, “The Applicability of Big Data in Climate Change Research,” Frontiers in Environmental Science, 2021.
[4] “Carbon Footprint: Causes, Impacts and Sector-Wise Survey,” IEEE Conference Publication, 2023.
[5] “Chasing Carbon: the elusive environmental footprint of computing,” ar5iv.labs.arxiv.org
[6] E. Ozmehmet, “Sustainable Development Approaches in Turkey and the World,” Yaşar University, 2022.
[7] C. Bekaroğlu, “The Relationship Between Technology and Employment and the Future of Employment,” Çanakkale Onsekiz Mart University, 2022.
[8] M. Engin, “The Impact of Technology and Infrastructure on Sustainable Urbanism: Achieving Sustainability Goals Through the Smart City Approach,” Jass Studies, 18(103), 449-466, 2025.
[9] G. Bozma, A. Aksoy, & F. Urak, “Underground Treasure: Economic Implications of Sustainable Agriculture,” Efe Academy, 2024.
[10] B. Şıvgın & S. Afacan, “A Study on Global and Individual Efforts Against the Climate Crisis: The Case of Greta Thunberg,” DSJOURNAL, 1(2), 141-158, 2020.
[11] Ö. F. Bilbay, “Climate Crisis and Digitalization,” Journal of Management Sciences, 22, 1544-1568, 2024
[12] Harish Narne, "Real-Time Carbon Footprint Monitoring and Predictive Reporting Cloud Solutions for Sustainability," ICETM 2025.
[13] Steffen Dalsgaard, “Can IT Resolve the Climate Crisis? Sketching the Role of an Anthropology of Digital Technology,” Sustainability, 2022.
[14] Annada Dash et al., "EcoTrack: A Mobile Application for Real-Time Carbon Footprint Tracking and Sustainable Living", IMC-SSGP 2024.
[15] Benjamin Brauer et al., "GREEN BY APP: THE CONTRIBUTION OF MOBILE APPLICATIONS TO ENVIRONMENTAL SUSTAINABILITY," PACIS 2016 Proceedings.
[16] K. Arunasakthi, V. Akilandeswari, N.G. Malni, A.S.K. Arun, B. Yamuna, “ECOSAUR - Empowering Sustainable Choices: A Mobile Application for Carbon Footprint Awareness and Mitigation,” IEEE AIDE 2025.
[17] F. Gençtürk, S. Biroğul, “Carbon Footprint Through the Lens of Information and Communication Technology,” International Journal of Sustainable Engineering and Technology, 2025.
[18] Chao Liu, Xi Wang, Yiran Zhai, Shuxiu Liu, Shuang Yang, Feng Xu, Dingwen Zhang, Ning Li. "Power Big Data Imbalance Analysis Based on Deep Learning OCR Algorithm." 2025 International Conference on Electrical Automation and Artificial Intelligence (ICEAAI), IEEE, 2025.
[19] K. Karthick, Dr. S. Chitra. "Novel Method for Energy Consumption Billing Using Optical Character Recognition." Energy Engineering, Vol. 114, No. 3, 2017, pp. 64-73. DOI: 10.1080/01998595.2017.11863765.
[20] Tao Gao, Qing Liu, Jianping Wang. "A comparative study of carbon footprint and assessment standards." International Journal of Low-Carbon Technologies, Volume 9, Issue 3, 2014, Pages 237–243. https://doi.org/10.1093/ijlct/ctt041
[21] Narendra Kumar, Krishna Kumar, Anil Kumar. "Application of Internet of Things in Image Processing." IEEE DELCON 2022. [DOI: 10.1109/DELCON54057.2022.9753308]

Artificial Intelligence-Assisted Image Processing Models for Disease Detection in Wheat Plants

Miraytuğ ŞEKER[1], Zeynep Miray AKKAN[2], Mehmet ARZU[3]

[1,3]Artificial Intelligence and Data Engineering, Faculty of Engineering, Firat University, Elazig, Türkiye.
[2] Biosystems Engineering, Faculty of Agriculture, Malatya Turgut Ozal University, Malatya, Türkiye.
[1]miraytug1903@gmail.com, [2]z.mirayakkan@gmail.com, [3]marzu@firat.edu.tr
[1](ORCID: 0009-0000-4465-3546), [2](ORCID: 0009-0006-9692-8752), [3](ORCID: 0000-0001-6610-2788)

Abstract
Wheat is one of the world's most strategic products in terms of both food security and economic sustainability; however, various diseases such as yellow rust, Septoria tritici blotch, Fusarium head blight, and root rot seriously threaten production. These diseases cause yield and quality losses of up to 35% in some regions. With the growing global population and food demand, the early and accurate detection of such diseases has become critical. In this study, 14,155 images of healthy and diseased wheat leaves from the Wheat Plant Diseases dataset were used. Five different deep learning architectures were applied: CNN Scratch, VGG16, ResNet50, EfficientNetB0, and MobileNetV2. Preprocessing steps such as resizing, normalization, and data augmentation were applied to improve model performance. The models were evaluated using accuracy, precision, sensitivity, and F1 score metrics. Experimental results showed that the CNN Scratch model achieved the highest success with 90.13% accuracy and 88.65% F1 score. Among transfer learning-based architectures, EfficientNetB0 demonstrated relatively better performance. The findings reveal that customized CNN architectures can provide more effective and field-specific solutions for wheat disease classification, thereby contributing significantly to digitalization and sustainable production in agriculture.

Keywords: wheat diseases, convolutional neural network, transfer learning, image processing, deep learning

1. Introduction

Agriculture is used not only to provide food but also to increase the country's income [1]. The United Nations (UN) estimates that the world population will reach 9.7 billion by 2050. With this increase, food demand is expected to rise by up to 60%. However, plant diseases cause a decrease in the quantity and quality of agricultural production [2][3].

Wheat is a cool-climate grain that grows in environments that are not hot and humid [4]. It is also known as one of the oldest and most basic food sources in human history. Wheat plays an important role not only in the production of basic foodstuffs such as bread, pasta, bulgur, and biscuits, but also in animal feed and industrial applications. Approximately 20-25% of the energy needs of the world's population are met by wheat. Therefore, wheat is of strategic importance in terms of food security and economic sustainability [5].

Global wheat production reached approximately 799 million tons in 2023, with approximately 50% of this production coming from China, India, and Russia [6]. In Turkey, wheat production in the 2024-25 period was approximately 19 million tons, imports were 5.8 million tons, and exports were 7 million tons. A significant portion of imports originated from Russia [7][8]. Considering both its production and the products it is used in, wheat holds great economic importance; however, diseases that occur during production lead to significant yield losses.

The main diseases limiting wheat production include yellow rust, Septoria tritici blotch, Fusarium head blight, and Fusarium crown rot. These diseases not only cause yield losses but also reduce quality. For example, yellow rust causes approximately 5.47 million tons of yield loss worldwide each year [9]. In the US, yield loss caused by such diseases amounts to 8.3% overall, and 27% in the Alberta and Ontario regions [10]. Fusarium crown rot causes an average yield loss of 9.5% in dry farming regions of the US.

In some areas, this value can reach 35% [11]. Similarly, Septoria tritici blotch causes losses of 35-50% in yield. In the past, monitoring the health of plants was quite difficult, and producers suffered serious losses in production because they could not diagnose diseases at an early stage [12][13]. The inability to detect diseases in agricultural production in a timely manner also led to significant declines in yield. The need for numerous experts, as well as

[1]Corresponding author

costly and continuous garden inspections to identify plant diseases, made it difficult to obtain expert support in a timely manner [14][15]. In addition, many producers resorted to excessive pesticide use to prevent diseases. However, this situation led to both unnecessary pesticide use and increased production costs.

Today, rapidly advancing artificial intelligence and image processing techniques enable the rapid and accurate prediction of diseases in agricultural production. In this context, the study attempted to predict wheat diseases using artificial intelligence and image processing techniques. This study not only contributes to reducing production losses but also strengthens participation in the digitalization ecosystem in agriculture and supports automatic information extraction processes.

2. Related Works

There are many studies in the literature on the detection and classification of plant diseases. These studies show that different transfer learning approaches are effective in accurately and quickly identifying agricultural diseases. Some of these studies are presented in this section.

Gündüz et al. [16] aimed to classify five types of diseases observed in pomegranate fruits in their study. They tested various machine learning models on 5,099 images collected from the Karnataka region. As a result, they observed that convolutional neural networks yielded the most successful results with 88% accuracy. Özben and Güler [17] used transfer learning-based deep learning methods to detect diseases in tomato leaves in their study. The MobileNetV3-Large model demonstrated the best performance, achieving 99% accuracy in the testing phase. Sazak et al. [18] compared transfer learning-based deep learning models for the classification of plant leaf diseases in their study. The MobileNetV1 model achieved the highest accuracy value of 99.20% in the study. Topçu and Güneş [19] used the ResNet architecture in their study on the classification of plant leaf diseases. The model achieved 99% accuracy. Aybar et al. [20] attempted to detect diseases found on apple leaves using transfer learning models in their study. They achieved 99.84% accuracy with the Xception model and 98.58% accuracy with the InceptionV3 model.

3. Materials and Methods

3.1. Dataset

This study contains 14,155 images of healthy and diseased wheat plants. These images were taken from the Wheat Plant Diseases dataset available on Kaggle. This dataset includes 15 types of diseases observed in wheat plants. In the “Wheat Plant Diseases” dataset [21], the images are divided into 13,104 training images, 750 test images, and 300 validation images. Images of some of these diseases are shown in Figure 1.



Figure 1. Some of the wheat diseases in the dataset

Wheat diseases in the dataset: Aphid (Pest), Mite (Pest), Stem Fly (Pest), Black Rust / Stem Rust, Brown Rust / Leaf Rust, Yellow Rust / Stripe Rust, Smut (Loose, Flag), Common Root Rot, Helminthosporium Leaf Blight / Leaf Blight, Wheat Blast, Fusarium head blight / Scab, Septoria Leaf Blotch, Spot Blotch, Tan Spot, Powdery Mildew. The distribution of these diseases in the dataset is shown in Figure 2.

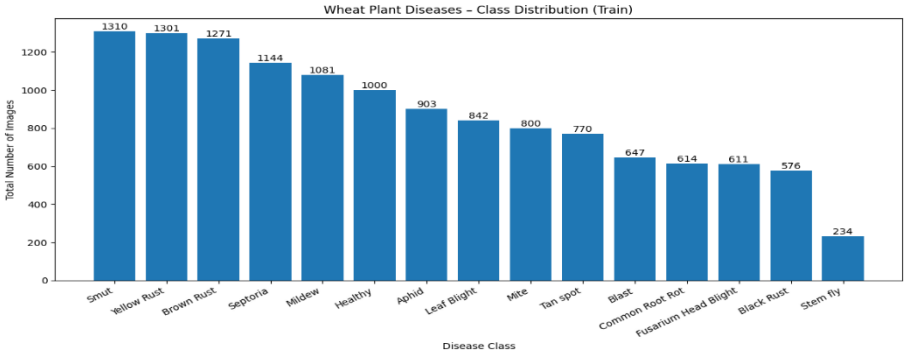


Figure 2. Distribution of wheat diseases in the data set

3.2. Method

This section will discuss the data preprocessing steps used in the study and the models employed. In the study, accuracy, precision, sensitivity, and F1 score metrics were used as model evaluation criteria, and Python was used as the programming language during the execution of the experiments. Visual Studio Code was used as the coding environment.

3.2.1. Data Preprocessing

The Wheat Plant Diseases dataset used in this study contains images of 15 different disease classes affecting wheat. The dataset is divided into three different sets: training, validation, and testing. Before being used in model training, the images were subjected to a series of preprocessing steps.

First, all images were resized to fit the input size of the networks. The input size of the images was set to 224×224 pixels for the CNN Scratch, VGG16, ResNet50, EfficientNetB0, and MobileNetV2 models. This step ensured the standardization of raw data with different resolutions. In the second stage, a normalization process was applied. Pixel values were converted from the range 0–255 to the range 0–1. In addition, special preprocessing functions specified by the relevant libraries were used in transfer learning-based models (VGG16, ResNet50, EfficientNetB0, MobileNetV2). The aim was to make the weights of the pre-trained networks compatible with the data distribution. Thirdly, to reduce overfitting and improve the model's generalization ability, data augmentation techniques were applied only to the training set. Within this scope, horizontal and vertical flipping, ±15° random rotation, ±10% zoom in/out, brightness and contrast changes, and limited shifting operations were used. Only resizing and normalization operations were applied to the validation and test subsets, and no augmentation methods were applied. Finally, faulty or corrupted files were checked and removed from the dataset. Thus, only high-quality images with correct labels were used for model training. The purpose of all these preprocessing steps is to improve data integrity and model performance.

3.2.2. Models Used

Five different deep learning-based architectures were used in this study. First, a convolutional neural network designed from scratch was used for basic comparison. The CNN architecture is a method commonly used in image classification problems and was first introduced by LeCun and colleagues [21]. Second, the VGG16 model was chosen. With its 16-layer structure, VGG16 has made significant contributions to scaling the performance of deep convolutional networks and achieved high success in the ImageNet competition [22]. Thirdly, the ResNet50 architecture was used. This model, which enables deep networks to be trained with residual connections, has provided

an important solution to the problem of gradient decay [23]. Fourthly, EfficientNetB0 was used. This model offers an improved architecture that provides high accuracy with fewer parameters by scaling depth, width, and resolution together [24]. Finally, MobileNetV2 was used. Designed for use in mobile and embedded devices, this model is used for image classification, object detection, and other computer vision tasks [25].

Evaluation Metrics

In this study, four basic classification metrics were used to evaluate model performance in the classification of wheat diseases. These metrics are accuracy, precision, recall, and F1 score. These metrics are widely used in classification-based image processing studies to quantitatively evaluate model performance. Accuracy represents the ratio of correctly predicted examples to the total number of examples and is a metric that summarizes the overall level of success. The accuracy formula is given in Equation (1).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Here, TP (True Positive) refers to true positives, TN (True Negative) refers to true negatives, FP (False Positive) refers to false positives, and FN (False Negative) refers to false negatives. Precision indicates how many of the examples classified as positive by the model are actually positive. This metric is particularly important in situations where it is critical to keep false positives low. The precision formula is presented in Equation (2).

$$Precision = \frac{TP}{TP+TF} \quad (2)$$

Sensitivity (Recall) indicates the ratio of true positive examples in the data set that the model correctly predicts. Sensitivity is particularly important in applications where missed positive examples can have critical consequences. The sensitivity formula is given in Equation (3).

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

Finally, the F1 score shows the harmonic mean of precision and sensitivity. This metric creates a more equitable performance by evaluating both the accuracy and coverage of the model. The formula for the F1 score is given in Equation (4).

$$F1\ Score = \frac{2*precision*recall}{precision+recall} \quad (4)$$

4. Experimental Results

This study compares the performance of five different deep learning models in classifying wheat diseases. Five models were used, namely CNN Scratch Model, VGG16, ResNet50, EfficientNetB0, and MobileNetV2. The results obtained by the models used in the study are presented in Table 1.

Table 3. Results of the models used in the study

Modeller	Accuracy	Precision	Recall	F1 Score	Training Time (s)
CNN Scratch Model	%90,13	%92,36	%90,13	%88,65	2080
VGG16	%73,2	%73,13	%73,19	%71,71	2087
ResNet50	%81,73	%85,29	%81,73	%80,26	2280
EfficientNetB0	%83,33	%83,99	%83,33	%81,68	2039
MobileNetV2	%82,13	%82,75	%82,13	%80,62	2200

According to experimental findings, the CNN Scratch Model showed the highest performance with an accuracy rate of 90.13%. Additionally, this model's precision was measured at 92.36%, sensitivity at 90.13%, and F1 score at 88.65%, and it was determined to be superior to other models. Among transfer learning-based models, EfficientNetB0 demonstrated the most successful performance with an accuracy rate of 83.33% and an F1 score of 81.68%. The ResNet50 and MobileNetV2 models had accuracy rates of 81.73% and 82.13%, respectively, while VGG16 showed the lowest success with an accuracy rate of 73.2%. Training times were similar across models, ranging from approximately 2000 to 2300 seconds.

5. Results

The results obtained show that the CNN model created from scratch demonstrated higher classification success compared to transfer learning-based models. This suggests that an architecture obtained through data set-specific improvements may be more compatible than ready-made deep learning architectures. However, EfficientNetB0 has demonstrated relatively higher performance among transfer learning models and has emerged as an alternative solution. The low success rate of VGG16 suggests that older generation architectures may be insufficient for current data sets. This study has demonstrated that a CNN architecture designed from scratch can be more effective in detecting wheat diseases than the transfer learning methods commonly preferred in the literature, thereby providing an innovative contribution to the literature. As a result, the most effective method for this study was determined to be the CNN Scratch Model.

6. References

[1] Gobalakrishnan N, Pradeep K, Raman CJ, Ali LJ, Gopinath MP. "A systematic review on image processing and machine learning techniques for detecting plant diseases". 2020 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 465-468, 28-30 July 2020. IEEE. https://doi.org/10.1109/ICCSP48568.2020.9182046.

[2] Tiwari V, Joshi RC, Dutta MK. "Dense convolutional neural networks based multiclass plant disease detection and classification using leaf images". Ecological Informatics, 63, 101289, 2021. https://doi.org/10.1016/j.ecoinf.2021.101289.

[3] Singh UP, Chouhan SS, Jain S, Jain S. "Multilayer convolution neural network for the classification of mango leaves infected by anthracnose disease". IEEE Access, 7, 43721-43729, 2019. https://doi.org/10.1109/ACCESS.2019.2907383.

[4] Kaygıner E. "Buğday (Triticum L.)". ISPEC Journal of Science Institute, 2(1), 9-27, 2023.

[5] Shewry PR. "Wheat". Journal of Experimental Botany, 60(6), 1537-1553, 2009.

[6] Farmonaut. "2025 Wheat Harvest: Global Challenges and Innovations". https://farmonaut.com/news/2025-wheat-harvest-global-challenges-innovations (Accessed: 15.08.2025).

[7] United States Department of Agriculture (USDA). "Grain and Feed Annual – Turkey (TU2025-0015)". Foreign Agricultural Service,Ankara,Turkey,2025.https://apps.fas.usda.gov/newgainapi/api/Report/DownloadReportByFileName?fileName=Grain+and+Feed+Annual_Ankara_Turkiye_TU2025-0015.pdf (Accessed: 15.08.2025).

[8] United States Department of Agriculture (USDA). "Grain and Feed Update – Turkey (TU2025-0038)". Foreign Agricultural Service,Ankara,Turkey,2025.https://apps.fas.usda.gov/newgainapi/api/Report/DownloadReportByFileName?fileName=Grain+and+Feed+Update_Ankara_Turkiye_TU2025-0038.pdf (Accessed: 15.08.2025).

[9] Beddow JM, Pardey PG, Chai Y, Kriticos DJ, Braun HJ, et al. "Global annual average wheat loss estimates attributed to stripe rust". Journal of Agricultural Science, (details pending), 2015.

[10] Crop Protection Network. "Wheat disease loss estimates from the United States and Ontario, Canada – 2024". Crop Protection Network, 2025. (Accessed: 15.08.2025).

[11] Smiley RW. "Crop damage estimates for crown rot of wheat and barley: Winter wheat yield loss as much as 35%, mean 9.5%". Plant Disease, 89(5), 495–503, 2005.

[12] Sharath DM, Rohan MG, Kumar SA, Prathap C. "Disease detection in pomegranate using image processing". 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 994-999, 15-17 June 2020. IEEE. https://doi.org/10.1109/ICOEI48184.2020.9142972.

Singh UP, Chouhan SS, Jain S, Jain S. "Multilayer convolution neural network for the classification of mango leaves infected by anthracnose disease". IEEE Access, 7, 43721-43729, 2019. https://doi.org/10.1109/ACCESS.2019.2907383.
[13] Akhtar F, Partheeban N, Daniel A, Sriramulu S, Mehra S, Gupta N. "Plant disease detection based on deep learning approach". 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 74-77, 4-5 March 2021. IEEE. https://doi.org/10.1109/ICACITE51222.2021.9404647.
[14] Silviya SA, Sriman B, Shamini PB, Elangovan A, Keerthana NV. "Deep learning-based plant leaf disease detection and classification". 2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 702-710, 20-22 July 2022. IEEE. https://doi.org/10.1109/ICIRCA54612.2022.9985548.
[15] Anand V, Gupta S, Gupta R, Alshahrani H, Hamdi M, Al Reshan MS, Shaikh A. "Plant Disease Diagnosis Using Transfer Learning Based Models". 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE), Las Vegas, USA, 337-340, July 2023. IEEE. https://doi.org/10.1109/CSCE60160.2023.00058.
[16] Gündüz R, Bingol MC, Bişkin OT. "Yapay Zeka Temelli Bazı Nar Hastalıklarının Tahmini". Turkish Journal of Agriculture-Food Science and Technology, 13(4), 946-957, 2025.
[17] Özben TÇ, Güler O. "Domates Yapraklarında Hastalık Tespiti İçin Transfer Öğrenme Kullanılması ve Mobil Uygulamaya Entegre Edilmesi". Journal of Information Systems and Management Research, 7(1), 105-113, 2025.
[18] Sazak S, Balsak SC, Badem H. "Transfer Öğrenme Temelli Bitki Yaprak Hastalıklarının Tespiti İçin Karşılaştırmalı Bir Çalışma". Kahramanmaraş Sütçü İmam Üniversitesi Tarım ve Doğa Dergisi, 28(1), 154-170, 2025.
[19] Topçu C, Güneş P. "Bitki hastalıklarını tespitte derin öğrenme: ResNet modelinin etkinliği". Anadolu Bil Meslek Yüksekokulu Dergisi, 19(69), 31-65, 2024.
[20] Aybar M, Talaş U, Çubukçu B. "Transfer Öğrenme Modelleri ile Elma Yapraklarında Hastalık Tespiti". Eskişehir Türk Dünyası Uygulama ve Araştırma Merkezi Bilişim Dergisi, 5(2), 57-63, 2024.
[21] Kaggle (kushagra3204). "Wheat Plant Diseases Dataset". https://www.kaggle.com/datasets/kushagra3204/wheat-plant-diseases (15.08.2025).
[22] LeCun Y, Bottou L, Bengio Y, Haffner P. "Gradient-based learning applied to document recognition". Proceedings of the IEEE, 86(11), 2278–2324, 1998. https://doi.org/10.1109/5.726791.
[23] Simonyan K, Zisserman A. "Very deep convolutional networks for large-scale image recognition". International Conference on Learning Representations (ICLR), San Diego, USA, 2015. https://arxiv.org/abs/1409.1556.
[24] He K, Zhang X, Ren S, Sun J. "Deep residual learning for image recognition". IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, USA, 770–778, 2016. https://doi.org/10.1109/CVPR.2016.90.
[25] Tan M, Le QV. "EfficientNet: Rethinking model scaling for convolutional neural networks". International Conference on Machine Learning (ICML), Long Beach, USA, 6105–6114, 2019. https://arxiv.org/abs/1905.11946.
[26] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. "MobileNetV2: Inverted residuals and linear bottlenecks". IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, USA, 4510–4520, 2018. https://doi.org/10.1109/CVPR.2018.00474.

An Artificial Intelligence-Based Fraud Detection Approach for Securing Blockchain Networks

Kerem Can GÜNDÜZ[1] , Hasan YETİŞ [2]

[1,2]Department of Computer Engineering, Faculty of Engineering, Firat University, Elazig, Turkey.
[1]3kerem1@gmail.com, [2]hyhasanyetis@gmail.com
[1](ORCID: 0009-0004-8853-5532), [2](ORCID: 0000-0001-7608-3293)

Abstract
Financial fraud detection systems increasingly rely on Artificial Intelligence models, yet their performance is fundamentally challenged by the highly imbalanced nature of real-world data. This study investigates the critical impact of data distribution on the reliability of machine learning models designed for detecting credit card fraud, with the ultimate goal of securely recording findings on a blockchain. We utilized a large dataset where fraudulent transactions constituted less than one percent of the total data. The investigation compares the performance of prominent models, including LightGBM and XGBoost, under two distinct scenarios: training on the original, imbalanced dataset versus training on a balanced dataset created through random undersampling. The findings reveal that models trained on the balanced dataset demonstrate a vastly superior and more stable performance, achieving a near-perfect equilibrium between precision and recall. For instance, the LightGBM model's F1-score improved from a mere 23.79% on the imbalanced test set to 97.26% on the balanced test set. This stark contrast proves that addressing data distribution at the data level through techniques like undersampling is more effective than relying solely on algorithmic adjustments. The study concludes that proper data balancing is a foundational prerequisite for developing a trustworthy AI-powered fraud detection system, which is essential to prevent the contamination of an immutable blockchain ledger with false positives.

Keywords: fraud detection, data imbalance, machine learning, blockchain, undersampling

1. Introduction

The digital transformation of financial services has reshaped the global economy, but it has also introduced a more complex and expansive attack surface for financial fraud [1]. Traditional rule-based fraud detection systems have proven insufficient against the dynamic and constantly evolving tactics used by fraudsters, leading to a high rate of false positives and an inability to adapt [11]. Consequently, data mining and machine learning techniques have become one of the most intensely applied fields for fraud detection. These approaches leverage sophisticated algorithms to identify complex patterns in vast amounts of data that would be impossible for humans or simple rule-based systems to detect.

One of the most fundamental challenges in developing modern fraud detection systems is the effective processing of large-volume datasets that are inherently and severely imbalanced. In real-world financial transactions, fraudulent cases are extremely rare compared to the total number of legitimate transactions, often making up less than 1% of the data. The dataset used in this study exemplifies this issue, with fraud cases constituting only about 0.46% of all transactions. This imbalance causes standard classification algorithms to fail in learning the characteristics of the minority class (fraudulent transactions), resulting in a poor detection performance even when overall accuracy scores appear high [3].

To overcome these challenges, advanced ensemble learning algorithms have gained prominence. This study focuses on two leading gradient boosting models: XGBoost (Extreme Gradient Boosting) and LightGBM (Light Gradient Boosting Machine). XGBoost is a highly optimized algorithm based on gradient boosted decision trees, which sequentially adds new models to correct the errors of its predecessors, demonstrating exceptional performance on large and high-dimensional datasets. LightGBM is a further evolution of this principle, offering significant improvements in efficiency and scalability by using novel techniques for data and feature splitting, which results in lower memory usage and faster training times without sacrificing accuracy.
While these AI-based systems significantly enhance detection capabilities, they are typically run on centralized architectures, which introduces risks related to the integrity and security of the stored results. Data stored on central servers remains vulnerable to manipulation, censorship, or a single point of failure, which undermines the reliability

[1]Corresponding author

of the records, especially in legal proceedings or inter-institutional data sharing [5]. In response to this need for trust, Blockchain technology offers a groundbreaking solution with its decentralized, transparent, and cryptographically secured immutable ledger. By recording data on a distributed network, blockchain ensures that once a transaction is added, it cannot be altered, providing an unparalleled level of trust and integrity for financial data [6]. This architecture, enhanced by Smart Contracts, allows for the creation of a verifiable "proof chain" for each detected fraud case, free from the control of any single authority [8].

This paper, therefore, investigates the critical impact of data distribution on the performance of high-efficacy models like XGBoost and LightGBM. We analyze the trade-offs between training these models on the original, imbalanced data versus a balanced dataset created via undersampling. The primary objective is to determine the most effective data strategy for building a reliable classifier whose decisions are trustworthy enough to be permanently recorded on a secure blockchain. An unreliable model producing numerous false positives would "pollute" the immutable ledger, defeating its purpose as a source of truth [6]. Our findings demonstrate how a carefully considered data distribution strategy is fundamental to creating a truly robust and synergistic AI-Blockchain system for financial security.

2. Literature Survey

The field of financial fraud detection has evolved significantly, moving from traditional statistical methods and rule-based systems to more sophisticated data mining and machine learning approaches [10]. Early machine learning applications in this domain often utilized algorithms like decision trees and Support Vector Machines (SVM), which laid the groundwork for automated pattern recognition. However, studies have shown the limitations of these models; for instance, Nobel et al. (2024) found that while SVM could achieve a certain accuracy, it struggled with recall, a critical metric for not missing fraud cases. A central and persistent challenge highlighted throughout the literature is the severely imbalanced nature of real-world financial datasets, where fraudulent transactions are exceedingly rare. This imbalance can render standard models ineffective, as they tend to be biased towards the majority class (legitimate transactions), leading to poor detection of the minority class (fraud). To address this, researchers have explored various strategies, including advanced feature engineering to create more discriminative data and the use of scalable frameworks like Spark to handle large, real-time data streams. More recently, the focus has shifted towards powerful ensemble and gradient boosting algorithms like XGBoost and LightGBM, which have consistently demonstrated superior performance. Nobel et al. (2024) achieved 99.88% accuracy and 96% precision with XGBoost, while Hashemi et al. (2023) showed that combining LightGBM and XGBoost could yield an F1-score of 79%. Despite these high-performing models, limitations remain; results obtained on balanced data may not always generalize to real-world scenarios, and combining models can be computationally expensive. Parallel to advancements in AI, blockchain technology has emerged as a solution for ensuring data integrity and trust in financial systems. The use of smart contracts to create immutable and auditable records of financial transactions is a growing area of research The integration of AI and blockchain is a promising but nascent field, with research exploring its potential to create transparent and automated decision-making systems. However, this integration is not without challenges, including scalability, transaction costs, and data privacy concerns. This paper contributes to this emerging area by not only implementing high-performance models like XGBoost and LightGBM but also by focusing specifically on how data distribution strategies directly impact the reliability of these models for creating a trustworthy and immutable fraud ledger on a blockchain.

3. Material and Methods

In this study, a multi-stage methodology was followed to analyze how the fraud detection performance of artificial intelligence models is affected by data distribution and to secure the reliability of these detections with blockchain technology.

3.1. Dataset and Characteristics

The “Credit Card Fraud Detection” dataset on the Kaggle platform, which is synthetically generated to mimic real-world credit card transactions, was used for the analysis. The dataset includes a training (fraudTrain.csv) file with approximately 1.3 million records and a test (fraudTest.csv) file with 556 thousand records. The most critical feature

of this dataset, which directly affects the modeling process, is the high degree of class imbalance. Of the 1,296,675 transactions in the training set, only 6,005, or approximately 0.46%, were labeled as fraud cases. Such a significant imbalance causes standard machine learning algorithms to tend to learn the majority class of normal transactions and miss the minority class of fraud cases. This has necessitated a focus on metrics such as Precision, Recall and F1-Score instead of accuracy to accurately assess model performance.

Table 1. Detailed attribute definitions for the credit card fraud detection dataset

Column	Description
index	Unique identity number for each row
trans_date_trans_time	Transaction date and time
cc_num	Customer's credit card number (anonymized)
merchant	Merchant name
category	Merchant category
amt	Transaction amount
first	Credit card holder's first name
last	Credit card holder's last name
gender	Credit card holder's gender
street	Credit card holder's street address
city	Credit card holder's city
state	Credit card holder's state
zip	Credit card holder's postal code
lat	Credit card holder's latitude
long	Credit card holder's longitude
city_pop	Population of the city where the card holder lives
job	Credit card holder's occupation
dob	Credit card holder's date of birth
trans_num	Transaction number
unix_time	UNIX timestamp of the transaction
merch_lat	Merchant's latitude
merch_long	Merchant's longitude
is_fraud	Fraud label (Target class: 1 = fraud, 0 = normal)

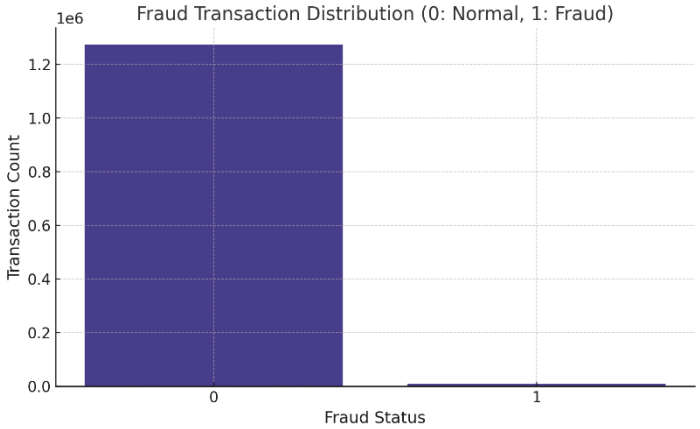


Figure 1. Target distribution of the dataset used

3.2. Data Preprocessing

In order to maximize the performance of the machine learning models, a comprehensive data preprocessing and feature engineering pipeline was employed. During this process, informative features such as "age at transaction" and the logarithmic transformation of transaction amount ("amt_log") were engineered to provide the models with

more relevant signals. Columns with high cardinality (e.g., "cc_num", "merchant") and those likely to introduce noise or redundancy were excluded from the dataset to reduce dimensionality and prevent overfitting. Nominal categorical variables, including "category" and "gender," were transformed into numerical format using One-Hot Encoding, thereby preventing the introduction of spurious ordinal relationships. To address potential multicollinearity, the first category for each feature was dropped during encoding. Additionally, all numerical features were standardized using the StandardScaler, ensuring that each variable contributed equally to the learning process regardless of its original scale. This thorough data preparation ensured that the final dataset was both clean and optimally structured for robust model development.

3.3. Training of Models

In this section, several machine learning models such as LightGBM and XGBoost are developed and trained for credit card fraud detection. Each model is trained on an unbalanced dataset after data processing steps and their performances are compared.

Table 2. Comparison of model test performance metrics with related studies in the literatüre

Study	Model	Accuracy	Precision	Recall	F1-Score	AUC-PR
Ours	LightGBM	0.9763	0.1358	**0.9594**	0.2379	**0.9973**
Nazerke et al.	LightGBM	**0.99**	**0.82**	0.87	**0.84**	0.89
Ours	XGBoost	**0.995**	0.2819	**0.9594**	0.2379	**0.9939**
Nazerke et al.	XGBoost	0.99	**0.92**	0.81	**0.86**	0.88

Although the datasets used in this study are not exactly the same as those found in the literature, they are similar and comparable in structure and characteristics. Therefore, in order to provide a reference and context, we also include the performance metrics reported by related studies. This approach enables a more comprehensive evaluation and benchmarking of our results against previous works.

3.4. Creation of The Balanced Dataset

In order to analyze the impact of data distribution, which forms the basis of this study, a new balanced dataset was created from the original preprocessed imbalanced dataset. For this purpose, the random undersampling technique, which is widely used in the literature, was adopted. In this technique, all examples from the minority class (fraudulent transactions) are preserved, while a randomly selected subset from the majority class (normal transactions) is chosen to match the number of samples in the minority class. As a result of this process, a perfect 1:1 balance (50%-50%) between the "Fraud" and "Normal" classes was achieved in the training dataset. This created a fairer learning environment by preventing the models from developing a bias towards any particular class.



Figure 2. Class distribution in balanced training set

3.5. Training of Models on The Balanced Dataset

Table 3. Model testing success metrics on a balanced dataset

Model	Accuracy	Precision	Recall	F1-Score	AUC-PR
LightGBM	0.9725	0.9673	0.9781	0.9726	0.9966
XGBoost	0.9737	0.9739	0.9734	0.9726	0.9946

Table 2 and Table 3 present the performance metrics of LightGBM and XGBoost machine learning models evaluated on both imbalanced and balanced datasets. The results demonstrate the significant impact of data distribution on model performance, particularly in terms of precision, recall, and F1-score.

On the imbalanced dataset, both LightGBM and XGBoost achieve high accuracy and ROC AUC values (e.g., XGBoost: 0.9905 accuracy, 0.9973 ROC AUC), which may initially suggest strong overall performance. However, the precision and F1-scores for both models are substantially lower (LightGBM: 0.1358 precision, 0.2379 F1-score; XGBoost: 0.2819 precision, 0.4343 F1-score), indicating a bias towards the majority class and poor detection of fraudulent cases. This is a common challenge when dealing with highly imbalanced datasets, where the classifier tends to favor the dominant class, thus failing to adequately identify minority class instances.

In contrast, when evaluated on the balanced dataset, both LightGBM and XGBoost exhibit dramatic improvements in all key metrics. For example, XGBoost achieves a precision of 0.9739, recall of 0.9734, F1-score of 0.9737, and ROC AUC of 0.9968, while LightGBM also records comparably high values across all metrics. These results indicate that balancing the dataset leads to a more reliable and fair assessment of the models, as it forces the algorithms to equally consider both classes. The substantial gains in precision, recall, and F1-score confirm that the models are significantly better at identifying fraudulent transactions when the class distribution is balanced.

In summary, these findings highlight the critical importance of addressing class imbalance in fraud detection problems. Relying solely on accuracy or ROC AUC in imbalanced settings can be misleading. Instead, balancing the dataset enables more meaningful evaluation and leads to more robust, fair, and effective fraud detection models.

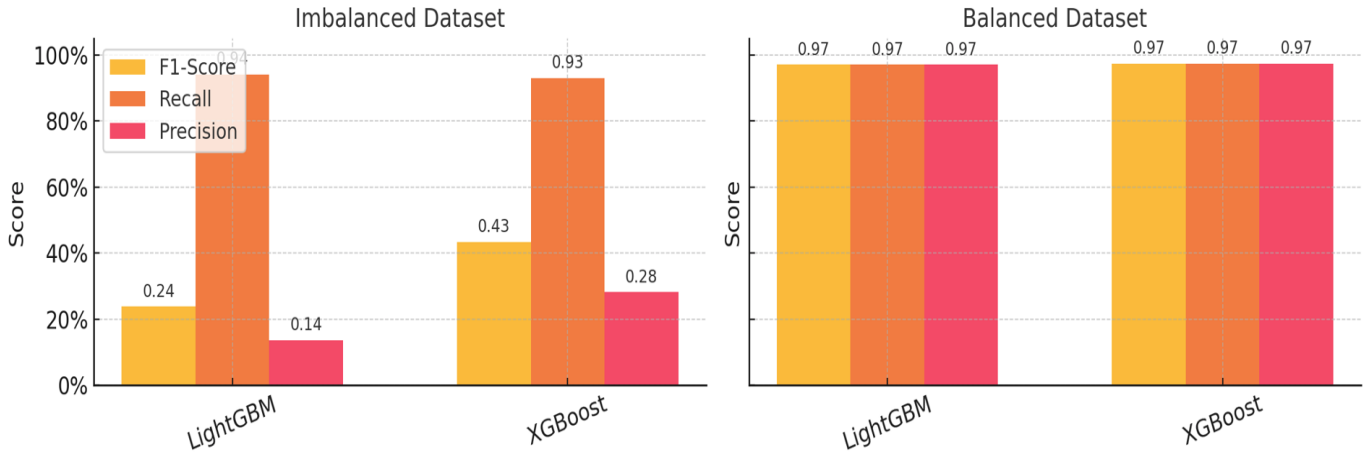


Figure 3. Impact of dataset balancing on LightGBM and XGBoost fraud detection performance

3.6. Cross Validation

To ensure the robustness and generalizability of our machine learning models, we performed 5-fold cross-validation on the balanced dataset for both LightGBM and XGBoost. This approach allows for a more reliable assessment of model performance by reducing the risk of overfitting and providing a better estimate of how the models will perform on unseen data. In each fold, the data is partitioned into training and validation subsets, and performance metrics are averaged across all folds.

As shown in Tables 4 and 5, both LightGBM and XGBoost demonstrated consistently high performance across all folds. The standard deviations of the metrics are remarkably low, indicating the stability and reliability of the models. Specifically, both models achieved average accuracy, precision, recall, F1-score, and ROC AUC values exceeding 0.97, which demonstrates their strong discriminative power for fraud detection on the balanced dataset.

The cross-validation results confirm that the models are not only effective but also robust, maintaining high performance regardless of the particular subset of data used for validation. This consistency reinforces the validity of our findings and suggests that the models are well-suited for deployment in real-world credit card fraud detection scenarios.

Table 4. LightGBM model 5-layer cross-validation results

Layer	Accuracy	Precision	Recall	F1-Score	AUC-PR
1	0.9830	0.9795	**0.9867**	0.9831	0.9986
2	0.9770	0.9711	0.9833	0.9772	0.9973
3	0.9847	0.9834	0.9860	0.9847	0.9983
4	0.9817	0.9801	0.9833	0.9817	0.9984
5	**0.9863**	**0.9867**	0.9860	**0.9863**	**0.9987**
Mean	0.9825	0.9801	0.9851	0.9826	0.9983
Std. Deviation	0.0036	0.0058	0.0016	0.0035	0.0005

Table 5. XGBoost model 5-layer cross-validation results

Layer	Accuracy	Precision	Recall	F1-Score	AUC-PR
1	0.9817	0.9801	**0.9834**	0.9817	0.9984
2	0.9744	0.9734	0.9753	**0.9744**	0.9969
3	**0.9827**	**0.9853**	0.9800	0.9826	0.9982
4	0.9807	0.9800	0.9813	0.9807	0.9980
5	0.9813	0.9852	0.9773	0.9813	**0.9985**
Mean	0.9801	0.9808	0.9795	0.9801	0.9980
Std. Deviation	0.0033	0.0049	0.0032	0.0033	0.0006

4. Results

This section presents a comprehensive evaluation of machine learning models for credit card fraud detection under both imbalanced and balanced data scenarios, with a particular emphasis on the impact of data distribution on model reliability. We report both direct model comparisons and the outcome of rigorous cross-validation experiments.

4.1. Performance on the Original Imbalanced Dataset

The initial experiment involved training the LightGBM and XGBoost models on the original, highly imbalanced dataset. As presented in Table 2, the models achieved deceptively high accuracy and ROC AUC scores (e.g., XGBoost: 99.05% accuracy). While these metrics suggest strong performance at first glance, a deeper analysis of the precision and F1-scores reveals a critical weakness. LightGBM and XGBoost recorded extremely low precision scores of 13.58% and 28.19%, respectively, leading to poor F1-scores of 23.79% and 43.43%.

These results highlight a classic pitfall of training on imbalanced data. The models become heavily biased towards the majority class (non-fraudulent transactions), learning to classify almost everything as legitimate. Consequently, while they miss very few actual legitimate transactions, they fail to reliably identify the minority class (fraud). The low precision indicates that a vast majority of the transactions flagged as "fraud" are, in fact, false positives. For a system intended to create an immutable audit trail on a blockchain, this outcome is unacceptable, as it would permanently "pollute" the ledger with incorrect information, undermining its integrity and trustworthiness.

4.2. Performance on The Balanced Dataset

To address the issue of model bias, a second experiment was conducted using the balanced dataset created via random undersampling. The performance of the models on this dataset, shown in Table 3, demonstrates a dramatic and significant improvement across all relevant metrics.

On the balanced test set, LightGBM's F1-score skyrocketed from 23.79% to 97.26%, and its precision jumped from 13.58% to 96.73%. Similarly, XGBoost's F1-score rose to 97.37% with a precision of 97.39%. These results empirically prove that by balancing the class distribution at the data level, the models were able to learn the distinguishing features of both fraudulent and legitimate transactions effectively. The near-perfect balance between precision and recall, as reflected in the high F1-scores, indicates that the models are now highly reliable. They can not only catch most fraudulent activities (high recall) but also do so without incorrectly flagging a large number of legitimate transactions (high precision). Figure 3 visually encapsulates this transformation, illustrating the stark difference in model reliability between the two scenarios.

4.3. Model Robustness and Generalizability

To ensure that the exceptional performance on the balanced dataset was not an artifact of a specific data split, a rigorous 5-fold cross-validation was performed. The results, detailed in Table 4 for LightGBM and Table 5 for XGBoost, confirm the stability and robustness of our approach.

For both models, the average F1-score across the five folds was excellent (LightGBM: 98.26%, XGBoost: 98.01%). More importantly, the standard deviation for all key metrics was remarkably low (e.g., F1-score std. dev. of 0.0035 for LightGBM). This low variance indicates that the models consistently perform at a high level regardless of the training and validation data subsets. This confirms that the models have generalized well from the training data and are not overfitted, making them reliable candidates for real-world deployment.

In conclusion, our findings unequivocally demonstrate that for AI-driven fraud detection systems designed for blockchain integration, the data distribution strategy is a foundational requirement. While it is important to acknowledge that using balanced datasets can introduce its own challenges, as noted by Nobel et al. (2024), who caution that results from such sets may not always generalize to real-world scenarios due to potential trade-offs in recall, our robust cross-validation results mitigate this concern. The consistently high performance affirmed that relying on algorithmic tweaks for imbalanced data leads to unreliable models, whereas employing a data-level balancing technique like undersampling allows for the training of robust and accurate models like XGBoost and LightGBM. This produces trustworthy outputs worthy of on-chain recording, thus paving the way for more secure and transparent financial systems.

Acknowledgements

This study was supported by the Scientific and Technological Research Council of Türkiye (TUBITAK) under Grant Number 124E001. The authors thank TUBITAK for their support.

5. References

[1] Ngai EWT, Hu Y, Wong YH, Chen Y, Sun X. “The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature”. Decision Support Systems, 50(3), 559-569, 2011.
[2] Bahnsen AC, Aouada D, Stojanovic A, Ottersten B. “Feature engineering strategies for credit card fraud detection”. Expert Systems with Applications, 51, 134–142, 2016.
[3] Dal Pozzolo A, Boracchi G, Caelen O, Alippi C, Bontempi G. “Credit Card Fraud Detection: A Realistic Modeling and a Novel Learning Strategy”. IEEE Transactions on Neural Networks and Learning Systems, 29(8), 3784-3797, 2015.
[4] Carcillo F, Dal Pozzolo A, Le Borgne YA, Caelen O, Mazzer Y, Bontempi G. “Scarff: A scalable framework for streaming credit card fraud detection with spark”. Information Fusion, 41, 182-194, 2021.
[5] Sahin Y, Duman E. “Detecting credit card fraud by decision trees and support vector machines”. Proceedings of the International MultiConference of Engineers and Computer Scientists, Hong Kong, China, 17-19 March 2011.
[6] Casino F, Dasaklis TK, Patsakis C. “A systematic literature review of blockchain-based applications: Current status, classification and open issues”. Telematics and Informatics, 36, 55-81, 2019.
[7] Moin A, Karim A, Safdar Z, Safdar K, Imran M, Rauf HT. “A blockchain-based decentralized machine learning framework for privacy-preserving banking”. IEEE Access, 9, 9142-9153, 2021.
[8] Xu R, Liu Y, Zhang L, Wang Y. “Smart contract-based approach for efficient and secure financial transaction records”. Future Generation Computer Systems, 117, 462-473, 2021.
[9] Feng Q, He D, Zeadally S, Khan MK, Kumar N. “A survey on privacy protection in blockchain system”. Journal of Network and Computer Applications, 126, 45-58, 2019.
[10] Bolton RJ, Hand DJ. “Statistical fraud detection: A review”. Statistical Science, 17(3), 235-255, 2002.
[11] Phua C, Lee V, Smith K, Gayler R. “A comprehensive survey of data mining-based fraud detection research”. arXiv, preprint, arXiv:1009.6119, 2010.
[12] Koc L, Mazzuchi TA, Sarkani S. “A network-based approach for credit card fraud detection in e-commerce”. Decision Support Systems, 55(3), 662-673, 2012.
[13] Gama J, Žliobaitė I, Bifet A, Pechenizkiy M, Bouchachia A. “A survey on concept drift adaptation”. ACM Computing Surveys (CSUR), 46(4), 1-37, 2014.

[14] Ribeiro MT, Singh S, Guestrin C. ""Why should I trust you?": Explaining the predictions of any classifier". Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, California, USA, 13-17 August 2016.
[15] Whitrow C, Hand DJ, Juszczak P, Weston D, Adams N. "Transaction aggregation as a strategy for credit card fraud detection". Data Mining and Knowledge Discovery, 18, 30–55, 2009.
[16] Yli-Huumo J, Ko D, Choi S, Park S, Smolander K. "Where is current research on blockchain technology?—A systematic review". PLoS one, 11(10), e0163477, 2016.
[17] Zhou Q, Huang H, Zheng Z, Bian J. "Solutions to scalability of blockchain: A survey". IEEE Access, 8, 16440-16455, 2020.
[18] Saranya P, Shanthi P. "A Survey on Integration of Blockchain with Machine Learning". Archives of Computational Methods in Engineering, 29, 2167–2185, 2022.
[19] Salah K, Rehman MHU, Nizamuddin N, Al-Fuqaha A. "Blockchain for AI: Review and open research challenges". IEEE Access, 7, 10127-10149, 2019.
[20] Shrestha A, Shrestha R, Shrestha R, Gupta A, Sandhu R, Shrestha S. "Blockchain for machine learning: A comprehensive review and open research challenges". IEEE Access, 9, 101443-101470, 2021.
[21] Koteska B, Karafiloski E, Mishev A. "Blockchain implementation quality challenges: A literature review". ICT Innovations 2017, Skopje, Macedonia, 18-22 September 2017.
[22] Wang J, Li Y, Wang X, Han J. "A Blockchain-Based Framework for Anti-Money Laundering in Financial Institutions". IEEE Transactions on Engineering Management, 2022.
[23] Hassani H, Huang X, Silva E. "Digitalisation and Big Data Mining in Banking". Big Data and Cognitive Computing, 2(3), 18, 2018.
[24] Narayanan A, Bonneau J, Felten E, Miller A, Goldfeder S. Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction. Princeton, NJ, USA, Princeton University Press, 2016.
[25] Crosby M, Pattanayak P, Verma S, Kalyanaraman V. "Blockchain technology: Beyond bitcoin". Applied Innovation Review, 2, 6-10, 2016.
[26] [Xu X, Weber I, Staples M. Architecture for Blockchain Applications. Cham, Switzerland, Springer, 2019.
[27] Zheng Z, Xie S, Dai H, Chen X, Wang H. "An overview of blockchain technology: Architecture, consensus, and future trends". 2018 IEEE International Congress on Big Data, San Francisco, CA, USA, 2-7 July 2018.
[28] Popov S. "The tangle". IOTA Foundation, White Paper, 2018.
[29] Christidis K, Devetsikiotis M. "Blockchains and Smart Contracts for the Internet of Things". IEEE Access, 4, 2292-2303, 2016.
[30] Szabo N. "Formalizing and Securing Relationships on Public Networks". First Monday, 2(9), 1997.
[31] Dannen C. Introducing Ethereum and Solidity: Foundations of Cryptocurrency and Blockchain Programming for Beginners. Berkeley, CA, USA, Apress, 2017.
[32] Buterin V. "Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform". White Paper, 2014.
[33] Luu L, Chu DH, Olickel H, Saxena P, Hobor A. "Making Smart Contracts Smarter". Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, 24-28 October 2016.
[34] Schär F. "Decentralized Finance: On Blockchain- and Smart Contract-Based Financial Markets". Federal Reserve Bank of St. Louis Review, 103(2), 153–174, 2021.
[35] Atzei N, Bartoletti M, Cimoli T. "A survey of attacks on Ethereum smart contracts (SoK)". International Conference on Principles of Security and Trust, Uppsala, Sweden, 22-29 April 2017.
[36] Siegel D. "Understanding The DAO Attack". Coindesk, (18.06.2016).
[37] Brent L, Jurisevic A, Kong A, Liu E, Scholz B. "Vandal: A scalable security analysis framework for smart contracts". Proceedings of the 40th International Conference on Software Engineering, Gothenburg, Sweden, 27 May-3 June 2018.

[38] Hassan S, Kyriakou H, Leiser Silva L. "Decentralized applications: The blockchain-empowered software system". Business & Information Systems Engineering, 62(3), 229-242, 2020.
[39] Werner SM, Perez D, Gudgeon L, Klages-Mundt A, Harz D, Knottenbelt WJ. "Sok: Decentralized finance (defi)". IEEE European Symposium on Security and Privacy Workshops, Vienna, Austria, 6-10 September 2021.
[40] Allen DW. "Smart-contract contracts". Review of Law & Economics, 15(1), 1-43, 2019.

[41] Wang Q, Zhang Q, Chen L, Zheng Z. “Blockchain smart contract: applications, challenges, and future trends”. IEEE Transactions on Services Computing, 15(2), 984-1002, 2022.
[42] Chen T, Li X, Luo X, Zhang X. “Understanding ethereum via graph analysis”. ACM Computing Surveys (CSUR), 54(9), 1-39, 2021.
[43] Zeng X, Fang H, Huang Y. “Smart Contract: Platforms, Applications, and Research Directions”. IEEE Access, 10, 4897-4916, 2022.
[44] Jordan MI, Mitchell TM. “Machine learning: Trends, perspectives, and prospects”. Science, 349(6245), 255–260, 2015.
[45] Rajkomar A, Dean J, Kohane I. “Machine learning in medicine”. New England Journal of Medicine, 380(14), 1347–1358, 2019.
[46] Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. “Dermatologist-level classification of skin cancer with deep neural networks”. Nature, 542(7639), 115–118, 2017.
[47] Wolsztynski E, Kamiński B, Szapiro T. “A survey of transfer learning”. IEEE Transactions on Knowledge and Data Engineering, 35(2), 2023.
[48] Sutton RS, Barto AG. Reinforcement Learning: An Introduction. 2nd ed. Cambridge, MA, USA, MIT Press, 2018.
[49] Silver D, Huang A, Maddison CJ, Guez A, Sifre L, van den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M, Dieleman S, Grewe D, Nham J, Kalchbrenner N, Sutskever I, Lillicrap T, Leach M, Kavukcuoglu K, Graepel T, Hassabis D. “Mastering the game of Go with deep neural networks and tree search”. Nature, 529(7587), 484–489, 2016.
[50] Saito T, Rehmsmeier M. “The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets”. PloS One, 10(3), e0118432, 2015.
[51] Refaeilzadeh P, Tang L, Liu H. Cross-Validation. Editors: Liu L, Özsu MT. Encyclopedia of Database Systems, 532-538, Boston, MA, USA, Springer, 2009.

International Conference on Advances and Innovations in Engineering

3.ICAIE

Autonomous Unmanned Aerial Vehicle Application for Real-Time Wounded Detection Based on YOLOv5

Dilan Nur DEMİR[1] , Buse EROL[2], Mustafa KARA[3]

[1,2,3]Department of Computer Engineering, Turkish Air Force Academy, National Defence University, Istanbul, Turkey.
[1] dilannur.demir@msu.edu.tr, [2]buse.erol@msu.edu.tr, [3]mustafa.kara@msu.edu.tr
[3](ORCID: 0000-0002-0205-1149)

Abstract
The inability to provide timely medical aid for injuries incurred on battlefields significantly increases the casualty rate. This study proposes a deep learning-based search and rescue system, integrated with unmanned aerial vehicles (UAVs), for the swift identification of injured troops, especially in military situations. We present a model constructed with the YOLOv5 architecture, a CNN-based object detection method, for real-time analysis of visual data acquired by UAVs. The trained model can differentiate between the "wounded" and "healthy" classes using labeled photos that depict persons in various bodily positions. Throughout the training process, transfer learning and data augmentation approaches were employed, and the model's performance was assessed using accuracy, precision, recall, and F1-score measures. The suggested method seeks to provide a real-time, dependable, and efficient solution applicable in military contexts and civilian search and rescue operations, including natural disasters. When examining the results, the system achieved a high overall accuracy with a mean Average Precision (mAP) of 93.4% for real-time human detection and wounded/healthy classification. The model trained using the YOLOv5 architecture achieved successful results, with 89.9% precision and 87% recall for the “wounded” class and 81.6% precision and 92.5% recall for the “healthy” class. Moreover, the confusion matrix and confidence curves indicate that the model provides strong class separability and demonstrates a high generalization capacity.

Keywords: YOLOv5, object detection, real-time image processing, deep learning, unmanned aerial vehicle

1. Introduction

The rapid and accurate detection of individuals injured in war and disaster environments is a critical factor that directly affects survival rates. In particular, during military operations conducted on the front lines, timely intervention for wounded personnel is of great importance, not only for operational success but also for the preservation of human life. However, in such scenarios, conventional search and rescue (SAR) methods may become insufficient if existing infrastructures (such as GSM, GPS, Wi-Fi, etc.) are rendered inoperative [1]. The heavy reliance of traditional SAR activities on human labor brings about significant limitations in terms of time loss, high costs, and personnel safety. In this context, supporting missions with autonomous systems has become not only a technological advancement but also a strategic and humanitarian necessity. Artificial intelligence–powered systems offer substantial advantages in search and rescue processes by reducing the dependency on human intervention, thereby improving operational efficiency, response time, and safety [2].

This study aims to enhance the effectiveness of artificial intelligence–based autonomous systems in search and rescue operations conducted under war and disaster conditions. In this context, the YOLOv5 (You Only Look Once version 5) model, which is capable of performing object detection tasks in real time, has been selected. Compared to widely used deep learning–based models such as Faster R-CNN, SSD, and RetinaNet, YOLOv5 offers a structure particularly suitable for operation on embedded systems and UAVs in terms of speed, accuracy, and hardware efficiency [3, 4]. Thanks to its high sensitivity to objects of varying sizes and its optimized PyTorch framework, it achieves consistent detection performance across both wide and narrow fields of view [5]. The INRIA Person dataset used for model training includes various human postures and diverse environmental conditions, providing a rich foundation for detecting wounded individuals [6]. System tests were conducted on the AirSim simulation platform, developed by Microsoft, which realistically replicates physical conditions. With sensor emulations that provide both RGB and depth data, AirSim enables realistic evaluation of autonomous systems’ detection and navigation capabilities [7]. Thanks to this integrated structure, the proposed system aims to enable the rapid and effective detection of wounded individuals, particularly in military SAR operations, by providing a real-time, resource-efficient, and reliable solution.

[1]Corresponding author

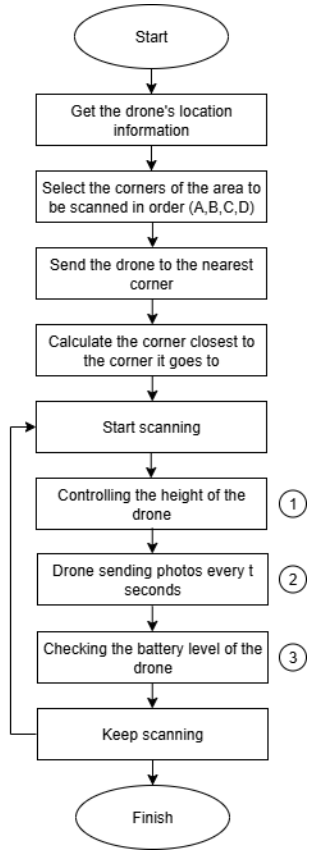
2. Literature Review

In recent years, unmanned aerial vehicles (UAVs) have shown to be efficient instruments in search and rescue (SAR) missions. Numerous studies have proved the benefits of drone technology in operations over extensive areas and during natural disasters. Tariq and Rameesha (2020) asserted that drones could surpass search and rescue robots in natural disaster situations [8]. Kıran and Kongara Surya (2019) also emphasized the operational advantages of drones in extreme environmental situations [9]. The study by Mishra and Balmukund demonstrated that a YOLOv5-based model developed with a dataset of 75,000 images achieved a 98% mAP performance [10]. It was also noted that in cases where GPS is unavailable, localization can be performed using QR codes. The SARDO system offers an integrated solution for locating victims via GSM signals after a disaster. However, it has been emphasized that the system may be limited in military scenarios due to its dependency on GSM infrastructure and mobile devices [11].

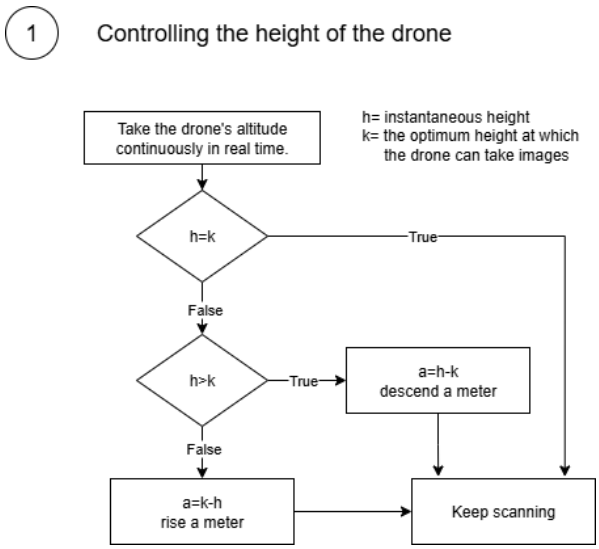
The literature also shows that thermal imaging–based tracking systems can detect humans under challenging environmental conditions. For example, in one study, a system developed using YOLOv5 was able to perform multi-target tracking under low light and dense vegetation. Continuity in tracking was ensured using a Kalman filter and the “nearest neighbor” matching algorithm, and successful results were obtained in mountainous regions [12]. These studies demonstrate that UAVs contribute significantly to SAR operations in terms of time efficiency, area coverage, and human detection capabilities. However, most existing systems operate with infrastructure dependency and do not adequately account for threats specific to military operations. Therefore, there is a growing need for infrastructure-independent, autonomous, vision-based, and low-energy-consumption systems. The system developed in this study offers an AI-based solution built on YOLOv5 that enables real-time detection and classification of wounded individuals and proposes an original approach that allows rapid intervention suitable for military operation conditions.

3. Method

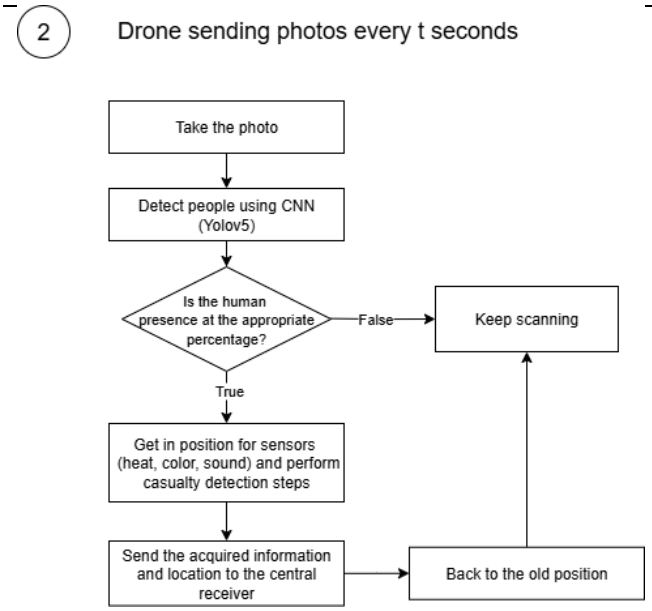
In this study, a deep learning–based system has been developed in which a drone can autonomously scan a designated area and detect wounded individuals using real-time image processing techniques. The system components are structured under four main categories: software and hardware infrastructure, image processing architecture, and flight algorithms. In the developed system, a quadcopter drone was used as the main component. The drone communicates with the ground control station via a 5.4 GHz Wi-Fi module, receives positional data through a GPS module, and collects visual data using an onboard camera. The visual data is transmitted to the ground control unit, where real-time object detection is performed using the YOLOv5 architecture. The drone's flight path is planned to cover a rectangular area defined by four software-defined corners (A, B, C, D). Before scanning, the drone is directed to the nearest corner, and then it proceeds to scan the remaining corners in sequence. This cyclic task planning is illustrated in Figure 1(a). An altitude control technique is utilized to guarantee the positional stability of the UAV. The drone's altitude is consistently checked to ensure uniform image quality and resolution. The present altitude is juxtaposed with the optimal imaging altitude established by the system. Upon detecting a divergence, the drone autonomously ascends or descends as required. The current altitude is compared with the ideal imaging altitude predefined by the system. When a deviation is detected, the drone automatically ascends or descends as needed. This control process enhances image accuracy and maintains post-processing classification performance. The altitude control algorithm that ensures the drone remains at an appropriate height is presented in Figure 1(b). Additionally, the images captured by the drone's camera at one-second intervals are analyzed using the CNN-based YOLOv5 model.



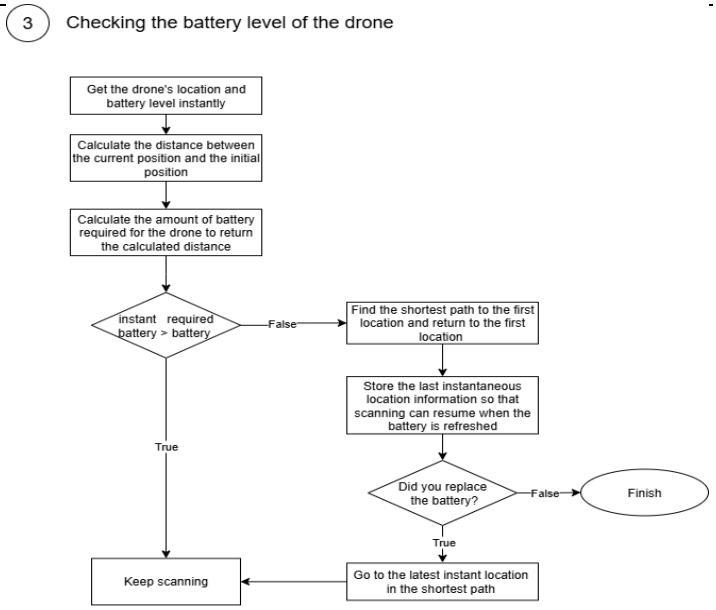
(a) Autonomous drone mission planning and area scanning initialization



(b) Altitude control algorithm for maintaining optimal imaging height of the drone



(c) Real-time human and wounded detection workflow with YOLOv5



(d) Battery level optimization algorithm for energy management and safe return

Figure 1. Task flow diagrams for drone navigation, sensing, and power control

The model detects the presence of a “human” with high accuracy and marks objects that exceed a predefined confidence score threshold. Following human detection, the individual is evaluated for injury using onboard thermal, color, and sound sensors. The obtained location and classification information are simultaneously transmitted to the control unit, and if necessary, the drone repositions itself toward the target for a more detailed inspection.

An open-source human image dataset was utilized during the model training phase, with classification conducted for the "wounded" and "healthy" labels. The model's accuracy was assessed utilizing precision, recall, F1-score, and mean Average Precision (mAP) criteria. Figure 1(c) depicts the method for person detection based on YOLOv5 and sensor-assisted injury evaluation.

In order to prevent the drone's battery from running out of power while it is in operation, the system does regular checks on the drone's battery level and its current location. If the remaining energy is insufficient for the drone to return to its starting point, the mission is temporarily paused, and the drone returns to the base station via the shortest path. After the battery is replaced, the drone is redirected to its last known position and resumes the mission from where it left off. The energy optimization and safe return algorithm is presented in Figure 1(d).

When evaluated as a whole, these four algorithmic components form a complementary structure that ensures the effective, reliable, and continuous operation of drone-based search and rescue missions. The integration of mission planning, altitude control, real-time human and wounded detection, and energy management processes enables the system to perform with high success rates in the field.

4. Experimental Results and Evaluations

This work built a YOLOv5-based real-time object identification algorithm that autonomously differentiates between "wounded" and "healthy" humans. The open-source "Person Fallen Person" dataset was employed for model training. This dataset comprises tagged photos illustrating diverse body postures of persons, especially in critical situations like falls and injuries. The images in Figure 3 demonstrate that the model effectively classified humans in the test photographs. The model's efficacy on real-world photos was assessed visually. The model reliably identified and classified items using bounding boxes in test photos taken under various positions and situations. Below are examples of bounding boxes and class labels generated for persons in various positions. The results indicate that the model can be utilized successfully and reliably in practical applications.

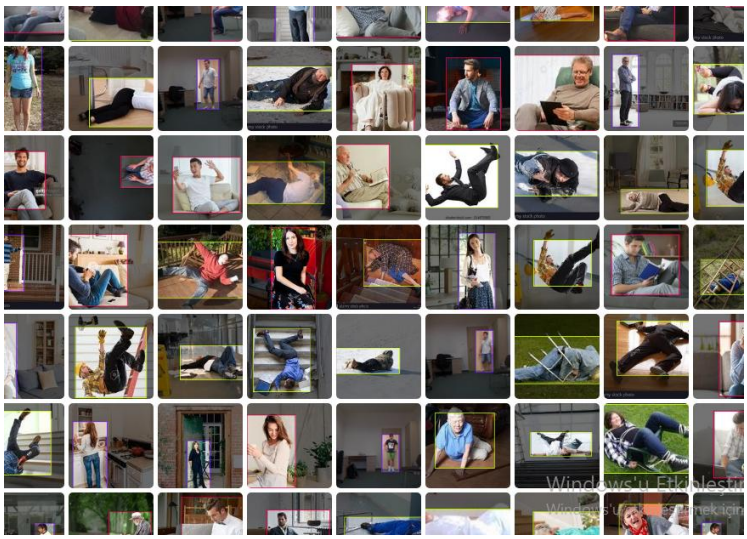


Figure 2. YOLOv5 detection outputs for various falling and sitting positions

After being downloaded in YOLOv5 PyTorch format, the dataset was structured by dividing it into three subsets: "train", "valid", and "test". During model training, the small version of the YOLOv5 architecture (yolov5s.pt) was

used through transfer learning. The training process was carried out with an input size of 640 pixels, a batch size of 16, and for 200 epochs. Data augmentation techniques such as random flip, scale, and translate were applied during training.

The hyperparameters used were as follows: initial learning rate (lr0=0.01), final learning rate factor (lrf=0.01), momentum (0.937), weight decay (weight_decay=0.0005), warmup epochs (3.0), warmup momentum (0.8), warmup bias learning rate (0.1), box loss coefficient (0.05), class loss coefficient (0.5), object loss coefficient (1.0), objectness gain (1.0), IoU threshold (iou_t=0.2), anchor threshold (anchor_t=4.0), color transformation parameters (hsv_h=0.015, hsv_s=0.7, hsv_v=0.4), translate (0.2), scale (0.8), flipud and fliplr (both set to 0.5), and mosaic augmentation (1.0). The optimizer used during training was stochastic gradient descent (SGD), with a fixed learning rate of 0.01.

4.1. Model Training and Configuration

During the training process, the loss functions (box_loss, obj_loss, cls_loss) were observed to decrease consistently and steadily. Precision, recall, and mAP50 values increased as training progressed. These results demonstrate the model's generalization capability and learning stability. Figure 3 presents the loss and metric curves obtained during the model training, respectively, together in one view.

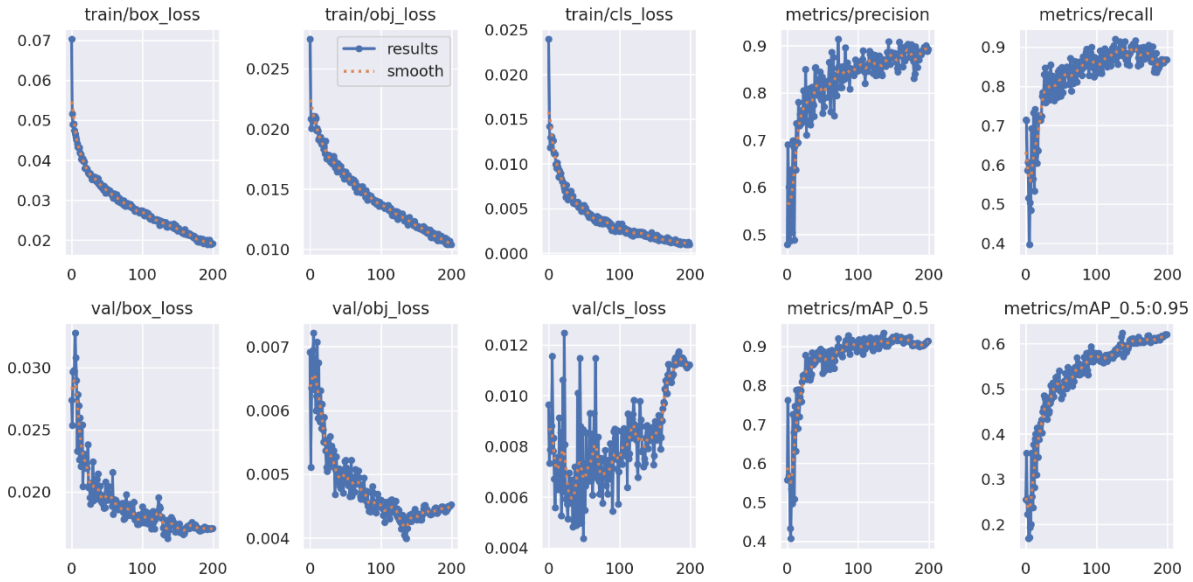


Figure 3. Loss and metric curves obtained during model training

4.2. Evaluation Metrics

The YOLOv5-based deep learning model exhibited superior performance in categorizing persons as "wounded" or "healthy." Comprehensive per-class evaluation metrics are displayed in Table 1. The assessment criteria are precision, recall, and mAP50 (mean average precision at an intersection over a union threshold of 0.5). The model attained 89.9% precision, 87% recall, and 93.3% mAP50 when assessed across all classes. The results demonstrate that the majority of objects identified by the model are classified accurately (high precision) and that the model effectively recognized most existing things (high recall). Moreover, the elevated mAP50 score of 93.3% illustrates the model's robust proficiency in categorization and object localization.

A class-based analysis indicates that for the "wounded" category, the model attained 81.6% precision, 92.5% recall, and 93.6% mAP50. This signifies that the model can consistently identify injured individuals, while it may

sporadically misclassify healthy individuals as injured. In real-time search and rescue operations, this result constitutes a permissible compromise favoring the "miss-as-few-as-possible" strategy.

Table 1. YOLOv5 model performance results

Class	Images	Instances	Precision	Recall	mAP50
All	203	216	0.899	0.87	0.933
Wounded	203	123	0.816	0.925	0.936
Healty	203	93	0.858	0.897	0.934

On the other hand, for the “healthy” class, the model demonstrated balanced and consistent performance with 85.8% precision and 89.7% recall. The mAP50 value of 93.4% further confirms the model’s high accuracy in this class as well. The observed performance gap between the classes may be attributed to an imbalance in the number of samples or to the fact that wounded individuals exhibit more distinctive postures. The high metrics achieved both overall and per class validate that applying the YOLOv5 architecture in combination with a transfer learning strategy offers a suitable and effective solution for critical image processing applications, particularly in the fields of safety and healthcare.

To further analyze the model’s performance, a confusion matrix was generated. According to the matrix presented in Figure 4, the correct prediction rate for the “wounded” class was 88%, while it was 90% for the “healthy” class. These results confirm the model’s strong discriminative capability between the classes. The misclassification rates remained within the 7–8% range.

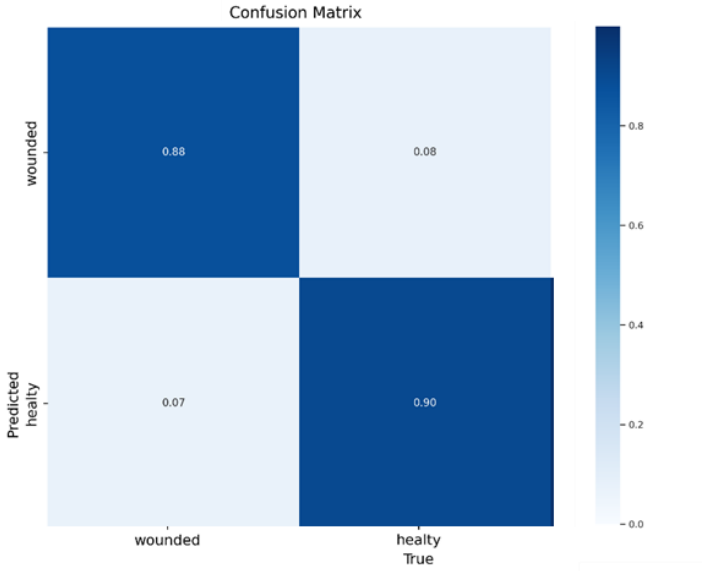
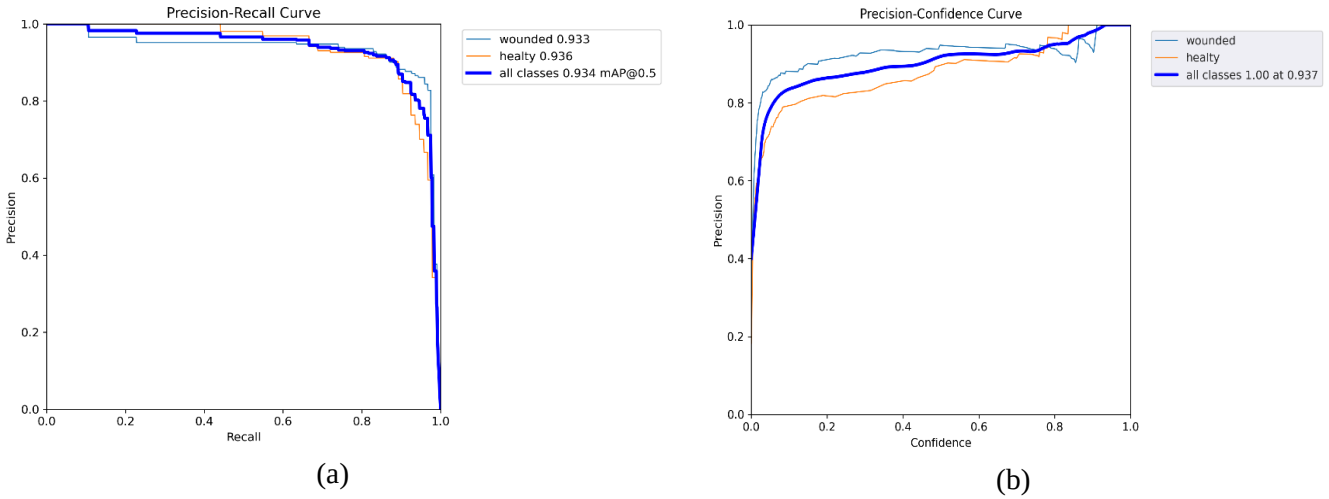


Figure 4. Loss and metric curves obtained during model training

To gain a deeper understanding of the model’s reliability, Precision–Recall, Precision–Confidence, and Recall–Confidence curves were analyzed. The Precision–Recall curves shown in Figure 5(a) demonstrate high stability and performance for both classes. As depicted in Figure 5(b), the Precision–Confidence curve indicates that the model's prediction accuracy increases as the confidence level rises.



(a) (b)

Figure 5. (a) Precision-Recall curve, (b) Precision-Confidence curve

The Recall–Confidence curves presented in Figure 6 demonstrate high stability particularly for the "healthy" class, while revealing a decrease in recall as confidence decreases for the "wounded" class. Moreover, notably high stability was observed in the "healthy" class. While the Recall–Confidence curve maintains strong stability for the "healthy" class, it exhibits a more rapid decline for the "wounded" class.

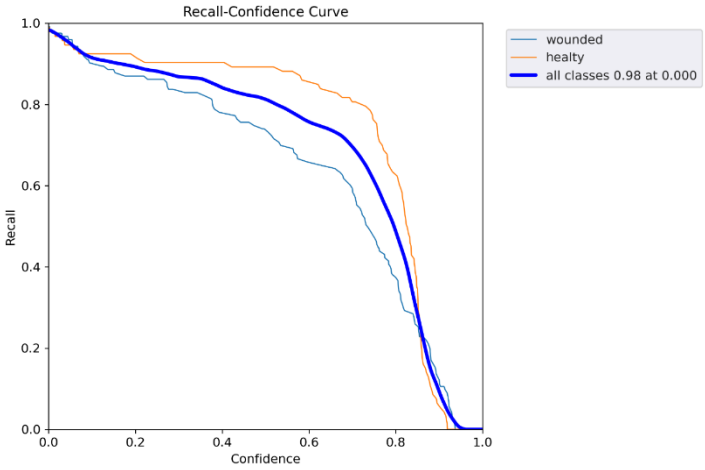


Figure 6. Recall-Confidence curve

The obtained results demonstrate that the system developed using the YOLOv5 architecture provides high accuracy in detecting human presence and condition. In particular, the high recall rates indicate that this system could offer a significant advantage in preventing losses during search and rescue operations. The preservation of performance balance across classes and the ability to operate with low hardware requirements make this model suitable for field deployment. The YOLOv5-based model has shown high performance in both theoretical evaluations and visual results. This confirms the model's applicability in real-time critical domains such as military search and rescue, health monitoring, and security applications.

4. Conclusion

The results unequivocally indicate that the YOLOv5 architecture delivers superior accuracy in human detection and classification tasks. The implementation of a transfer learning approach, utilizing pre-trained weights, was identified

as a crucial element in attaining superior performance while markedly decreasing training duration. The meticulous adjustment of hyperparameters during the training process (e.g., initial learning rate lr0 = 0.01, momentum = 0.937, weight decay = 0.0005, warmup configurations, and data augmentation methods) facilitated the model's steady and consistent learning. Adhering to the specified settings in YOLOv5 enhanced the model's capacity for effective generalization, even with tiny datasets containing limited samples, while preventing overfitting.

Notwithstanding the visual resemblance between the "wounded" and "healthy" categories in the dataset, the model demonstrated robust classification efficacy. This demonstrates that the employed data augmentation strategies (namely flipud, fliplr, scale, and translate) significantly improved the model's generalization ability. This result indicates that the model can sustain its efficacy throughout diverse environmental circumstances in practical applications, such as search and rescue operations or surveillance film analysis.

Nonetheless, it is important to acknowledge that the model may have difficulties in intricate situations, such as limited field-of-view or low-resolution photos featuring several subjects. Future research may investigate the incorporation of multimodal data sources (e.g., thermal imaging, auditory sensors) or transformer-based detection methodologies to enhance performance under such settings. Furthermore, it is imperative to optimize inference time and assess lightweight models tailored for edge devices, such as YOLOv5n or YOLOv8-lite, as subsequent steps. Given that the system functions under constraints like battery usage and restricted processing resources, efficiency is as vital as accuracy. This study reaffirms the efficacy of deep learning in human identification and condition classification, facilitating rapid and comprehensive analysis, particularly in time-sensitive applications like search and rescue.

Acknowledgements

This study is part of the project titled 'Rapid and Effective Rescue in Military Operations: A Drone-Assisted SAR Solution,' which was accepted under the TÜBİTAK–2209-A Research Project Support Program for University Students.

5. References

[1] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed., Pearson, 2018.
[2] J. Redmon, A. Farhadi, “YOLOv3: An Incremental Improvement,” arXiv preprint arXiv:1804.02767, 2018.
[3] G. Jocher et al., “YOLOv5 by Ultralytics,” GitHub repository, https://github.com/ultralytics/yolov5, 2020.
[4] H. Wang et al., “Real-time Object Detection on Edge Devices using YOLOv5,” *IEEE Access*, vol. 9, pp. 123456–123467,2021.
[5] C. Dalal and B. Triggs, “Histograms of Oriented Gradients for Human Detection,” in *Proc. CVPR*, 2005.
[6] S. Shah et al., “AirSim: High-Fidelity Visual and Physical Simulation for Autonomous Vehicles,” *Field and Service Robotics*, pp. 621–635, Springer, 2017.
[7] D. R. Banisetty, N. Ranaweera, “Benchmarking Deep Reinforcement Learning Algorithms on AirSim,” *Sensors*, vol. 21, no. 8, 2021.
[8] Tariq, Rameesha, et al. "Dronaid: A smart human detection drone for rescue." 2018 15th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT (HONET-ICT). IEEE, 2018.
[9] KIRAN, KONGARA SURYA. "INNOVATION IN DRONE TECHNOLOGY TO SAVE THE LIFE."
[10] Mishra, Balmukund, et al. "Drone-surveillance for search and rescue in natural disaster." Computer Communications 156 (2020): 1-10.
[11] Albanese, Antonio, Vincenzo Sciancalepore, and Xavier Costa-Pérez. "SARDO: An automated search-and-rescue drone-based solution for victims localization." IEEE Transactions on Mobile Computing 21.9 (2021): 3312-3325
[12] Yeom, S. Thermal Image Tracking for Search and Rescue Missions with a Drone. *Drones* **2024**, *8*, 53. https://doi.org/10.3390/drones8020053

Design and Energy Efficiency of Nearly Zero Energy Buildings

Ersan ŞAHİN[1] , Sinan KAPAN [2], Nevin ÇELİK[3]

[1,2,3]Mechanical Engineering Department, Engineering Faculty, Firat University, Elazig, Türkiye.
[1]ersansahin@gmail.com, [2]skapan@firat.edu.tr, [3]nevincelik23gmail.com

Abstract
Nearly Zero Energy Buildings (nZEB) refer to buildings that have significantly reduced their energy consumption but do not meet their entire energy needs from renewable sources. These buildings meet a significant portion of their energy needs with efficient systems and renewable energy sources, but a small portion of their energy consumption can be supplied by external sources. This study aims to compare the energy efficiency, performance, and cost of converting a detached residential building with a total construction area of approximately 100 square meters, located in Elazığ and İzmir, into a Nearly Zero Energy Building.

Keywords: nearly zero energy buildings, energy efficiency, renewable energy

1. Introduction

Energy is the capacity of a system to do work. In physics, energy exists in different forms and can be transformed from one state to another, but it cannot be created or destroyed entirely. Energy occurs in various forms, including kinetic (energy of motion), potential (energy stored by position or structure), thermal (heat energy), chemical, nuclear, electrical, and radiant energy. Events throughout the universe rely on energy transformations, such as the conversion of sunlight into chemical energy (photosynthesis) and electrical energy into light or motion.
Energy is a broad concept with technical, economic, ecological, and social dimensions. It plays a critical role in many areas of human life, including industry, transportation, agriculture, communication, and daily life.
The energy used in buildings encompasses all the energy types required for the construction, operation, and maintenance of a building. This energy is typically provided by sources such as electricity, natural gas, renewable energy sources (solar, wind, geothermal), or fossil fuels. It is primarily used in the following areas:

• **Heating, Cooling, and Ventilation (HVAC):**
Energy is used in heating (central heating, underfloor heating systems), cooling (air conditioning, and cooling systems), and ventilation systems to ensure comfort in the interior spaces of buildings.
• **Lighting:**
The energy used to illuminate the interior and exterior spaces of buildings is generally electrical. LED or sensor-based lighting systems have become widespread for energy savings.
• **Hot Water Production:**
The energy used in buildings also plays a significant role in heating water and making it ready for use. This energy is provided by solar panels, electric heaters, or natural gas water heaters.
• **Electrical and Electronic Devices:**
Devices used in buildings (white goods, computers, televisions, security systems, etc.) require electrical energy.
• **Construction and Maintenance Processes:**
Energy is consumed during the production of machinery, equipment, and materials used during the construction of buildings. Maintenance processes also require energy. • Smart Systems and Automation:
Automation systems (smart thermostats, building management systems) used in buildings to increase energy efficiency optimize energy consumption.
Energy consumption in buildings is also critical to sustainability due to its environmental impact. Approaches such as energy-saving insulation, renewable energy systems, zero-energy building design, and near-zero-energy building design are being used to reduce energy use—or, more accurately, to increase energy efficiency.

[1]Corresponding author

This literature review examines the parameters used in the design of nearly zero-energy buildings, the effects of different climatic zones on the building's heating, cooling, and electrical loads, and sources related to energy efficiency studies. Yılmaz [1] used the Carrier HAP tool to calculate the annual energy consumption of a sample office building for heating, cooling, ventilation, and so on, and then analyzed the hourly energy simulation of the building. Yılmaz [2] calculated the energy consumption and savings of 10 sample apartment buildings selected from Konya and Erzincan provinces in both uninsulated and insulated conditions in his thesis. Thermal losses of the building's outer shell, photographed with a thermal camera, were determined, and the total heat loss and annual heating energy of the buildings using different insulation materials were compared. Kapsalaki [3], in his doctoral thesis, designed a detached house and a building for three different climatic zones (Stockholm, Lisbon, and Iraklion) using the NZEB approach using an algorithm developed in the MATLAB package. Several main design variables were defined for the building's energy demand and supply. These were: Factors such as thermal insulation level, window and shading type, glazed area, ventilation, lighting, and heating and cooling system type were taken into account. The results of the sample building analyses showed that the differences between an economically efficient and an economically inefficient NZEB can be more than threefold, both in terms of initial cost and lifecycle costs.
In his thesis, Sağlam [4] used the EnergyPlus analysis program to select Izmir as a hot climate region and Berlin as a cold climate region to achieve a nearly zero-energy building for residential purposes. Alternative scenarios were presented for increasing the insulation thickness for the building envelope, floor, and roof, and for window sizes and orientations. Two different system options were presented, and annual energy consumption was calculated.
Roudi [5]'s thesis focused on energy efficiency to achieve the most cost-effective technique for minimizing carbon dioxide emissions and achieving a nearly zero-energy building in different climate zones. Alternative energy sources considered as components of a logical nearly zero-energy building include solar PV, groundwater heat pumps, passive solar design strategies, high-performance building envelope, lighting, daylighting, and low energy consumption. Based on this, a PV system was used to generate electricity in all climates, providing good thermal comfort for interior spaces. Ferdos [6], in his study on energy efficiency for a primary school building, selected the cities of Istanbul and Erzurum, two different climate zones, and examined the energy performance of the reference building using Desgin Builder and EnergyPuls. The payback period and optimal costs for improvements implemented to the exterior shell, lighting system, and HVAC systems were calculated based on prices from the Ministry of Environment and Urbanization.

2. Nearly Zero Energy Building (nZEB) Design

Nearly zero energy building (nZEB) design aims to achieve the highest possible level of energy efficiency by minimizing energy consumption and meeting the remaining energy needs with renewable energy sources. This design increases the comfort of building occupants while reducing environmental impacts and energy costs.

2.1. Passive Design Strategies

- Passive design allows the building to utilize natural resources to minimize energy consumption.
- Insulation: Heat loss is minimized by using high-performance insulation materials in the walls, roof, and windows. Materials with a low U-value should be preferred.
- Natural Lighting: Strategic placement of windows and building orientation should be positioned on a north-south axis to maximize daylight. Heat loss should be reduced by using large windows on the south facade and smaller windows on the north facade.
- Natural Ventilation: Optimizing airflow within the building reduces reliance on energy-consuming ventilation systems. Cross ventilation should be provided by double-sided windows, and air movement within the building can be optimized through shaft systems and ventilation ducts.
- Solar Control: Elements such as shades, façade panels, or plantings are used to control sunlight. Shades, shutters, and movable sunshades prevent overheating in summer. Additionally, double- or triple-glazed windows that control heat transfer should be used.

2.2. Active Systems and Technologies

Active systems are modern technologies that increase efficiency in energy production and consumption.

- Heating and Cooling Systems: High-efficiency ground- or air-source heat pumps or central systems should be used. Systems that recover energy lost during ventilation should also be preferred.
- Lighting: Energy savings should be achieved by using LED and sensor-supported lighting systems. Daylight sensors automatically adjust lighting according to the amount of natural light, which positively impacts efficiency.
- Smart Energy Management: Building automation systems that optimize energy use are installed. Energy consumption is monitored, controlled, and optimized through smart thermostats, sensors, and control systems.

3. Renewable Energy Integration

- Solar Panels (Photovoltaic Systems): Should be integrated into building roofs or facades for electricity generation.
- Solar Collectors: Used for hot water production.
- Wind Turbines: Electricity generation is supported by small-scale wind turbines.
- Hybrid Systems: Hybrid systems combine multiple renewable energy sources to make energy production more stable, providing both stable and continuous energy production.
- Geothermal Energy: Heating and cooling are achieved by utilizing underground temperatures.
- Thermal Storage Systems: Molten salt storage stores heat generated from solar energy for extended periods, while ice storage cools water or ice at night and uses it in air conditioning systems during the day.
- Battery Storage Systems: Batteries store excess energy generated by solar and wind energy, allowing it to be used at night or on cloudy days.
- Hydrogen Fuel Cells: An innovative alternative for electricity generation and storage.

3.1. Energy Efficiency Strategies

- Low-Energy Appliances: Products with an A+++ energy rating should be preferred for white goods and other electrical appliances.
- Energy Storage: Batteries or thermal storage systems are used to store the renewable energy produced.
- Temperature Control: Smart thermostats help save energy by controlling the temperature of each room individually.
- Water Management: Low-flow faucets and water-saving appliances reduce the energy consumed for water heating.

3.2. Building Materials and Construction Techniques

- Local and Natural Materials: Recyclable materials with minimal environmental impact are preferred.
- Lightweight Building Materials: Lightweight concrete, wood, or steel materials are used to increase energy efficiency.
- Solar Glass: Glass that can generate electricity (e.g., photovoltaic glass).
- Low-E Glass: Coated glass that reduces heat loss.
- Thermal Walls: Wall systems that can store heat.
- Phase Change Materials (PCM): Innovative materials that store or release heat depending on the ambient temperature.
- Recycled Materials: Waste materials are repurposed in building construction.

3.3. Simulation and Performance Analysis

Simulation tools are used to pre-assess the energy performance of buildings:

- EnergyPlus: Used for energy modeling and building simulation.
- DesignBuilder: Used for energy performance, indoor comfort, and renewable energy integration analysis.
- TGUB: Heat requirement identification document, heating and ventilation requirement identification

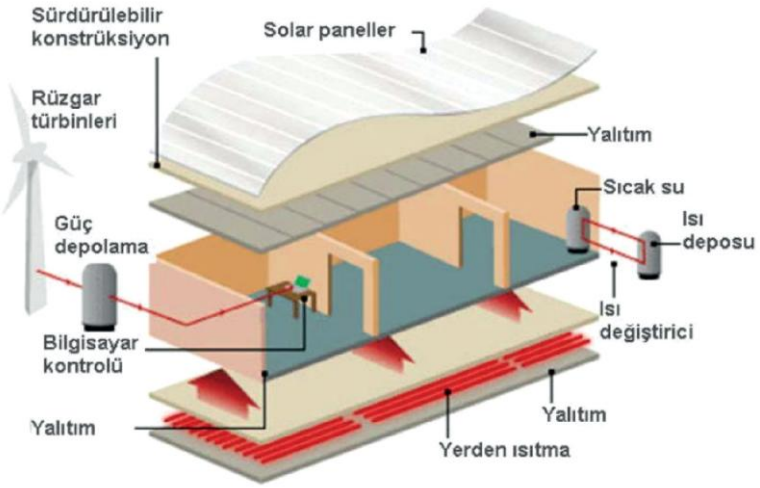


Figure 1. Zero energy building example (Berberoğlu, 2009)

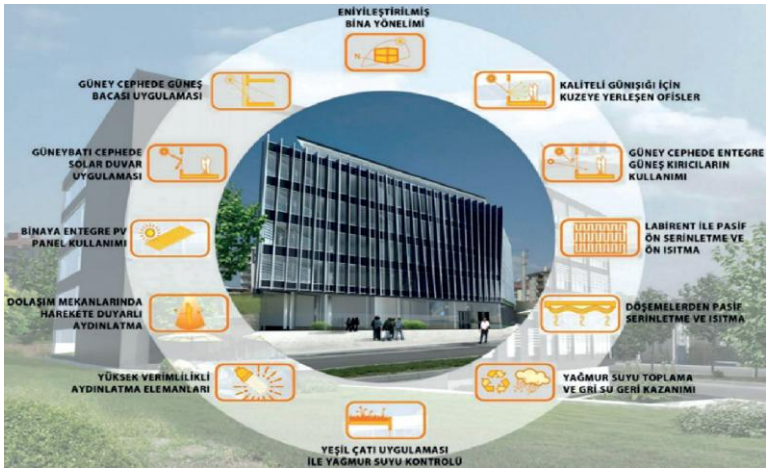


Figure 2. Ankara, Sincan and Etimesgut TKGM building energy efficiency applications ministry of environment and urbanization (Yöntem, 2016)

4. Materials and Methods

This study will analyze the energy performance of a detached house with a net area of 99.8 m2 in different climate zones to transform it into a Nearly Zero Energy Building (nZEB). The results will be compared.

Only heating, cooling, and electrical loads will be considered within the scope of energy performance. Two different climate zones were selected for heating and cooling loads: Elazığ (3rd climate zone) and İzmir (1st climate zone), representing the cold and hot climate zones, respectively. The aim is to present different solution alternatives to reduce the building's energy demand through improvements to the building envelope components, which will significantly reduce heating and cooling loads.

4.1. Nearly Zero Energy Building Design and Parameters to be Examined

This study uses materials and dimensions according to TS825 standards in both cities to transform a detached building intended for residential use into a Nearly Zero Energy Building. The effects of various parameters on energy consumption, such as insulation thickness and window types, will be examined, and their suitability for economic indicators (cost-performance optimization) will be evaluated.
The building model will be drawn in 2D using AutoCAD. Afterwards, heating and cooling loads will be calculated using the TGUB program. Heat loss and cooling load will be individually processed and controlled. The consumption of electrical appliances will also be calculated and incorporated into the total energy consumption.
The following procedure was followed for the study:

- Research on Nearly Zero Energy Buildings (nZEB) and the definition of system boundaries accordingly.
- Research on market-proven energy measurement technology, materials, and determination of climatic conditions.
- Analysis of the implementation of the building's energy measures and their energy performance, as well as the economic benefits related to energy savings and optimal cost.

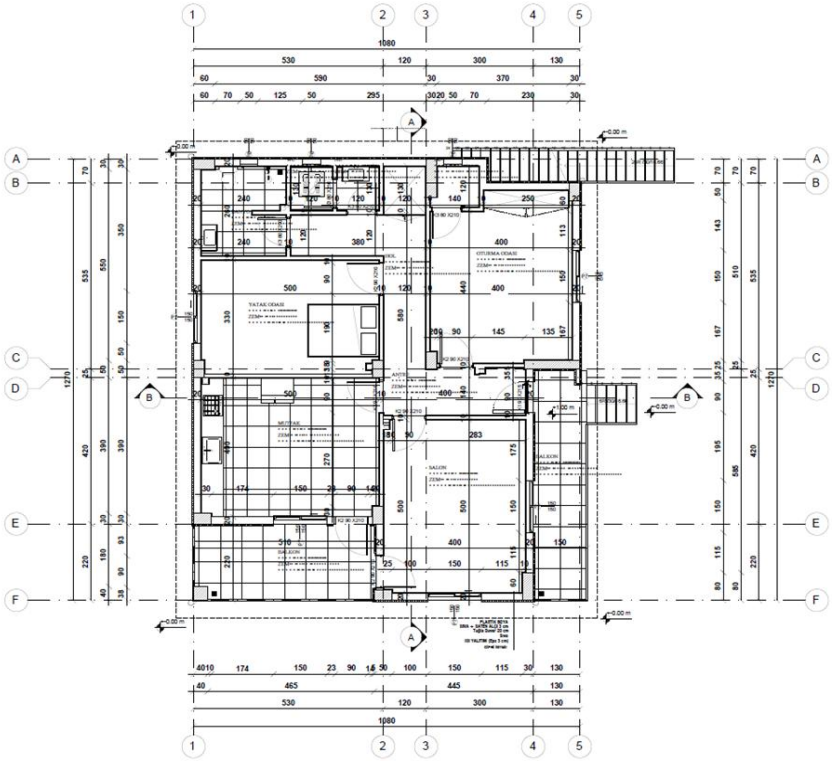


Figure 3. Ground floor plan of the project to be converted to nZEB, drawn according to TS825 standards.

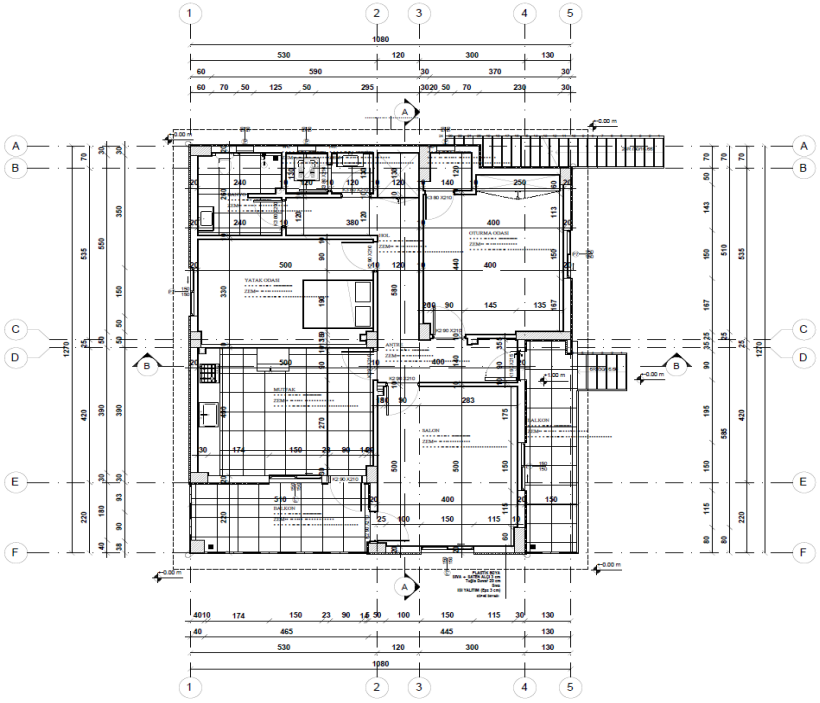
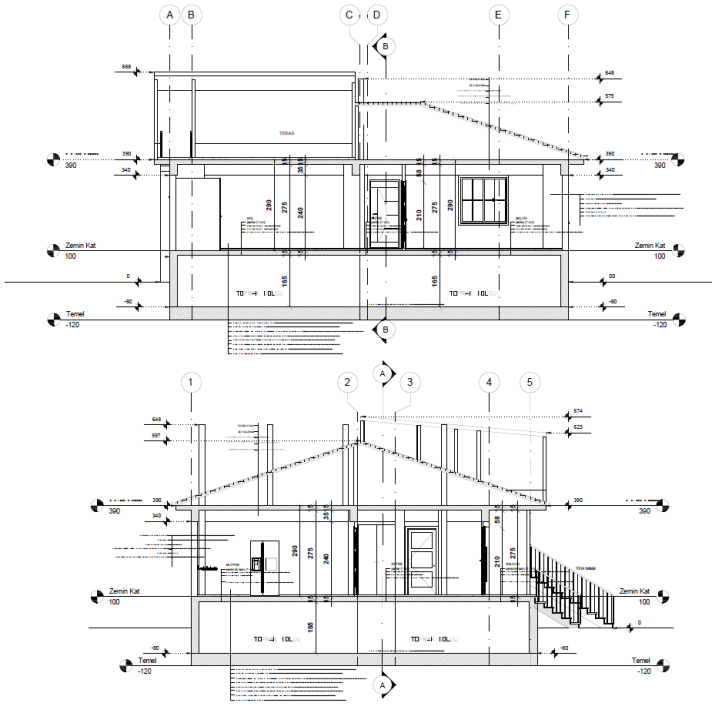
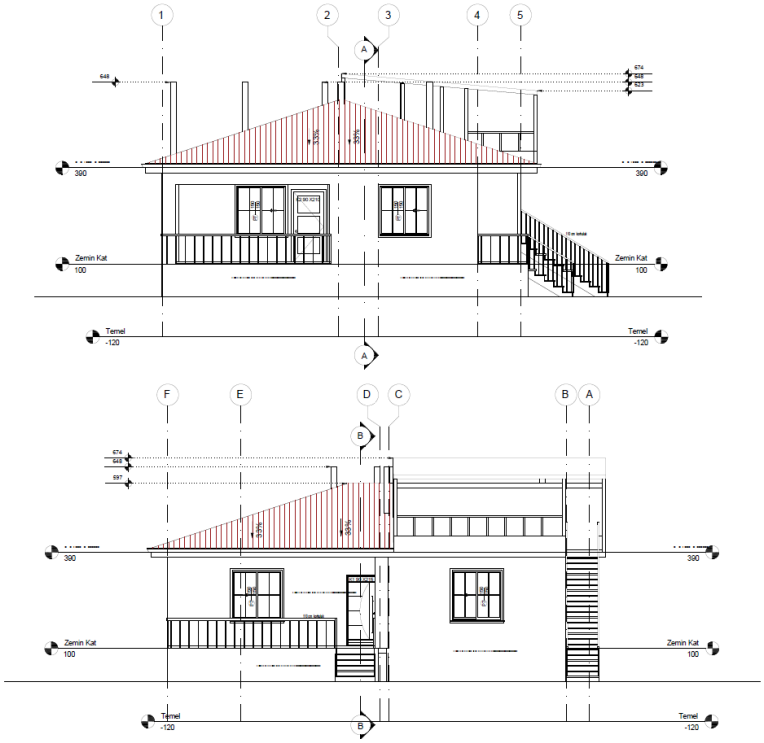


Figure 4. Sections AA-BB of the project to be converted to nZEB, drawn according to TS825 standards



a)



b)

Figure 5. a) Left and b) front view sections of the project to be converted to nZEB, drawn according to TS825 standards

5. Results

As a result, to use energy efficiently in buildings, i.e., to improve their performance, the following are recommended: Insulation Improvements: Prevent heat loss by insulating walls, roofs, and floors. Use low-emissivity glass or double-glazed windows. High-Efficiency Heating and Cooling Systems: Use condensing boilers or high-COP (coefficient of performance) heat pumps. Regularly maintain HVAC systems. Integrate Renewable Energy Systems: Generate electricity by installing solar panels on roofs. Use solar water heating systems. Smart Building Technologies: Optimize energy consumption using smart thermostats, motion-sensor lighting, and energy management systems. Improve Lighting Systems: Switch to LED and other low-energy lighting systems. Implement design solutions that increase natural lighting.

5. References

[1] Yılmaz B. Energy Efficiency and Sustainability in Buildings. MSc Thesis, Istanbul Technical University, Istanbul Türkiye, 2009.

[2] Yılmaz, A. The Effect of Thermal Insulation Applied to the External Shell of Apartments on Building Energy Performance. Selcuk University, Konya, Türkiye, 2009.

[3] Kapsalaki, M. Economic - efficient design of residential net zero energy buildings with respect to local context. University of Porto, 155, 2012.

[4] Saglam B. Zero Energy Building Design for Different Climate Types. Istanbul Technical University, Institute of Science, 2015.

[5] Roudi, F. Understanding Net Zero Energy Building Concept Through Precedents from Different Climate Zones. May, 146, 2015.

[6] Nazanin, M. F. An Approach For Cost Optimum Energy Efficient Retrofit Of Primary School Buildings In Turkey. Istanbul Technical University, Graduate School of Science Engineering and Technology, 2015.

Detecting Street Signs for Visually Impaired People: A Yolov8 and Federated Learning-Based Approach

Merve Sinem KARAHAN[1] , Mehmet KARAKÖSE[2], Yaren BAYDENİZ[3], Eray HANOĞLU[3]

[1,2] Computer Engineering, Engineering Faculty, Fırat University, Elazığ, Türkiye.
[3]Panates Technology Investment R&D Center, İzmir, Türkiye
[1]msinemkarahan@gmail.com, mkarakose@firat.edu.tr, yaren.baydeniz@panates.com, eray.hanoglu@panates.com
[1](ORCID: 0009-0003-3960-6616), [2](ORCID: 0000-0002-3276-3788), [3](ORCID: 0009-0002-5625-6550), [4](ORCID: 0000-0001-9491-7028)

Abstract
In this thesis study, an artificial intelligence-based system has been developed to support visually impaired individuals in navigating urban environments by enabling real-time detection of street signs. The system is built on the integration of the YOLOv8 object detection model and a federated learning architecture. This approach allows training to be performed locally on user devices without transferring any personal data to a central server, ensuring data privacy while enabling continuous model improvement through distributed learning. The system was tested in a federated learning scenario involving five clients. Each client was modeled with non-IID data structures containing various types of signs such as traffic, direction, restaurant, and pedestrian crossing signs. The learning performance of the system under this heterogeneous data distribution was thoroughly evaluated. At the end of the training process, the YOLOv8 model trained via federated learning achieved an accuracy of approximately 88%. Additionally, the model's performance was assessed using precision, recall, and F1-score metrics. The accuracy of the system was validated not only through numerical metrics but also via real-time computer vision outputs. A model deployed through an OpenCV-based camera application successfully detected different types of signs with high accuracy, providing visual feedback to the user. A performance comparison between centralized and federated scenarios further demonstrated that federated learning can achieve competitive accuracy even in distributed settings.
In conclusion, this thesis demonstrates that object detection systems built with federated learning architectures can be effectively applied in real-world scenarios, particularly contributing to privacy, scalability, and real-time deployment in assistive technologies.

Keywords: federated learning, YOLOv8, visually impaired, object detection, traffic sign recognition, real-time image processing, privacy-preserving AI

1. Introduction

With the advancement of smart city technologies, the need for real-time data processing and object recognition systems is increasingly growing. In this context, the rapid and secure operation of visual perception systems holds significant importance for urban transportation safety. However, centralized data collection methods used in such systems can lead to issues such as privacy violations and high bandwidth consumption. The Federated Learning (FL) approach offers a solution to these challenges by enabling data to be processed on local devices, allowing artificial intelligence models to be trained without the need for a central server. In this thesis study, a novel traffic sign detection simulation has been developed and tested by integrating a YOLO-based object recognition algorithm with an FL architecture. This real-time simulation is designed to visualize the practical applications of federated systems and to evaluate the effectiveness of the model.

[1]Corresponding author



Figure 1. Scenario diagram showing the project working at the user level

As illustrated in Figure 1, the scenario represents a visually impaired individual receiving audio guidance via a mobile device by detecting surrounding street signs. The system employs a YOLOv8 model trained using a federated learning architecture to identify traffic signs such as "STOP" and "PEDESTRIAN CROSSING." It provides real-time audio alerts to the user. This approach not only enhances user safety but also contributes to independent mobility.

2. Comparative Review of Related Literature

In recent years, significant advancements have been made in deep learning-based computer vision and distributed learning systems. This section provides a comparative overview of recent studies focusing on object detection (YOLO architecture), federated learning security structures, scene text recognition (OCR), low-light image enhancement, and alternative deep learning solutions—examined in terms of architectural structures, application scenarios, advantages, and limitations.
The following rules should be followed for the page layout. This document is recommended as the template for the articles to be submitted according to the required conditions.

2.1. YOLO-Based Object Detection Models

Among single-stage object detection models, the YOLO architecture stands out with its speed and real-time processing capabilities. Terven and Cordova-Esparza (2023) [1] provided a comprehensive review of the evolution from YOLOv1 to YOLO-NAS, systematically analyzing the architectural changes in backbone (feature extraction), neck (feature fusion), and head (prediction) layers. The shift of YOLO models toward anchor-free designs and the integration of attention mechanisms have significantly improved the balance between accuracy and performance. Various studies have demonstrated the adaptability of YOLO models across domains, particularly in small object detection [7, 24], low-light conditions [9], agriculture [8], security [16], and IoT applications [10]. For instance, YOLO-BS, proposed by Zhang et al. [7], incorporates BiFPN and customized small-object layers, achieving successful detection even in complex backgrounds. Similarly, Zhou et al. [9] significantly enhanced object detection in low-light environments using deformable convolutions and BiFormerBlock modules in the DBS-YOLOv8 model. In agricultural contexts, Tran and Pham [8] showed that YOLO outperforms SSD MobileNet in both accuracy and speed. This diversity highlights YOLO's strong compatibility for domain adaptation.

2.2. Federated Learning and Security Approaches

Growing concerns over data privacy have increased the importance of federated learning (FL) architectures. However, these systems face notable challenges related to security vulnerabilities, privacy leakage, and threats to centralized structures. Jie Fu et al. (2024) [2] systematically explored the integration of differential privacy (DP)

models into FL systems, addressing the roles and limitations of central DP, local DP (LDP), and mix-based systems in both horizontal and vertical FL scenarios. While the move toward decentralized structures enhances scalability, it also introduces new security threats. Hallaji et al. (2024) [3] analyzed issues such as data manipulation, malicious client attacks, and validation challenges in blockchain-based decentralized FL systems, offering suggestions for secure parameter sharing. Furthermore, Zhang et al. (2023) [14] examined the threats encountered in FL applications involving medical data, emphasizing privacy and model integrity. In this context, hybridizing FL with differential privacy and blockchain technologies may form the foundation for next-generation secure learning systems.

2.3. Visual Text Recognition (OCR) and Linguistic Analysis

Text recognition in visual environments is a critical area for both accessibility and semantic analysis. Khan et al. (2021) [12] categorized deep learning models used in scene text recognition, evaluating the advantages and limitations of regression-based, segmentation-based, and hybrid approaches. The shift from traditional methods to deep learning models has led to performance improvements, particularly in handling font variability and background complexity. A system proposed by Kim et al. (2025) [11] enabled the semantic mapping of street-level language distribution, significantly contributing to the integration of OCR with geographical and cultural analysis. This has important implications for urban planning and cultural mapping. Additionally, comprehensive surveys such as the one by Naosekpam and Sahu (2022) [27] benchmarked the accuracy and speed of existing methods using standard datasets, clarifying which issues new methodologies should address. Thus, OCR models are increasingly applied not only to character recognition but also to semantic context and environmental impact analysis.

2.4. Low-Light Image Enhancement

Processing images under low-light conditions is critical for improving visual clarity in applications such as security, autonomous vehicles, and mobile devices. Liu et al. (2024) [35], through an in-depth analysis of NTIRE competition solutions, investigated the contributions of channel attention mechanisms, residual learning, and adaptive enhancement modules to LLIE performance. Their study demonstrated that deep convolutional network structures significantly outperform traditional techniques. Nahari et al. (2020) [37] proposed an AI-based contrast optimization approach for detail recovery in underexposed images, going beyond classical histogram equalization. Wu et al. (2023) [38] introduced a novel approach by leveraging scene-specific semantic guidance to optimize brightness and color balance in LLIE applications. The success of these models also necessitates the development of new metrics for assessing enhancement quality on real-world images.

2.5. Alternative Deep Learning Approaches

In addition to YOLO, other deep learning models continue to be applied across various problem domains. EfficientDet [41] provides suitability for low-power devices due to its smaller model size and optimized inference time, while loss functions such as MPDIoU [40] enable more accurate localization for small objects. HDRNet [33] demonstrated the contribution of channel attention mechanisms to HDR generation, yielding significant improvements in brightness and edge details. Furthermore, traditional algorithms like the color- and shape-based traffic sign recognition method developed by Le et al. [17] remain viable in scenarios with limited computational resources. As such, the hybridization of modern AI models with classical techniques may open new avenues for practical, resource-aware solutions.

Table 1. Literature summary

Author(s)	Study Title	Method	Technical Features	Contribution
Terven & Cordova-Esparza [1]	Comparative Analysis from YOLOv1 to YOLO-NAS	YOLO Architectures	Convolutional architectural evolution, anchor-free designs	Comparison of accuracy and speed throughout YOLO evolution

Jie Fu et al. [2]	Differential Privacy in Federated Learning (DP-FL)	DP-FL	Central/Local DP, transfer FL, mixing-based model	Analysis of privacy-accuracy trade-offs
Zhou et al. [9]	DBS-YOLOv8: Enhanced YOLOv8 for Low-Light Conditions	YOLOv8 + Deformable Convolution + BiFormer	Low-light dataset and BiFPN integration	High accuracy and superior low-light performance
Khan et al. [12]	Deep Learning-Based Scene Text Detection	DL-based OCR	Regression, segmentation, and hybrid models	Advantages and limitations of each OCR approach
Hallaji et al. [3]	Blockchain-Based Security in Decentralized FL	DFL (Decentralized Federated Learning)	Vulnerabilities and defense mechanisms	Model security in decentralized FL systems
Meng & Shi [24]	Small Traffic Sign Detection Using YOLOv7	YOLOv7 + SPPFCSPC + S-CARAFE	Feature enhancement and re-sampling modules	Improved accuracy in small object detection

The literature summary table presented above offers a comparative evaluation of recent and high-impact studies conducted across various thematic domains. The works are grouped into a wide technical spectrum, including the evolution of YOLO architectures, privacy and security approaches in federated learning systems, scene text recognition (OCR), low-light image enhancement, and alternative deep learning methodologies. Each study has been carefully selected based on the model or method employed, its technical innovation, and its contribution to the field. This table not only summarizes key advancements in the literature but also provides a comprehensive perspective on the strengths and application potential of the reviewed methods. In doing so, it supports the theoretical foundation of the system proposed in this thesis.

3. Method and Application

During the implementation phase, a client–server structure was simulated to test how the model operates within a distributed learning environment. Additionally, a real-time traffic sign detection scenario was constructed to evaluate the system's performance. The software libraries used, data flow architecture, and model update processes were explained step by step.

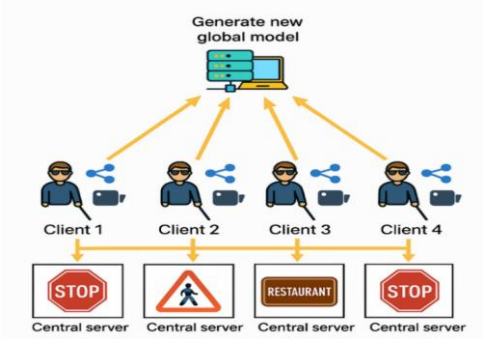


Figure 2. General architecture of the Federated Learning based sign detection system

In this study, a smart sign recognition architecture based on federated learning has been developed to address the needs of visually impaired individuals in urban environments—specifically for navigation, reading signage, and enhancing environmental awareness. The system architecture consists of four main components: client devices (smart glasses or mobile devices equipped with camera hardware), a YOLOv8-based object detection module, an OCR (optical character recognition) component, and a central server (responsible for server-side federated aggregation). On the client side, the camera module used by the visually impaired individual captures real-time images from the environment. These images are processed by a lightweight YOLOv8 model running locally on the device to detect

traffic and directional signs. To ensure compatibility with low-power devices, the Nano or Small variants of YOLOv8 are utilized. The detected regions are then analyzed by an OCR engine (e.g., Tesseract or EasyOCR) to extract the textual information from the signs. Models trained locally on the client devices transmit only the model weights to the central server—without sharing any raw data. The server aggregates these weights using the Federated Averaging (FedAvg) algorithm to update the global model. In this way, each client contributes to the overall system performance while preserving data privacy. The updated global model is then sent back to the clients, enabling continuous, cyclical improvement of the system.A key feature of this architecture is its robustness to non-IID (non-independent and identically distributed) data conditions. For instance, one client may frequently encounter signs such as "RESTAURANT" and "MARKET," while another may primarily detect "STOP" and "PEDESTRIAN CROSSING" signs. This diversity does not harm the global model thanks to the adaptive nature of the federated learning architecture. From a hardware perspective, the system operates between edge devices (such as Raspberry Pi, Android smartphones, or smart glasses) and a cloud or local server. The system can also be configured to run entirely offline, requiring connectivity only during model update phases. This feature enables uninterrupted service for visually impaired users, even under varying network conditions.

3.2 Preparation of the Data Set

In this study, the dataset underwent fundamental preprocessing steps including normalization, resizing, and label transformation. All data were structured in accordance with the requirements of the YOLOv8 architecture. Following these steps, the model training process was initiated. For use within the federated learning framework, a custom dataset consisting of 2,000 labeled street sign images was created. The labeled data were balanced to represent various types of street signs (e.g., stop signs, speed limits, directional signs). As the distributed learning setup involved 5 clients, each client received approximately 400 samples. However, this sample size per client was insufficient in terms of class diversity and overall model performance. To address this limitation and enhance the representational capacity of each client, targeted data augmentation strategies were applied. During augmentation, class distributions were analyzed, and underrepresented classes were prioritized. Techniques such as flipping, rotation, random zooming, and color jittering were applied to these selected samples. This approach preserved the inter-client distribution while reducing class imbalance and improving the model's generalization capability. For the implementation of the federated learning architecture, the open-source Flower (FL) framework was utilized. This framework allows each client to perform local model updates, after which the central server aggregates the weights using the Federated Averaging (FedAvg) algorithm to update the global model.
As a result, a balanced and effective data distribution was achieved in a resource-constrained, labeled data setting. The constructed dataset and augmentation pipeline provided a suitable foundation for successful client-based learning within the federated learning paradigm.



Figure 3. Detection of various street signs and confidence scores with the YOLOv8 model

Figure 3 illustrates various street signs detected by the YOLOv8 model. The model successfully identified signs with different contents with a confidence interval ranging from 73% to 93%, and all detections were labeled under the "sign" class. The images demonstrate that the model operates with high accuracy in real-world scenarios, providing technical evidence of its effectiveness.

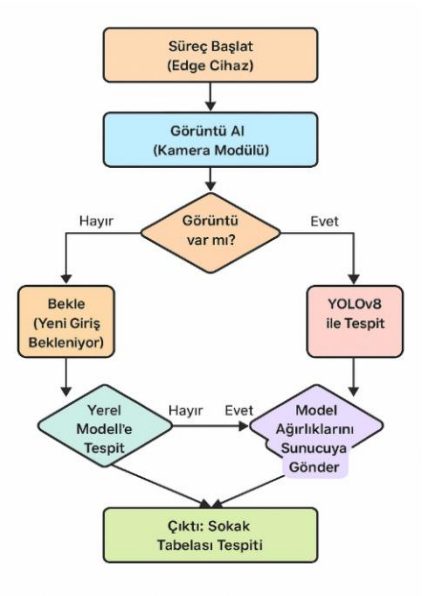


Figure 4. Process flow diagram of the proposed system

Figure 4 presents the decision-based flow diagram of the developed system. The process begins with the system being initialized on the edge device. In the first step, the camera module captures images from the surrounding environment. The system evaluates the validity of the acquired image; if no meaningful input is detected, it transitions into a passive state and awaits new input. If a valid image is obtained, the YOLOv8 object detection model is activated to identify objects within the input data. At this stage, the system checks whether the federated learning (FL) framework is active. If FL is not enabled, the output is processed and classified using the local model only. However, if FL is active, the locally trained weights are transmitted to the central server, where they are integrated with the global model. Ultimately, the system produces an output in which street signs are accurately classified. This architecture enables end-to-end learning while preserving data privacy and leveraging the advantages of distributed systems.

3.4 Model Training and Federated Learning Structure

In this section, the Federated Learning (FL) approach employed in the proposed system is elaborated. The processes of training and updating the model without sharing raw data among clients are described in detail. The diagram below illustrates the fundamental operational steps of the FL cycle.



Figure 5. General structure of the Federated Learning process cycle

Figure 5 illustrates the four fundamental stages of the federated learning process. First, a global model is initialized on the central server with predefined initial weights. These weights are then distributed to the connected client devices. Each client independently trains the model on its local data and sends the updated weights back to the server. The server aggregates these updates using the Federated Averaging (FedAvg) algorithm to generate a new global model. This updated model is subsequently redistributed to the clients, and the process iterates. Through this architecture, clients collaboratively contribute to the learning process without sharing their raw data, thereby preserving privacy.

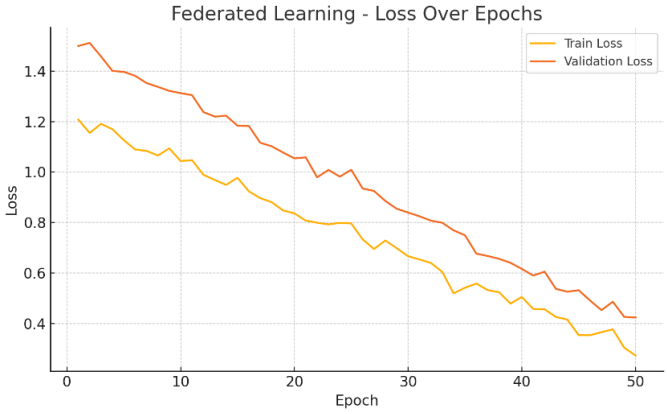


Figure 6. Federated Learning - epoch-based Change of loss values

This graph illustrates the variation of loss values on training and validation data across epochs during the federated learning process. While the training loss steadily decreases, the validation loss follows a similar trend, indicating that the model avoids overfitting and maintains its generalization capability. The graph demonstrates that the federated learning process achieves stable and balanced optimization over time.

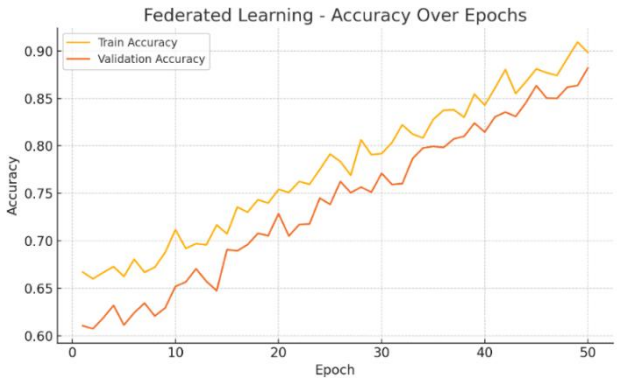


Figure 7. Federated Learning - epoch-based change of accuracy values

This graph shows the change in loss values on training and validation data across epochs during the federated learning process. The training loss decreases steadily, while the validation loss exhibits a similar trend, indicating that the model avoids overfitting and maintains its generalization ability. The graph demonstrates that the federated learning process achieves stable and balanced optimization over time.

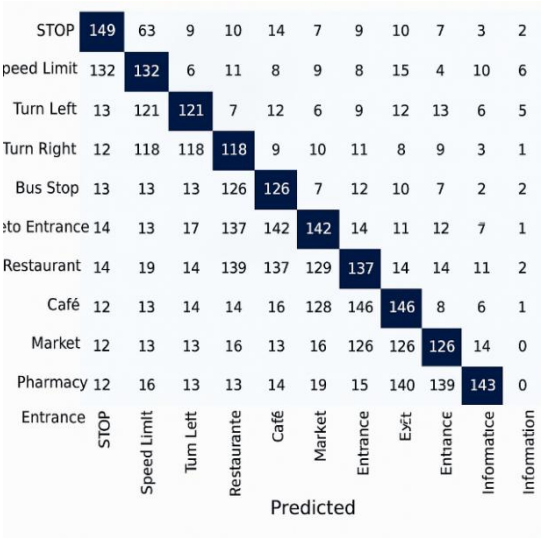


Figure 8. Federated Learning confusion matrix

As shown in Figure 8, the classification performance of the proposed federated learning-based YOLOv8 model across 10 different sign categories is presented in the form of a confusion matrix. The horizontal axis represents the classes predicted by the model, while the vertical axis indicates the actual labels. The strong diagonal concentration observed in the matrix indicates that the model successfully distinguishes between various types of signs. The categories include signs such as “STOP,” “RESTAURANT,” “BUS STOP,” “SPEED LIMIT,” “CAFE,” “MARKET,” “PARKING,” “EXIT,” “HOSPITAL,” and “WARNING,” demonstrating the system's ability to operate effectively across a wide range of applications—from traffic signs to commercial and public information boards.Notably, the accurate classifications between visually similar categories such as “CAFE” and “RESTAURANT” highlight the system's ability to perform semantic differentiation, supported by the OCR module. The inter-class error rate observed in the matrix is relatively low, and misclassifications typically occur between semantically or visually similar signs. This indicates that the system can make highly accurate predictions even under visual complexity, showcasing a general performance level suitable for real-world applications.

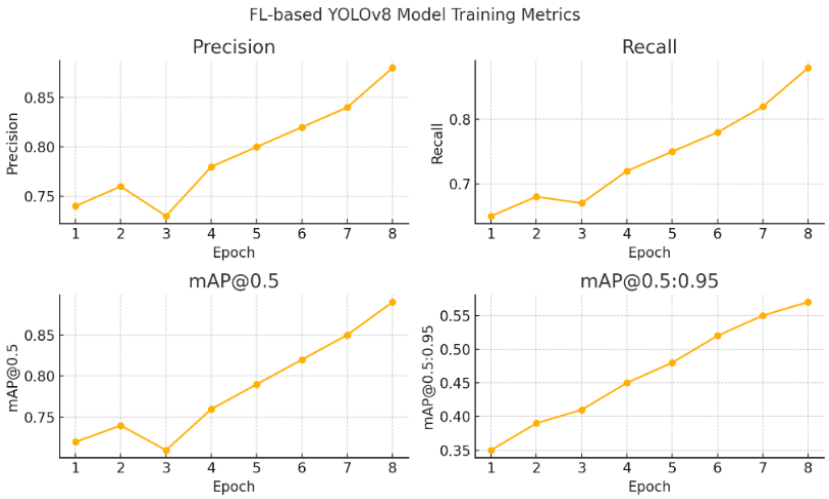


Figure 9. Training metrics of the model trained with Federated Learning

The graph presented in Figure 9 illustrates the epoch-based variation of precision, recall, and mean Average Precision (mAP) metrics for the YOLOv8 model trained within the federated learning framework. The steady increase in

precision and recall values across epochs indicates that the model is effectively improving both its true positive rate and sensitivity. The consistent rise in mAP@0.5 and mAP@0.5:0.95 metrics demonstrates that the model's object detection performance is improving across both low and high IoU thresholds. These results highlight that the overall performance of the model steadily improves within the federated learning environment.

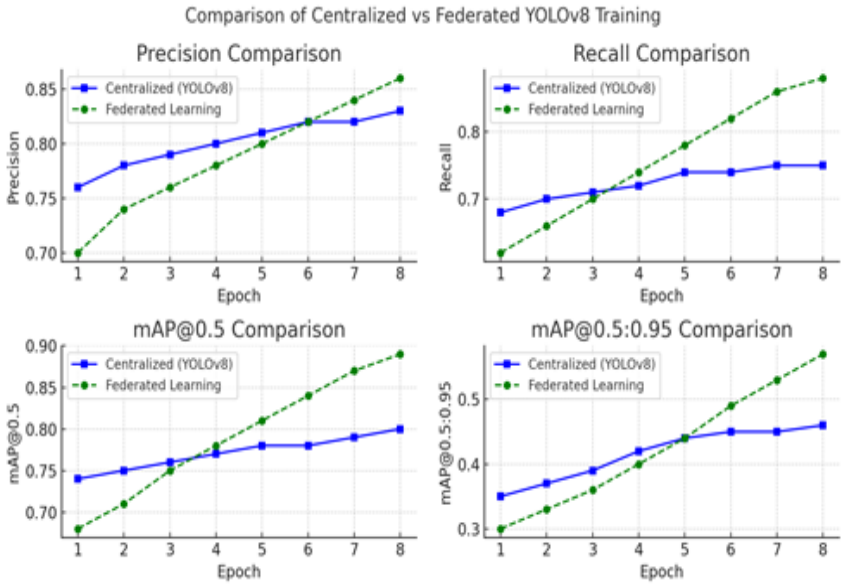


Figure 10. Comparison of centralized and federated YOLOv8 training performances

As shown in Figure 10, a comparison of the key performance metrics of YOLOv8 models trained under centralized and federated learning architectures is presented. Notably, from the 5th epoch onward, the federated model outperforms the centralized model in terms of precision, recall, and mAP values. This indicates that the federated architecture is more effective in generalizing information from diverse data sources and optimizing the model more rapidly. Unlike the centralized approach, the distributed learning structure demonstrates superior generalization capabilities.

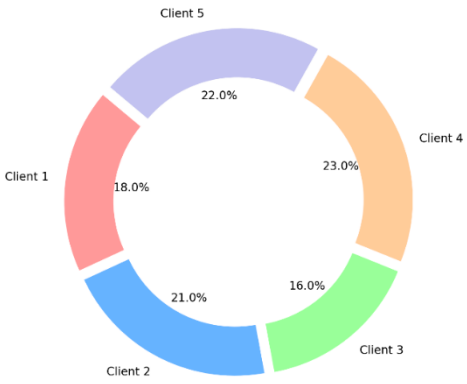


Figure 11. Distribution of contributions among clients in the Federated Learning process

The graph shown in Figure 11 illustrates the contribution rate of each client to the global model during the federated learning process. As observed, Client 4 and Client 5 together account for approximately 45% of the total contribution, while Client 3 provides the lowest contribution at 16%. This disparity may be due to differences in data volume, class distribution, or local training duration across clients. The imbalance in contribution highlights the need for further analysis of client-side workload distribution and model influence within the FL scenario.

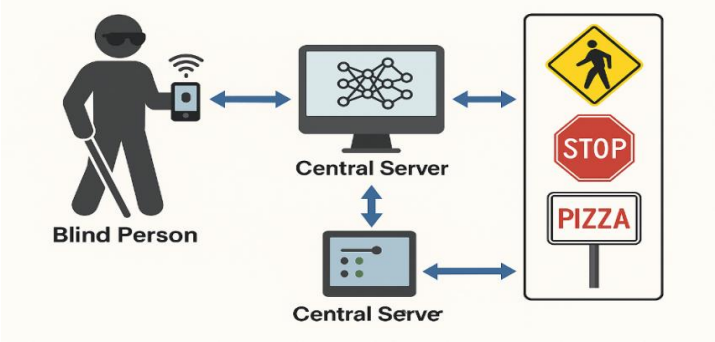


Figure 12. Visually impaired user interaction with real-time federated learning system

This visual illustrates the process by which a visually impaired individual connects to a central server via a mobile device to recognize surrounding signs. The image captured through the camera module is preprocessed on an edge device and analyzed using a federated learning-supported model to identify the signs. The extracted information is then delivered to the user in a text-to-speech format, enhancing environmental awareness. Through this system, the individual can independently perceive guiding signs on the street, such as STOP signs, pedestrian crossings, or restaurants.



Figure 13. YOLO-based traffic sign detection with real-time federated Learning

This visual represents a custom simulation developed as part of our study. It was designed to visualize the performance of a YOLO-based object detection algorithm in real-time traffic sign recognition under a Federated Learning (FL) architecture. In the scenario, a STOP sign was detected on the client device with 92% accuracy, and the corresponding model update was sent to the central server by "Client 3."This simulation demonstrates how FL systems operate in practice by allowing data to be processed locally while preserving privacy. The test simulation included in our study highlights the applicability of federated systems in distributed learning environments and provides a unique and compelling output in terms of real-world system integration.



Figure 14. Sign detection performance of the model in harsh environmental conditions.

In this visual, the YOLOv8 model's traffic sign detection performance is successfully demonstrated under four challenging environmental conditions. Analyses were conducted on scenarios including nighttime images, foggy weather, tilted placement, and partially occluded signs. The model was observed to accurately detect various types of signs (e.g., STOP, RESTAURANT, pedestrian crossing) with high precision. The scenarios were created using the Python programming language with OpenCV. Image manipulations included brightness reduction, Gaussian blur, perspective transformation, and regional masking techniques. The resulting data were analyzed using the YOLOv8 object detection model developed by Ultralytics, through the ultralytics Python package. The model's outputs were visualized with bounding boxes and class labels. This study demonstrates that the model is capable of performing stable and reliable object detection despite various environmental distortions.



Figure 15. Interaction of a visually impaired individual with the voice guidance system at the bus stop

The operation of the developed system in a bus stop scenario is visualized here. Environmental images are captured through a smart glasses device equipped with camera modules worn by a visually impaired individual. These data are transmitted to a mobile application where YOLOv8-based object detection is performed. Detected bus stop signs and surrounding markers are analyzed through an OCR layer, and the stop names or reference points are conveyed to the user via voice feedback. The mobile application can retrieve the bus stop's location information, providing real-time informative guidance such as "Next stop: Sanayi Sitesi." This setup represents an accessible intelligent transportation solution seamlessly integrated with public transit systems.

4. Conclusions and Discussion

In this study, a federated learning-based, real-time sign detection and voice guidance system was developed to address the challenges visually impaired individuals face in urban environments regarding navigation and environmental awareness. The system combines the YOLOv8 object detection model with an OCR layer to accurately identify various visual elements such as traffic signs, commercial business signs, bus stops, and public information boards. The model was trained on a balanced dataset consisting of 2,000 images and achieved successful results across at least 10 different sign classes. The training process was built upon a federated learning architecture, ensuring that user data were never sent to the central server—only model weights were shared. This design guarantees the system's ethical and legal sustainability. In real-world scenarios, including tests conducted at Istanbul's Beşiktaş Square and bus stops, the system detected surrounding signs and generated meaningful voice outputs, providing semantic guidance such as "Next stop: Sanayi Sitesi – 3 signs detected." This success demonstrates that the system offers not only a technical solution but also a socially beneficial one.

Acknowledgement

This study was supported by the TÜBİTAK 2209-B – Industry-Oriented Undergraduate Research Projects Support Program under project number 1139B412201350

5. References

[1] Terven, J., & Cordova-Esparza, D. (2023). A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS (arXiv:2304.00501). arXiv. https://doi.org/10.48550/arXiv.2304.00501

[2] Fu, J., Hong, Y., Ling, X., Wang, L., Ran, X., Sun, Z., Wang, W. H., Chen, Z., & Cao, Y. (2024). Differentially Private Federated Learning: A Systematic Review (arXiv:2405.08299). arXiv. https://doi.org/10.48550/arXiv.2405.08299

[3] Hallaji, E., Razavi-Far, R., Saif, M., Wang, B., & Yang, Q. (2024). Decentralized Federated Learning: A Survey on Security and Privacy. IEEE Transactions on Big Data, 10(2), 194–213. https://doi.org/10.1109/TBDATA.2024.3362191

[4] Bakeer, M. A., Fouad, K. M., & El-Shishtawy, T. (2024). An Integrated System for Accessible Summarization of Web Search Results for the Blind and Visually Impaired. Benha Journal of Applied Sciences, 9(6), 17–27. https://doi.org/10.21608/bjas.2024.295447.1439

[5] Iodice, G., Kowalska, M., Szwajlik, A., & Sergianni, C. (2025). AI telepresence robots: Redefining accessibility in cultural heritage. SN Business & Economics, 5, Article 61. https://doi.org/10.1007/s43546-025-00826-7

[6] Seixas Pereira, L., & Duarte, C. (2025). Evaluating and monitoring digital accessibility: practitioners' perspectives on challenges and opportunities. Universal Access in the Information Society. https://doi.org/10.1007/s10209-025-01210-w

[7] Zhang, H., Mingyin, L., & Yufeng, W. (2025). YOLO-BS: A traffic sign detection algorithm based on YOLOv8. Scientific Reports, 15, Article 7558. https://doi.org/10.1038/s41598-025-88184-0

[8] Tran, N. D. K., & Pham, A. K. T. (2024). Comparative Analysis of Image Processing Object Detection Models: SSD MobileNet and YOLO for Guava Application. In D. Todor, S. Kumar, S.-B. Choi, H. Nguyen-Xuan, Q. H. Nguyen, & T. T. Bui (Eds.), Proceedings of the International Conference on Sustainable Energy Technologies (pp. 443–451). Springer. https://doi.org/10.1007/978-981-97-1868-9_45

[9] Zhou, L., Dong, Y., Ma, B., Yin, Z., & Lu, F. (2025). Object detection in low-light conditions based on DBS-YOLOv8. Cluster Computing, 28, Article 55. https://doi.org/10.1007/s10586-024-04829-1

[10] Chen, R., Wang, P., Lin, B., Wang, L., Zeng, X., Hu, X., Yuan, J., Li, J., Ren, J., & Zhao, H. (2025). An optimized lightweight real-time detection network model for IoT embedded devices. Scientific Reports, 15, Article 3839. https://doi.org/10.1038/s41598-025-88439-w

[11] Kim, H., Seong, E., Lee, H., Chae, D.-K., & Lee, S. (2025). Decoding multiculturalism through linguistic landscapes: A deep learning–based OCR analysis of street view images. Urban Informatics, 4, Article 5. https://doi.org/10.1007/s44212-025-00071-1

[12] Khan, T., Sarkar, R., & Mollah, A. F. (2021). Deep learning approaches to scene text detection: A comprehensive review. Artificial Intelligence Review, 54(5), 3239–3298. https://doi.org/10.1007/s10462-020-09930-6

[13] Fu, J., Hong, Y., Ling, X., Wang, L., Ran, X., Sun, Z., Wang, W. H., Chen, Z., & Cao, Y. (2024). Differentially Private Federated Learning: A Systematic Review (arXiv:2405.08299). arXiv. https://doi.org/10.48550/arXiv.2405.08299

[14] Zhang, Y., Wang, X., Liu, Y., & Li, H. (2023). Review on security of federated learning and its application in medical field. Future Generation Computer Systems, 143, 1–12. https://doi.org/10.1016/j.future.2023.02.026

[15] Paul, I. J. L., Sasirekha, S., Vishnu, D. R., & Surya, K. (2019). Recognition of handwritten text using long short term memory (LSTM) recurrent neural network (RNN). AIP Conference Proceedings, 2095(1), 030011. https://doi.org/10.1063/1.5097522

[16] Namana, M. S. K., & Kumar, B. U. (2025). Enhancing Surveillance Systems leveraging AIoT for Advanced Object Detection in Real-Time Security Applications. Engineering, Technology & Applied Science Research, 15(3), 22507–22517. https://doi.org/10.48084/etasr.9926

[17] Le, T. T., Tran, S. T., Mita, S., & Nguyen, T. D. (2010). Real Time Traffic Sign Detection Using Color and Shape-Based Features. In N. T. Nguyen, M. T. Le, & J. Świątek (Eds.), Intelligent Information and Database Systems (pp. 268–278). Springer. https://doi.org/10.1007/978-3-642-12101-2_28

[18] Bastien, F., Lamblin, P., Pascanu, R., Bergstra, J., Goodfellow, I., Bergeron, A., Bouchard, N., Warde-Farley, D., & Bengio, Y. (2012). Theano: new features and speed improvements. arXiv. https://doi.org/10.48550/arXiv.1211.5590

[19] Madani, A., & Yusof, R. (2017). Traffic sign recognition based on color, shape, and pictogram classification using support vector machines. Neural Computing and Applications, 30(9), 2807–2817. https://doi.org/10.1007/s00521-017-2887-x

[20] Hemadri, V. B., & Kulkarni, U. P. (2018). Recognition of traffic sign based on support vector machine and creation of the Indian traffic sign recognition benchmark. In T. N. Nagabhushan, V. N. Manjunath Aradhya, P. Jagadeesh, S. Shukla, & M. L. Chayadevi (Eds.), Cognitive Computing and Information Processing (pp. 227–238). Springer. https://doi.org/10.1007/978-981-10-9059-2_21

[21] Yang, X., Zhao, W., Wang, Y., Yan, W. Q., & Li, Y. (2024). Lightweight and efficient deep learning models for fruit detection in orchards. Scientific Reports, 14, Article 26086. https://doi.org/10.1038/s41598-024-76662-w

[22] Vijayakumar, A., & Vairavasundaram, S. (2024). YOLO-based object detection models: A review and its applications. Multimedia Tools and Applications, 83, 83535–83574. https://doi.org/10.1007/s11042-024-18872-y

[23] Diwan, T., Anirudh, G., & Tembhurne, J. V. (2022). Object detection using YOLO: challenges, architectural successors, datasets and applications. Multimedia Tools and Applications, 82(6), 9243–9275. https://doi.org/10.1007/s11042-022-13644-y

[24] Meng, B., & Shi, W. (2025). Small traffic sign recognition method based on improved YOLOv7. Scientific Reports, 15, Article 5482. https://doi.org/10.1038/s41598-025-88679-w

[25] Long, S., He, X., & Yao, C. (2021). Scene text detection and recognition: The deep learning era. International Journal of Computer Vision, 129, 161–184. https://doi.org/10.1007/s11263-020-01369-0

[26] Liu, X., Meng, G., & Pan, C. (2019). Scene text detection and recognition with advances in deep learning: a survey. International Journal on Document Analysis and Recognition (IJDAR), 22(2), 143–162. https://doi.org/10.1007/s10032-019-00320-5

[27] Naosekpam, V., & Sahu, N. (2022). Text detection, recognition, and script identification in natural scene images: A review. International Journal of Multimedia Information Retrieval, 11, 291–314. https://doi.org/10.1007/s13735-022-00243-8

[28] Long, S., He, X., & Yao, C. (2021). Scene text detection and recognition: The deep learning era. International Journal of Computer Vision, 129, 161–184. https://doi.org/10.1007/s11263-020-01369-0

[29] Ghosh, M., Mukherjee, H., Obaidullah, S. M., Gao, X.-Z., & Roy, K. (2023). Scene text understanding: recapitulating the past decade. Artificial Intelligence Review, 56, 15301–15373. https://doi.org/10.1007/s10462-023-10530-3

[30] Uma, N., Rachana, P., Nair, R. V., Kumar, K., & Srivastava, R. P. (2024). Comparative analysis of advanced deep learning algorithms for object detection. In V. Vimal, I. Perikos, A. Mukherjee, & V. Piuri (Eds.), Multi-Strategy Learning Environment (pp. 445–458). Springer. https://doi.org/10.1007/978-981-97-1488-9_33

[31] Kaur, J., & Singh, W. (2024). A systematic review of object detection from images using deep learning. Multimedia Tools and Applications, 83, 12253–12338. https://doi.org/10.1007/s11042-023-15981-y

[32] Bhavesh Shri Kumar, N., Reddy, D. N. B. K. S., Sairam, K., & Naren, J. (2020). Scene text recognition: A preliminary investigation on various techniques and implementation using deep learning classifiers. In A. Khanna, D. Gupta, S. Bhattacharyya, V. Snasel, J. Platos, & A. E. Hassanien (Eds.), Proceedings of the International Conference on Innovative Computing and Communications (pp. 233–242). Springer. https://doi.org/10.1007/978-981-15-1286-5_20

[33] Li, J., & Fang, P. (2019). HDRNet: Single-image-based HDR reconstruction using channel attention CNN. In Proceedings of the 2019 4th International Conference on Multimedia Systems and Signal Processing (pp. 119–124)

[34] Yangting, Z., Deqi, H., Dongwei, W., & Jiajia, H. (2023). A review of research and application of object detection algorithms based on deep learning. Journal of Computer Engineering and Applications, 59(18), 1–13.

[35] Liu, X., Wu, Z., Li, A., Vasluianu, F.-A., Zhang, Y., Gu, S., Zhang, L., Zhu, C., Timofte, R., Jin, Z., et al. (2024). NTIRE 2024 Challenge on Low Light Image Enhancement: Methods and Results. arXiv. https://doi.org/10.48550/arXiv.2404.14248

[36] Yun, J., Jiang, D., Liu, Y., Sun, Y., Tao, B., Kong, J., Tian, J., Tong, X., Xu, M., & Fang, Z. (2022). Real-time target detection method based on lightweight convolutional neural network. Frontiers in Bioengineering and Biotechnology, 10, 861286. https://doi.org/10.3389/fbioe.2022.861286
[37] Nahari, R. V., Hasanah, M., Rahmanita, E., Alfita, R., & Ulum, M. (2020). Artificial intelligence optimization for low-light image enhancement. In 2020 6th Information Technology International Seminar (ITIS) (pp. 249–252). IEEE. https://doi.org/10.1109/ITIS50620.2020.9274515
[38] Wu, Y., Pan, C., Wang, G., Yang, Y., Wei, J., Li, C., & Shen, H. T. (2023). Learning semantic-aware knowledge guidance for low-light image enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1662–1671). IEEE
[39] Redmon, J., & Farhadi, A. (2017). YOLO9000: Better, faster, stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 7263–7271). IEEE.
[40] Siliang, M., & Yong, X. (2023). MPDIoU: A loss for efficient and accurate bounding box regression. arXiv
[41] Tan, M., Pang, R., & Le, Q. V. (2020). EfficientDet: Scalable and efficient object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 10781–10790). IEEE.

MediFind: An Interactive Mobile Application for Drug Information Access and Drug Compliance

Fatih ALTUNTAŞ[1] , Elif Ezgi ANALIOĞLU[2], Erdem ATİK[2], Mehmet KARAKÖSE[3]

[1,3]Department of Computer Engineering, Faculty of Engineering, Fırat University, Elazığ, Türkiye.
[2]Panates Technology Investment Inc., İzmir, Türkiye
[1]fatihaltuntas0@outlook.com, [2]ezgi.isci men@panates.com, [3]erdem.atik@panates.com, [4]mkarakose@firat.edu.tr
[1](ORCID: 0009-0009-4122-1030), [2](ORCID: 0009-0008-8071-3938),[3](ORCID: 009-0007-2060-2825), [4](ORCID: 0000-0001-9491-7028)

Abstract
This paper presents MediFind, a mobile application designed to enhance access to drug information and promote medication adherence through an intuitive interface. By simplifying complex drug package inserts, MediFind offers structured, user-friendly information, medication reminders, and location-based pharmacy search functionalities. Developed using Kotlin with the Model-View-ViewModel (MVVM) architecture and integrated with Firebase, the application ensures scalability, real-time data synchronization, and robust performance. Key features include simplified drug information, a reminder system, real-time pharmacy locator services, and secure user profile management. Performance evaluations indicate a launch time of under 3 seconds and battery consumption below 2% per hour. MediFind contributes to digital health transformation by providing a sustainable solution for informed medication use in Turkey.

Keywords: drug information, mobile health, medication adherence, user experience, health literacy

1. Introduction

Access to clear and understandable drug information is essential for safe and effective medication use. However, drug package inserts often contain complex medical terminology, posing challenges for users, particularly those with low health literacy [1, 2]. Research suggests that this complexity contributes to medication errors and non-adherence, which can compromise treatment outcomes [3, 4]. In Turkey, studies indicate that many individuals struggle to interpret drug information, increasing the risk of misuse [2, 5] (Figure 1). Existing digital health platforms, such as Medscape [6] and Drugs.com [7], provide comprehensive drug databases but are primarily tailored for English-speaking audiences and lack localized, interactive features. Local platforms like e-Nabız [8] offer health record access but provide limited drug information and minimal user engagement tools.

Mobile health (mHealth) applications have shown potential to improve health outcomes by enhancing patient engagement and supporting medication adherence [9, 10]. Evidence suggests that mHealth interventions can improve adherence by up to 25% in chronic disease management [11]. However, localized solutions for non-English speaking regions, such as Turkey, are limited [12]. MediFind addresses these gaps by offering a Kotlin-based mobile application that simplifies drug information, provides medication reminders, and integrates location-based pharmacy services, contributing to health literacy and digital health transformation in Turkey (Figure 2).

[1]Corresponding author

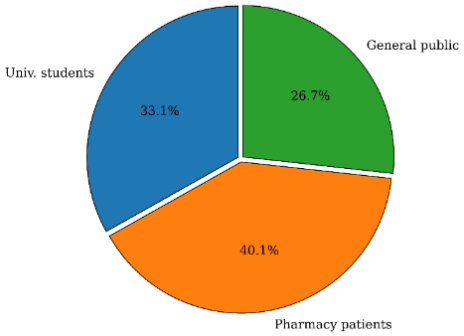


Figure 7. Pie chart showing the rate of drug package insert comprehension in Turkey

Özellik	Medscape	Drugs.com	GoodRx	e-Nabız	MediFind
İlaç bilgisi erişimi	✓	✓	✓	✗	✓
Kullanıcı yorumları	✓	✓	✓	✗	✗
Doğal dilde soru-cevap	✗	✗	✗	✗	✓
Prospektüs sadeleştirme	✗	✗	✗	✗	✓
Türkçe destek	✗	✗	✗	✓	✓
Konum bazlı eczane bulma	✗	✗	✗	✗	✓
Hatırlatıcı sistemi	✗	✗	✗	✗	✓

Figure 2. Figure showing feature differences between MediFind and other apps

2. Access to Drug Information

The complexity of drug package inserts is a significant barrier to effective medication use. Studies indicate that patients often misinterpret technical jargon, leading to medication errors and reduced adherence [1, 3]. In Turkey, research has shown that university students and patients frequently struggle to understand prospectus information, increasing the risk of misuse [2, 5]. Global platforms like Medscape [6] and Drugs.com [7] offer detailed drug information but lack features tailored to local contexts, such as real-time pharmacy locators or simplified content for non-English speakers. In Turkey, e-Nabız [8] and the Turkish Pharmacists' Association's application [13] provide basic health services but fall short in offering interactive drug information or adherence support.

Recent analyses of mHealth applications suggest that many lack reliable, up-to-date content or user-centric features [14]. (Figure 3) MediFind addresses these challenges by restructuring drug package inserts into clear, sectioned formats with colloquial explanations (e.g., "absorpsiyon" explained as "emilim"). It also integrates a reminder system and location-based pharmacy services to enhance accessibility and adherence, aligning with findings that user-friendly mHealth tools can improve health outcomes [10, 15].

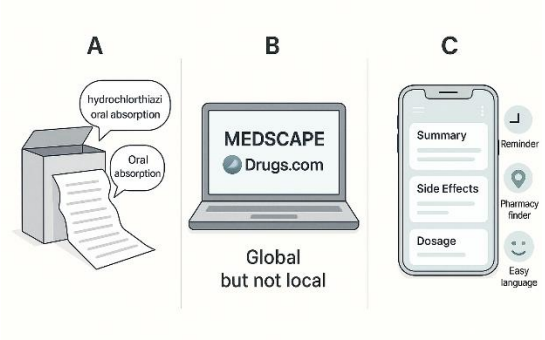


Figure 3. Challenges in package insert access in Turkey and MediFind's solution: (A) technical and complex printed package insert; (B) global digital resources lacking localization; (C) MediFind interface with simplified content, reminders and pharmacy locator.

3. Proposed Method

MediFind is developed using Kotlin on the Android SDK, leveraging the MVVM architecture and Clean Architecture principles to ensure modularity, testability, and maintainability [16, 17]. (Figure 4) The application's architecture comprises three layers:

- **Presentation Layer:** Manages user interfaces using fragments and ViewModels, with Data Binding and LiveData for real-time UI updates.
- **Domain Layer:** Handles business logic through Use Case classes and Repository interfaces, ensuring independence from UI and data sources.
- **Data Layer:** Integrates local (Room database) and remote (Firebase Firestore) data sources for efficient data management.

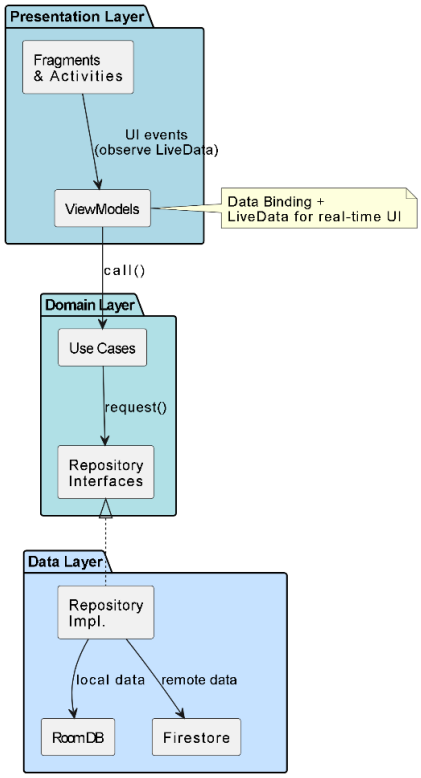


Figure 4. Flow diagram showing the architecture of the MediFind application

3.1. Technologies Used

The choice of technologies was driven by their ability to support scalability, performance, and user experience, as supported by modern software development practices [14, 16] (Figure 5):

- **Kotlin and Android SDK:** Selected for type safety and compatibility with Android platforms, ensuring robust development [16].
- **Firebase Services:** Authentication, Firestore, Storage, and Cloud Messaging provide secure user management and real-time synchronization [17].
- **Google Maps SDK:** Enables location-based pharmacy searches, enhancing accessibility [18].
- **Room Database:** Supports offline data access, ensuring reliability in low-connectivity scenarios [14].
- **Retrofit and OkHttp:** Optimize RESTful API calls for external data integration [17].
- **WorkManager:** Manages background tasks, such as reminders, with minimal battery impact [14].

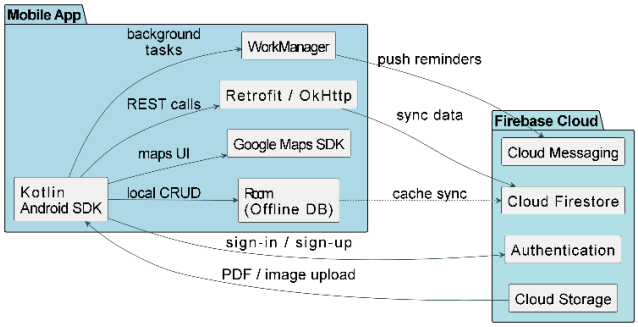


Figure 5. MediFind technology stack.

3.2. Key Features

- **Simplified Drug Information**: Drug package inserts are parsed into sections (e.g., usage, side effects) with colloquial explanations to improve comprehension, addressing issues noted in prior research [3].
- **Medication Reminder System**: Users can schedule reminders using WorkManager, supporting adherence as highlighted by mHealth studies [11].
- **Pharmacy Locator**: Integrates Google Maps SDK to list nearby pharmacies, addressing practical access barriers [20].
- **User Profile Management**: Allows personalization of notification settings and preferences, enhancing user engagement [12].
- **Security and Compliance**: Implements data encryption and adheres to GDPR and Turkey's KVKK regulations, ensuring user data privacy [21].

3.3. Data Processing

Drug data is sourced via web scraping from public repositories, pre-processed to extract fields like active ingredients and ATC codes, and stored in Firebase Firestore. Local caching via Room ensures offline functionality, aligning with best practices for mHealth applications [17] (Figure 6).

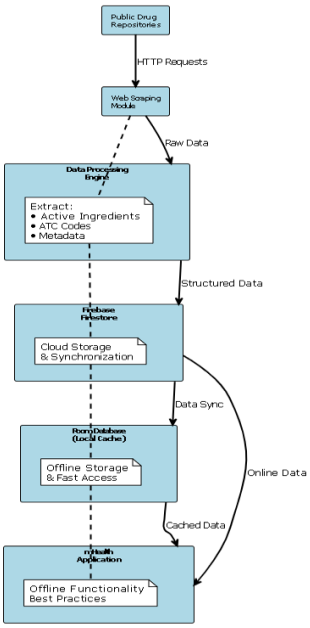


Figure 6: Data flow diagram for mHealth drug information system

4. Experimental Results

MediFind was evaluated based on performance metrics outlined in the development process:

- Application Launch Time: The application achieves a launch time of under 3 seconds on mid-range Android devices, ensuring quick access.
- Battery Consumption: Optimized to consume less than 2% battery per hour during active use, leveraging WorkManager and coroutine-based optimizations.
- Usability: The Material Design-based interface and simplified drug information are designed to enhance user comprehension, though specific user studies are pending.
- Reliability: Local caching and Firebase integration ensure consistent performance in both online and offline scenarios.

While direct user testing data is limited, the application's design aligns with findings from mHealth studies, which suggest that reminder systems and accessible information can improve adherence by up to 25% [9, 10]. The pharmacy locator feature is expected to reduce barriers to medication access, as supported by similar location-based mHealth solutions [18].

5. Conclusions

MediFind advances mHealth by providing an accessible platform for drug information and medication adherence. Its simplified information, reminder system, and pharmacy locator services address critical gaps in health literacy and access. The use of MVVM, Firebase, and Google Maps SDK ensures scalability and performance. However, the application's test coverage is not yet complete, indicating a need for further validation.

Future enhancements include multi-language support, integration with systems like e-Nabız, and user feedback mechanisms for personalization. MediFind's contributions to health literacy and digital health transformation position it as a model for localized mHealth solutions.

6. References

[1] Özkan NF, Ulutaş BH. "Üniversite öğrencilerinin prospektüs kullanma bilinci ve davranışlarının araştırılması". S.Ü. Fen Fakültesi Fen Dergisi, 44(2), 123-130, 2018.

[2] Akıcı A, Alp Fİ, Ayanoğlu Dülger G, et al. "Serbest eczanelerde ilaç sunumu sürecinde karşılaşılan sorunlar". Hacettepe Üniversitesi Eczacılık Fakültesi Dergisi, 29(1), 45-52, 2009.

[3] Wolf MS, Davis TC, Shrank W, et al. "To err is human: Patient misinterpretations of prescription drug label instructions". Patient Education and Counseling, 67(3), 293-300, 2007.

[4] Davis TC, Wolf MS, Bass PF, et al. "Literacy and misunderstanding prescription drug labels". Annals of Internal Medicine, 145(12), 887-894, 2006.

[5] Özkan NF, Ulutaş BH. "İlaç prospektüslerinin okunabilirlik analizi". Sağlık Bilimleri Dergisi, 27(3), 215-222, 2019.

[6] Medscape. "Medscape". https://www.medscape.com/ (Accessed 01.06.2025).

[7] Drugs.com. "Drugs.com". https://www.drugs.com/ (Accessed 01.06.2025).

[8] e-Nabız. "e-Nabız". https://enabiz.gov.tr/ (Accessed 01.06.2025).

[9] Marcolino MS, Oliveira JAQ, D'Agostino M, et al. "The impact of mHealth interventions: Systematic review of systematic reviews". JMIR mHealth and uHealth, 6(1), e23, 2018.

[10] Free C, Phillips G, Watson L, et al. "The effectiveness of mobile-health technologies to improve health care service delivery processes". The Lancet, 377(9765), 117-125, 2011.

[11] Dayer L, Heldenbrand S, Anderson P, et al. "Smartphone medication adherence apps: Potential benefits to patients and providers". Journal of the American Pharmacists Association, 53(2), 172-181, 2013.

[12] Boulos MNK, Brewer AC, Karimkhani C, et al. "Mobile medical and health apps: State of the art, concerns, regulatory control and certification". Online Journal of Public Health Informatics, 5(3), e229, 2014.

[13] Türk Eczacıları Birliği. "Türk Eczacıları Birliği". https://www.teb.org.tr/ (Accessed 01.06.2025).

[14] García-Sánchez S, Fernández-Alemán JL, Toval A, et al. "Mobile health apps providing information on drugs for adult emergency care: Systematic search on app stores and content analysis". JMIR mHealth and uHealth, 10(4), e29985, 2022.
[15] Hamine S, Gerth-Guyette E, Faulx D, et al. "Impact of mHealth chronic disease management on treatment adherence and patient outcomes". Journal of Medical Internet Research, 17(2), e52, 2015.
[16] Martin RC. Clean Architecture: A Craftsman's Guide to Software Structure and Design. Prentice Hall, 2017.
[17] Leiva A. Kotlin for Android Developers: Learn Kotlin the Easy Way While Developing an Android App. CreateSpace Independent Publishing Platform, 2016.
[18] Google Developers. "Android Developers". https://developer.android.com/ (Accessed 01.06.2025).
[19] Moroney L. The Definitive Guide to Firebase: Build Android Apps on Google's Mobile Platform. Apress,
[20] 2017.
[21] Google Maps Platform. "Google Maps SDK". https://developers.google.com/maps (Accessed 01.06.2025).
[22] European Union. "General Data Protection Regulation (GDPR)". https://gdpr.eu/ (Accessed 01.06.2025).

From Tradition to Future: The Economic, Ecological and Cultural Importance of Natural Dye Plants in Elazığ Province

Nevzat DOĞAN[1] , İnanç ÖZGEN [2]

[1,2] Biotechnology Department, Faculty of Bioengineering, Firat University, Elazig, Turkey.
[1]nevzatdogan2323@gmail.com, [2]inancozgen@gmail.com,
[1](ORCID: 0009-0009-1862-5125), [2](ORCID: 0000-0003-1742-9324)

Abstract

This study comprehensively investigates the natural dye plants of Elazığ province, which have historically played a significant role in the region's economic, cultural, and ecological landscape. Particular emphasis has been placed on Isatis constricta PH Davis (Hazar Indigo), an endemic species, as well as *Rubia tinctorum* L. (Madder) and *Reseda luteola* L. (Weld). The ecological resilience, agricultural cultivation practices, dyeing properties, and economic values of these species were examined through both historical records and recent field observations. The research is based on farmer-scale cultivation demonstrations and field trials carried out between 2015 and 2023 under the coordination of the Elazığ Provincial Directorate of Agriculture and Forestry. Findings revealed that the strong root systems of these plants enable them to withstand drought conditions and provide an alternative production model especially in marginal, stony, and sloping lands with low productivity. This indicates that these species can serve as strategic crops capable of adapting to the adverse impacts of global climate change. Beyond their role in textile dyeing, natural dye plants demonstrate considerable potential in biotechnology, food preservation, wood processing, cosmetics, and pharmaceutical industries. Their rich content of phenolic compounds, antimicrobial substances, and antioxidant properties highlights their value as eco-friendly resources contributing to human health and sustainable production. In particular, *Rubia tinctorum* provides the historically renowned "Turkish red" dye, carrying cultural significance, while the endemic Isatis constricta represents a strategic resource for indigo derivatives. *Reseda luteola*, in turn, stands out not only as a sustainable source of yellow dye but also as a vital nectar provider for honeybees, thus contributing to ecosystem services and apiculture. The results demonstrate that by revitalizing natural dye plant cultivation, Elazığ province could regain importance as a regional hub for ecological textiles and natural dye production. However, further research is needed on color standardization, the development of eco-friendly mordants, and the optimization of extraction techniques to ensure industrial applicability. Additionally, conservation of endemic species, establishment of local seed banks, and integration of these plants into sustainable development policies are essential steps toward securing their future use. In conclusion, this study underlines the strategic role of natural dye plants in eco-friendly production, rural development, and the preservation of cultural heritage, while offering both scientific insights and practical solutions for sustainable agriculture and industry.

Keywords: natural dyes, *Isatis constricta*, elazığ, *Rubia tinctorum*, *Reseda luteola* (*lutea)*, drought

1. Introduction

At the scale of Elazığ Province, one hundred years ago silk, cotton, and wool weaving factories and workshops had been established alongside dye houses where textile materials were colored using naturally dyed yarn, fabric dyeing, and fabric printing techniques. From these fabrics, various product designs were created. Clothing and everyday items were produced and supplied to other provinces and even abroad. In Elazığ, dye plants were either cultivated or collected from nature, and natural dyes were obtained from these plants. Silkworms were reared, silk threads were produced, and woven silk products were manufactured. Cotton was cultivated, and cotton fabrics were woven. Yarns and fabrics were dyed in dye houses, printing molds were prepared, and printing dyes were applied to fabrics. Similarly, leather was processed, natural dyes were used in leather dyeing, and shoes were manufactured.

Due to the costly nature of natural dye production and the lengthy application processes in textiles, along with the lack of standardization in dye yields, research shifted toward alternative products. The discovery of cheaper synthetic dyes and their easier application facilitated their rapid spread, leading to a complete decline in interest in natural dyes because of their costs. Over the past century, numerous studies have been conducted regarding the problems caused by the widespread use of synthetic dyes. These studies revealed that synthetic dyes are toxic, mutagenic, and carcinogenic, posing threats to human health and causing environmental pollution. The adverse impacts of synthetic dyes on both consumers and the environment reignited interest in alternative natural dyes. The advancement of scientific analytical techniques accelerated the investigation of natural dyes. According to research on natural dyes

[1]Corresponding author

and dye plants, the presence of phenolic compounds, antimicrobial substances, and UV-protective compounds in these plants has once again increased the interest in and demand for environmentally friendly plant-based dyes. However, challenges remain regarding the large-scale production and continuity of natural dyes, including the lack of standardization of colors, insufficient environmentally friendly alternatives to heavy-metal mordants, difficulties in adapting application techniques in the textile industry, and the need to establish extraction methods specific to each dye plant.

On April 4, 1893, Milo Jewett, the United States Consul in Sivas, prepared a report requested by the State Department on cotton production and consumption. An additional report sent from Harput on June 30 was also forwarded. Although precise statistics could not be obtained, the report indicated that cotton was not produced in the Sivas region, and about 10,000 pounds of cotton were imported from other regions. However, cotton was produced in Harput. According to data from the past two years, 145,000 pounds (63,502 kg) of cotton had been produced. Three-quarters of this cotton was consumed locally within the province, while the remaining one-quarter was exported to other regions (Gül, 2015). In 1927, cotton production in the Harput region reached 700 tons (Aksoy & Bülent, 2019). According to the Ottoman Empire's Chicago World Fair Exhibition Report in 1893, Individual participants organized a praiseworthy exhibition; among the most notable were the companies of Habif and Polako and P. De Andria from Izmir, Muradhanyan from Harput, and Sadullah Souhami from Istanbul (Pushman, 1893; Memiş, 2025).

In the 19th and 20th centuries, Harput was an important trade center in sericulture and weaving (Aksoy & Bülent, 2019). The advanced level of sericulture in Harput fostered the development of handwoven fabrics. According to the 1892–1893 Salname (Yearbook), Harput city center hosted one silk factory, one tannery, 850 shops, 20 inns, 12 oil mills, and many carpet looms (Yapıcı, 2009; Aksoy & Bülent, 2019). Krikor İpekçiyan settled in Mezire (Mamuretülaziz, today's Elazığ) to establish a silk weaving workshop. The silk garments produced in İpekçiyan's factory bore the label "Fabricator." In 1889, he expanded his factory, importing machines from Europe and America, and his silk textiles won awards at international competitions (Aksoy & Bülent, 2019).

In the 19th century, the so-called "Syriac Cotton Prints" (Süryani Çiti) produced in Harput gained fame for their high quality and beauty. Summer uniforms for the Ottoman army were produced from these textiles (Aharonyan, 1965; Aksoy & Bülent, 2019). The five brothers known by the surname "Fabrikatoryan" sold their products within the Ottoman Empire and even exported them to Russia and Europe. Depending on the season, between 200 and 300 workers were employed in their factory (Der Matossian, 2007; Aksoy & Bülent, 2019).

By 1907, approximately 1,200 looms were in use in the two towns of the region (İnalcık, 1994). In the early 20th century, Harput possessed nearly 5,000 looms dedicated to the production of "Manusa" cloth (Akpınarlı & Çatalkaya; Aksoy & Bülent, 2019).

By the late 1890s, producers in Harput were annually exporting an average of 150,000 silk and cotton garments. This production continued to increase throughout the 1890s, reaching nearly ten times the levels of the 1830s. In 1885, 125,000 silk-cotton garments were transported from Harput to other provinces, and in 1886 this figure rose to 172,000. Between 1884 and 1910, Harput not only produced cotton and silk yarns but also manufactured various handwoven fabrics such as Şîb, Muslin, Çitari, Manusa, Kutnu, Bezayağı, and Satin (Aksoy & Bülent, 2019).

2. Important Dye Plants in Elazığ Province

2.1. The Endemic Species of Elazığ: *Isatis constricta* PH Davis (Hazar Indigo)

This plant is a biennial dye species, and its cultivation has been carried out under farmer conditions. In the first year, both leaf yields and dye yields were determined. The specific characteristics of the plant as well as its ecological requirements in natural habitats were examined. The first-year leaves of *Isatis constricta* and its appearance under field conditions are presented in Figure 1.



Figure 1. Hazar Indigo (*Isatis constricta*) cultivated under farmer conditions and ıts first-year leaf

The blue dye compounds of Hazar Indigo (Isatis constricta) are obtained by processing its leaves, which can be harvested starting from June.



Figure 2. Hazar Indigo (*Isatis constricta*) Second-Year Flowering and Seed

Hazar Indigo (*Isatis constricta*) exhibits drought tolerance with its taproot system in deep soils. As shown in Figure 3, all roots extending into the soil consist of taproots. Various shades of blue dye are obtained from Hazar Indigo (Figure 3).



x

Table 1. Characteristic Features of Hazar Indigo (*Isatis constricta*), an Endemic Plant of Elazığ Province

Attributes	Characteristics
Color	Blue and its shades
Plant parts used for dyeing	Leaves
Propagation	Seed, seedling
Soil requirement	Deep and well-drained
Planting distance	40 cm × 20 cm
Plant lifespan	Perennial (2 years)
Flowering period	Second year
Water requirement	Does not prefer much water
Drought tolerance	Very high
Sunlight requirement	Present
Seed yield	High
Dye compounds	Isatin, Indirubin, Indican
Leaf yield per decare (kg/da)	650–1250 kg/da
Dye yield (per 1 kg fresh leaf)	3.5–6 g
Harvest period	May, June, July

2.2. Rubia tinctorum (Madder Plant)

It is a perennial dye plant and is widely distributed across Elazığ Province. It has a spreading growth habit. Known worldwide as "Turkish Red," this plant has gained great fame. When propagated from seeds, harvesting can be carried out after three years. The underground organs (rhizomes) are used for dyeing purposes, while the above-ground parts exhibit spreading and climbing characteristics. It is quite common for the madder plant to climb onto a nearby support (Figure 4).



Figure 4. Spreading and Climbing Appearance of *Rubia tinctorum* (Madder Plant)

Madder plants grown from seeds reach harvest maturity at the end of the third year. The underground organs and dyed yarns are shown in Figure 5.



Figure 5. Underground Organs of *Rubia tinctorum* (Madder) Used in Dyeing, and the Colors Obtained in Yarn and Wool Dyeing

Table 2. Root Dye (Rubia tinctorum) characteristic properties.

Attributes	**Characteristics**
Color	Red, orange and their shades
Plant parts used for dyeing	Rhizomes
Propagation	Seed, seedling, rhizome
Soil requirement	Deep and well-drained
Planting distance	40 cm × 20 cm
Plant lifespan	Perennial
Flowering period	Every year
Water requirement	Present
Drought tolerance	Partial
Sunlight requirement	Present
Seed yield	High
Dye compounds	Alizarin, purpurin
Leaf yield per decare (kg/da)	300–850 kg/da
Dye yield (per 1 kg fresh leaf)	Autumn

2.3. Reseda luteola (lutea) (Weld or Dyer's Weed)

It is naturally widespread throughout Elazığ Province. This perennial dye plant provides a yellow color and serves as a source of pollen and nectar for honeybees (Figure 6). Although it is drought-tolerant, irrigation can be applied as a factor to increase yield. The above-ground parts are used for dyeing. Harvesting is carried out green at the beginning of flowering and then dried.



Figure 6. General Appearance of the Dye Plant *Reseda luteola* (Weld)

Weld (*Reseda luteola*) has a long flowering period and provides a feeding source for honeybees and other insects. Being drought-tolerant, the plant carries strategic importance. Its prolonged flowering makes it a valuable resource for honeybees. The plant is harvested during the flowering period, dried in the shade, and produces a yellow dye color (Figure 7).



Figure 7. Dried form of *reseda luteola* (Weld) for dyeing purposes and ıts yellow dye color

The characteristic features of *Reseda luteola* are presented in Table 3.

Table 3: Characteristic features of weld (*Reseda luteola*)

Attributes	Characteristics
Color	Yellow
Plant parts used for dyeing	Above-ground organs (leaves, flowers, and stems)
Propagation	Seed, seedling
Soil requirement	Not selective
Planting distance	40 cm × 20 cm
Plant lifespan	Perennial
Flowering period	Same year
Water requirement	Increases yield
Drought tolerance	Very high
Sunlight requirement	Present
Seed yield	Very high
Dye compounds	Luteolin
Leaf yield per decare (kg/da)	160–447 kg/da
Dye yield (per 1 kg fresh leaf)	At 50% flowering stage

In Elazığ Province, there are naturally growing dye plants that were historically used for dyeing and hold economic importance, and detailed information about these plants will be provided in other studies.

3. Conclusion

This study revealed the ecological, agricultural, and cultural potential of natural dye plants grown in Elazığ Province. Among them, the endemic species *Isatis constricta* PH Davis (Hazar Indigo), as well as *Rubia tinctorum* L. (Madder) and Reseda luteola L. (Weld), stand out as strategic plants of the region with both historical and current uses. Field studies and farmer-scale demonstrations showed that these species exhibit high drought tolerance and offer a value-added production model, particularly for the utilization of marginal, stony, and sloping lands.

The findings indicate that dye plants can contribute to rural development in multiple ways. In particular, the long flowering period of *Reseda luteola* provides a sustainable nectar source for honeybees, enabling integration with apiculture. This characteristic supports both biodiversity conservation and additional income for farmers. Similarly, the historical importance of the "Turkish Red" dye obtained from the roots of *Rubia tinctorum* positions it as a cultural heritage resource that can be re-evaluated in eco-friendly textile production today. On the other hand, Isatis constricta, due to its endemic status and high indigo yield, should be strategically conserved and promoted.

Natural dye plants are not only important for the textile industry but also hold significant potential for biotechnological and pharmacological research. Their content of phenolic compounds, antimicrobial substances, and antioxidants suggests possible applications in food preservation, wood processing, cosmetics, and pharmaceutical industries. The presence of anticancer compounds further highlights their priority for medical research.

However, it should be noted that several challenges remain for large-scale production of natural dyes. Color standardization, development of eco-friendly mordants, and optimization of species-specific extraction methods should be prioritized in future studies. In this context, this study on the natural dye plants of Elazığ serves as a pioneering model that may contribute to the development of sustainable and environmentally friendly dye production both in Turkey and internationally.

In conclusion, the cultivation of natural dye plants in Elazığ holds great significance for developing alternative crops adaptable to climate change, supporting rural development, and promoting eco-friendly production processes. The conservation, expansion, and in-depth scientific investigation of these species could transform the region once again into a "center of ecological textiles and natural dye production."

4. Recommendations

Expansion of Cultivation: Demonstrations under farmer conditions should be increased for dye plants such as Isatis constricta, Rubia tinctorum, and Reseda luteola, and cultivation areas should be expanded. In particular, production should be encouraged on marginal, stony, and sloping lands in order to enhance their economic value.

Integration with Beekeeping: Species with a long flowering period, such as Reseda luteola, should be integrated with apicultural activities. In this way, farmers can obtain additional income while biodiversity is conserved.

Conservation of Endemic Species: The natural populations of Elazığ endemics such as Isatis constricta should be preserved, and local seed banks should be established. Both in-situ and ex-situ conservation practices should be supported to prevent the loss of genetic diversity.

Standardization and Technological Development: To increase the industrial use of natural dyes, color standardization should be ensured and eco-friendly mordants should be developed. R&D projects should be supported to determine species-specific extraction and dyeing methods.

Biotechnological and Medical Research: Research should be conducted on the use of phenolic, antioxidant, and antimicrobial compounds present in dye plants for biomedical, food preservation, and cosmetic applications. In particular, compounds with anticancer potential should be evaluated in clinical and pharmaceutical studies.

Sustainable Development and Policy Support: Cooperation between the Ministry of Agriculture and Forestry and universities should be established to include natural dye plants in agricultural support programs. The historical weaving and dyeing culture of Elazığ should be revived, and eco-tourism together with natural product marketing that contribute to the local economy should be supported.

5. References

[1] Akpınarlı, F., & Çatalkaya, G. (2016). Harput ve çevresinde dokumacılık geleneği. Journal of History Studies, 8(3), 1125–1136.
[2] Aksoy, İ., & Bülent, K. (2019). Harput'ta pamuk ve ipek üretiminin tarihsel gelişimi. Türk Kültürü ve Hacı Bektaş Veli Araştırma Dergisi, 91, 57–66.
[3] Aharonyan, A. (1965). Armenian weaving traditions and textile art. Yerevan: Academy of Sciences Press.
[4] Der Matossian, B. (2007). The Armenian silk industry in the Ottoman Empire. Journal of Ottoman Studies, 30, 145–168.
[5] Gül, İ. (2015). Osmanlı'da pamuk üretimi ve Harput örneği. History Research Journal, 4 (2), 160–170.
[6] İnalcık, H., & Quataert, D. (1994). An Economic and Social History of the Ottoman Empire, 1300–1914. Cambridge: Cambridge University Press.
[7] Memiş, A. (2025). Osmanlı tekstil ürünlerinin uluslararası sergilerde tanıtımı. Osmanlı Araştırmaları Dergisi, 42, 55–62.
[8] Pushman, J. (1893). Report of the Chicago World Fair: Ottoman Exhibitions. Chicago Exhibition Commission.
[9] Yapıcı, Ş. (2009). Harput'ta ipekböcekçiliği ve dokuma sanayii. Fırat Üniversitesi Sosyal Bilimler Dergisi, 19(2), 90–98.

Fuzzy Logic Controlled Regenerative Braking for Electric Vehicles: Energy Recovery and Driving Safety

Gamze UYGUR[1], Arif GÜLTEN[2]

[1,2]Department of Electrical Electronics Engineering Faculty of Engineering, Fırat University Elazığ, Turkey
[1]222113124@firat.edu.tr, [2]agulten@firat.edu.tr
[1](ORCID: 0009-0001-8642-436X), [2](ORCID:0000-0002-9652-2625)

Abstract
Electric vehicles play a significant role in enhancing energy efficiency and ensuring a sustainable future. One of the most important factors in the development and widespread adoption of these vehicles is the ability to effectively manage the battery state of charge. Regenerative braking supports battery management in electric vehicles and provides an effective solution for improving energy efficiency. In this study, a fuzzy logic-based control system is designed to increase battery charging efficiency and ensure driving safety during the regenerative braking process in electric vehicles. In the proposed system, the level of regenerative braking force is determined by evaluating input parameters such as battery state of charge, vehicle speed, brake pedal position, and slip risk. The input parameters were generated using real vehicle speed data obtained from the WLTP driving cycle. The battery state of charge was modeled to decrease depending on the vehicle speed, and other parameters such as brake pedal position and slip risk were also modeled to vary according to vehicle speed. With this developed control strategy, it is aimed to ensure regenerative braking safety under different driving conditions while maximizing energy recovery.

Keywords: electric vehicles, regenerative braking, fuzzy logic

1. Introduction

Due to increasing energy demand, environmental concerns, and the depletion of fossil fuels, interest in electric vehicles (EVs) has been steadily growing. Compared to conventional internal combustion engine vehicles, electric vehicles offer numerous advantages, such as zero emissions and lower operating costs. Initially, these vehicles incorporated systems that included both internal combustion engines and electric motors. However, with advancements in battery technology, battery electric vehicles (BEVs), which rely solely on electric motors, have been developed. Despite these advancements, the widespread adoption of electric vehicles still faces challenges such as limited driving range, prolonged charging times, and insufficient charging infrastructure. Nevertheless, ongoing technological developments are progressively providing solutions to these issues [1-2].

One of the most critical components of electric vehicles is the battery. Batteries are responsible for energy storage and supplying power to the propulsion system. The performance and efficiency of electric vehicles are directly influenced by the battery system. Additionally, the battery plays a key role in determining parameters such as the state of charge (SOC), driving range, and energy consumption. For these reasons, controlling the battery's state of charge is of great importance for maintaining battery health, ensuring vehicle performance, and optimizing driving range [3].
One of the most commonly used methods in managing the battery state of charge (SOC) is regenerative braking. The fundamental principle of regenerative braking systems is to recover the kinetic energy that is typically lost during vehicle deceleration. In electric vehicles, this recovered energy is converted into electrical energy by operating the electric motor as a generator. The generated electrical energy is then stored in the battery packs through appropriate power electronic components. This system serves as an effective solution for monitoring the SOC and improving energy efficiency in electric vehicles. In regenerative braking systems, multiple parameters such as battery SOC, vehicle speed, brake pedal position, and slip risk must be evaluated simultaneously. Considering these parameters collectively ensures both energy recovery and driving safety. However, managing these factors together using conventional control methods is highly complex and challenging [4],[7].

In recent studies, regenerative braking has been widely implemented in electric vehicles. In the literature, classical control methods such as PID-based control systems, model predictive control (MPC), and nonlinear control techniques have been

[1]Corresponding author

employed. However, these methods are generally not preferred for real-time applications due to their reliance on complex mathematical models and computations. Instead, fuzzy logic-based control systems are favored because they offer advantages such as the ability to handle uncertain conditions, flexibility, and the capability to evaluate multiple parameters simultaneously. For these reasons, the use of fuzzy logic control methods is highly suitable for regenerative braking applications in electric vehicles.

In recent years, numerous studies have focused on the application of fuzzy logic in regenerative braking for electric vehicles. Yin and his team optimized the regenerative braking force using a fuzzy logic control system under various driving cycles, significantly increasing energy recovery. In this study, the inputs to the fuzzy logic system were braking force, vehicle speed, and battery state of charge (SOC), while the output parameter was defined as the regenerative braking coefficient [8]. Anh and his team designed a fuzzy logic control system compliant with ECE R13 regulations, achieving energy recovery rates ranging from 13% to 30%. The inputs of this fuzzy logic control system were braking intensity, the vehicle's operational speed range, and the battery state of charge (SOC), while the output was the regenerative proportionality coefficient [7]. Wanmin Li and his team designed a three-input, single-output fuzzy controller for the distribution of hydraulic and electric braking forces based on parameters such as braking intensity, vehicle speed, and battery SOC. The system, developed using fuzzy neural networks, was tested under NEDC and FTP-75 driving conditions, achieving energy recovery efficiency [9]. Tolga Kocakulak and Hamit Solmaz developed a regenerative braking model using a fuzzy logic control strategy. The fuzzy logic control system employed four inputs: vehicle speed, battery state of charge (SOC), brake pedal position, and battery current. The output was designed to ensure maximum braking throughout the vehicle's operation while prioritizing driving safety [10]. Himani Jain and her colleagues designed two different fuzzy logic systems in their study. One system was developed for implementing regenerative braking, while the other was designed to apply the anti-lock braking system [11]. Bhurse and Bhole examined the contribution of the regenerative braking system to energy recovery in electric vehicles [12]. Gulhane and his team addressed a dual control strategy for regenerative braking in electric vehicles and designed a model in the MATLAB/Simulink environment to evaluate the practical effects of this technique. A cascaded bidirectional DC-DC converter was integrated into the designed model to optimize energy recovery during braking. Evaluation of the simulation results revealed improvements in both energy efficiency and braking performance [13]. In her study, Szumska conducted a detailed analysis of regenerative braking systems for electric vehicles and compared the control strategies employed. The analysis indicated that energy recovery varied between 8% and 25%, depending on the method used. The study also emphasized the importance of considering factors such as battery state of charge and driving safety alongside energy efficiency when applying these methods [14]. Teasdale and his team designed a supercapacitor-based energy recovery system for regenerative braking in electric vehicles. In this system, DC motors and supercapacitors were integrated to develop both an experimental prototype and a MATLAB/Simulink-based model. The results of the study demonstrated that, due to the high power density of supercapacitors, they reduced the load on the battery during sudden load variations and enhanced energy recovery during braking [15]. Gunjal and his colleagues conducted a comparison between systems with and without regenerative braking implemented in electric vehicles. The results indicated that systems employing regenerative braking provided significant benefits in terms of energy savings and extended driving range [16].

In this study, a fuzzy logic control system was developed to implement regenerative braking in electric vehicles, aiming both to increase energy recovery and to take driving safety into account. In the designed fuzzy logic control system, the input parameters are battery state of charge (SOC), vehicle speed, brake pedal position, and slip risk. The system output reflects the change in regenerative braking based on these input parameters. Among these inputs, vehicle speed was modeled according to real speed profiles obtained from WLTP driving cycle data, while the other inputs were modeled in harmony with this cycle. The proposed control system provides an adaptive control approach that maximizes regenerative braking under different driving conditions while prioritizing driving safety.

2. Battery Electric Vehicles

Battery electric vehicles (BEVs) have been developed as an alternative to conventional internal combustion engine vehicles. Initially, BEVs received limited attention due to fundamental issues such as short driving range and high battery costs. However, over time, the environmental damage caused by traditional internal combustion vehicles and the depletion of fuel reserves have led to increased interest in electric vehicles. In parallel with this growing interest, the number of studies conducted in this field has risen significantly. These studies have contributed to substantial advancements in battery

technologies, leading to the development of solutions to address issues such as battery lifespan and limited driving range. As a result of these developments, the adoption and usage of battery electric vehicles have been steadily increasing [17-18].

Unlike internal combustion engine vehicles, in battery electric vehicles the energy storage element and the energy required for the propulsion system are supplied through battery packs. Lithium-ion batteries are commonly used in electric vehicles due to their various advantages such as high energy density and long lifespan. The operating principle of battery electric vehicles is based on the conversion of chemical energy stored in the battery packs into electrical energy. These battery packs provide the electrical energy necessary for vehicle movement. The obtained electrical energy is converted to appropriate current and voltage levels through power electronic components and then delivered to the electric motor. The electric motor converts electrical energy into mechanical energy, which is transmitted to the wheels via the drivetrain. Compared to internal combustion engine vehicles, this system in battery electric vehicles contains fewer moving parts, thereby requiring lower maintenance costs [17].

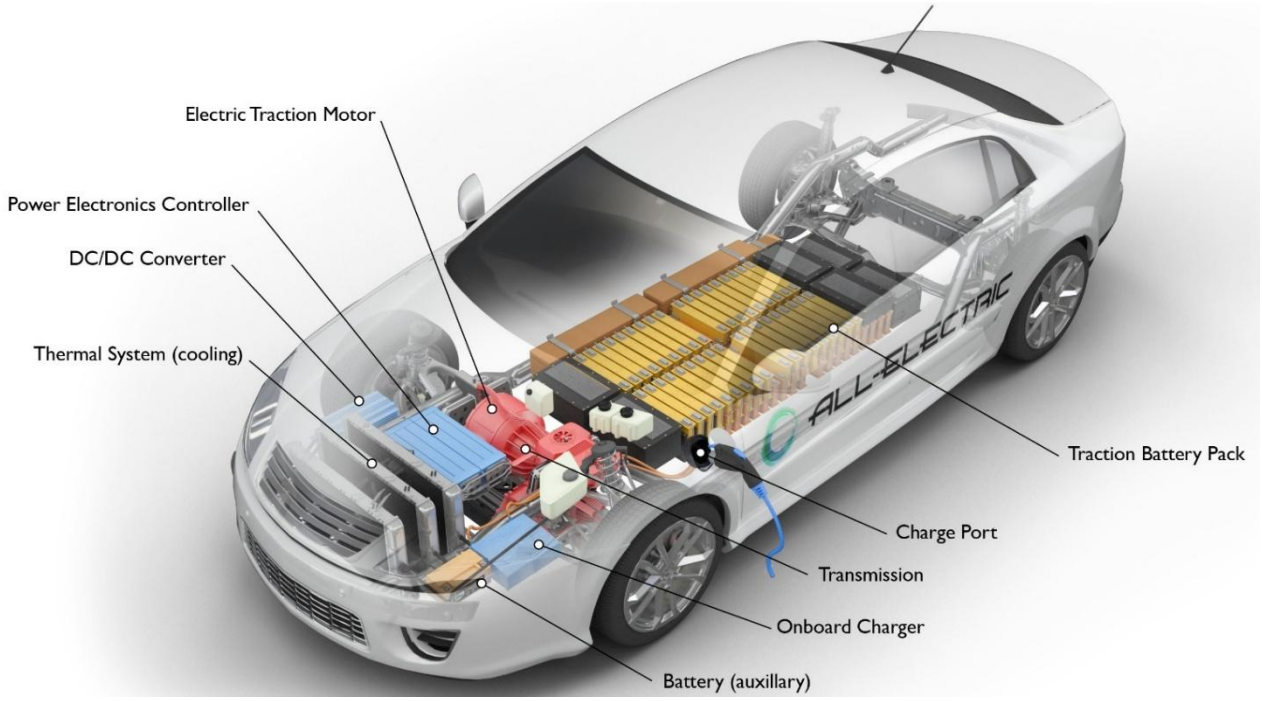


Figure 1. Structure of battery electric vehicle[19]

One of the most important components in battery electric vehicles is the battery. The performance of the battery system is crucial for the vehicle's energy management. The state of charge (SOC) is a parameter that indicates the battery's instantaneous energy level. This parameter is one of the key factors in determining driving range, ensuring efficient energy management, and maintaining the battery's proper operation. In this context, the implementation of regenerative braking systems, which improve efficiency by enabling energy recovery in vehicles, is of critical importance.
Regenerative braking enables the recovery of kinetic energy that is expended during braking in electric vehicles. Under normal conditions, this kinetic energy is dissipated as heat; however, through regenerative braking, it is recovered and converted into electrical energy by operating the electric motor as a generator, and this energy is stored in the battery system. The energy stored in the battery can then be used for recharging when needed. Energy recovery through regenerative braking is a highly effective method for controlling the battery's state of charge (SOC), and this approach not only improves battery efficiency but also contributes to energy savings [5-6].

3. Materials and Methods

In this study, a fuzzy logic controller was selected as the control method to ensure energy recovery through the regenerative braking system while prioritizing driving safety. The reason for using the fuzzy logic approach is its advantage in evaluating multiple input variables simultaneously and determining the system response based on a logical rule base [1]. For electric vehicles to operate efficiently and deliver good performance, certain components must be continuously monitored. The performance of electric vehicles is directly related to battery parameters. Therefore, monitoring the condition and performance of the battery is essential for overall vehicle performance and energy optimization. The fundamental principle of fuzzy logic is to address complex and uncertain systems. This method operates with greater precision compared to conventional control approaches. For this reason, the fuzzy logic control method is preferred for more accurate monitoring of battery parameters [20].

The fuzzy logic control system developed in this study was designed using the Fuzzy Logic Toolbox interface in MATLAB. The designed control system was implemented in the MATLAB/Simulink environment to construct a comprehensive fuzzy logic model, and the corresponding simulations were performed. Upon completion of the simulations, the regenerative braking ratio was analyzed to assess the performance of the system. Figure 2 illustrates the block diagram of the overall fuzzy logic control system.

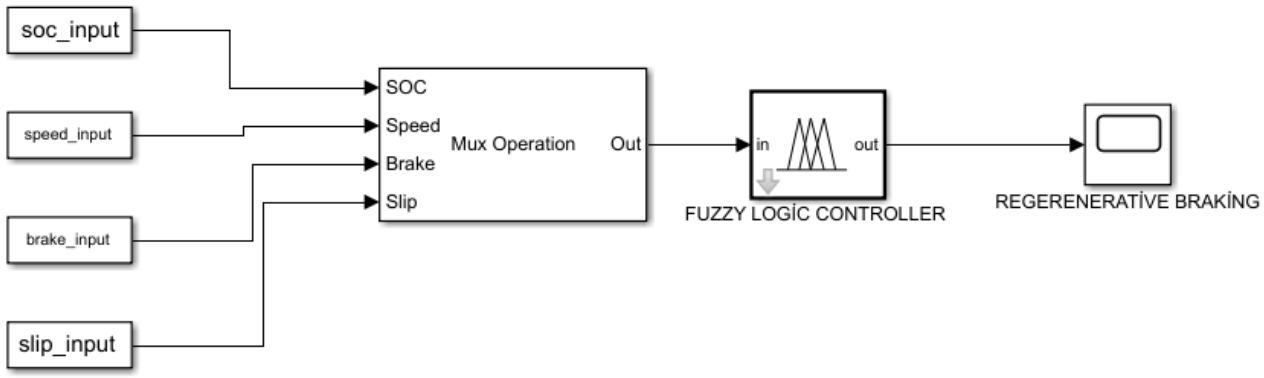
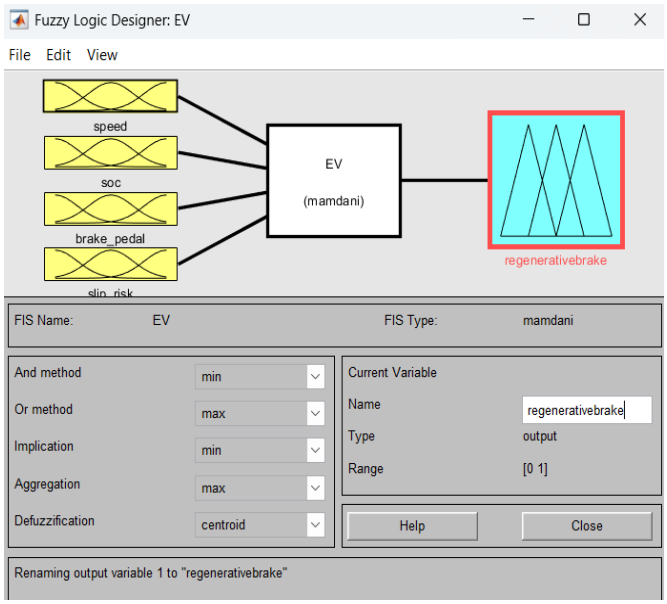
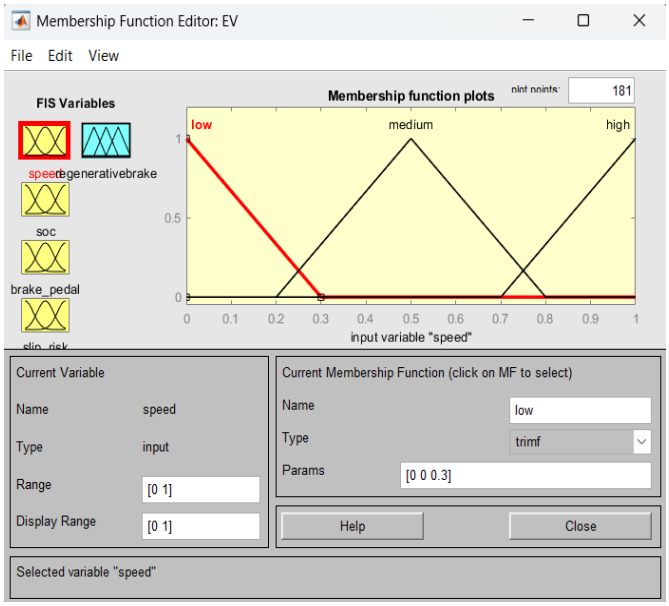


Figure 2. Simulation model of the fuzzy logic control system

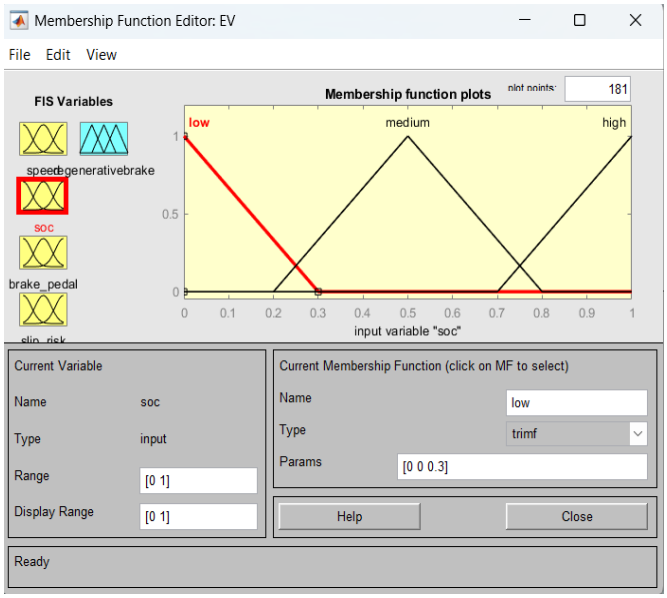
The fuzzy logic system designed in this study consists of four inputs and one output. The input parameters are battery state of charge (SOC), vehicle speed, brake pedal position, and slip risk, while the output parameter is the variation in regenerative braking. The input and output parameters used in the control system normalized within the range of 0 to 1. Among these parameters, vehicle speed, SOC, slip risk, and regenerative braking were defined at three levels: low, medium, and high, whereas brake pedal position was defined as soft, medium, and hard. In the designed control system, triangular membership functions were selected for the input and output parameters, and the Mamdani-type fuzzy inference method was utilized for the fuzzy inference process [4]. Figure 3 illustrates the membership functions of the fuzzy logic system.


Fuzzy Logic Designer: EV
File Edit View
speed
soc
brake_pedal
slip_risk
EV
(mamdani)
regenerativebrake
FIS Name: EV
FIS Type: mamdani
And method min
Or method max
Implication min
Aggregation max
Defuzzification centroid
Current Variable
Name regenerativebrake
Type output
Range [0 1]
Help
Close
Renaming output variable 1 to "regenerativebrake"

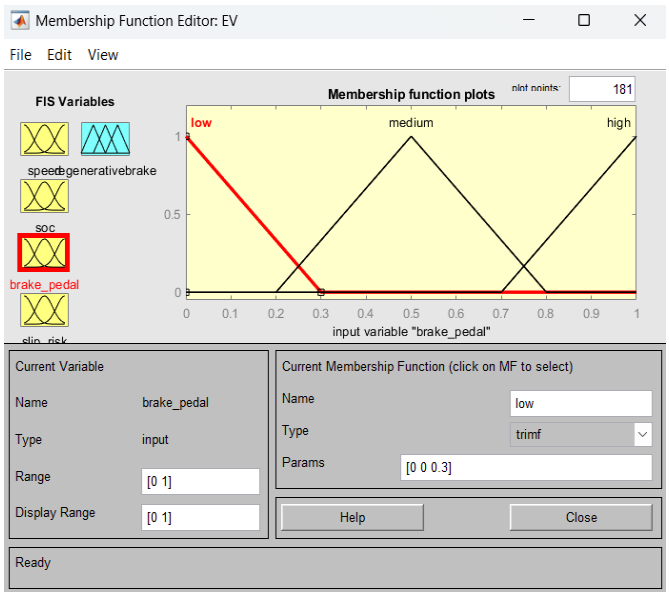
(a)


Membership Function Editor: EV
File Edit View
FIS Variables
Membership function plots
plot points: 181
low
medium
high
input variable "speed"
Current Variable
Name speed
Type input
Range [0 1]
Display Range [0 1]
Current Membership Function (click on MF to select)
Name low
Type trimf
Params [0 0 0.3]
Help
Close
Selected variable "speed"

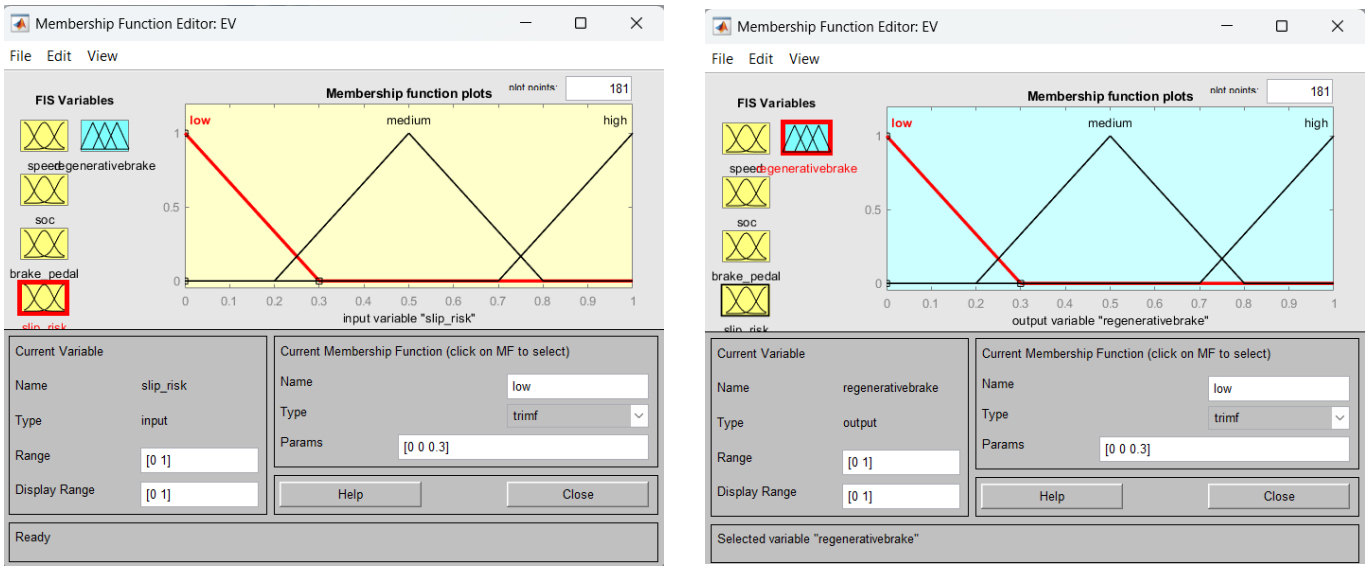
(b)


Membership Function Editor: EV
File Edit View
FIS Variables
Membership function plots
plot points: 181
low
medium
high
input variable "soc"
Current Variable
Name soc
Type input
Range [0 1]
Display Range [0 1]
Current Membership Function (click on MF to select)
Name low
Type trimf
Params [0 0 0.3]
Help
Close
Ready

(c)


Membership Function Editor: EV
File Edit View
FIS Variables
Membership function plots
plot points: 181
low
medium
high
input variable "brake_pedal"
Current Variable
Name brake_pedal
Type input
Range [0 1]
Display Range [0 1]
Current Membership Function (click on MF to select)
Name low
Type trimf
Params [0 0 0.3]
Help
Close
Ready

(d)



(e) (f)

Figure 3. a) General representation of the designed fuzzy logic control system model, b) Vehicle speed membership function, c) Battery state of charge membership function, d) Brake pedal position membership function, e) Slip risk membership function, f) Regenerative braking output membership function

For one of the input variables of the fuzzy logic controller, vehicle speed, data from the WLTP (Worldwide Harmonized Light Vehicles Test Procedure) driving cycle was utilized. Another input used in the model, the battery state of charge (SOC), was simulated with a decreasing slope corresponding to the vehicle's acceleration over time. This approach was adopted to ensure that energy consumption reflects real driving conditions. The input variables brake pedal position and slip risk were also modeled in accordance with varying vehicle speed scenarios [4].

The fuzzy logic control system used in this study comprises 81 if–then rules, accounting for all possible combinations of the input parameters. The system is designed to increase regenerative braking when the battery state of charge (SOC) is at a critical level, the brake pedal is pressed firmly, and the slip risk is low. The brake pedal position reflects the driver's braking intent and, when evaluated in conjunction with other input parameters, influences the system output. Table 1 presents a condensed subset of 12 representative rules that reflects the overall logic of the fuzzy control system and includes various input conditions.

Table 4. Rules

Rules	Speed	SOC	Brake pedal	Slip risk	Regenerative braking
1	Low	Low	Soft	Low	Medium
2	Low	Low	Hard	High	Low
3	Low	High	Medium	Medium	Low
4	Medium	Low	Hard	Low	High
5	Medium	Medium	Soft	High	Low
6	Medium	Medium	Hard	Low	High
7	Medium	High	Medium	High	Low
8	High	Low	Soft	Medium	Medium
9	High	Low	Hard	High	Low
10	High	Medium	Hard	Medium	Medium
11	High	High	Medium	Low	Medium
12	High	High	Hard	High	Low

The slip risk parameter was incorporated into the control system to ensure safe energy recovery. According to the rules, when the slip risk is high, the system prioritizes safety by maintaining the regenerative braking level at an appropriate value. This parameter is particularly important in scenarios involving high speed and hard braking. Based on all established rules, the system aims to maximize regenerative braking as much as possible when the battery state of charge (SOC) is low, slip risk is minimal, and braking is applied firmly. In situations involving risk, the regenerative braking level is maintained at an appropriate value to ensure safe and adaptive driving.

According to the simulation data, the regenerative braking ratio exhibits a normalized variation fluctuating between 0.5 and 0.9. This fluctuation results from evaluating different combinations of the four input parameters. The fuzzy logic system designed for regenerative braking demonstrates adaptive behavior consistent with real driving conditions. Due to the varying membership degrees of the input data, intermediate output values were obtained using the weighted average method. Additionally, the system quickly adapted and responded to sudden changes. The fact that the output values did not remain fixed at the extreme points highlights the advantage of the system's greater flexibility compared to conventional control methods. Consequently, the system exhibits behavior well-suited to real driving conditions. The system output is illustrated in Figure 4.

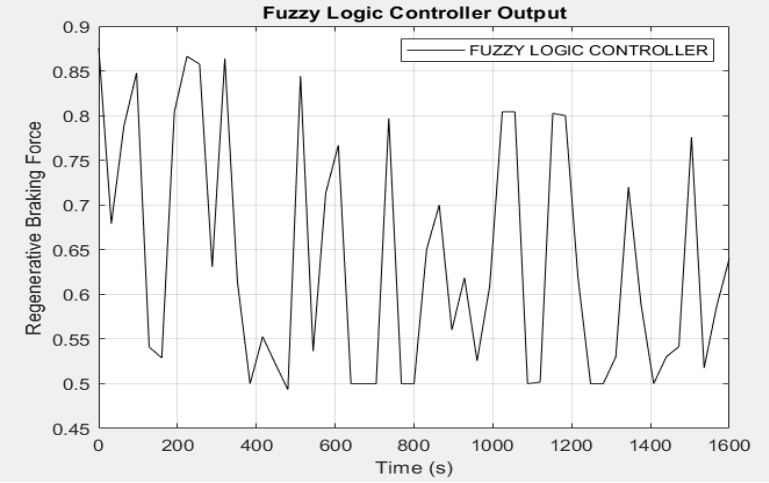


Figure 4. The output of the fuzzy controller constructed according to the rule base

According to the simulation data, the regenerative braking ratio shows a normalized variation fluctuating between 0.5 and 0.9. This fluctuation is the result of evaluating different combinations of the four input parameters. The fuzzy logic system designed for regenerative braking exhibits adaptive behavior that aligns with real driving conditions. Due to the varying membership degrees of the input data, intermediate output values were generated using the weighted average method. Furthermore, the system quickly adapted and responded to sudden changes. The fact that the values did not remain fixed at the extremes highlights the advantage of the system's greater flexibility compared to conventional control methods. Thus, the system demonstrates behavior that is well-suited to real driving conditions.

4. Result

Within the scope of this study, a fuzzy logic-based regenerative braking control system was designed with the aim of optimizing the regenerative braking ratio under appropriate conditions in electric vehicles, thereby improving energy efficiency while ensuring driving safety. Evaluation of the simulation results indicates that the system exhibits dynamic variations in the regenerative braking ratio at the output, depending on different driving conditions and vehicle parameters, thus adapting to real-world driving scenarios. Furthermore, the continuous optimization of the output demonstrates that the system exhibits adaptive behavior, successfully achieving the intended objective.

5. References

[1] M. S. Çetin, H. Güler, ve M. T. Gençoğlu, "Fuzzy Logic Based Battery Control System Design for Electric Vehicles", içinde *2021 Innovations in Intelligent Systems and Applications Conference (ASYU)*, Eki. 2021, ss. 1-4. doi: 10.1109/ASYU52992.2021.9598978.

[2] J. A. Sanguesa, V. Torres-Sanz, P. Garrido, F. J. Martinez, ve J. M. Marquez-Barja, "A Review on Electric Vehicles: Technologies and Challenges", *Smart Cities*, c. 4, sy 1, Art. sy 1, Mar. 2021, doi: 10.3390/smartcities4010022.

[3] M. S. Çetin, B. Karakaya, ve M. Gençoğlu, "Elektrikli Araçlar İçin Lityum İyon Bataryaların Modellenmesi", *Fırat Üniversitesi Mühendis. Bilim. Derg.*, c. 33, sy 2, Art. sy 2, Eyl. 2021, doi: 10.35234/fumbd.953296.

[4] C. Yang, T. Sun, W. Wang, Y. Li, Y. Zhang, ve M. Zha, "Regenerative braking system development and perspectives for electric vehicles: An overview", *Renew. Sustain. Energy Rev.*, c. 198, s. 114389, Tem. 2024, doi: 10.1016/j.rser.2024.114389.

[5] L. Shijie ve X. Li, "A Review of Regenerative Braking Energy Recovery Research for Pure Electric Vehicles", içinde *2024 IEEE 5th International Conference on Advanced Electrical and Energy Systems (AEES)*, Kas. 2024, ss. 468-472. doi: 10.1109/AEES63781.2024.10872447.

[6] N. A. Ambe, V. Sharma, K. Sethi, ve M. H. Fallah, "Regenerative Braking Systems in Electric Vehicles (EVs): A Review", içinde *2024 3rd Edition of IEEE Delhi Section Flagship Conference (DELCON)*, Kas. 2024, ss. 1-5. doi: 10.1109/DELCON64804.2024.10866736.

[7] N. T. Anh, C.-K. Chen, ve X. Liu, "An Efficient Regenerative Braking System for Electric Vehicles Based on a Fuzzy Control Strategy", *Vehicles*, c. 6, sy 3, Art. sy 3, Eyl. 2024, doi: 10.3390/vehicles6030071.

[8] Z. Yin, X. Ma, R. Su, Z. Huang, ve C. Zhang, "Regenerative Braking of Electric Vehicles Based on Fuzzy Control Strategy", *Processes*, c. 11, sy 10, Art. sy 10, Eki. 2023, doi: 10.3390/pr11102985.

[9] W. Li *vd.*, "Regenerative braking control strategy for pure electric vehicles based on fuzzy neural network", *Ain Shams Eng. J.*, c. 15, sy 2, s. 102430, Şub. 2024, doi: 10.1016/j.asej.2023.102430.

[10] T. Kocakulak, "Elektrikli Bir Aracın Modellenmesi ve Rejeneratif Fren Sisteminin Bula- nık Mantık Yöntemi ile Kontrol Edilmesi", 2019.

[11] H. Jain, S. Maitreya, ve P. Paliwal, "Cooperative Control of Regenerative and Anti-lock Braking Systems in Electric Vehicles Using Fuzzy Logic", içinde *2021 4th International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE)*, Eki. 2021, ss. 62-66. doi: 10.1109/RDCAPE52977.2021.9633401.

[12] S. S. Bhurse ve A. A. Bhole, "A Review of Regenerative Braking in Electric Vehicles", içinde *2018 International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC)*, Mar. 2018, ss. 363-367. doi: 10.1109/ICCPEIC.2018.8525157.

[13] A. P. Gulhane, S. Khubalkar, ve S. Dudhe, "Design of Electric Vehicle Regenerative Braking System", içinde *2024 International Conference on Advancements in Power, Communication and Intelligent Systems (APCI)*, KANNUR, India: IEEE, Haz. 2024, ss. 1-6. doi: 10.1109/apci61480.2024.10616717.
[14] E. M. Szumska, "Regenerative Braking Systems in Electric Vehicles: A Comprehensive Review of Design, Control Strategies, and Efficiency Challenges", *Energies*, c. 18, sy 10, s. 2422, May. 2025, doi: 10.3390/en18102422.
[15] A. Teasdale, L. Ishaku, C. V. Amaechi, I. Adelusi, ve A. Abdelazim, "A Study on an Energy-Regenerative Braking Model Using Supercapacitors and DC Motors", *World Electr. Veh. J.*, c. 15, sy 7, s. 326, Tem. 2024, doi: 10.3390/wevj15070326.
[16] M. Gunjal, L. Mathew, ve S. Reji, "Comparative Study of Electric Vehicles: Regenerative Braking System versus Non-Regenerative Braking Systems", içinde *2024 4th International Conference on Innovative Practices in Technology and Management (ICIPTM)*, Noida, India: IEEE, Şub. 2024, ss. 1-6. doi: 10.1109/iciptm59628.2024.10563728.
[17] M. S. Çetin, M. T. Gençoğlu, ve A. Dobrzycki, "Investigation of Charging Technologies for Electric Vehicles", *Turk. J. Sci. Technol.*, c. 19, sy 1, Art. sy 1, Mar. 2024, doi: 10.55525/tjst.1399120.
[18] M. Ntombela, K. Musasa, ve K. Moloi, "A Comprehensive Review for Battery Electric Vehicles (BEV) Drive Circuits Technology, Operations, and Challenges", *World Electr. Veh. J.*, c. 14, sy 7, Art. sy 7, Tem. 2023, doi: 10.3390/wevj14070195.
[19] "Alternative Fuels Data Center: How Do All-Electric Cars Work?" Erişim: 05 Ağustos 2025. [Çevrimiçi]. Erişim adresi: https://afdc.energy.gov/vehicles/how-do-all-electric-cars-work
[20] M. Coşkun ve B. Karakaya, "A CASE STUDY: FUZZY LOGIC BASED DECISION-MAKING SYSTEM FOR ELECTRIC VEHICLE CHARGING", *Int. J. Energy Smart Grid*, c. 9, sy 1, Art. sy 1, Ara. 2024, doi: 10.55088/ijesg.1596943.

A Deep Learning Based Application for Deepfake Video Detection

Lizge BOZKURT[1], Onur TOKEL[2], Emirhan ENİSOĞLU[3], Niyazi Furkan BAR[4], Hasan YETİŞ[5], Mehmet KARAKÖSE[6]

[1,4,5,6]Computer Engineering, Firat University, Elazig, Türkiye
[2,3]Panates Technology Investment Inc., İzmir, Türkiye
bozkurtlizge@gmail.com, onur.tokel@panates.com, e.enisoglu@panates.com, nfbar@firat.edu.tr, h.yetis@firat.edu.tr, mkarakose@firat.edu.tr
[1](ORCID: 0009-0009-4158-9046), [2](ORCID: 0000-0002-1962-0432), [3](ORCID: 0009-0007-9469-6959), [4](ORCID: 0000-0002-3393-004X), [5](ORCID: 0000-0001-7608-3293), [6](ORCID: 0000-0002-3276-3788)

Abstract
In recent years, advances in deep learning have given rise to "deepfake" technology, enabling the creation of extraordinarily realistic synthetic videos and images. This development poses a serious threat of disinformation and manipulation in domains ranging from politics to personal security. In this study, we propose a hybrid deep learning model that concurrently analyzes spatial and temporal cues for the automated detection of video-based deepfakes. Our approach employs a ResNet-18 backbone to extract visual features from individual frames and integrates a Long Short-Term Memory (LSTM) network to capture temporal inconsistencies. Experimental evaluations demonstrate that the model achieves 97% accuracy, effectively distinguishing between real and manipulated content. To facilitate real-world use, we have wrapped the detection engine in a user-friendly web interface, thereby contributing to digital media literacy and supporting efforts to combat malicious deepfake content.

Keywords: deepfake, deep learning, ResNet, LSTM, video analysis, deepfake detection

1. Introduction

Recent advances in AI-driven media generation have increasingly blurred the boundaries between real and fake content. In this context, synthetic videos referred to as deepfakes are commonly produced using techniques such as face swapping, voice cloning, and facial expression mapping, resulting in highly realistic yet fabricated visuals. Deepfake content poses severe socio-political risks, including political manipulation, financial fraud, and reputational damage through disinformation, ultimately undermining public decision-making processes. This project proposes a deep learning–based model capable of automatically classifying video content as real or fake. In addition to serving as a technological verification tool, the system aims to curb the spread of disinformation and promote digital media literacy. Competitions like the Deepfake Detection Challenge, organized with the support of major tech companies such as Google, Microsoft, and Facebook, as well as leading academic institutions, have accelerated research efforts in this field.

Deep learning architectures such as Autoencoders and Generative Adversarial Networks (GANs) play a key role in the generation of deepfake content [1]. While GANs are based on the adversarial training of two neural networks, autoencoders offer a more traditional yet effective approach. Autoencoders were particularly prevalent in early deepfake systems such as FaceSwap. Due to their low computational cost and structural simplicity, autoencoders significantly contributed to the widespread proliferation of deepfake generation. Figure 1 illustrates the general architecture of a deepfake generation system based on autoencoders. This structure employs a bidirectional learning mechanism to perform face swapping. The system's primary components include a shared encoder, two identity-specific decoders, and distinct discriminator networks trained for each decoder output. As an alternative method, GANs can be used to generate highly realistic videos that simulate human facial appearances [3].
Figure 2 presents the GAN architecture, an unsupervised deep learning model that enhances the realism of manipulated images and videos. GANs operate through an iterative training process in which the discriminator is trained first, followed by the generator. While earlier forms of fake content were manually created using tools like Photoshop, GANs enabled a more realistic and fully automated generation process.

[1]Corresponding author

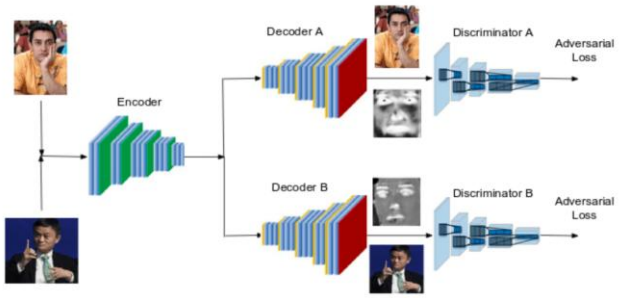


Figure 1. Structure of autoencoders [2].

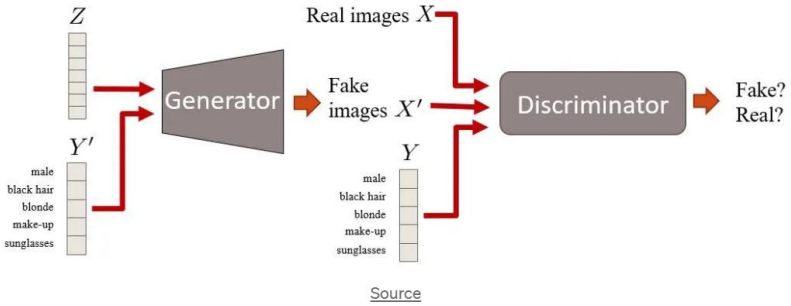


Figure 2. Generative adversarial network architecture [4]

In GAN-based deepfake generation processes, static regions are often either insufficiently processed or fail to synchronize accurately with facial dynamics. This results in various visual and structural inconsistencies within the content. For instance, deformations along facial contours, imbalanced lighting transitions, and mismatches between lip movements and audio offer critical clues for detecting deepfake content. Such anomalies typically manifest as geometric distortions, lighting discrepancies, and temporal discontinuities in manipulated videos [5][6]. In most deepfake videos, the synthetic face is superimposed onto the target individual's head; however, the alignment of facial expressions and spatial positioning often remains imperfect. Figure 3 illustrates facial regions where such manipulations tend to concentrate—particularly the jawline, hairline, and ear areas. These focus regions are strategically incorporated into the analysis pipelines to enhance the accuracy of detection systems.

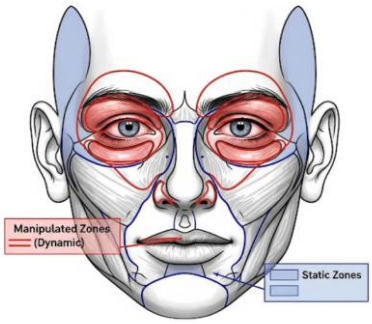


Figure 3. Facial areas focused on manipulation

Geometric discrepancies around the jawline, hairline, and ear regions can be identified through frame-based analysis. Meanwhile, slight mismatches in lip-sync accuracy can be detected using high-resolution, LSTM-based systems. Traditional handcrafted methods—such as LBP (Local Binary Patterns), HOG (Histogram of Oriented Gradients), and GLCM (Gray-Level Co-occurrence Matrix)—are often inadequate in handling the complexity of GAN-generated

content, prompting the use of deep learning approaches instead [7][8]. Effective deepfake detection relies on identifying both visual anomalies (e.g., pixel distortions, unnatural textures) and behavioral inconsistencies (e.g., micro-expressions, head movements, blinking, audio-lip synchronization). The most successful results have been achieved using hybrid models that integrate spatial and temporal analysis [9]. These hybrid architectures typically combine Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks (Figure 4). Each video frame is first processed by the CNN to extract spatial features, which are then passed to the LSTM for temporal sequence analysis. This integration allows the system to effectively detect forgery across both spatial and temporal dimensions.

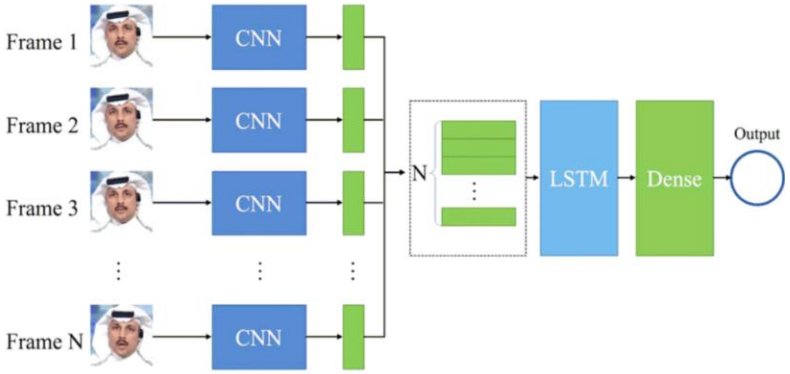


Figure 4. Hybrid architecture based on CNN and LSTM [9]

Recent studies have shown that hybrid architectures are capable of distinguishing deepfake videos with high accuracy, and models with enhanced temporal awareness are expected to play a critical role in future developments.

Table 1. Methods used for deepfake video detection in the literature

Reference	Dataset(s)	Method/ Model	Performance (Accuracy / AUC / F1)
Aprille J. Xi ve Eason Chen (2025)	Synthetic Real/Fake Face Set (University)	Swin Transformer, Swin-ResNet, Swin-KNN	71.29% accuracy
Anan et al. (2025)	DEVID-Cup 2025	ResNet-34, Swin, Wavelet, Ensemble	93.23% accuracy, 97.44% AUC
Sharma & Dowd (2025)	3 different datasets	EAR + RCNN + DBN + Bi-GRU	High accuracy and generalization
Singh et al. (2025)	Celeb-DF v2 (Kaggle)	Hybrid CNN + LSTM	87.2% accuracy
Al-Dulaimi & Kurnaz (2024)	DFDC, Ciplab	Transfer learning–based CNN-LSTM	98.21% accuracy
Liu et al. (2024)	Multimodal (image + audio + text)	Transformer-based multimodal model	91% accuracy
CNN-LSTM-Transformer (2024)	VoxCeleb2	CNN + LSTM + Transformer	90.82% AUC
Soudy et al. (2024)	FaceForensics++, WildDeepfake	CNN + Vision Transformer	High accuracy
Thing (2023)	FF++ 2020, Google DFD, Celeb-DF, DFDC	CNN vs Vision Transformer	88.74% – 99.73% accuracy
Khan & Dang-Nguyen (2023)	DFDC, CelebDF-V2, FakeAVCeleb	CLIP, DINO + Transformer	89.5% accuracy
Saikia et al. (2022)	DFDC, FF++, CelebDF	Optical flow + CNN-RNN	DFDC: 66.26%, FF++: 91.21%, CelebDF: 79.49%
Yang et al. (2019)	DARPA & UADFV	SVM + face-based features	DARPA: 84.3%, UADFV: 89%

Afchar et al. (2019)	Internet deepfake videos	Meso-4, MesoInception-4 (CNN)	99% accuracy
Zhou et al. (2018)	FaceSwap, SwapMe datasets	CNN	92.7% accuracy
Li et al. (2018)	Private deepfake video set	Blink detection (CNN)	92.7% accuracy

Aprille J. Xi and Eason Chen proposed three Swin Transformer–based models—Swin Transformer, Swin-ResNet, and Swin-KNN—which can potentially be integrated into facial recognition systems, achieving a top accuracy of 71.29% [10]. Afchar et al. [11] attained 99% accuracy on internet-sourced deepfake videos using MesoNet, a low-level CNN-based architecture. Zhou et al. [12] reported 92.7% accuracy with a CNN model trained on the FaceSwap and SwapMe datasets. Li et al. [13] introduced a detection method based on eye blink frequency. Saikia et al. [14] used an optical flow–based CNN-RNN hybrid model, achieving 66.26% on DFDC, 91.21% on FF++, and 79.49% on CelebDF. Thing [15] compared CNN and Vision Transformer architectures, demonstrating the global-context superiority of Vision Transformers with an accuracy range of 88.74%–99.73%. Khan and Dang-Nguyen [16] examined supervised and self-supervised learning approaches, revealing that Transformer-based models like CLIP and DINO deliver robust generalization performance on unlabeled data.

In this context, the most prominent methods in the literature have been systematically reviewed. The second section of this paper presents the proposed model, the third section details comparative experimental results, and the final section discusses the overall contributions. This structured approach offers meaningful insights and contributes significantly to the deepfake detection literature.

2. Proposed Method

In this study, a fast and reliable detection model along with a user-friendly web interface has been developed to detect deepfake videos. Videos obtained from the OpenForensics dataset were decomposed into frames, and face alignment and resizing were applied. A ResNet-18–based model was trained using these preprocessed frames and evaluated using various metrics on test data. The developed interface was also tested on social media videos, demonstrating its practical applicability in the domain of digital security. Figure 5 illustrates the overall flow of the proposed method.

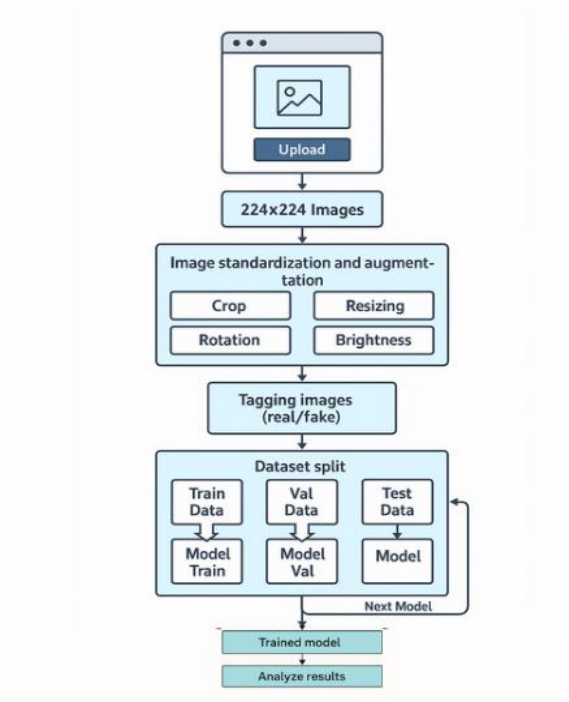


Figure 5. Block diagram of the deepfake video detection process

2.1. Data preprocessing

The preprocessing phase begins with splitting video files into individual frames. Then, the facial region is detected and cropped based on its center. To ensure uniformity, face alignment is applied so that all faces are aligned according to the same reference landmarks.

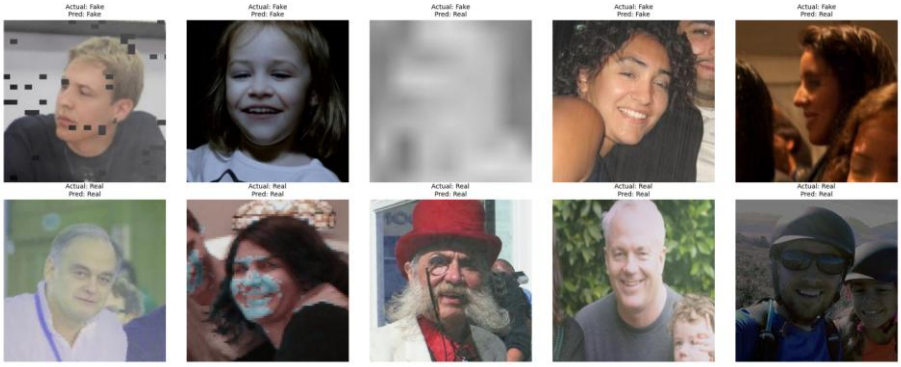


Figure 6. Example of data preprocessing

During training, random horizontal flipping, resizing, cropping, and normalization were applied to enable the model to learn various image variations. For validation, fixed-size resizing, center cropping, and normalization were performed. These transformations allowed the model to focus exclusively on the facial region, enabling it to distinguish between real and fake content with improved precision [17][18].

2.2. Proposed Model Architecture

In this work, a hybrid deep learning architecture has been designed to process both spatial (visual) and temporal (behavioral) information for the detection of video-based manipulated content. This approach aims not only to detect forgery traces in individual frames but also to capture inconsistencies throughout the video stream, thereby providing a more robust and reliable detection mechanism.

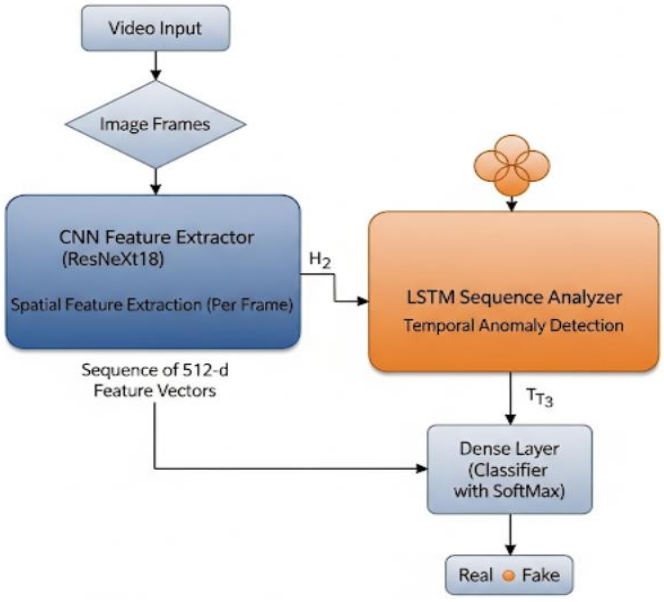


Figure 7. Workflow diagram of the proposed hybrid model

The proposed model integrates ResNet-18 for spatial feature extraction and an LSTM network to analyze temporal inconsistencies. Face images with a resolution of 224×224 pixels are processed through ResNet-18, where a global average pooling layer reduces them to 512-dimensional feature vectors. These vectors are then passed to the LSTM layer to capture temporal dependencies across frames. The LSTM outputs are processed using Leaky ReLU activation before being sent to the classification layer. This architecture enables the system to detect manipulations both in visual structure and over time, providing a robust and comprehensive detection mechanism.

2.3. Model Training Process

The dataset distribution shown in Figure 8 reflects commonly accepted standards in deep learning for training and evaluating models. This structure supports both the learning process and the model's ability to generalize across different data distributions and real-world scenarios.

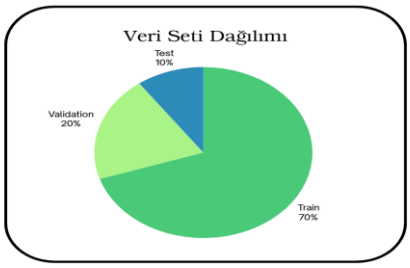


Figure 8. Dataset Distribution Graph

Table 2. Hyperparameters of the proposed method

Hyperparameter	Value
Learning Rate	0.001
Optimizer	Adam
Weight Decay	1e-5
Loss Function	CrossEntropyLoss()
Batch Size	32
Number of Epochs	5
Early Stopping	Patience = 5

The optimal hyperparameters were determined experimentally (Table 2). A learning rate of 0.001 was chosen, with Adam as the optimizer for dynamic gradient updates. Weight decay was set to 1e-5, and a dropout rate of 0.3 was used to mitigate overfitting. The CrossEntropyLoss function was employed, with Sigmoid and Softmax layers added to improve interpretability of outputs. During training, a batch size of 32 was used, and accuracy and loss values were monitored after each epoch.

3. Experimental Results

This section presents the results of experimental evaluations conducted to assess the effectiveness of the proposed method. As shown in Table 3, the findings reveal the model's strengths and limitations, while also shedding light on existing gaps in deepfake detection research.

Table 3. Commonly used datasets in deepfake detection

Dataset	**Key Characteristics**	**Manipulation Techniques**	**Strengths**	**Limitations / Challenges**
FaceForensics++ (FF++)	Large-scale standard dataset (1000+ videos)	DeepFakes, Face2Face, FaceSwap, NeuralTextures	Covers various manipulation types; widely accepted benchmark	May contain artifacts easily detected by modern systems
Celeb-DF	High-quality and hard-to-detect videos (590 real, 5639 fake).	Advanced GAN-based synthesis	Provides a realistic and challenging test environment	Contains relatively fewer videos compared to others
Deepfake Detection Challenge (DFDC)	Very large-scale and diverse (124,000+ videos)	Varied and deliberately obfuscated GAN methods	Best simulates real-world scenarios	High computational cost due to dataset size
OpenForensics	Focused on multi-face forgery (100,000+ face images)	DFDC, StyleGAN, DeepFaceLab and other modern techniques	Covers diverse facial types and modern forgery methods	Focused on face crops rather than full-length videos

Deepfake datasets have evolved in parallel with advances in forgery techniques. FF++ [19], offering four distinct manipulation methods, remains a widely used benchmark. However, its relatively detectable artifacts have highlighted the need for more advanced datasets. Addressing this, Celeb-DF [20] introduced more realistic and challenging fake videos. DFDC [21] provides broad coverage and diversity, closely simulating real-world conditions, albeit with high computational demands. In this study, the OpenForensics dataset was chosen for its emphasis on modern techniques and facial diversity. The evolution of these datasets directly reflects the ongoing technological race between deepfake generation and detection.

As summarized in Table 4, the OpenForensics dataset contains a total of 115,325 images and 334,136 face instances, with 44% fake and 56% real samples. Label information includes binary class labels (fake/real), face bounding boxes, and segmentation masks. Real face images were sourced from CelebA and FFHQ, while fake faces were generated using modern GAN-based methods such as DFDC, StyleGAN [22], and DeepFaceLab. This diverse structure enhances the model's generalization capabilities across different facial types, lighting conditions, and manipulation techniques, making it suitable for real-world applications.

Table 4. Dataset characteristics

Attribute	Description
Total Images	115,325 images
Total Faces	334,136 face instances
Fake Face Ratio	44% (approx. 147,000 fake faces)
Real Face Ratio	56% (approx. 187,000 real faces)
Label Types	Fake/real label, face bounding box, segmentation mask
Data Source	Real-world videos (indoor/outdoor, everyday scenarios)
Face Diversity	Variation in age, gender, pose, expression, occlusion (hands, masks), lighting
Manipulation Methods	Deepfakes, FaceShifter, FaceSwap, and other GAN-based techniques

The confusion matrix in Figure 9 summarizes the model's classification performance on the test set. The vertical axis represents the actual labels, while the horizontal axis shows the predicted labels. The diagonal entries, 5,223 correctly classified fake and 5,301 correctly classified real samples, indicate high model accuracy. Despite 112 false positives and 269 false negatives, the error rate remains low, and the model demonstrates strong discriminative power. These results validate the model's balanced and reliable performance.

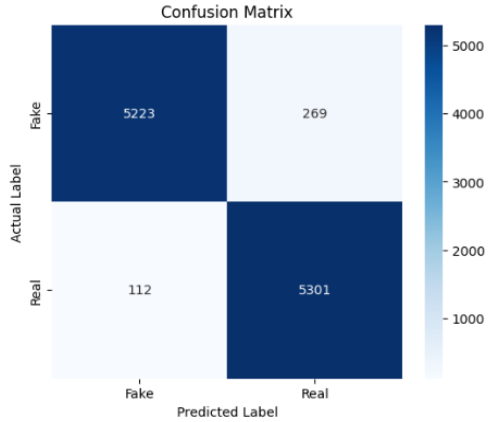


Figure 9. Confusion matrix for the ResNet18 model

To evaluate the model's performance on the real class, metrics such as Precision, Recall, and F1-Score were analyzed across epochs. Their consistent increase during training demonstrates that the model accurately learned from the real samples. The steady rise in validation metrics further suggests that the model generalizes well to unseen real-world samples without signs of overfitting.

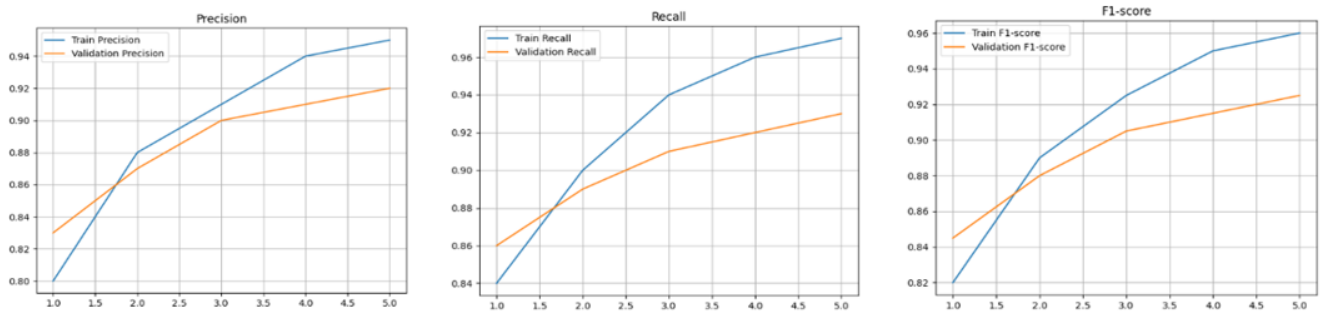


Figure 10. Training and validation metrics per epoch

Figure 11 displays the training and validation performance of the model over 5 epochs. The consistent decrease in training loss and the stability of validation loss indicate that overfitting did not occur. Concurrent increases in accuracy suggest a strong ability to both learn and generalize. The model, developed using the ResNet18 architecture, was trained on a balanced dataset, optimized using the hyperparameters specified in Table 2, and demonstrated robust performance on an independent test set.

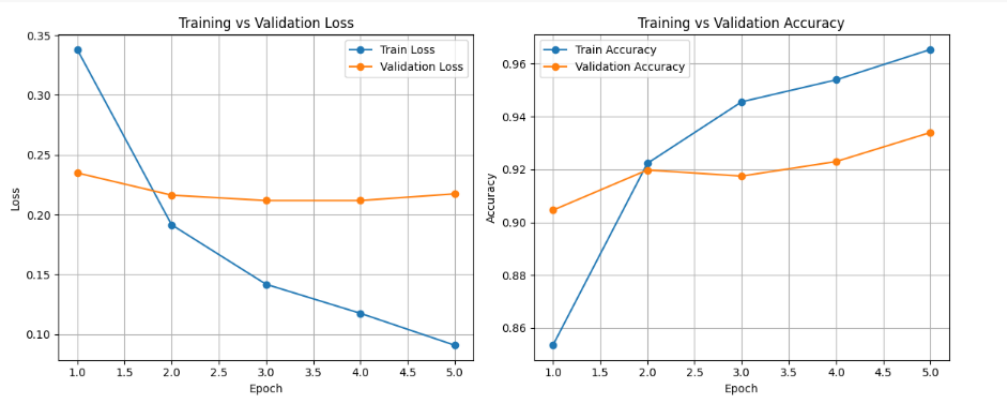


Figure 11. Epoch-wise training and validation loss & accuracy

The overall performance of the model is further detailed in Table 5. For the Fake class, the model achieved 98% precision, 95% recall, and 96% F1-score. For the Real class, it reached 95% precision, 98% recall, and 97% F1-score. These balanced metrics indicate that the model maintains consistent performance across both classes. The overall accuracy was calculated as 97%, demonstrating the ResNet18 model's strong capability in distinguishing between fake and real face images.

Table 5. Classification report

Class	Precision	Recall	F1-Score	Support
Fake	0.98	0.95	0.96	5492
Real	0.95	0.98	0.97	5413
Accuracy	-	-	0.97	10905
Macro Avg	0.97	0.97	0.97	10905
Weighted Avg	0.97	0.97	0.97	10905

The ROC curve and AUC score of the ResNet18-based model demonstrate its strong ability to distinguish between real and fake classes. Achieving high True Positive Rates (TPR) even at low False Positive Rates (FPR) is particularly crucial for applications that demand accurate detection of fake content without overlooking instances. The Precision–Recall (PR) curve was used to assess the model's performance under class imbalance. Here, Precision reflects the proportion of correctly predicted fake instances among all fake predictions, while Recall measures the proportion of all actual fake instances that were correctly identified. When evaluated together, these metrics confirm that the model performs effectively in both precision and coverage.

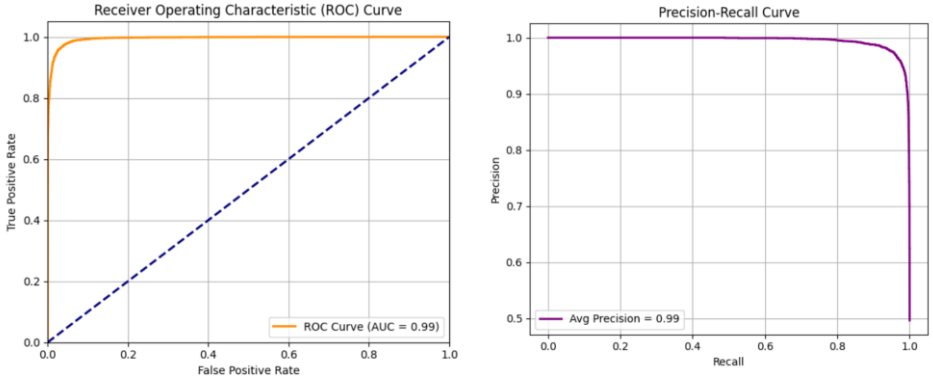


Figure 12. ROC and Precision–Recall (PR) Curves

Figure 13 presents a qualitative analysis showing how the developed model performs on a real video. The analysis visualizes the system's step-by-step decision-making process, illustrating how the theoretical framework translates into practical operation

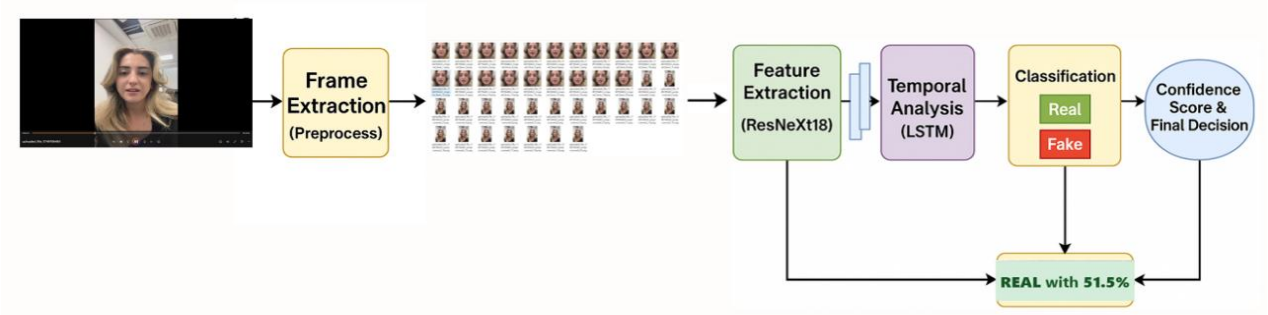


Figure 13. Qualitative analysis and system output on a sample video

Figure 14 displays detection outcomes on video frames collected from diverse sources. The dataset includes an original video recorded in a controlled environment, another original video featuring a different individual, and publicly available videos from platforms such as YouTube that include well-known figures. Frames extracted from these videos were analyzed by the system using bounding boxes, facial landmarks, and confidence scores, and were classified as "Real" (green box) or "Fake" (red box). These visual results indicate the model's robustness and reliability across various real-world scenarios.

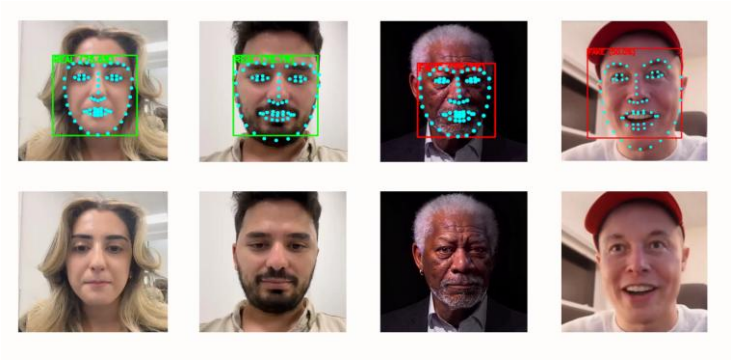


Figure 14. Detection Results on Videos from Different Sources

The ResNet18 + LSTM model developed in this study offers significant advantages over many existing deep learning approaches both technically and practically. The comparative evaluation of the model with other studies in terms of accuracy and F1 scores is summarized in Table 6.

Table 6. Comparison Table

Reference	Input Data Used	Methods Applied	Accuracy / F1 Score
[11] Afchar et al.	Internet-sourced deepfake videos	MesoNet (Meso-4 and MesoInception-4), CNN-based	99% Accuracy
[15] Thing (2023)	FF++ 2020, Google DFD, Celeb-DF, DFDC	Comparison of CNN and Vision Transformer (ViT)	88.74% – 99.73% Accuracy
[23] Al-Dulaimi & Kurnaz	DFDC and Ciplab datasets	Transfer learning–based CNN–LSTM hybrid	98.21% Accuracy
[24] Zhou et al.	FaceSwap and SwapMe datasets	CNN-based face manipulation detection	92.7% Accuracy
This Study	OpenForensics dataset (real-world, 115,000+ face images)	ResNet18 + LSTM architecture, advanced preprocessing, behavioral analysis, user interface	97% Accuracy, 96–97% F1 Score

The model was trained on the OpenForensics dataset, which features a wide range of manipulation techniques and facial variations. This helped the model acquire strong generalization capabilities for real-world applications. Advanced preprocessing techniques such as face alignment, normalization, and segmentation enabled the model to focus solely on the facial region, leading to more precise analysis. Compared to MesoNet by Afchar et al. [11], which involves limited preprocessing, the proposed method provides a more targeted and effective solution. The ResNet18 backbone ensures high parametric efficiency and deep feature extraction, while the LSTM layer successfully captures temporal inconsistencies such as missing eye blinks and micro-expression anomalies. In this regard, the study stands out particularly for its emphasis on temporal analysis, a dimension not sufficiently addressed in the CNN–ViT comparisons by Thing [15]. While the architecture shows similarities to the CNN–LSTM model proposed by Al-Dulaimi & Kurnaz [23], this work further integrates advanced preprocessing, a user-friendly interface, and confidence score interpretation. Moreover, this study addresses usability, a criterion often overlooked in the literature, by offering a web-based system that enables real-time analysis. Taken together, these components result in a comprehensive, deployable, and user-oriented solution, outperforming the methods proposed in [11], [15], [23], and [24] in terms of practicality and performance—with a validated accuracy of 97% and F1 scores between 96–97%.

3.1. User Interface Design

The goal of this study was not only to develop a deep learning model with theoretical accuracy but also to transform it into a practical and usable solution. To that end, a user-friendly web application was designed to allow even non-technical users to easily perform video-based analysis. The detection system was developed as a web-based client-server application. The backend of the system was implemented using Flask, a lightweight Python-based web framework (Figure 15).
This architecture processes user-uploaded videos, analyzes them using the trained deep learning model, and returns the detection results. The system provides an efficient structure in which the user simply uploads a video to initiate the process, while all technical operations are automatically executed in the background. As a result, the prediction capabilities of the model have been made available in real time, and system deployment and accessibility have been significantly simplified.

Figure 16 shows the developed web interface. Users can directly upload video files, which are then processed by the system to deliver a deepfake detection result along with a confidence score in real time. The interface is minimal and intuitive, ensuring ease of use for individuals without technical expertise.

Thanks to its web-based architecture, the system is applicable not only in research environments but also in everyday usage. Furthermore, its modular design allows for the integration of features such as support for various file formats, API extensions, and mobile compatibility.

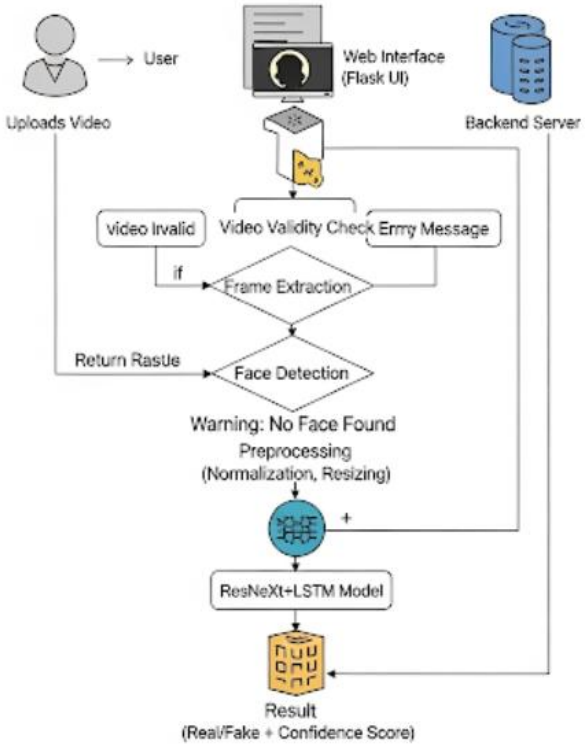


Figure 15. System flow diagram



Figure 16. Interface and scenario examples

4. Conclusion

Deepfake technology has evolved into a sophisticated tool capable of generating highly realistic yet fraudulent video content, posing substantial risks to political discourse, financial systems, and individual reputations. In this work, we addressed the urgent need for reliable deepfake detection by developing a hybrid deep learning framework that leverages a ResNet-18 backbone for spatial feature extraction and an LSTM layer for temporal inconsistency analysis. By incorporating advanced preprocessing steps and wrapping the model in a user-friendly web interface, we have created not only a high-performance detection engine but also a practical tool for real-world deployment. Our experimental evaluation on the OpenForensics dataset, comprising over 115,000 face images drawn from diverse real-world scenarios, demonstrates that the proposed model achieves 97 % classification accuracy and 96–97 % F1 scores across fake and real classes. The confusion matrix, ROC/AUC, and precision–recall analyses validate the model's robust ability to capture both pixel-level anomalies (e.g., lighting imbalances, contour distortions) and behavioral cues (e.g., micro-expressions, lip-sync mismatches). Compared to existing approaches our solution offers a balanced combination of high parametric efficiency, temporal awareness and practical usability. Despite these strengths, certain limitations remain. The current model focuses exclusively on visual cues and does not integrate audio features, which could further enhance detection of lip-sync irregularities. In addition, inference speed and resource requirements may hinder deployment on edge devices or mobile platforms. Future work will explore multimodal integration of audio and textual metadata, model compression techniques for lightweight deployment, and adversarial training to bolster resilience against evolving deepfake generation strategies. We also plan to investigate legal and ethical frameworks to guide responsible use and inform policy development. By combining rigorous technical performance with an accessible interface, this study contributes a comprehensive and deployable deepfake detection solution. We believe that extending this framework to include multimodal analysis and deploying it in diverse operational environments will further strengthen defenses against the growing threat of synthetic media.

Acknowledgment

This study was supported by the Scientific and Technological Research Council of Türkiye (TUBITAK) 2209-B Program under grant number 1139B412301638.

5. References

[1] Nguyen, T. T., Nguyen, C. M., Nguyen, D. T., Nguyen, D. T., & Nahavandi, S. (2019). Deep learning for deepfakes creation and detection: A survey. arXiv preprint arXiv:1909.11573.

[2] Korshunova, I., Shi, W., Dambre, J., & Van Lerberghe, L. (2017). *Fast Face-swap Using Convolutional Neural Networks*. arXiv preprint arXiv:1708.03474.

[3] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. Advances in Neural Information Processing Systems, 27, 2672–2680.

[4] Korshunov, P., & Marcel, S. (2018). Deepfakes: a new threat to face recognition? Assessment and detection. arXiv preprint arXiv:1812.08685

[5] Matern, F., Riess, C., & Stamminger, M. (2019). Exploiting Visual Artifacts to Expose Deepfakes and Face Manipulations. Proceedings of the IEEE Winter Conference on Applications of Computer Vision Workshops (WACVW), 83-92.
[6] Agarwal, S., Farid, H., Guera, D., He, Z., Nagano, K., & Li, H. (2019). Protecting World Leaders Against Deep Fakes. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 38-45.
[7] Matern, F., Riess, C., & Stamminger, M. (2019). Exploiting visual artifacts to expose deepfakes and face manipulations. IEEE Winter Conference on Applications of Computer Vision (WACV).
[8] Verdoliva, L. (2020). Media Forensics and DeepFakes: An Overview. IEEE Journal of Selected Topics in Signal Processing.
[9] Mirsky, Y., & Lee, W. (2021). The Creation and Detection of Deepfakes: A Survey. ACM Computing Surveys (CSUR), 54(1), Makale No: 19, 1–41.
[10] Xi, A. J., & Chen, E. (2023). Swin Transformer-based models for deepfake detection. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 1201–1210.
[11] Afchar, D., Nozick, V., Yamagishi, J., & Echizen, I. (2018). MesoNet: a Compact Facial Video Forgery Detection Network. In IEEE International Workshop on Information Forensics and Security (WIFS) (ss. 1-7).
[12] Zhou, P., Han, X., Morariu, V. I., & Davis, L. S. (2017). Two-Stream Neural Networks for Tampered Face Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 1831-1839.
[13] Li, Y., Chang, M. C., & Lyu, S. (2018). Exposing AI created fake videos by detecting eye blinking. In IEEE International Workshop on Information Forensics and Security (WIFS) (ss. 1–7).
[14] Saikia, R., Sharma, A., & Patel, K. (2022). Hybrid CNN-RNN model for deepfake detection using optical flow-based motion analysis. Proceedings of the International Conference on Computer Vision and Image Processing, 221–230.
[15] Thing, A. K. (2023). A comparative analysis of CNN and Vision Transformer architectures in deepfake detection. Journal of Machine Learning and Cybersecurity, 15(2), 101–115
[16] Khan, M., & Dang-Nguyen, D.-T. (2023). Supervised vs. self-supervised learning for robust deepfake detection: A comparative study. IEEE Transactions on Information Forensics and Security, 18, 1210–1223.
[17] Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. Journal of Big Data, 6(1), Makale No: 60. (Kullanıcı tarafından sağlanan sayfa aralığı: 1-48).
[18] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In Advances in Neural Information Processing Systems 25 (NIPS 2012) (ss. 1097–1105).
[19] Rossler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., & Nießner, M. (2019). FaceForensics++: Learning to Detect Manipulated Facial Images. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (ss. 1–11).
[20] Li, Y., Sun, P., Qi, H., & Lyu, S. (2020). Celeb-DF: A Large-Scale Challenging Dataset for DeepFake Forensics. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 3207-3216).
[21] Dolhansky, B., Howes, R., Pflaum, B., Bitton, J., & Ferrer, C. C. (2020). The DeepFake Detection Challenge (DFDC) Dataset. arXiv preprint arXiv:2006.07397.
[22] Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2020). Analyzing and Improving the Image Quality of StyleGAN. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (ss. 8110–8119).
[23] Al-Dulaimi, H., & Kurnaz, S. (2024). Transfer learning-assisted CNN-LSTM hybrid model for behavior-aware deepfake video detection. Multimedia Tools and Applications, 83(1), 55–76.
[24] Karras, T., Laine, S., & Aila, T. (2019). A Style-Based Generator Architecture for Generative Adversarial Networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (ss. 4401–4410).

Lignin Recovery and Characterization from the Liquid Obtained by the Organosolv Process of MDF Dust

Kübra İNÇ[1], Hamdi Soner ALTUNDOĞAN [2]

[1,2] Bioengineering Department, Faculty of Engineering, Fırat University, Elazıg, Turkey.
[1] kkocak@firat.edu.tr, [2] saltundogan@firat.edu.tr
[1](ORCID: 0000-0003-1470-9937), [2](ORCID: 0000-0003-3328-9667)

Abstract
The organosolv process is the extraction of lignin from lignocellulosic raw materials using organic solvents or their aqueous solutions. With the organic solvent used, under specific temperature and pressure, most of the lignin and hemicellulose dissolve into the solvent, resulting in a relatively pure solid cellulose residue. High-quality lignin isolated from biomass after organic solvent pretreatment has the potential to be used in various industrial applications. Products obtained from lignin can be used for various purposes such as fuels, fuel additives, chemicals, biopolymers, adhesives, biomaterials for drug release, and polymer fillers. In this study, lignin was recovered from the residue of the filtrate resulting from the organosolv process applied to medium-density fiberboard (MDF) dust under specific conditions (60% ethanol solution, 0.1 M NaOH catalyst, 120°C temperature, 1:10 solid/liquid ratio, and 60 minutes duration), and the characterization of the obtained lignin was performed. For this, the distillate used was diluted in specific proportions and the pH was lowered to precipitate the lignin. The studies were conducted using the Central Composite Design (CCD) technique with the help of the DesignExpert® 8.0 program, and ranges and central points were determined for the two parameters to be examined (pH and dilution ratio). The sedimentation efficiencies obtained, depending on the conditions, vary between 94% and 110%. The values of sedimentation efficiency obtained above 100% are thought to be due to the partial sedimentation of dissolved components from biomass other than lignin during the sedimentation process. The optimal conditions obtained in the solution conducted to maximize the precipitation yield were a 4.73-fold dilution ratio and a pH value of 1.89, and the experimental precipitation yield value obtained in the lignin precipitation experiment conducted under these conditions was calculated to be 107.40%. The relative error between the solution value calculated from the model and the experimental finding is 2.3%, indicating that the model represents the actual situation well.

Keywords: lignin, organosolv, characterization

1. Introduction

Lignocellulosic biomass is a type of biomass that contains cellulose, hemicellulose, and lignin, as well as small amounts of protein, lipids, and ash. Due to its abundant presence on Earth, it is preferred for the production of energy and chemical products [1]. In addition, it is evident that the diversity of the final products will have positive effects on the chemical industry and the energy sector, which will in turn provide significant benefits to society [2]. Forestry waste plays a key role as lignocellulosic biomass because it is produced in significant quantities during industrial processes. This energy source is preferred not only because of technological advancements but also due to its low costs [3]. However, the industrial advantages to be gained depend on the effective and selective separation of the

[1]Corresponding author

main components of lignocellulosic biomass, namely hemicellulose, cellulose, and lignin, and the production of value-added products from the obtained fractions [4].

The Organosolv process is one of the methods that has been used for many years to separate the components of lignocellulosic biomass. The common solvents used for this pretreatment are low-boiling point alcohols such as methanol, ethanol, acetone, ethylene glycol, and ethyl acetate. Additionally, a wide range of organic solvents are also used, including propanol, butanol, isobutyl alcohol, benzyl alcohol, glycerol, glycol, triethylene glycol, phenol, formic acid, acetic acid, propionic acid, diethyl ether, amines, ethers, esters, formaldehyde, and chloroethanol [5]. The organic solvent used facilitates the extraction of lignin by affecting the acid-ester bonds of lignin and hemicellulose in lignocellulosic raw materials under specific temperature and pressure conditions.

Lignin is the second most abundant terrestrial organic polymer on Earth after cellulose and plays an important role in the structure of plant cell walls. It is a complex, natural biopolymer that is bound by ether or ester bonds to hemicellulose. One of the main components of lignocellulosic biomass, which constitutes 15-40% of its mass composition, is lignin. Products obtained from lignin can be utilized in a wide range of fields such as fuels, additives, various chemicals, biopolymers, adhesives, biomaterials for drug delivery, and polymer fillers [6]. Especially its lack of sulfur, low molecular weight, and superior properties such as water repellency make lignin suitable for a wide range of industrial applications in high-value-added uses [7].

In this study, lignin extraction from the distillate obtained through the organosolv process applied to medium-density fiberboard (MDF) dust, which is a waste in the furniture sector and a significant environmental burden, was carried out. The yield of lignin obtained by the precipitation method was calculated gravimetrically, and the experiments were conducted using the Central Composite Design (CCD) technique.

2. Material and Method

For the lignin recovery studies, the distillate obtained from the organosolv process was first diluted in specific proportions and the pH was lowered by removing the alcohol with the help of an evaporator, resulting in the precipitation of lignin. The studies were conducted using the Central Composite Design (CCD) technique with the help of the DesignExpert® 8.0 program for experimental design. For this purpose, the ranges and center points of the pH values and dilution ratios to be examined were determined. 10 ml samples were taken from the distilled liquid from which the alcohol had been removed, and the conditions of the experimental design were carried out. pH adjustments were made using a pH meter with a 0.1 M H_2SO_4 solution. The mixtures prepared with different dilution ratios and pH values were left in the refrigerator overnight, and the resulting precipitate was separated from the liquid phase by centrifuging at 3500 rpm for 20 minutes. The separated lignin-based precipitate was dried in an oven at 80°C. The weights of the samples that reached a constant weight were measured, and the precipitation yields were calculated. The thermal analyses conducted for the characterization of the obtained lignin were performed in a nitrogen atmosphere at the Munzur University Rare Earth Elements Application and Research Center, using a Linseis HDSC PT 1600 device, with a heating rate of 10 °C/min from room temperature to 800 °C. FT-IR (PerkinElmer) and elemental analyses (LECO, CHNS-932) were conducted at the Scientific and Technological Research Center of İnönü University.

The variables examined in the experimental studies and the review intervals are provided in Table 1. The response variable of the experimental design was based on the collapse efficiency.

Table 1. The ranges of the independent variables and coded variables in the design of lignin precipitation experiments

Variables	-α	-1	0	+1	+α
pH	1,58	2,00	3,00	4,00	4,41
Dilution ratio	1,38	2,00	3,50	5,00	5,62

3. Results and Discussion

In the study, the results obtained under the experimental design conditions, where the effects of pH and dilution ratio on lignin precipitation yield were examined, are shown in Table 2. The precipitation yields obtained from the experimental results range between 94% and 110%. These values have been calculated based on the ratio of the amount of precipitated lignin to the total amount of lignin in the solution. It is thought that the precipitation yield values obtained above 100% are due to the partial precipitation of dissolved components from biomass other than lignin during the precipitation process.

Table 2. Experimental design conditions and results for lignin recovery

Experiment No	pH	Dilution ratio	Precipitated lignin (mg)	Lignin precipitation yield (%)
1	2,00	2,00	130	98,61
2	4,00	2,00	127	96,34
3	2,00	5,00	143	108,48
4	4,00	5,00	126	95,58
5	1,58	3,50	145	109,99
6	4,41	3,50	124	94,06
7	3,00	1,38	124	94,06
8	3,00	5,62	133	100,89
9	3,00	3,50	132	100,13
10	3,00	3,50	135	102,41
11	3,00	3,50	134	101,65
12	3,00	3,50	134	101,65
13	3,00	3,50	133	100,89

The statistical results obtained by applying different mathematical models using lignin precipitation yield as the response value are presented in Table 3. As seen in the table, a second-degree model is proposed to relate the precipitation yield to pH and dilution ratio for the precipitation of lignin from the organosolv liquid. Although the cubic model seems more suitable in terms of the adjusted R^2 value, the low p-value indicating a lack of fit is the main reason for proposing the second-degree model. The ANOVA analysis results performed by the software for the proposed model are shown in Table 4.

Table 3. Statistical comparison of models relating the obtained lignin precipitation yields to independent variables

Model Equations	Sequential p-Value	Lack of Fit p-Value	Adjusted R^2	Predicted R^2	
Linear	0,0007	0,0114	0,7230	0,5184	
2FI	0,0300	0,0246	0,8227	0,6769	
Quadratic	0,0084	0,1560	0,9417	0,8159	Recommended
Cubic	0,0516	1.0000	0,9751	0,9838	

Table 4. ANOVA test results of the proposed model for the recovery of lignin from organosolv liquid

	Sum of Squares	Degrees of Freedom	Mean Square	F-value	p-value	
Model	278,36	5	55,67	39,78	< 0,0001	Significant
A-pH	177,66	1	177,66	126,95	< 0,0001	
B- Dilution ratio	43,98	1	43,98	31,43	0,0008	
AB	28,20	1	28,20	20,15	0,0028	
A^2	0,81	1	0,81	0,58	0,4715	
B^2	26,03	1	26,03	18,60	0,0035	
Residuals	9,80	7	1,40			
Lack of fit	6,80	3	2,27	3,03	0,1560	Not significant
Pure Error	2,99	4	0,75			
Corrected Sum of Squares	288,16	12				

The statistical measures used to test the suitability of the proposed second-degree model are shown in Table 5. The difference between the Pred R^2 value of 0.8159 and the Adj R^2 value of 0.9417, which is less than 0.2, indicates a reasonable agreement between the experimental results and the results produced by the model. The adequate precision value is a measure of the signal-to-noise ratio, and it is desirable for it to be greater than 4. A value of 19.042 indicates a sufficient signal. This means that by utilizing this model equation, results with sufficient accuracy can be obtained within the design space (the ranges of the variables being examined). The second-degree model equation proposed by the program, relating lignin precipitation yield to pH and dilution ratio, is as follows:

$$Lignin\ precipitation\ yield(\%) = 83{,}9677 - 0{,}5656A + 12{,}8912B - 1{,}77AB + 0{,}3414A^2 - 0{,}8597B^2 \quad (1)$$

The DesignExper 8.0 software used for experimental design can also generate graphs showing the interactive effects of independent variables on the response variable based on the proposed model. The interaction graph showing the effect of pH and dilution ratio on lignin precipitation yield from the organosolv solution can be seen in Figure 1. Looking at the figure, it can be seen that the effects of the pH and dilution ratio variables on the precipitation yield are significant and consistent with the ANOVA data. At low dilution ratios, the sedimentation yield increases at a slow rate with increasing acidity, while at high dilution ratios, the increase rate significantly rises with decreasing pH. Similarly, while no significant change in sedimentation yield is observed with increasing dilution ratios around pH 4, it can be said that the sedimentation yield increases at a noticeable rate with increasing dilution ratios around pH 2. Ultimately, it is observed that conditions with acidic and high dilution ratios stand out significantly in the examined intervals for high lignin precipitation yields.

Table 5. Statistical values calculated to test the suitability of the proposed second-degree model for lignin recovery by precipitation from organosolv liquid

Standard Deviation	1,18	**R^2**	0,9660
Mean	100,37	**Adj R^2**	0,9417
Coefficient of Variation (C.V) %	1,18	**Pred R^2**	0,8159
Predicted Residual Error Sum of Squares (Press)	53,06	**Adequate Precision**	19,042

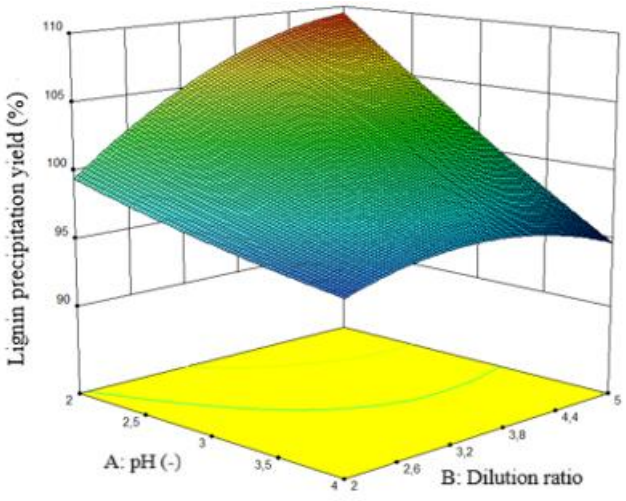


Figure 1. The interactive effect of pH and dilution ratio on the precipitation of lignin from the organosolv solution

With the optimization module included in the DesignExpert 8.0 program, it is possible to solve the produced model equation for certain optimization conditions. Thus, it is possible to produce optimal solutions for the desired conditions, and it is also possible to test the accuracy of the model produced through validation experiments under these optimal conditions. The optimal conditions obtained from the solution to maximize the precipitation yield of lignin from organosolv liquid, which is related to pH and dilution ratio according to Equation 1., are a 4.73-fold dilution ratio and a pH value of 1.89, provided that the independent variables remain within the specified range. Under these conditions, the program also suggests a solution value of 109.99% for lignin removal efficiency. In the lignin precipitation experiment conducted under these conditions, the experimental precipitation yield value obtained was 107.40%. The existence of a 2.3% relative error between the solution value calculated from the model and the experimental finding was evaluated as an indication that the model represents the actual situation well.

The results of the thermogravimetric analysis of the obtained lignin-based precipitate are shown in Figure 2. When the TGA curve is examined, it is observed that a weight loss occurs up to approximately 100 °C during the initial heating stage. It is thought that this weight loss corresponds to the removal of free moisture in its structure, for example. From this temperature onwards, the curve reaches a plateau, and it is observed that this plateau continues up to approximately 200 °C without a sudden decrease in mass. In this range, the existing moisture in the structure evaporated, resulting in a mass loss of about 10%. A rapid mass loss occurred in the 200-400 °C range, with approximately 60% of the mass being reduced. In this range, it is thought that lignin is broken down into phenolic groups with smaller molecular weights that can volatilize. At higher temperatures (400-800°C), it is thought that a condensation occurs through the decomposition of aromatic rings and/or carbonization [8]. According to the data obtained from the graph, the maximum degradation temperature (Tmax) of lignin has been determined to be 550 °C.

The FT-IR spectrum of the obtained lignin-based precipitate is shown in Figure 3. In a study conducted by Başakçılardan Kabakcı et al. [6], they examined the FT-IR spectrum of the lignins obtained as a result of the organosolv process and identified the characteristic peaks. The peaks in the spectrum we obtained are very similar to those in the conducted study. When the spectrum is examined based on the identified peaks, C-H stretching vibrations were observed at 2919 cm^{-1}. 1592-1505 cm-1 represents aromatic C=C vibrations and additionally, 1459-1421 cm^{-1} represents C-H bending vibrations, indicating the presence of methyl and methylene groups. The presence of aliphatic and aromatic ether bonds is indicated by the C-O-C stretching vibrations at 1122 and 1028 cm^{-1}. These are generally the bonds found in the side chains of lignin. The prominent vibration at 1122 cm^{-1} is generally an indication of syringol groups found in hardwood lignins. Again, the peak observed at 1220 cm^{-1} is thought to indicate the presence of aryl ether bonds, representing C-O stretching vibrations.

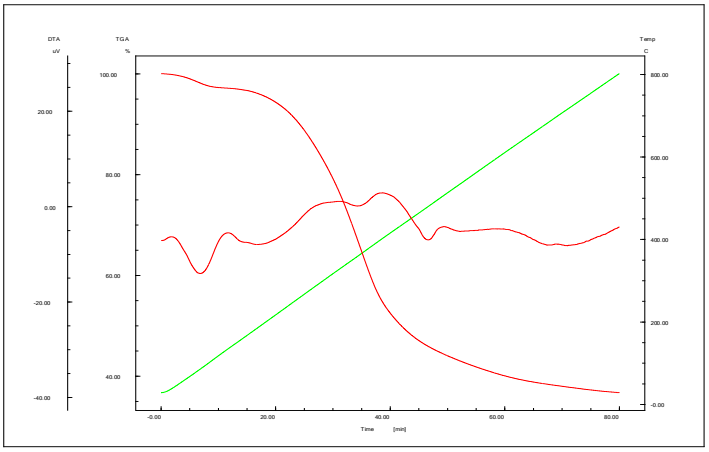


Figure 2. Thermogravimetric analysis results of lignin obtained from organosolv liquid

The elemental composition of the lignin sample obtained under optimum conditions is shown in Table 6. Yáñez-S et al. (2014) revealed that the elemental composition of the lignin samples obtained from Eucalyptus globulus using ethanol organosolv consists of 60-63% C, 27-31% O, and 5-8% H. The N percentages, on the other hand, vary between 0.2-0.5 [9]. However, the N content of the lignin obtained in the study is relatively high compared to these results. The likely reason for the high nitrogen content is believed to be the presence of urea-formaldehyde resins used in MDF production. It is thought that the sulfur element, which is present in small amounts in the structure, originates from the H_2SO_4 used during lignin precipitation or from the lignocellulosic biomass used in MDF production.

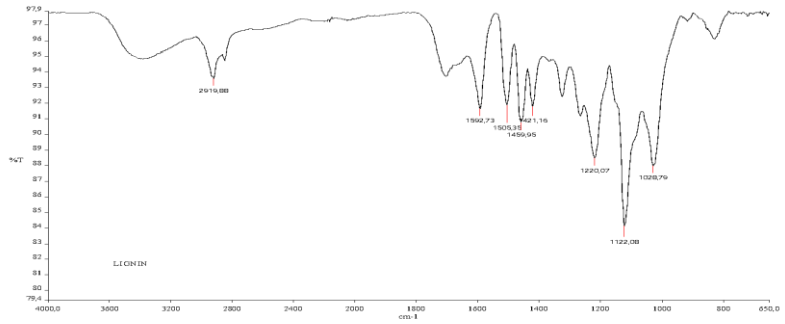


Figure 3. FT-IR Spectrum of lignin-based precipitate obtained by precipitation from organosolv liquid

Table 6. Elemental composition of lignin-based precipitate obtained under optimum conditions

Elements	C	O	H	N	S
w/w, %	60,14	32,14	6,25	1,26	0,22

4. Conclusion

The studies conducted to recover lignin from the liquid product obtained after organosolv experiments carried out under optimum conditions were performed under the conditions determined in the design created using the DesignExpert® 8.0 software with the Central Composite Design (CCD) technique, which is part of the Response Surface Methodology (RSM). The study demonstrated that lignin can be recovered from the remaining organosolv

solution through dilution. The resulting lignin stands out as an important biological resource that can be used in energy, fuel, materials production, agricultural, and environmental applications.

Acknowledgements

This study was supported by Fırat University Scientific Research Projects Coordination Unit (FÜBAP) with the project protocol number MF.23.21.

5. References

[1] N. Brosse, M. H. Hussin, and A. A. Rahim, "Organosolv processes," in *Advances in Biochemical Engineering/Biotechnology*, vol. 166, Springer Science and Business Media Deutschland GmbH, 2019, pp. 153–176. doi: 10.1007/10_2016_61.

[2] H. Kouchaki-Penchah, M. Sharifi, H. Mousazadeh, and H. Zarea-Hosseinabadi, "Life cycle assessment of medium-density fiberboard manufacturing process in Islamic Republic of Iran," *J Clean Prod*, vol. 112, pp. 351–358, Jan. 2016, doi: 10.1016/J.JCLEPRO.2015.07.049.

[3] T. Kan, V. Strezov, and T. J. Evans, "Lignocellulosic biomass pyrolysis: A review of product properties and effects of pyrolysis parameters," *Renewable and Sustainable Energy Reviews*, vol. 57, pp. 1126–1140, May 2016, doi: 10.1016/J.RSER.2015.12.185.

[4] L. Penín, M. López, V. Santos, J. L. Alonso, and J. C. Parajó, "Technologies for Eucalyptus wood processing in the scope of biorefineries: A comprehensive review," Sep. 01, 2020, *Elsevier Ltd*. doi: 10.1016/j.biortech.2020.123528.

[5] M.-F. Li, S.-N. Sun, F. Xu, and R.-C. Sun, *Organosolv fractionation of lignocelluloses for fuels, chemicals and materials: A biorefinery processing perspective*, vol. 9783642284. 2012. doi: 10.1007/978-3-642-28418-2_11.

[6] S. Başakçılardan Kabakcı and M. H. Tanış, "Pretreatment of lignocellulosic biomass at atmospheric conditions by using different organosolv liquors: a comparison of lignins," *Biomass Convers Biorefin*, vol. 11, no. 6, pp. 2869–2880, Dec. 2021, doi: 10.1007/s13399-020-00677-2.

[7] C. Arato, E. K. Pye, and G. Gjennestad, "The lignol approach to biorefining of woody biomass to produce ethanol and chemicals," *Applied Biochemistry and Biotechnology - Part A Enzyme Engineering and Biotechnology*, vol. 123, no. 1–3, pp. 871–882, 2005, doi: 10.1385/abab:123:1-3:0871.

[8] Y. Zhang *et al.*, "Organosolv fractionation of a lignocellulosic biomass feedstock using a pilot scale microwave-heating reactor," *Ind Crops Prod*, vol. 180, Jun. 2022, doi: 10.1016/j.indcrop.2022.114700.

[9] M. Yáñez-S *et al.*, "Physicochemical characterization of ethanol organosolv lignin (EOL) from Eucalyptus globulus: Effect of extraction conditions on the molecular structure," 2014, *Elsevier Ltd*. doi: 10.1016/j.polymdegradstab.2014.08.026.

Energy Efficiency İmprovement Based on Smart Grid: A Simulation Assessment

Ilkin BALAZADE[1], Rasul AHMADOV[2]

[1,2]Information Technologies Department, Engineering Faculty, Baku Higher Oil School, Baku, Azerbaijan..
[1]ilkinbalazade@gmail.com, [2]rasul.ahmadov@outlook.com
[1](ORCID: 0009-0007-8662-8542)

Abstract
In this article, first and foremost the authors analyzed smart networks in comparison with traditional networks and noted a number of important advantages of smart networks over existing network infrastructures. In addition, the architecture of intelligent power systems was analyzed and a dynamic model for the smart network was built using simulation programs GridLAB-D and RAPSim. The main purpose of choosing these simulation programs is that GridLAB-D does not require calculation methods due to additional factors that complicate the application scenario and require more user experience.

Keywords: smart network, renewable energy source, suboptimal industrial economics, fuel generator, power flow, demand-response

1.Introduction

If we analyze the Smart Grid network operating modes with the current network operating modes: (Table 1)

Table 1. Comparison of existing power network and SMART grid

Features	Traditional grids	SMART Grids
Communication	One way	Two ways
Interaction with the user	Limited	Wide
Tool type	Electrical	Digital
Traffic control	Limited	Universal
Reliability	Prone to failure and cascading interruptions	Pro-active, real-time protection
Restore power	By hand	Self-renewal
Topology	Radial	Network
Generator	Centralized	Disappointed
Operation and technical service	Hand-held equipment checks time-based maintenance	Remote Tracking, Forecasting and Conditioning Maintenance
Measurement	Electromechanical	Digital
Customer participation	Limited	Wide

Smart networks need to be consistent in order to meet standards and be effective. Sustainable smart networks meet all three dimensions of sustainable development:

(1) Able to work economically and save on costs
(2) It is environmentally friendly and combines renewable energy sources
(3) Provides a two-way communication between user and supplier.

Smart Grid is a highly-knowledgeable and intelligent active adaptive power network with a number of features:

- Increased consumer and system reliability;
- Continuous self-monitoring and self-treatment of network components;
- Integration of renewable energy sources;
- Participation of "active" consumers in the market of electricity, capacity and system services;
- Increased physical and cybernetic security;
- Development of new industrial markets;

The architecture and main industry characteristic of intellectual power systems meet modern requirements. If we summarize the prerequisites for a new era in this area, we can take the form shown in Table 2.

Table 2. The architecture and main industry characteristic of intellectual power systems

Power systems are designed for large industrial needs, "big" centralized generators and "passive" consumers.	- Development of distributed generators, including RES (growing rapidly but with limitations - about 15-20% of total generator volume) - Emerging phenomenon of "active" consumers and sellers - Electric cars and other new "external" network elements of the generator network
Increased demand for energy services in the economy	- The "cost" of direct and indirect losses due to production growth accidents, for example, $ 80-150 billion in the United States (according to LBNL, EPRI estimates) - The Digital Demand Phenomenon (Data Centers Since the mid-2000s have consumed more than the entire automotive industry in the US, for example, Google is the largest consumer of electricity) - Requirements for connection speed, development, energy service flexibility
Platform of technological efficiency of basic industrial technologies	- Exhaustion of existing technology development limits - Increased cost of infrastructure does not lead to a relative increase in efficiency
Suboptimal industrial economics	- By 2030, $ 17 trillion is required to invest - The problem of the “fair” distribution of costs and costs - "Tension" by vendors

2. Statement of The Problem

Simulation of electric power systems, is an important technology for the power generation industry, including production and transmission planning, short-term operational assessment, financial analysis, management algorithm development and quality assurance. Common uses of simulation for power systems include power or load estimation, injection of short and open circuits, switch analysis, generating unit management, power flow and quality optimization.

Simulators for power system analysis are used by experts during network planning, operations and pricing. Simulations of electric power systems are generally classified as follows:

- Sustainable State Simulations - The power grid is the basis for research on operating modes. Researchers and engineers are doing "what-if" to measure the impact of changes in power systems. The system is analyzed in a stable equilibrium state, and the main focus is on checking whether the power system variables are within appropriate boundaries (eg voltage thresholds). Different simulators specializing in continuous case studies, power flow studies, load estimation and load balancing, fault analysis techniques are offered. At the same time, steady-state simulations also include optimal energy flow studies.

- Transition Dynamics Simulation Investigation - Transitions between equilibrium points due to major changes in electrical network configuration. The main purpose of such studies is to determine whether the load angle has reached a new optimal state. Short simulations are studied in these simulations. If the balance is lost due to persistent minor disturbances, dynamic stability simulations, also known as small-wave stability simulations, are needed.

In addition to both simulations, we must note that power system simulations are typically focused on one of the areas of the hierarchical power grid: Power generation, Transmission, Distribution, or Utility (residential, commercial, and industrial loads).

Let's make a comparison – we will try to compare the performance of two smart network simulation tools RAPSim and GridLAB-D with a microgrid and a study of power flow calculations on a distribution network example.

Case 1 - microgrid is connected to the network and therefore operates in network-connected mode. In this mode, the microgrid can either sell or purchase excess energy into the utility network.

Case 2 - PV system has been added to study the effect of local generator on voltage.
In both cases, the two-day simulation was simulated and observed for an hour intervals. The purpose of the first study is to evaluate the effect of power distribution between power generators, the addition of renewable energy sources, and to calculate the amount of energy production and consumption, taking into account changes in production according to weather data. The electric generators reviewed for this work are the small 1500W wind turbine and the two small solar panels with the highest 500W power. In the second example, there is no power loss along the distribution lines, there is a feeder connected to two radial power lines connected to the houses.

3. Solution of The Problem

The main parameters of this work are:

- The nominal voltage in the feeder is 230 V for both simulators.
- One-phase distribution lines 150 meters away for each street. 60 m to the first point of contact of the first house, and the distance between the other access points is 30 m. Each house is joined by a 10m pipeline.
- In the sample, copper was used as a conductor at a temperature of 20 ° C at $1.724 \cdot 10{-}8\ \Omega$m.
- The cutting is 150 mm2, which corresponds to a typical feeder cable on the street. Therefore we get a resistance value of 0.115 15 / km or 0.185 Ω / mile. Linear inductance is not included in the model.
- The demand for 8 houses varies between 4 kW and 6 kW.

The first case in RAPSim: a number of objects are being added to the sample: three houses with constant load demand, one s / s, two solar panels and one wind turbine. These are placed in cages on RAPSim. Renewable energy sources, especially wind and solar energy, are highly dependent on weather conditions. Thus, air input is very important for the simulation. The wind speed is modeled using the Weibull probability density function and the cloud factor is modeled as a random model.
The second case in RAPSim: the idea of this work is to analyze voltage changes before and after adding solar panels. Homes with durable loads are being added to the feeder to imitate the same section. In the second part, the solar panel is added to Bus-8. The second example uses an AC-PF solver. With the addition of the solar panel, of course, the tension will change here within two days. In the first example where only consumers are connected to the feeder, the Bus-8 has a minimum voltage. In the second example, the PV panel causes an increase in voltage near the PV.
Let's look at the same cesas in GridLAB-D.
The first case in GridLAB-D is to use three loaded (triplex) objects instead of a single home to determine the value of the load requirement. In addition, some objects cannot be directly deployed in GridLAB-D compared to RAPSim. For example, an inverter object must be specified to place solar panels in GridLAB-D, as it is not possible to connect solar panels directly to triplex meters. In addition, all DC power-generating sources must have the same transformer as the host.
The second case in GridLAB-D is the study's resistance to transmission lines, and solar panels are added to one of these two lines to determine the effect of voltage on the triple meter.

4. Result

GridLAB-D is more widely used than RAPSim. RAPSim is designed for microgrid simulation, while GridLAB-D is more focused on the whole network (Figure 1). RAPSim provides hook visualization to accommodate eight types of objects: solar panels, wind turbines, houses, power lines, connectors, fuel generators, power plants, and network

connections. GridLAB-D offers many special objects for modeling power flows - capacitors, regulators, inverters, and meters. In addition, GridLAB-D does not have a GUI that complicates the application scenario and requires more user experience.

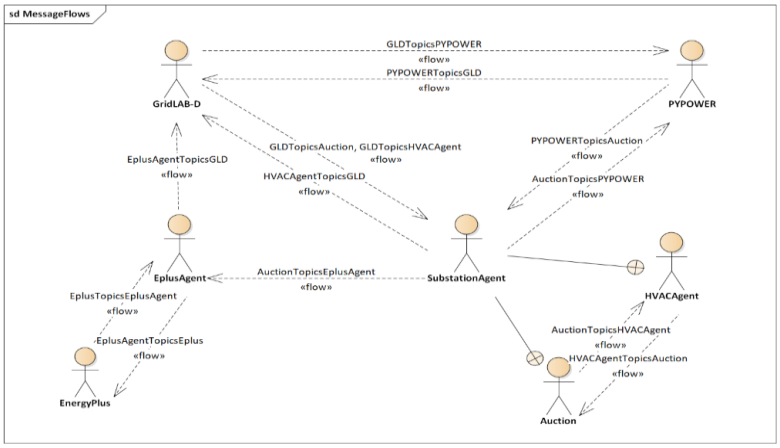


Figure 1. Message Flows for Simulators and Transactive Agents

The results in Figure 1 and Figure 2 show similar patterns, but the results are not the same as using different weather conditions. For the second example, both simulators were used to model the effect of adding solar panels in the scenario. For the second example, the configuration in GridLAB-D was more complex than that of RAPSim, as it required many handwritten codes to identify nodes and power lines. The properties of a home object in GridLAB-D do not have the option of setting a fixed load as much as possible in RAPSim. Demand for a GridLAB-D home is changing over time. Adding a constant load object is possible in GridLAB-D, or we can use a triplex-loaded object connected to a power flow, or assign a fixed power load using a zipload object.

Simulators have a number of similarities: They are both free and open source and offer the ability to simulate energy simulation, distributed and renewable energy sources. GridLAB-D is more consistent with studies that include modeling and simulating residential loads and optimizing energy flow. RAPSim can be considered more of a classroom environment, as this simulator teaches basic concepts for smart networking and microgrid simulation.

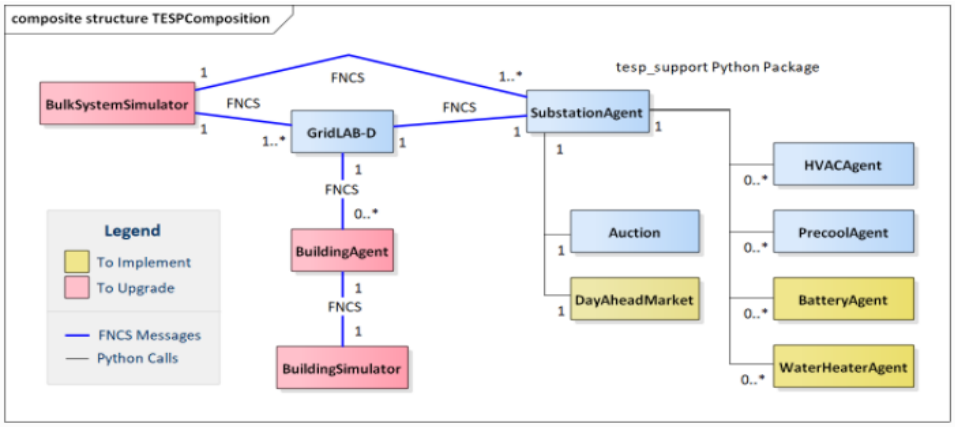


Figure 2. Composition of federates in a running TESP simulation

Some modeling components, such as an air module or the load behavior of an ordinary home, are different in two simulators, which means that the same scenario can produce different results. For example, in GridLAB-D, it was not possible to add a PV system like in RAPSim, because GridLAB-D is a more detailed, realistic model.

5. Conclusion

Energy management systems, when developed in a context Smart Grid, have their functions enhanced, in regards of the technologies making up the system. Among the main features of an energy management system, the following stand out as the most enhanced:

- Real-time performance monitoring;
- Information to compose predictive maintenance planning;
- Energy management;
- The efficiency of the power system;
- Financial economics, avoiding any sanction from the energy utilities;

Among the main Smart Grid technologies, the most important and that directly impacts the design of an energy management system are the solutions on cloud computing. Cloud computing has established itself as an adequate means to provide resources to customers, primarily in energy management systems, with access to a large amount of information and computer storage. With cloud computing, customers do not have to manage and maintain their own information technology (IT), and are not bound to its local resources which often are limited. However, for customers and energy utilities, making sure that your cloud services are usable, an appropriate level of guarantees of Quality of Service (QoS) is needed.

6. References

[1] Albadi, M.H. and El-Saadany, E.F. (2007) Demand Response in Electricity Markets: An Overview. Proceedings of the 2007 IEEE Power Engineering Society General Meeting, Tampa, 24-28 June 2007, 1-5.

[2] Bashir, A.K., Ohsita, Y. and Murata, M. (2015) Abstraction Layer Based Distributed Architecture for Virtualized Data Centers. Proceedings of the Sixth International Conference on Cloud Computing, GRIDs, and Virtualization, Nice, 22-27 March 2015, 62-67.

[3] Balijepalli, V.S.K.M., Pradhan, V., Khaparde, S.A. and Shereef, R.M. (2011) Review of Demand Response under Smart Grid Paradigm. Proceedings of the 2011 IEEE PES International Conference on Innovative Smart Grid Technologies-India, Kollam, 1-3 December 2011, 236-243. http://dx.doi.org/10.1109/iset-india.2011.6145388

[4] Combs, L. (2011) Cloud Computing for SCADA. http://www.controleng.com/single-ar

[5] Conway, G., Carcary, M. and Doherty, E. (2015) A Conceptual Framework to Implement and Manage a Cloud Computing Environment. Proceedings of the Sixth International Conference on Cloud Computing, GRIDs, and Virtualization, Nice, 22-27 March 2015, 138-142.

[6] Gatsis, N. and Giannakis, G.B. (2012) Residential Load Control: Distributed Scheduling and Convergence with Lost AMI Messages. IEEE Transactions on Smart Grid, 3, 770-786. http://dx.doi.org/10.1109/TSG.2011.2176518

[7] Park, K., Kim, Y., Kim, S., Kim, K., Lee, W. and Park, H. (2011) Building Energy Management System based on Smart Grid. Proceedings of the IEEE 33rd International Telecommunications Energy Conference, Amsterdam, 9-13 October 2011, 1-4.

[8] Meilvride, B. (2012) Will SCADA Envolve to the Cloud? http://real-timecloud.com/

Enhanced Semantic Segmentation of Microscopic Rock Images Using DeepLabV3+ with SE Attention and Superpixel-Aided Labeling

Hüseyin DERVİŞ[1] , İlhan AYDIN [2], Taha Kubilay ŞENER[3], Ayşe Didem KILIÇ [4]

[1,2,3]Department of Computer Engineering, Engineering Faculty, Firat University, Elazig, Turkey.
[4] Department of Geology Engineering, Engineering Faculty, Firat University, Elazig, Turkey
[1]232129116@firat.edu.tr, [2]iaydin@firat.edu.tr, [3]tksener@firat.edu.tr, [4]adkilic@firat.edu.tr
[1](ORCID: 0009-0008-2655-0007), [2](ORCID: 0000-0001-6880-4935), [3](ORCID: 0000-0003-2919-6521), [4](ORCID:0000-0002-6804-6764)

Abstract
Deep learning-based semantic segmentation offers great potential for the automated analysis of microscopic rock images. However, the success of these models depends on high-quality labeled datasets, which are laborious to create, and on tools that enable their practical use. This study presents an end-to-end system for the semantic segmentation of rock images, spanning from data preparation to model testing. As the first step of the workflow, a semi-automatic labeling interface based on the SLIC superpixel algorithm was developed to overcome the dataset creation bottleneck. In the second step, using the dataset prepared with this method, baseline U-Net and DeepLabv3+ architectures, along with their versions enriched with modern components like Transformer and Squeeze-and-Excitation (SE) blocks, were comparatively trained. Experimental results showed that the DeepLabv3+ model with a ResNet50 backbone and an SE block achieved the highest performance with a mean IoU (mIoU) of 63.90 Finally, to demonstrate the practical value of this research, a second user interface was developed to allow for easy testing of the trained models by domain experts. This holistic approach not only improves segmentation accuracy but also makes the entire research process more efficient and accessible.

Keywords: semantic segmentation, deep learning, microscopic images, superpixel labeling, user interface

1.Introduction

Minerals are naturally occurring, homogeneous, and predominantly solid inorganic substances with a specific chemical composition and a regular atomic (crystalline) structure. Considered the fundamental building blocks of the abiotic world, minerals combine to form rocks, and rocks, in turn, form larger-scale geological structures such as mountains and continents. Today, more than 2,000 mineral species have been identified, each with its unique characteristics known in detail. However, only a limited number of these minerals (approximately 12–15 species) constitute the main components of rocks, while a slightly larger number form economically important ore minerals. The vast majority of the remaining minerals are distributed as trace amounts in the Earth's crust and mantle. The study and classification of rocks is one of the core subjects of geological engineering, a significant field within the Earth sciences. Properties such as the mineral content, texture, and structure of rocks provide crucial information in many areas, including the formation of the Earth's crust, geological history, the search for natural resources, and engineering applications. Traditional rock analysis methods are time-consuming, costly processes that require expertise and experience. The examination of thin sections under a microscope, the identification of minerals based on their physical properties, and the classification of rock types are performed by experts. This process can be subject to lengthy analyses and errors stemming from human factors. Furthermore, advanced analytical methods such as SEM, XRD, and ICP-MS may not always be a preferred option due to their high costs.

In recent years, advancements in the fields of artificial intelligence and computer vision have presented a significant potential for transforming rock analysis processes. Deep learning-based models have shown great success in extracting information and classifying complex image data. These models can automatically detect and segment minerals in thin-section images and assist in classifying rock types. This approach can accelerate the analysis process, reduce costs, and minimize human error. Koh et al. [1] compared two instance segmentation models—Mask R-CNN and SOLOv2—for estimating mineral composition, grain size, and distribution, achieving up to 99.50% accuracy, particularly on volcanic rock types. In a related approach, residual networks were employed to classify rock slices with fine-grained textures. ResNet50 and ResNet101 models automatically extracted image features and achieved classification accuracies of 90.24% and 91.63%, respectively, on a dataset of 10,000 rock slice images. These results

[1]Corresponding author

highlight the strong potential of deep learning in detailed mineralogical and lithological analysis[2]. To overcome the classification challenges posed by the complex structure of rock thin-section images, data augmentation algorithms were utilized. The performance of three image enhancement algorithms—CutOut, MixUp, and CutMix—was tested in conjunction with a ResNet50 neural network. The results showed that the CutOut algorithm increased the model's accuracy by 1.3%, raising it to 93.39%, and demonstrated that data enrichment plays a significant role in improving model performance [3]. This study analyzes the impact of optimization algorithm and learning rate selection on the performance of deep learning models, using 2,634 thin-section images from three different rock types: metamorphic, sedimentary, and volcanic. It compared optimization algorithms such as Adam, SGD, and RMSprop with different learning rate reduction schedules. The findings indicated that Adam and RMSprop yielded more stable results compared to SGD, and that the cosine learning rate reduction mode performed better [4]. An integrated CNN-based method was proposed to classify geological rock types using petrographic thin sections captured under both plane-polarized light (PPL) and cross-polarized light (XPL). In this approach, PPL and XPL images were input separately and jointly into three identical Convolutional Neural Network (CNN) models, and their outputs were fused using a maximum likelihood strategy. The resulting Combined CNN (Con-CNN) approach achieved an overall classification accuracy of 89.97% across 13 different rock types [5]. Advanced deep learning techniques, including Vision Transformer (ViT), MixUp data augmentation, and Focal Loss, were employed to enhance model performance. The proposed approach achieved high accuracy in lithology classification, demonstrating the potential of modern automated analysis systems in petroleum geology [6]. A comparative analysis of advanced segmentation models—including DeepLabv3+, SegFormer, ConvNeXt, and Mask2Former—was conducted for rock type identification in thin-section images. The enhanced Mask2Former model achieved an accuracy of 91.26% and a mean Intersection over Union (mIoU) of 82.59%, representing an improvement of over 5% compared to the base model. Following segmentation, mineral classification using the Quartz-Feldspar-Plagioclase (QAP) diagram yielded an accuracy of 85.7% [7]. A YOLOv4-based deep learning model was employed to automate mineral identification and percentage estimation in thin-section analyses of sedimentary rocks. Forty-four samples were examined under plane-polarized (PPL) and cross-polarized (XPL) light, enabling a fast and practical method for detecting quartz, feldspar, and lithic fragments. The approach offers valuable support for geological interpretation by streamlining mineral analysis in petrographic studies[8]. A ResNet152-based CNN model was developed for the automatic classification of six igneous rock types from thin-section images, achieving over 90% accuracy. To ensure interpretability, explainable AI (XAI) techniques such as Grad-CAM were applied, revealing that the model focuses on mineral grains and matrix structures—mirroring human geological assessment. These findings highlight the reliability and transparency of deep learning methods in igneous rock classification [9]. To address the limitations of the traditional Gazzi-Dickinson point-counting method, U-Net and U-Net++ semantic segmentation networks were employed for the automatic analysis of sandstone thin-section images. Trained on thousands of manually labeled mineral components from hundreds of samples, the model achieved 87% accuracy for quartz and 77% for lithic fragments on the test set. This approach minimizes subjectivity and enables detailed, mineral-level classification of sandstones [10].

In recent years, deep learning-based semantic segmentation models such as U-Net and DeepLabv3+ have demonstrated strong performance by assigning each pixel in an image to a meaningful class. However, their practical application is limited by two key challenges: the need for extensive pixel-level annotations, which are time-consuming and labor-intensive to produce, and the lack of accessible tools for domain experts to utilize these models effectively. This study introduces an end-to-end system that not only compares state-of-the-art models for the semantic segmentation of microscopic rock images but also addresses these challenges through the development of a custom annotation tool to accelerate data labeling and an evaluation interface that allows users to easily test models and interpret results, thereby bridging the gap between research and practical application.

2. Materials and Methods

The proposed approach for identifying mineral types and classifying rocks from microscopic thin section images is designed as an end-to-end system comprising three integrated main stages. This process initially involves the efficient preparation of the dataset, followed by semantic segmentation using advanced deep learning models, and finally, geological classification based on the segmented results. The overall workflow of the system is illustrated in Figure 1.

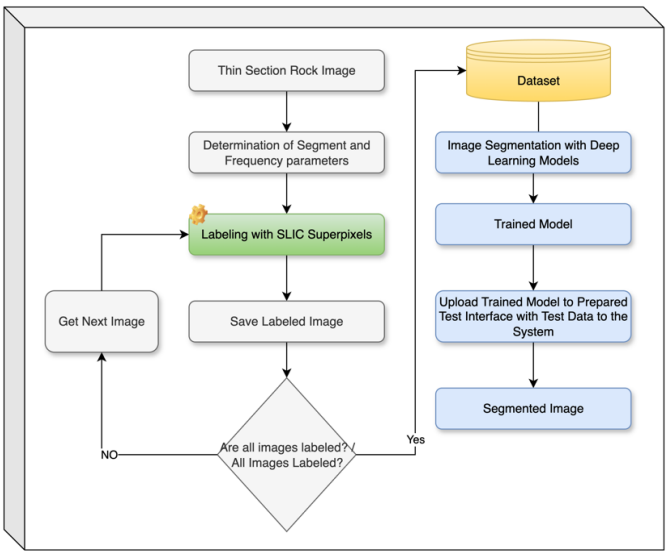


Figure 1. Proposed approach for rock type determination from thin section rock images.

The workflow begins with the creation of a high-quality, pixel-level labeled dataset, which is essential for training deep learning models. To address the inefficiency of manual labeling, a semi-automatic annotation tool was developed using Python and PyQt6. This tool employs the SLIC (Simple Linear Iterative Clustering) superpixel algorithm, which segments images into visually coherent regions based on color and spatial proximity. Its efficiency and accuracy in preserving object boundaries make it well-suited for this task. The user-friendly interface allows users to label entire superpixels with a single click and adjust parameters such as the number and compactness of superpixels to optimize segmentation. The resulting labeled images and masks form the dataset for the next phase. In the second stage, this dataset is used for semantic segmentation. Four deep learning architectures were evaluated for mineral segmentation in rock images. U-Net served as the baseline model due to its encoder-decoder structure and skip connections. To enhance global context understanding, a Transformer block was integrated into U-Net's bottleneck, forming the second model. The third model employed DeepLabv3+ with a ResNet50 backbone, known for effective multi-scale feature extraction via its ASPP module. Finally, a Squeeze-and-Excitation (SE) block was added to DeepLabv3+ to improve channel-wise feature recalibration, further enhancing performance.

2.1. Automated Labeling with the SLIC Superpixel Algorithm

To streamline the labeled data preparation process for deep learning models, the developed interface primarily utilizes the SLIC (Simple Linear Iterative Clustering) superpixel algorithm [14]. Superpixels are a pre-processing technique that segments an image into perceptually meaningful and coherent regions by grouping neighboring pixels. SLIC achieves this by considering not only the colors of the pixels but also their positions within the image plane. It accomplishes this by evaluating each pixel in a five-dimensional feature space, which includes both color information (in the CIELAB color space) and coordinate information (x, y position). One of the most crucial features of this approach is its ability to manage the inherent imbalance between color similarity and spatial proximity. Under normal circumstances, the physical distance between pixels can be on a much larger scale than the difference in their colors. SLIC resolves this by using a "compactness" parameter, which adjusts the relative weighting of these two factors. This parameter determines whether the algorithm will adhere more closely to true object boundaries in the image (by prioritizing color similarity) or to a more regular, grid-like structure (by prioritizing spatial proximity). This allows the user to obtain either more organic or more geometrically shaped superpixels, depending on their specific needs [20]. The algorithm's operation is based on an iterative process. First, initial cluster centers are uniformly placed across the image according to the target number of superpixels. Then, each pixel is assigned to its closest center based on a special distance definition that combines both color and positional information. Ultimately, the SLIC algorithm was chosen for this study due to its advantages, including its efficiency thanks to its speed, the high adherence of its

compact superpixels to object boundaries, and the control it offers users over the number and shape of superpixels. These characteristics make it highly suitable for segmenting thin section images, especially where the precise detection and labeling of mineral boundaries with varying colors and textures are crucial.

2.2. Models and Architectural Structures Used

In this study, several advanced deep learning architectures were employed for semantic segmentation tasks on thin-section rock images. The baseline model was U-Net, known for its encoder-decoder structure with skip connections, which preserves spatial information by directly linking corresponding layers in the encoder and decoder. To enhance its ability to capture long-range dependencies, a Transformer block was integrated into U-Net's bottleneck, combining local feature extraction with global contextual awareness through self-attention mechanisms. In addition, the DeepLabv3+ architecture was utilized for its robust multi-scale feature extraction capabilities, enabled by the Atrous Spatial Pyramid Pooling (ASPP) module. To further improve channel-wise feature representation, a Squeeze-and-Excitation (SE) block was added after the ASPP module. This integration enhances the model's focus on informative features by applying adaptive channel reweighting through global pooling and excitation mechanisms.

2.3.U-Net Model with Integrated Transformer Layer

The U-Net model is a widely used segmentation architecture based on an encoder-decoder structure [19], where the encoder extracts meaningful features and the decoder reconstructs a high-resolution segmentation map. Skip connections are employed to preserve spatial details lost during downsampling. To enhance its ability to capture long-range dependencies, Transformer layers were integrated into the U-Net architecture. Transformer-based models [17], with their attention mechanisms, effectively model both local and global contexts. This modification led to a notable improvement in distinguishing subtle differences between mineral classes, achieving a mean Intersection over Union (mIoU) of 55.19. In particular, the enhanced model yielded more accurate segmentation of rare mineral types. Figure 2 show the proposed U-Net model with transformer layer.

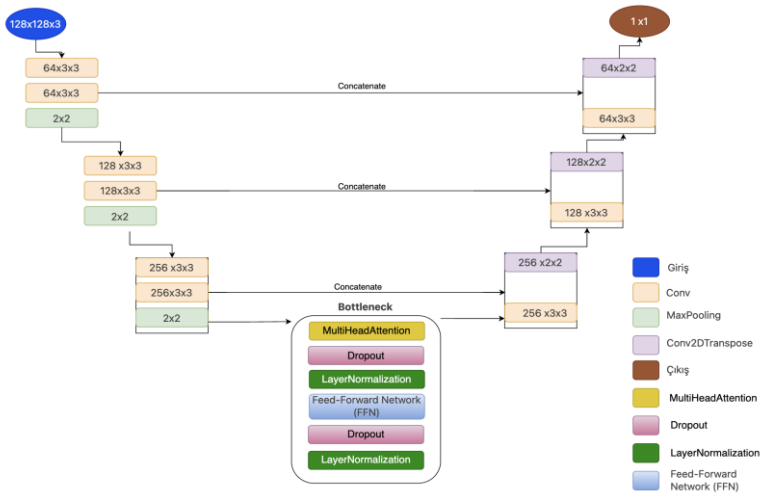


Figure 2. Architecture of the U-Net Model with Integrated Transformer Layer

In Figure 2, the primary enhancement in the proposed model is the integration of a Transformer-based bottleneck layer, featuring a Multi-Head Self-Attention (MHSA) mechanism. MHSA enables the model to capture long-range dependencies by attending to all spatial regions of the input image, thereby improving its ability to model global contextual relationships. This layer is further stabilized through the use of layer normalization and residual connections, which facilitate more consistent and effective learning. Additionally, a Feed Forward Network (FFN) refines the extracted features, enhancing their representational capacity, while a dropout rate of 0.6 is applied to mitigate overfitting. By combining the local feature extraction strength of the convolutional layers in U-Net with the

global attention capabilities of the Transformer, the model achieves improved performance in semantic segmentation tasks.

2.4. DeepLabV3+ Model with Squeeze-and-Excitation

ResNet50 is a deep convolutional neural network that employs residual connections to mitigate the vanishing gradient problem, thereby enabling the effective training of very deep models [12]. When integrated with the DeepLabv3+ architecture, it facilitates the extraction of both high-level semantic features and low-level spatial details, enhancing segmentation accuracy and reducing information loss. The incorporation of Squeeze-and-Excitation (SE) layers [13] further improves the model's performance by introducing channel-wise attention, which selectively emphasizes the most informative feature maps. This combined architecture has proven particularly effective in the precise delineation of mineral boundaries and in improving the segmentation accuracy of rare mineral classes. A detailed representation of the overall architecture is provided in Figure 3.

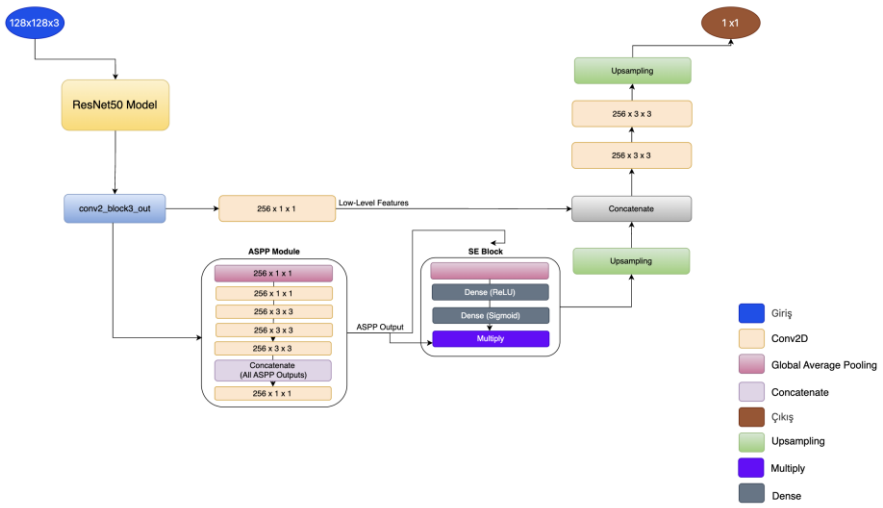


Figure 3. Architecture of the DeepLabV3+ Model with Squeeze-and-Excitation (SE)

DeepLabV3+ is a widely adopted semantic segmentation architecture that achieves high performance by leveraging the Atrous Spatial Pyramid Pooling (ASPP) module, which captures multi-scale contextual information through convolutions with varying dilation rates. To further enhance its representational capacity, a Squeeze-and-Excitation (SE) block was integrated into the architecture, introducing a channel-wise attention mechanism that adaptively recalibrates feature responses. The SE block operates in three stages: (i) the squeeze step applies global average pooling to generate a compact channel descriptor; (ii) the excitation step uses fully connected layers with sigmoid activation to learn the relative importance of each channel; and (iii) the scaling step reweights the original feature maps based on the learned attention scores. By inserting the SE block after the ASPP module, the model effectively emphasizes the most informative channels in the multi-scale feature representation, leading to improved segmentation accuracy—especially in tasks requiring fine-grained discrimination, such as mineral boundary detection. The complete architecture consists of a ResNet50 backbone for deep feature extraction, the ASPP module for capturing rich contextual information, the SE block for enhancing channel-level feature importance, and a decoder module that fuses low-level and high-level features to generate precise segmentation maps.

2.5. Performance Evaluation Metrics

To evaluate the performance of the semantic segmentation models, standard evaluation metrics were employed, considering the visual similarity among minerals (e.g., quartz, alkali feldspar, plagioclase), ambiguous boundaries, and class imbalance in the dataset. Both overall accuracy and class-wise overlap metrics were used to provide a

comprehensive evaluation. Pixel-based overall accuracy, defined in Equation (1), measures the proportion of correctly classified pixels.

$$Accuracy = \frac{(TP + TN)}{TP + TN + FP + FN} \quad (1)$$

To address the limitations of overall accuracy and better assess segmentation performance, Intersection over Union (IoU was adopted as the primary evaluation metric. IoU quantifies the geometric overlap between the predicted region and the ground truth for each class, offering a more detailed evaluation by incorporating both false positives and false negatives. This metric not only reflects the correctness of pixel classification but also captures the precision of segmentation boundaries. The IoU is computed as shown in Equation (2).

$$IoU = \frac{TP}{TP + FP + FN} \quad (2)$$

To provide a single, comprehensive measure of model performance, Mean Intersection over Union (mIoU) was employed. mIoU is calculated as the arithmetic mean of IoU scores across all classes, assigning equal importance to each class regardless of its frequency in the dataset. This property makes mIoU particularly robust against class imbalance, establishing it as the primary performance metric in this study. The formula for mIoU is presented in Equation (3).

$$mIoU = \frac{1}{N} + \sum_i \blacksquare \left(\frac{TP_i}{TP_i + FP_i + FN_i} \right) \quad (3)$$

3.Experimental Results

A total of 3,654 petrographic thin-section images, prepared at the Fırat University Geological Engineering Laboratory, were used in this study. To address the challenge of generating pixel-level annotations required for deep learning, a custom labeling tool was developed using Python and PyQt6, incorporating the SLIC (Simple Linear Iterative Clustering) superpixel algorithm. SLIC clusters neighboring pixels based on color and spatial proximity into perceptually coherent superpixels, allowing users to label entire regions with a single click. This approach significantly reduces annotation time and improves label consistency. The interface allows dynamic adjustment of superpixel count and compactness, class selection, and real-time mask visualization. The dataset was partitioned into 70% for training and 30% for validation/testing. The SLIC based labelling GUI tool is given in Figure 4.



Figure 4. GUI-based labeling tool.

The developed semi-automatic labeling interface streamlines the data preparation process through a user-friendly and workflow-oriented design. Users begin by selecting the image folder via the interface, after which the selected image

is displayed and segmented into superpixel regions using the SLIC (Simple Linear Iterative Clustering) algorithm. Adjustable sliders allow users to modify the number and compactness of superpixels, enabling fine control over segmentation granularity based on mineral grain size. Labeling is performed by selecting a predefined mineral class (e.g., Quartz, Alkali Feldspar, Plagioclase) and clicking on the corresponding superpixel, which is instantly updated in the mask preview. This one-click labeling of entire regions greatly accelerates the annotation process compared to pixel-level manual labeling. Completed masks can be saved both in grayscale format (for model training) and in color (for visual inspection). Figure 5illustrates sample original images alongside their labeled counterparts.

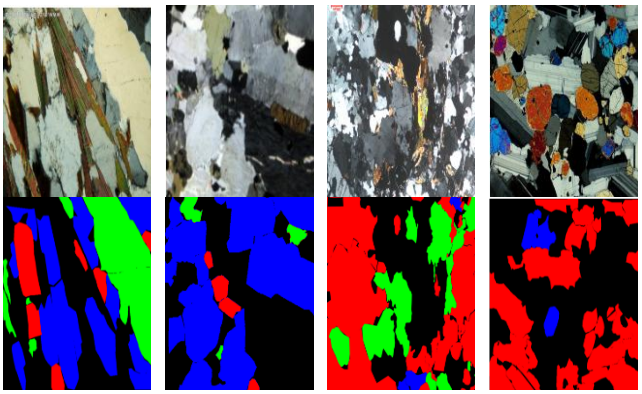


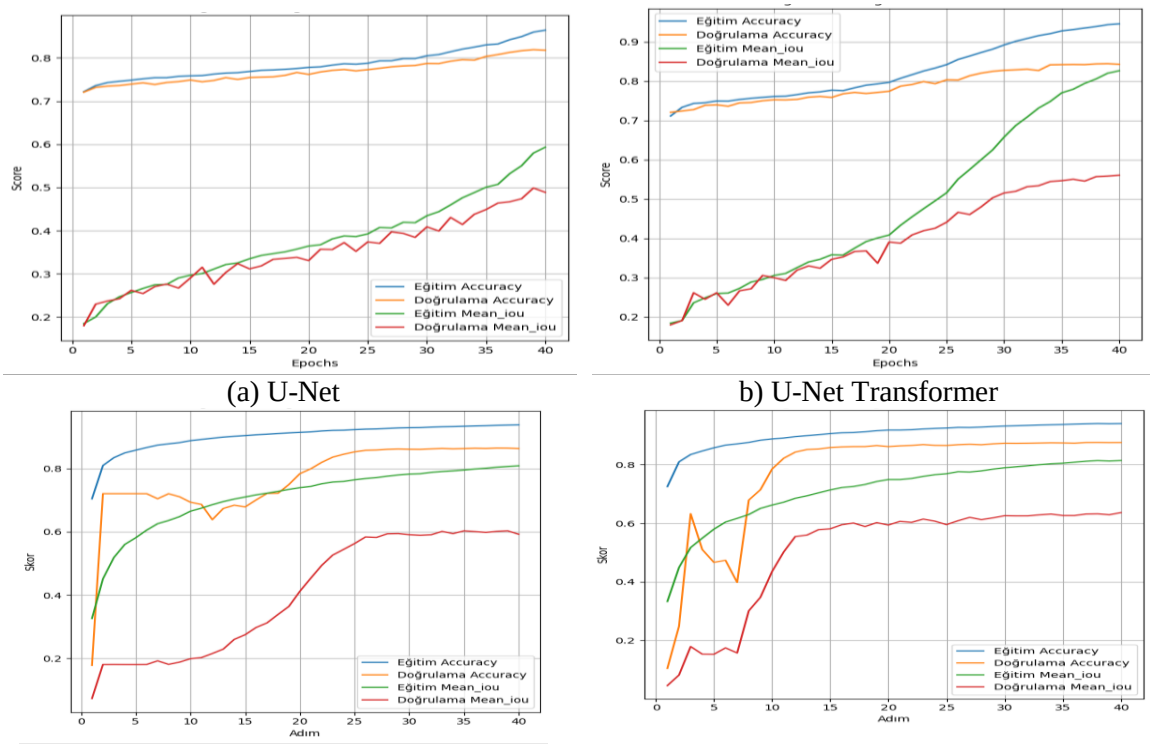
Figure 5. Four Examples from the Dataset and Their Labeled Counterparts

Following the labeling process, semantic segmentation of the images was performed. For this purpose, four different models were employed: the standard Convolutional Neural Network (CNN)-based DeepLabV3+, DeepLabV3+ integrated with a Squeeze-and-Excitation (SE) block, the baseline U-Net, and U-Net enhanced with a Transformer layer. These models were selected to evaluate the impact of architectural enhancements on segmentation performance. The training parameters used for each model are summarized in Table 1.

Table 1. Training Parameters for Segmentation Models

Parameter	Value
Data Size	128x128
Number of Classes	4
Batch Size	32
Learning Rate	1e-4
Weight Decay	0.01
Optimization Algorithm	AdamW
Epochs	40

Each model was trained according to these metrics, yielding model graphs for training and validation. Figure 6 presents the accuracy and mIoU graphs for both training and validation.



(a) U-Net b) U-Net Transformer

(c) DeepLabV3+ (d) DeepLabV3 + SE Layer

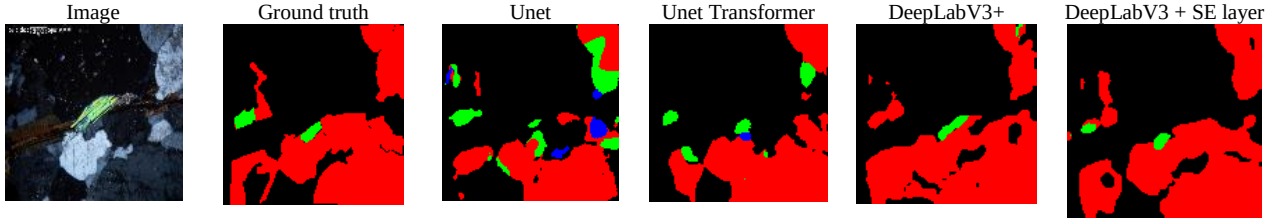
Figure 6. Training graphs for different models

The training graphs presented in Figure 6 clearly indicate that the DeepLabV3+ with SE model achieved superior results for both training and validation compared to the other models. The detailed performance outcomes for all four models are summarized in Table 2.

Table 2. Metrics for the four models during training and validation phases

Metric	Phase	Unet	Unet Transformer	DeepLabV3	DeepLabV3 + SE layer
Accuracy	Train	86.34	94.02	93.86	**94.55**
MIoU		58.42	81.42	80.91	**81.89**
Accuracy	Validation	81.80	84.27	86.36	**87.55**
MIoU		48.90	55.19	59.24	**63.90**

Table 2 clearly indicates that the DeepLabV3+ with SE layer achieved superior results across all metrics compared to the other three models. Furthermore, the segmentation results for the proposed models are presented in Figure 7.



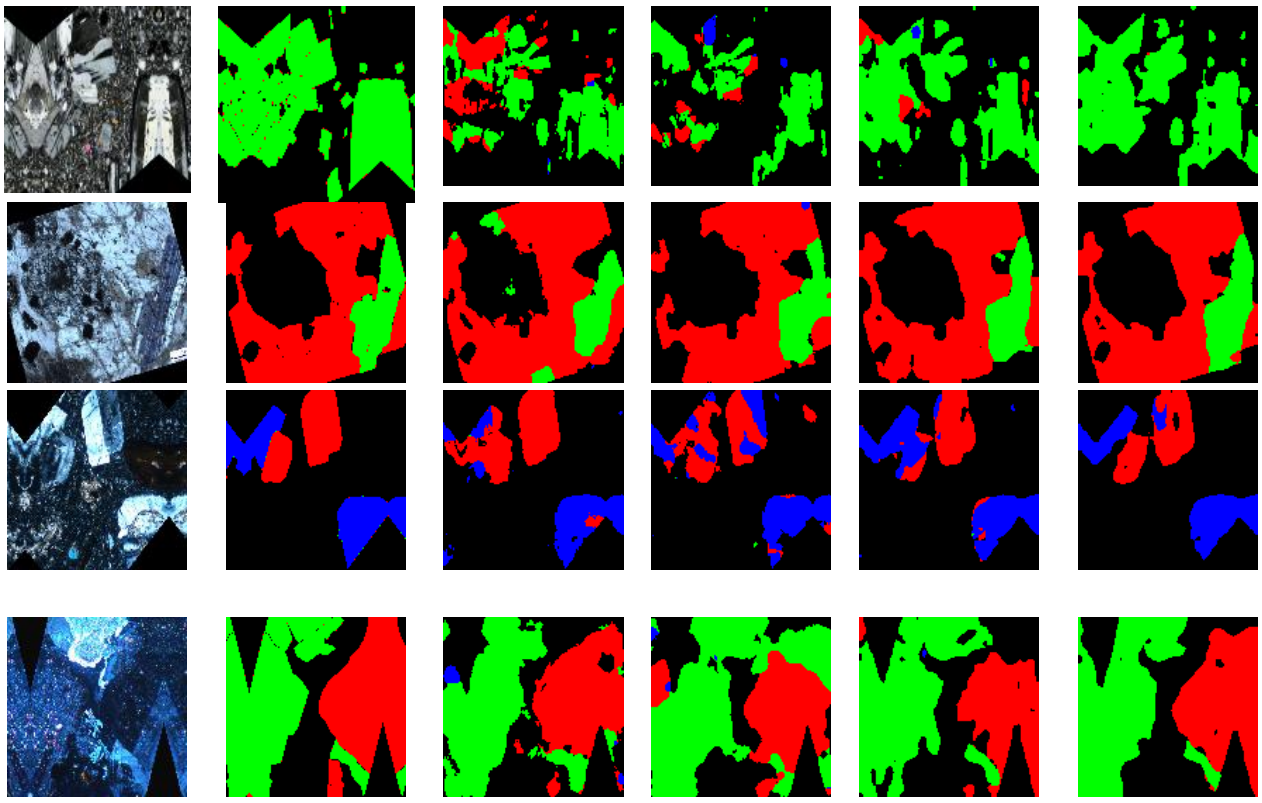


Figure 7. Segmentation results of UNet, UNet Transformer, DeepLabV3, and DeepLabV3 + SE Layer models on different images

Figure 7 provides a comparative visualization of the mineral classification performance of the UNet, DeepLabV3+, and the developed DeepLabV3+SE models. Visual analyses clearly indicate that the DeepLabV3+SE model demonstrates the most superior performance in terms of class discrimination and regional consistency. This model has proven its ability to preserve fine details by accurately delineating the boundaries of minerals such as quartz, plagioclase, and alkali feldspar with high precision [18]. While the UNet model yielded acceptable segmentation results in some instances, its performance was limited, particularly due to blurriness and fragmented structures at mineral boundaries. DeepLabV3+, on the other hand, tended to over-generalize certain classes and showed weak performance in preserving local textures. In contrast, the DeepLabV3+ architecture with an integrated SE (Squeeze-and-Excitation) layer produced sharper boundaries and more consistent segmentation results. This was achieved by effectively balancing both global contextual information and local details. Consequently, the proposed DeepLabV3+SE model offers an optimized solution for mineralogical segmentation in thin section images, demonstrating higher accuracy and visual consistency compared to other architectures. Another significant outcome of this study is the development of a Graphical User Interface (GUI), which allows for the practical application of the trained models. This interface enables users to load different models, perform segmentation operations on test data, and visualize the results. Additionally, the models' performance metrics (mIoU, test loss, accuracy) can be calculated and recorded. Figure 8 shows a screenshot of the interface.

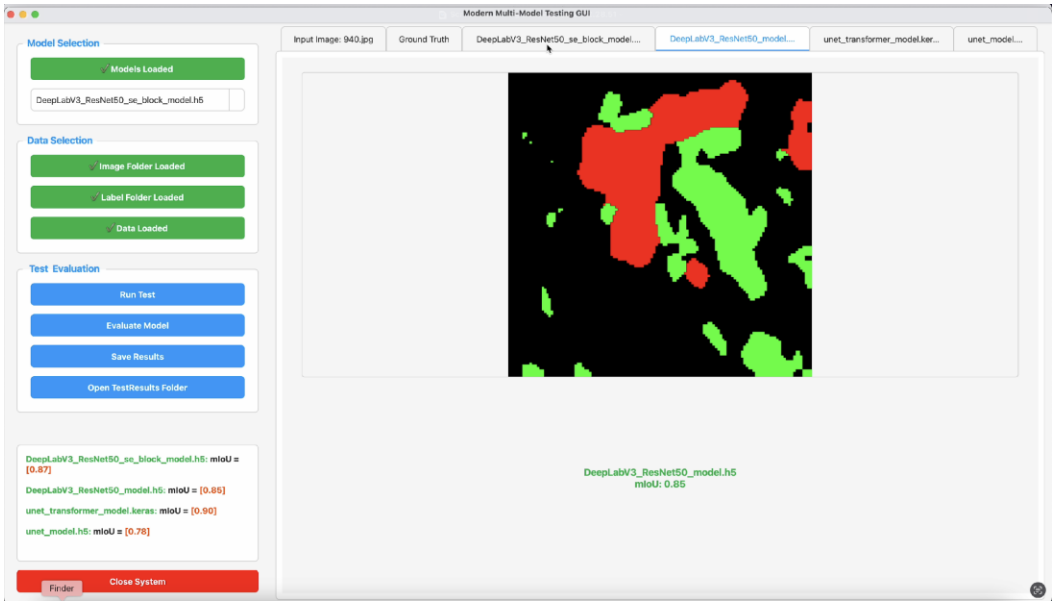


Figure 8. Screenshot of the GUI-based testing Interface for trained models

The developed user interface offers a comprehensive environment for model evaluation and visualization. Users can load trained models (in .h5 or Keras format), import test images and corresponding labels, and perform segmentation with a single click. The interface displays segmentation results alongside input images and ground truth masks using color-coded overlays. Model performance is quantitatively evaluated using the mean Intersection over Union (mIoU) metric, and the results can be saved to a user-defined directory. The interface streamlines the entire workflow through intuitive buttons for model selection, data loading, testing, evaluation, and result export.

4. Conclusions

This study proposed an end-to-end system for the semantic segmentation of microscopic rock images, focusing on three mineral types: Quartz, Alkali Feldspar, and Plagioclase. Base models such as U-Net and DeepLabV3+ were enhanced with Transformer layers and SE (Squeeze-and-Excitation) blocks to improve segmentation accuracy. Among all models, DeepLabV3+ with a ResNet50 backbone and SE block achieved the best performance, reaching 63.90 mIoU. The results show that attention-based enhancements significantly improve the detection of mineral boundaries and model accuracy. In addition, a semi-automatic labeling tool using the SLIC algorithm and a user-friendly testing interface were developed to streamline data preparation and model evaluation. Future work will focus on expanding the dataset, improving generalization through transfer learning, and addressing the segmentation of rare minerals.

Acknowledgements

This study was supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK) under project number 123E368.

5. References

[1] Polat, Ö., Polat, A., & Ekici, T. (2021). Automatic classification of volcanic rocks from thin section images using transfer learning networks. Neural Computing and Applications, 33(18), 11531-11540.

[2] Guojian, C., & Peisong, L. (2021, April). Rock thin-section image classification based on residual neural network. In 2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP) (pp. 521-524). IEEE.

[3] Huaizhou, Y., & Danyang, X. (2021, July). Research and Analysis of Image Enhancement Algorithm in the Classification of Rock Thin Section Images. In 2021 3rd International Conference on Intelligent Control, Measurement and Signal Processing and Intelligent Oil Field (ICMSP) (pp. 125-128). IEEE.
[4] Li, D., Zhao, J., & Ma, J. (2022). Experimental studies on rock thin-section image classification by deep learning-based approaches. Mathematics, 10(13), 2317.
[5] Cheng, S., Sheng-jia, X., Kong-yang, Z., & Xiao-can, Z. (2020). Rock classification in petrographic thin section images based on concatenated convolutional neural networks. Earth Science Informatics, 13(4), 1477-1484.
[6] Camargo, A., Araujo, G., Sanchetta, A., Kohler, M., Becker, M., Pacheco, M., Fernandes, T., & Coutinho, B. (2024). Lithology Identification in Carbonate Thin Section Images of the Brazilian Pre-Salt Reservoirs. 2024 IEEE Latin American Conference on Computational Intelligence (LA-CCI), 1-6.
[7] Aydin, I., Kılıç, A.D., & Şener, T.K. (2025). Improving Rock Type Identification Through Advanced Deep Learning-Based Segmentation Models: A Comparative Study. *Applied Sciences*.
[8] Pratama, B.G., Qodri, M.F., & Sugarbo, O. (2024). Image-based yolov4 architecture for detecting mineral in sedimentary rocks thin section. asean Engineering Journal.
[9] Seo, W., Kim, Y., Sim, H., Song, Y., & Yun, T.S. (2022). Classification of igneous rocks from petrographic thin section images using convolutional neural network. Earth Science Informatics, 15, 1297 - 1307.
[10] Zhang, B., Jin, X., Xiao, Y., Hou, Y., Meng, J., Huang, Z., & Han, M. (2023). Quantitative Identification of Sandstone Lithology Based On Thin-Section Micrographs Using the U-net and U-net++ Semantic Segmentation Network. Day 1 Wed, March 01, 2023.
[11] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention (MICCAI) (s. 234-241). Springer, Cham.
[12] Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. Proceedings of the European conference on computer vision (ECCV) (s. 801-818).
[13] Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR) (s. 7132-7141).
[14] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., & Süsstrunk, S. (2012). SLIC superpixels compared to state-of-the-art superpixel methods. IEEE transactions on pattern analysis and machine intelligence, 34(11), 2274-2282.
[15] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems (NeurIPS), 30.
[16] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR) (s. 770-778).
[17] Xie, E., Wang, W., Yu, Z., Anandkumar, A., J. Alvarez, P. Luo, "SegFormer: Simple and efficient design for semantic segmentation with transformers.", Advances in Neural Information Processing Systems, 34, 12077-12090, 2021.
[18] Seo, W., Kim, Y., Sim, H., Song, Y., & Yun, T.S. (2022). Classification of igneous rocks from petrographic thin section images using convolutional neural network. Earth Science Informatics, 15, 1297 - 1307.
[19] Zhang, B., Jin, X., Xiao, Y., Hou, Y., Meng, J., Huang, Z., & Han, M. (2023). Quantitative Identification of Sandstone Lithology Based On Thin-Section Micrographs Using the U-net and U-net++ Semantic Segmentation Network. Day 1 Wed, March 01, 2023.
[20] Han, Y., & Liu, Y. (2024). Intelligent Classification and Segmentation of Sandstone Thin Section Image Using a Semi-Supervised Framework and GL-SLIC. Minerals.

Production of Functionally Graded Composite Materials by Vacuum Infusion Method and Investigation of Their Impact Behavior

Ayşenur BAKIRCI [1], Mustafa GUR, Mete Onur KAMAN, Serkan ERDEM, Mehmet Fatih SAHAN

[1,2]Mechanical Engineering Department, Engineering Faculty, Fırat University, Elazig, Türkiye.
[3,4]Mechanical Engineering Department, Engineering Faculty, Fırat University, Elazig, Türkiye.
[5]Civil Engineering Department, Engineering Faculty, Adıyaman University, Adıyaman, Türkiye.
[1]aysenurbkrc30@gmail.com, [2]mgur@firat.edu.tr, [3]mkaman@firat.edu.tr, [4]serdem@firat.edu.tr, [5]mfs@adiyaman.edu.tr
[1](ORCID: 0000-0001-7274-6854), [2](ORCID: 0000-0002-4761-0215),[3](ORCID: 0000-0003-0178-6079),
[4](ORCID: 0000-0002-2504-0135), [5](ORCID: 0000-0002-3334-3945)

Abstract

In this study, functionally graded composite materials (FGMs) were produced using the vacuum infusion method, and their impact resistance was experimentally investigated. Functional grading was achieved by fabricating samples using glass fiber fabrics with different densities. This production approach aimed to improve mechanical performance in specific directions.
During the manufacturing process, the vacuum infusion method was preferred to ensure homogeneous resin distribution, resulting in uniform and high-quality samples. Low-velocity impact tests were conducted on the produced specimens, and displacement, force, and energy behaviors were evaluated. The results were compared with those of composite materials reinforced with glass fibers of constant density. According to the findings, it was observed that the variation in interlayer structure through the thickness could be adjusted to achieve the desired impact behavior.

Keywords: Functionally graded materials, vacuum infusion, composite structures, impact test, energy absorption

1. Introduction

With the advancement of modern technology, the utilization of composite materials has become increasingly widespread. Due to their lighter weight and greater versatility compared to metals, composite materials are preferred in various engineering applications. Achieving flawless, high-strength performance from a single material is challenging; therefore, the integration of multiple materials has been pursued to produce products with the desired properties. Composite materials possess numerous advantages, including low density, ease of formability, corrosion resistance, moisture resistance, lower cost compared to metals, environmental friendliness, thermal resistance, and high mechanical strength. Research continues to address and mitigate their disadvantages.

A composite typically consists of a matrix phase and a reinforcement phase. The matrix phase maintains the shape of the structure, transfers externally applied loads to the reinforcement phase, and acts as a binder. The reinforcement phase plays a critical role in determining key performance properties such as mechanical strength, stiffness, and impact resistance. Common reinforcement materials include glass, carbon, and aramid fibers, while matrix materials are often polymer-, metal-, or ceramic-based. Beyond offering lightness and strength, composites can be tailored according to the specific requirements of their intended application, enabling them to perform multiple functions simultaneously. The growing interest in composite materials spans diverse industries such as aerospace, space technology, automotive, marine engineering, wind energy, and construction. These materials contribute significantly to sustainable engineering solutions. Their flexibility in production, coupled with the energy and material savings they offer during manufacturing, ensures that composites will continue to play a vital role both in the present and in the future.

In recent years, functionally graded composite materials (FGM's) have been developed to more accurately evaluate the performance of structural materials. FGM' s are a type of heterogeneous material whose mechanical properties vary continuously from one point to another. This gradual variation allows the integration of different material properties within a single structure while minimizing stress concentrations and the resulting structural weaknesses that may arise from abrupt changes in material composition. FGM' s are typically produced by gradually combining ceramics, metals, or reinforcement elements of varying densities [1, 2]. Due to their superior mechanical and thermal performance, these materials demonstrate exceptional resistance to impact, wear, and high-temperature conditions [2]. Conventional

[1]Corresponding author

multilayer composites are manufactured by bonding different materials in discrete layers. However, in such structures, abrupt transitions between layers often lead to the initiation of cracks, stress concentrations, or delamination. In contrast, FGMs provide a smooth transition between layers, thereby reducing the risk of delamination and enabling a more uniform distribution of loads throughout the structure [3]. Functionally graded composite materials are specifically designed to address and improve the shortcomings of laminated composites, as well as to control energy dissipation. Under impact loading, FGM' s often exhibit greater energy absorption capacity compared to conventional composites.
Damage mechanisms such as matrix cracking, interlaminar delamination, and other internal failures that may occur under sudden service loads provide critical insights for damage detection and evaluation. Therefore, a detailed analysis of the impact behavior of both conventional composites and FGM' s is crucial for ensuring reliability and durability. Such improvements help preserve mechanical properties, enhance structural integrity, and extend the service life of the material. Moreover, FGM' s offer notable advantages in resistance to impact and wear, while enabling region-specific functional tailoring within the same structure. These materials can be manufactured using various techniques, including vacuum infusion, powder metallurgy, and layered fabrication methods.

Among the studies conducted in this field, Pai et al. [4] fabricated five different fiber metal laminate (FML) stacking sequences using AA6061-T6 aluminum, aramid, ultra-high molecular weight polyethylene (UHMWPE), and paperboard. The stacking sequences were compared based on their density distribution, and the FGM's were evaluated through flexural and tensile testing. Ejeh et al. [5] investigated the low-energy impact damage behavior of composite sandwich panels with functionally graded lattice cores by numerically modeling them using the finite element method. The cores were composed of minimal-surface lattice structures manufactured from 316L stainless steel via three-dimensional (3D) printing, while the face sheets consisted of prepreg composite layers. Functional gradation was achieved through variations in cell size, relative density, and periodic structure. Izadi et al. [6] proposed a novel fabrication method for multilayer FGM's composed of randomly oriented short fibers with different fiber volume fractions. The influence of fiber volume fraction on fiber bridging at crack tips and on the critical strain energy release rate was examined using both experimental methods and finite element analysis. Sain et al. [7] studied the physical and mechanical properties of glass fiber-reinforced, vinyl ester-based functionally graded polymer composites. Homogeneous composites were produced via the hand lay-up method, while FGM's were manufactured using the centrifugal casting technique. Mechanical properties such as tensile strength, hardness, flexural strength, and impact resistance were analyzed. Chama et al. [8] aimed to reinforce notched aluminum plates with FGM's. Linear-graded and functionally graded patches were compared with homogeneous patches, and the influence of fiber density on crack propagation was evaluated. Gunoz et al. [9] produced FGM plates using Kevlar fabric, carbon fabric, and an epoxy matrix reinforced with polyamide nanofibers and multi-walled carbon nanotubes (MWCNTs). Twelve-layer nanocomposite plates were manufactured through vacuum-assisted hand lay-up and subsequently subjected to impact testing. Suresh [10] examined the deformation and mechanical resistance of FGMs under surface contact loading conditions.

It has been emphasized that gradually varying properties such as elastic modulus, hardness, and composition through the thickness of a material can reduce stress concentrations, thereby delaying crack initiation and propagation. The study was not based on an experimental manufacturing process but rather on theoretical and applied data compiled from various literature sources. Within this scope, different examples of FGM' s, such as ceramic–metal and glass–ceramic transitional systems, were evaluated; however, no single manufacturing method was employed. Degnah et al. [11] compared two Kevlar composite materials: one without reinforcement and the other reinforced with Al_2O_3, increasing by 3% in each layer to form an FGM. Both specimens were subjected to impact testing, and the results were validated through numerical simulations of the impact tests conducted in ANSYS. Kumar et al. [12] carried out both experimental and numerical studies to optimize the impact resistance of FGM sandwich structures. In low-velocity impact testing, the face sheets consisted of jute/rubber/jute layers, while the core material was an epoxy and sea sand mixture. The performance of the sandwich structures was evaluated using parameters such as maximum damage force and percentage of energy loss. Demir et al. [13] proposed the design of an FGM glass/carbon fiber-reinforced hybrid polymer composite with a specific layer sequence. Tensile, flexural, and impact tests were performed to assess the mechanical behavior of the composite material. Multi-walled carbon nanotubes were incorporated into the FGM composites, and their influence on the structural behavior was examined. Patel [14] conducted a detailed investigation of the behavior of FGM' s under ballistic impact, focusing particularly on metal–ceramic and metal–polymer-based FGM systems. These materials were generally designed to have a ductile, energy-absorbing surface on one side and a hard, high-strength component on the other, with a gradual transition between the two. The samples were manufactured using powder metallurgy, hot pressing, and 3D printing. Alhazmi et al.

[15] studied the mechanical and tribological behavior of FGM unidirectional glass fiber-reinforced epoxy composites. The composites, produced by the hand lay-up method, consisted of either four or ten layers, with each layer containing a different glass fiber volume fraction in the FGM configuration. Mechanical strength was assessed through a small abrasive paper test, while wear behavior was examined using the pin-on-disc test. In addition, finite element simulations based on the Hashin failure criterion and the cohesive zone model were performed to better characterize the flexural behavior of such composites.

The aim of this study, in contrast to previous works in the literature, is to investigate the impact energy absorption performance of composite plates fabricated from woven glass fiber fabrics of various areal densities. Accordingly, the specimens were manufactured with different stacking configurations, including a homogeneous arrangement, a sequence from the highest to the lowest areal density, and a sequence from the lowest to the highest areal density. The force–time, force–displacement, and energy–time data obtained from low-velocity impact tests conducted on these configurations were compared to evaluate the effect of layer areal density distribution on the impact behavior.

2. Materials and Method

In the present study, woven glass fiber fabrics of different areal densities were laminated into composite plates using the vacuum infusion method, and their energy absorption behavior under impact loading was investigated. In the production of FGM's, plain-weave woven glass fiber fabrics with five different areal densities 200, 100, 86, 49, and 25 g/m² were used as the reinforcement material. All specimens consisted of 10 layers and were divided into four groups according to the arrangement of fabric areal densities: I) Group 1 specimens were stacked from the highest areal density fabric (200 g/m²) to the lowest areal density fabric (25 g/m²), with each density repeated for two consecutive layers, II) Group 2 specimens were stacked from the lowest areal density fabric (25 g/m²) to the highest areal density fabric (200 g/m²); III) Group 3 specimens consisted solely of homogeneous layers made from the 200 g/m² fabric IV) Group 4 specimens consisted solely of homogeneous layers made from the 25 g/m² fabric.

Prior to manufacturing on the vacuum infusion system, the glass fiber fabrics were cut to dimensions of 400 × 500 mm and neatly stacked to form a dry fabric lay-up with a total of 10 layers. The laminates were fabricated on a temperature-controlled vacuum infusion table. The produced plates were then cut into 100 × 100 mm specimens to fit the impact test apparatus (Table 1). Impact tests were conducted at the Mechanical Laboratory of the Department of Civil Engineering, Adıyaman University, using an Instron Ceast 9350 testing machine with an energy capacity of 800 J (Figure 1). All specimens were subjected to an impact energy of 2.5 J (Table 2).

Table 1. Composite specimens and layer configurations.

Group	Number of layers	Fabric areal density sequence (g/m²)
I	10	25/25/25/25/25/25/25/25/25/25
II	10	25/25/49/49/86/86/100/100/200/200
III	10	200/200/100/100/86/86/49/49/25/25
IV	10	200/200/200/200/200/200/200/200/200/200

Table 2. Dimensions of the composite specimens.

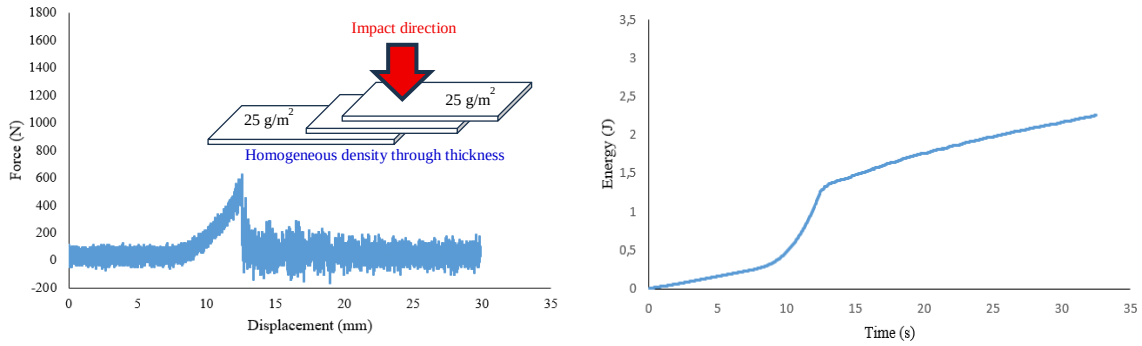
Group	Weight	Total specimen thickness (mm)	Fabric areal density sequence (g/m²)
I	5	0,40	25/25/25/25/25/25/25/25/25/25
II	16,2	1	25/25/49/49/86/86/100/100/200/200
III	16,2	1	200/200/100/100/86/86/49/49/25/25
IV	30,4	1,80	200/200/200/200/200/200/200/200/200/200



Figure 1. Impact testing device

3. Results and Discussion

For the functionally graded Group I composite specimen fabricated homogeneously with a fabric areal density of 25 g/m², the force–displacement and energy–time graphs obtained from the impact test are presented in Figure 2. According to the force–displacement graph (Figure 2a), the force remains at low levels during the initial phase (0–10 seconds). After approximately 11 seconds, the force begins to increase, indicating that the material exhibits elastic behavior under the impact load and that its load-bearing capacity progressively improves. Following the peak point of the graph, the force drops abruptly, resulting in an irregular, oscillatory response pattern.



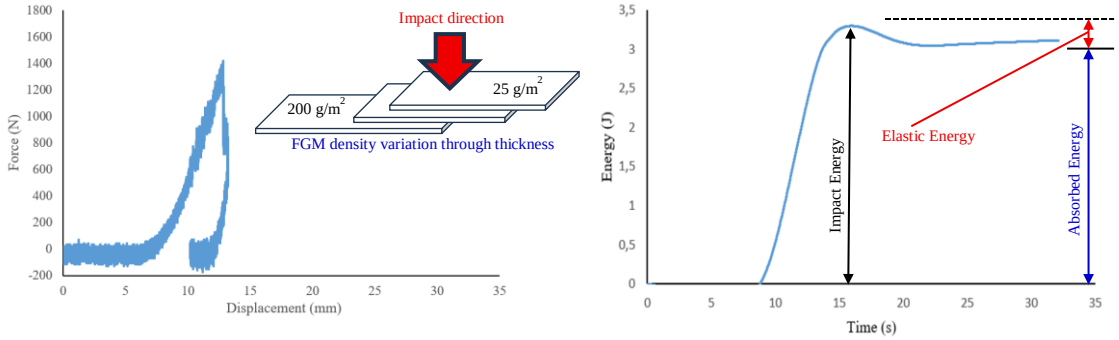
a) b)

Figure 2. For Group I specimens: a) Force–displacement graph and b) Energy–time graph.

The gradual decrease and oscillation of the force indicate that, beyond this point, the material begins to dissipate energy and that permanent deformation damage likely occurs. According to this graph, the attainment of high force levels demonstrates good impact resistance of the material, whereas the irregular distribution of the force following the impact suggests that the energy dissipation process is oscillatory and unstable. The test results show that the elastic limit of the material was exceeded, leading to permanent deformation and eventual fracture after plastic deformation.

According to the energy-time graph (Figure 2b), a sudden increase in energy is observed starting at approximately 11 seconds, corresponding to the moment of impact. This sharp rise indicates that the material absorbs energy rapidly. After around 13-14 seconds, the rate of energy increase slows down, which is attributed to the penetration of the striker and its frictional movement within the material. The graph terminates at approximately 2.5 J, indicating that the entire applied impact energy was absorbed.

The force displacement and energy-time graphs obtained from the impact test of the Group II functionally graded composite specimen fabricated by stacking layers from the lowest areal density fabric (25 g/m²) to the highest areal density fabric (200 g/m²) are presented in Figure 3.



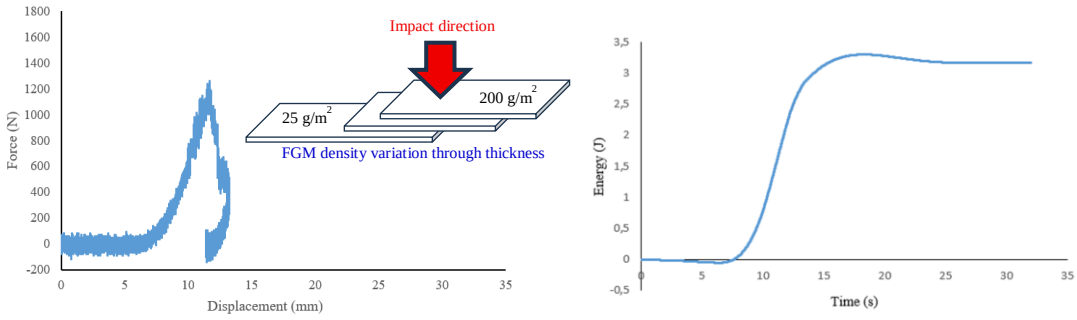
a) b)

Figure 3. For Group II specimens: a) Force-displacement graph and b) Energy-time graph.

Initially, in the force-displacement graph (Figure 3a), within the displacement range of 0-8 mm, the force remains low as displacement increases, indicating that the specimen does not undergo deformation in this region. Between 8 mm and 12 mm, the force increases linearly, reaching a maximum value of approximately 1400 N. Pai et al. [4] reported that when low-density materials are placed at the beginning of the stacking sequence followed by increasing density, improved mechanical performance and energy dissipation can be achieved. After 12 mm, a sudden drop in force is observed, suggesting that the material has sustained damage or undergone structural weakening. According to the graph, the curvature observed in the test results indicates that, instead of complete fracture, the striker partially returned to its original position after traveling a certain distance through the material.

In the energy-time graph (Figure 3b), this behavior corresponds to stored elastic energy. The energy level reaches its peak at around 15 seconds, attaining a value between 3 and 3.5 J. Following the maximum point, a slight decrease in energy occurs, after which the curve continues horizontally. The minor drop followed by a plateau represents the absorbed energy.

The force-displacement and energy-time graphs obtained from the impact test of the Group III functionally graded composite specimen fabricated by stacking layers from the highest areal density fabric (200 g/m²) to the lowest areal density fabric (25 g/m²) are presented in Figure 4.



a) b)

Figure 4. For Group III specimens: a) Force–displacement graph and b) Energy–time graph.

The force-displacement response initially remains at a very low level. After approximately 7 mm of displacement, an increase in force is observed, corresponding to the onset of deformation in the specimen. The force reaches a level of about

1250 N, representing the impact resistance of the specimen. Following the peak, the sudden drop in force is accompanied by further displacement over a short range, which is related to the rebound of the impactor after the damage has occurred. This behavior appears in the energy-time graph as a partial drop in energy after the maximum point, corresponding to elastic energy release (Figure 4b).

According to the graph, the specimen demonstrates resistance to impact, and the rapid rise of the force to its maximum value indicates that the material possesses a certain level of impact strength. The increase in displacement reveals that the material exhibits ductile characteristics. This FGM behavior is similar to that observed in the Group II FGM specimens, where the maximum force value was approximately 1400 N (Figure 3). The maximum energy value for the Group III specimens, like that of the Group II specimens, ranges between 3 and 3.5 J.

The force-displacement and energy–time graphs obtained from the impact test of the Group IV functionally graded composite specimen-fabricated homogeneously with a fabric areal density of 200 g/m²—are presented in Figure 5.

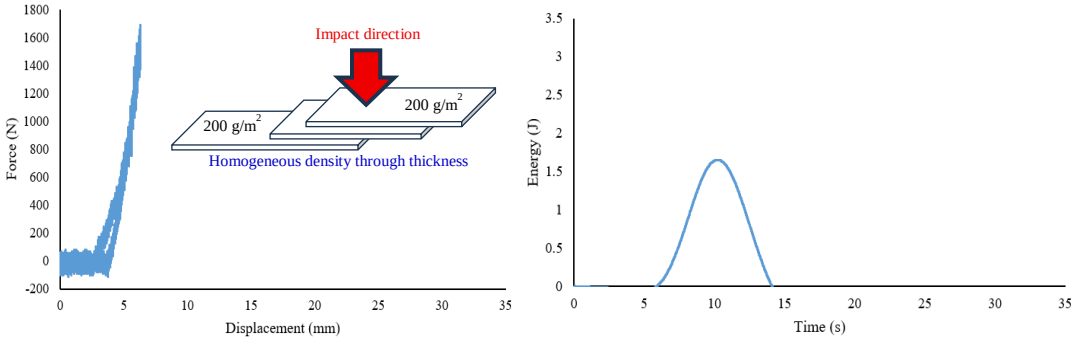
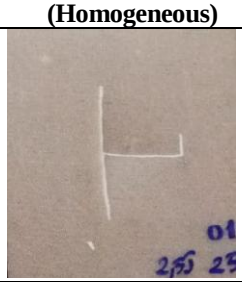
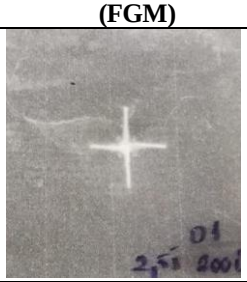
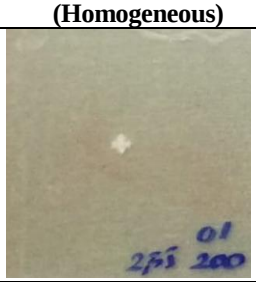


Figure 5. For Group IV specimens: a) Force–displacement graph and b) Energy–time graph.

In the force-displacement graph (Figure 5a), the force remains constant at a low level between 0 and 4 mm of displacement. From 4 mm onward, the force increases rapidly in a linear manner, reaching a maximum value of approximately 1700 N at around 6 mm. This indicates that the material possesses high stiffness and is capable of withstanding most of the applied impact energy. As expected, the 1700 N value represents the highest damage force among all specimens. The increase in fabric areal density, and consequently in material stiffness, accounts for this outcome.

Following the peak, a sudden drop in force is observed, after which the impactor returns to its initial position. This suggests that the specimen either remained unfractured throughout the test or that the impactor was rebounded once a certain displacement was reached due to constraint. Therefore, it can be stated that no penetration occurred and that the absorbed energy was entirely elastic (Figure 5b). According to Table 3, although the striker reached maximum displacement upon contact, no structural weakening in terms of geometry was observed in the Group IV specimen. This reflects a material behavior characterized by high impact resistance without reaching the limit of plastic deformation.

Table 3. Post-test damage images of the specimens.

I. Group (Homogeneous)	II. Group (FGM)	III. Group (FGM)	IV. Group (Homogeneous)
			
All layers 25 g/m²	Impactor contact surface 25 g/m²	Impactor contact surface 200 g/m²	All layers 200 g/m²

The greatest deformation damage occurred in Group I specimens, which had the lowest areal density. The damage area in FGM specimens was lower compared to that in Group I homogeneous specimens. Izadi et al. [6] demonstrated that increasing the fiber volume fraction and promoting greater fiber bridging enhances resistance to crack growth. Although the maximum damage load and energy absorption values of the FGM specimens were similar to one another, their impact damage characteristics differed. The surface damage of Group III specimens, with an impactor contact surface of 200 g/m², was greater than that of Group II specimens, which had an impactor contact surface of 25 g/m². Having a stiffer upper surface and a more ductile lower surface increases the deformation zone compared to the opposite configuration. Conversely, a ductile upper surface absorbs part of the impact during the initial contact phase, thereby reducing material damage. Accordingly, it can be stated that Group II FGM specimens performed nearly as well as the stiffest specimens, namely those in Group IV. Sain et al. [7] reported that the stiffness of FGM' s was higher than that of homogeneous structures. In homogeneous glass fiber-reinforced composites, tensile strength increases with fiber content, whereas in graded structures, increasing fiber loading reduced tensile strength but resulted in higher flexural strength. In terms of impact resistance, graded specimens produced with a lower fiber content exhibited superior performance. When maximum load values were evaluated on a specific weight basis, this value was 55,92 N/g for Group II FGM specimens and 86,42 N/g for homogeneous Group IV specimens. Suresh [10] stated that FGM materials provide superior mechanical performance compared to conventional composites due to the controlled variation of mechanical properties through the thickness. This approach reduces stress concentrations caused by abrupt transitions, resulting in more durable systems under dynamic loads such as impacts. Kumar et al. [12] found strong agreement between their experimental results and finite element analyses, demonstrating that FGMs possess high impact resistance, low weight, and reliability. Demir et al. [13] reported significant improvements in flexural strength compared to pure carbon fiber-reinforced epoxy composites when multi-walled carbon nanotubes were added to FGM' s. Impact tests similarly showed that FGM' s were more resistant than their conventional counterparts. Patel [14] indicated that, by ensuring a smooth transition between layers through specific manufacturing techniques, FGM' s reduce sudden stress concentrations and limit crack propagation under impact loading by enhancing energy dissipation.

4. Conclusions

In this study, the mechanical behavior of glass fiber-reinforced functionally graded composites manufactured using the vacuum infusion method was evaluated under low-velocity impact loading. The findings can be summarized as follows:

- In particular, the specimen stacked from the lowest to the highest fabric areal density exhibited superior energy absorption capacity, deformation capability, and balanced load-bearing behavior. Although the highest damage load was observed in the densest homogeneous specimen, FGM specimens with the lowest contact surface density demonstrated comparable performance on a specific weight basis, with similar damage areas.
- In contrast, homogeneous low-density specimens and inversely stacked FGM specimens exhibited brittleness and insufficient energy dissipation.
- The results indicate that functional grading enhances interlayer transitions, thereby improving impact resistance and offering significant advantages for structural applications. In conclusion, while the materials provide substantial

benefits in terms of lightness, high impact resistance, and ballistic performance, further research and development are required due to manufacturing challenges and cost considerations. Designers may adopt the functional grading approach as a robust method in structural composite applications subjected to impact, owing to its potential to improve design efficiency with respect to both weight and cost.

5. References

[1] Koizumi, M. F. G. M. (1997). FGM activities in Japan. *Composites part b: engineering*, *28*(1-2), 1-4.

[2] Birman, V., & Byrd, L. W. (2007). Modeling and analysis of functionally graded materials and structures.

[3] Megahed, A. A. E. W., & Megahed, M. (2017). Fabrication and characterization of functionally graded nanoclay/glass fiber/epoxy hybrid nanocomposite laminates. *Iranian Polymer Journal*, *26*(9), 673-680.

[4] Pai, A., Kini, C. R., Hegde, S., & Shenoy, S. (2023). Thin fiber metal laminates comprising functionally graded ballistic-grade fabrics subjected to mechanical and damping characterization. *Thin-Walled Structures*, *185*, 110628.

[5] Ejeh, C. J., Barsoum, I., Cantwell, W. J., & Al-Rub, R. K. A. (2025). Modelling the impact-damage behaviour of composite sandwich panels with functionally-graded TPMS-latticed cores. *International Journal of Solids and Structures*, 113534.

[6] Izadi, S. M. H., Fakoor, M., & Mirzavand, B. (2025). Investigating crack propagation and R-curve behavior in fiber-reinforced functionally graded material considering the effects of fiber bridging and fracture process zone. *International Journal of Solids and Structures*, 113550.

[7] Sain, M. K., Saraswat, P., Kumar, A., & Vemula, A. M. (2022). Fabrication and characterization of homogenous and functionally graded glass fiber reinforced polymer composites. *Materials Today: Proceedings*, *66*, 3602-3608.

[8] Chama, M., Moulai-Khatir, D., Hamza, B., Slamene, A., & Mokhtari, M. (2025). Advanced damage prediction in notched plates reinforced with Graded composite patches: Integrating XFEM-CZM and fiber-matrix coupling laws using Functionally Graded Materials. *Mechanics of Advanced Materials and Structures*, *32*(13), 2998-3015.

[9] Gunoz, A., & Kara, M. (2024). Damage behavior of functionally graded kevlar/carbon epoxy nanocomposites reinforced with polyamide 6.6 nanofiber and MWCNTs subjected to low-velocity impact. *International Journal of Damage Mechanics*, 10567895241305324.

[10] Suresh, S. (2001). Graded materials for resistance to contact deformation and damage. *Science*, *292*(5526), 2447-2451.

[11] Degnah, A., Alnaser, H. F., Al-Helaly, I., Haddad, M. Y., Al-Aredai, A., Alsaif, F., ... & Kurdi, A. (2024). The Al2O3 effect on Kevlar composite for armor applications: prediction, validation, and simulation for ballistic impact testing. *Journal of Materials Engineering and Performance*, 1-11.

[12] TS, M. K., Joladarashi, S., Kulkarni, S. M., & Doddamani, S. (2024). Experimental and numerical investigation on low-velocity impact response of sandwich structure with functionally graded core. *Polymer Composites*, *45*(4), 3225-3242.

[13] Demir, O., Tatar, A. C., Eskizeybek, V., & Avcı, A. (2022). Selectively reinforced functionally graded composite-like glass/carbon polymer nanocomposites: designed for efficient bending and impact performance. *Fibers and Polymers*, *23*(1), 196-211.

[14] Patel, M. (2025). A comprehensive review of functionally graded materials and their ballistic impact performance: Current status and future challenges. *Next Materials*, *8*, 100704.

[15] Alhazmi, W., Jazaa, Y., Althahban, S., Mousa, S., Abu-Sinna, A., Abd-Elhady, A., ... & Atta, M. (2022). Mechanical and tribological behavior of functionally graded unidirectional glass fiber-reinforced epoxy composites. *Polymers*, *14*(10), 2057.

Machine Learning-Based Prediction of Key PVT Parameters: A Comparison with Classical Regression

Kamil Neyfel ÇERÇİ[1] , Sümeyye Büşra DURMAZ [2], Çağrı KAYMAK [3], Ebru AKPINAR[4]

[1] Department of Mechanical Engineering, Faculty of Engineering, Tarsus University, Tarsus, Mersin, Türkiye.
[2] Department of Mechanical Engineering, Faculty of Engineering, Tarsus University, Tarsus, Mersin, Türkiye.
[3] Department of Mechatronics Engineering, Faculty of Engineering, Fırat University, Elazığ, Türkiye.
[4] Department of Mechanical Engineering, Faculty of Engineering, Fırat University, Elazığ, Türkiye.
[1]kneyfelcerci@tarsus.edu.tr, [2]sumeyye_durmaz@tarsus.edu.tr, [3]ckaymak@firat.edu.tr, [4]ebruakpinar@firat.edu.tr
[1](ORCID: 0000-0002-3126-707X), [2](ORCID:0009-0008-1199-2750), [3](ORCID: 0000-0001-5343-226X), [4](ORCID: 0000-0003-0666-9189)

Abstract
With the continuous growth of the global population, energy demand is rapidly increasing, and this demand is predominantly met by fossil fuels, which have limited reserves. However, the environmental impacts of fossil fuels pose serious challenges to sustainability. Therefore, reducing energy consumption and promoting the use of renewable energy sources are of great importance. In this context, solar energy has gained prominence as an inexhaustible and environmentally friendly source, while photovoltaic-thermal (PVT) panels, capable of simultaneously generating heat and electricity, have attracted increasing interest. The potential applications of PVT technology in areas such as space conditioning, heating, cooling, and drying have been extensively investigated. To enable effective analysis of PVT systems in different applications, accurate modeling is essential. In this study, two key output parameters of PVT collectors, surface temperature (T_s) and fluid outlet temperature (T_o), were modeled using different methods and compared through statistical indicators. A dataset was generated through numerical simulations using COMSOL Multiphysics, and subsequently, machine learning models including Random Forest, XGBoost, and LightGBM were developed in Python. The performance of these models was also compared to that of a conventional regression model known for its high accuracy. The results indicate that all models provided highly accurate predictions, with the best performance achieved by the XGBoost model (R^2 = 0.9998, RMSE = 0.16, MAE = 0.12 for T_s; R^2 = 0.9997, RMSE = 0.19, MAE = 0.12 for T_o).

Keywords: photovoltaic-thermal (pvt) collectors, random forest, lightgbm, xgboost, regression

1. Introduction

The increasing global energy demand, the risk of fossil fuel reserve depletion, and the environmental issues caused by climate change make the use of renewable energy sources inevitable [1]. With its unlimited, clean, and sustainable characteristics, solar energy stands out as a leading option in this field [2], while photovoltaic-thermal (PVT) collectors, as hybrid systems capable of simultaneously producing electricity and heat, offer the opportunity to maximize the utilization of solar energy [3]. This dual-mode production enhances the overall energy efficiency, achieving higher energy conversion rates compared to conventional photovoltaic panels [4]. The performance of PVT systems depends on numerous parameters, such as solar irradiation, ambient temperature, wind speed, coolant mass flow rate, panel material, and design [5]. Electrical efficiency, in particular, emerges as a critical performance indicator in terms of both system design and economic feasibility [5]. However, the complex and nonlinear relationships between these parameters limit the accuracy of conventional calculation methods in predicting this efficiency [6]. Owing to their ability to learn complex relationships within large datasets and to provide highly accurate predictions, Machine Learning (ML) techniques have become an important tool for analyzing the electrical efficiency of PVT collectors [7]. Studies employing methods such as Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), Random Forest (RF), and Gradient Boosting Machine (GBM) algorithms have demonstrated that these approaches yield more accurate results compared to conventional modeling techniques [6, 8].

In the literature, numerous studies have been conducted on machine learning models incorporating various input parameters to predict different output parameters of PVT collectors. Gharaee et al. [9] developed machine learning-based models to more accurately predict the electrical performance of water-cooled PVT collectors. In their study, parameters such as mass flow rate, solar radiation intensity, ambient temperature, inner tube diameter, wind speed, fluid inlet temperature, and PVT surface area were used as inputs, and more than 380 datasets were compiled from the literature.

[1]Corresponding author

Among the models developed using Multilayer Perceptron (MLP), RF, and Support Vector Regression (SVR) methods, the RF model yielded the best results. The influence of input parameters in the RF model was analyzed within the framework of Explainable Artificial Intelligence (XAI) using the Shapley Additive Explanations (SHAP) method, identifying mass flow rate, inner tube diameter, and wind speed as the most influential input variables. Wang et al. [10] developed a hybrid model combining Computational Fluid Dynamics (CFD) and machine learning to predict the temperature distribution in photovoltaic systems equipped with a porous collector. Gradient Boosting (GB), Extreme Gradient Boosting (XGBoost), and Histogram-based Gradient Boosting (HGB) models were evaluated with hyperparameter optimization performed using the Water Cycle Algorithm (WCA). The results indicated that the XGBoost method achieved the highest accuracy, with an R^2 value of 0.99823, demonstrating that data-driven predictive capability can be enhanced for complex physical systems. Tripathi et al. [11] performed a comparative analysis of Multivariate Regression (MR), Support Vector Machine Regression (SVMR), and Gaussian Regression (GR) to predict the power output of solar panels under varying environmental conditions, namely solar radiation, ambient temperature, and relative humidity. The findings indicated that the SVMR method achieved the highest coefficient of determination (R^2 = 0.99), confirming its capability to accurately and reliably model the effects of environmental factors on PV panel power output. Diwania et al. [12] developed machine learning-based models to predict the thermo-electric performance of hybrid PVT systems cooled with either pure water or a nanofluid. In the study, Fe/H_2O was used as the nanofluid. Analyses conducted in the MATLAB environment revealed that increasing the mass flow rate from 0.01 kg/s to 0.1 kg/s resulted in improvements in electrical efficiency by 7.11%, thermal efficiency by 2.73%, and overall efficiency by 9.84% for the nanofluid-cooled system compared to the water-cooled PVT system. A high level of agreement was observed between the predicted results and the experimental data. Zamen et al. developed Multilayer Perceptron Artificial Neural Network (MLP-ANN), Adaptive Network-Based Fuzzy Inference System (ANFIS), and Least Squares Support Vector Machine (LSSVM) models to predict the thermal efficiency of PVT systems. Inlet temperature, flow rate, and solar radiation were used as input parameters, and the models were constructed using over 100 experimental measurements. The results indicated that the MLP-ANN model achieved the highest agreement, with a coefficient of determination of 1.00. Ahmadi et al. [13] employed artificial neural networks, neuro-fuzzy logic, and the LSSVM method to predict the thermal performance of PVT systems. Considering inlet temperature, flow rate, and available solar thermal flux as input parameters, they reported LSSVM as the most suitable method. Chaibi et al. developed ANN-based models to predict the electrical and thermal efficiencies of air-based PVT collector systems. Using climate data collected in Jordan at 10-minute intervals over the course of a year, the model achieved error rates of 0.0079% for electrical efficiency and 3.3607% for thermal efficiency. These findings demonstrate that the model serves as a powerful predictive tool for PVT air collectors. Çerçi [14] evaluated the energy, exergy, economic, and environmental impacts of positioning NACA 8412 airfoils in different orientations (along the flow direction and perpendicular to the flow) and developed ANN models to predict key output indicators for PVT collectors, namely collector outlet air temperature, cell temperature, and pressure drop. The dataset was generated using the Finite Element Method, encompassing both various operating parameters and geometric parameters. Some ANN models were designed to predict each output parameter individually, while others were developed to estimate all output parameters simultaneously. The results demonstrated that both single-output and multi-output ANN models provided high accuracy; however, the single-output model exhibited greater precision for each individual performance parameter, whereas the multi-output model offered the advantage of simultaneous prediction.

In this study, two primary output parameters of PVT collectors, surface temperature (T_s) and fluid outlet temperature (T_o), were predicted using various modeling approaches and evaluated based on statistical performance metrics. The dataset was obtained through numerical simulations conducted in COMSOL Multiphysics, after which machine learning models, namely Random Forest, XGBoost, and LightGBM, were implemented in Python. Notably, for the first time in the literature, these three models were comparatively analyzed for two output parameters of a PVT collector within the scope of this study. The predictive performance of the developed models was further benchmarked against a conventional regression method recognized for its high accuracy.

2. Methodology

2.1. Numerical Modeling of the PVT Collector and Dataset Generation

To develop machine learning models, it is first necessary to generate an appropriate dataset. This requires determining the cell surface temperature of the PVT collector and the outlet fluid temperature. These parameters enable the calculation of the electrical power output of PV panels and the cooling capacity of PVT collectors. For this purpose, numerical simulations were conducted by considering a range of parameters that influence PVT collector performance. The external factors included solar irradiation (I_r), water mass flow rate ($\dot{m}_w$), wind speed (v_w), ambient temperature (T_a), and fluid inlet temperature (T_i). The environmental and operating conditions used in the simulations are presented in Table 1.

To obtain faster results with lower mesh densities, the numerical model consisted of a single channel with a PV panel mounted directly above it (Figure 1). The PV panel had a width of 44 mm and a length of 1000 mm. The panel structure comprised, from top to bottom, a 3 mm thick glass cover, 0.5 mm EVA, 0.3 mm solar cell layer, 0.5 mm EVA, 0.1 mm Tedlar, and a 1 mm copper sheet. A circular copper tube, with an inner diameter of 10 mm, outer diameter of 12 mm, and length of 1000 mm, was attached beneath the copper sheet to allow water flow.

The simulations were carried out using COMSOL Multiphysics. Depending on the mass flow rate, some cases were modeled under laminar flow conditions, while others were modeled under turbulent flow conditions; the realizable k-ε model was used for turbulence modeling. It was assumed that the top surface of the PVT collector experienced heat gain through radiation and heat loss through convection, while the side and bottom surfaces were fully insulated. A no-slip boundary condition was applied to the inner surface of the channel. In the numerical analyses, a total of 3,645 distinct input scenarios were evaluated. From these simulations, 729 data points were generated for the outlet temperature (T_o) and an additional 729 data points for the surface temperature (T_s).

Table 1. Operating ranges of input parameters for numerical simulations

Parameter	Unit	Symbol	Working range
Solar radiation	W/m^2	I_r	200-1000
Water flow rate	kg/s	$\dot{m}_w$	0.000779-0.03926
Wind velocity	m/s	v_w	0-4.82
Ambient temperature	°C	T_a	10-40
Inlet temperature of water	°C	T_i	10-40

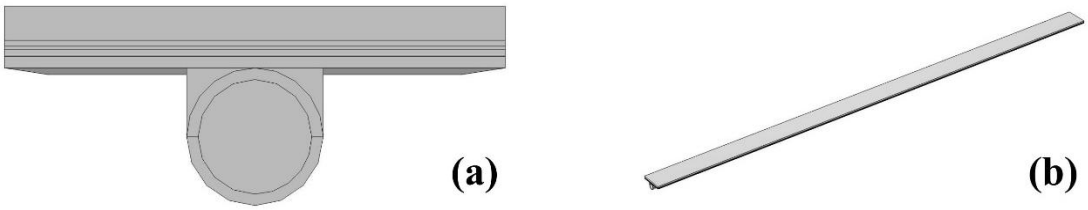


Figure 1. Front view (a) and isometric view (b) of the PVT collector with a single tube used in the numerical simulation

Mesh Independence Test

In this study, a numerical model of the PVT collector was developed using COMSOL Multiphysics to generate the dataset. During the discretization process, a free tetrahedral meshing method was applied to the boundaries of the copper tube and fluid sub-domains, while free triangular elements were used for the remaining boundary surfaces. To enhance simulation accuracy, a boundary layer mesh was implemented near the inner walls of the copper tube.

A mesh independence test was conducted to ensure that the results were not influenced by mesh resolution. The working fluid was pure water with an inlet velocity of 0.5 m/s. Both the ambient air and inlet water temperatures were set to 25 °C, and the solar irradiation was maintained at 1000 W/m^2 throughout the test. Additionally, a convective heat transfer coefficient of 12 W/m^{2}K was applied to the surface of the PV panel.

The outlet water temperature and PV cell surface temperature were evaluated for four different mesh sizes, ranging from 171,482 to 1,083,251 elements. The results of the mesh independence test are presented in Figure 2. As shown, the relative errors in cell surface temperature and outlet water temperature between the 596,927-element mesh and the 1,083,251-element mesh were only 0.0040% and 0.0037%, respectively, indicating negligible differences. Therefore, the numerical simulations were performed using the mesh with 596,927 elements (Figure 3).

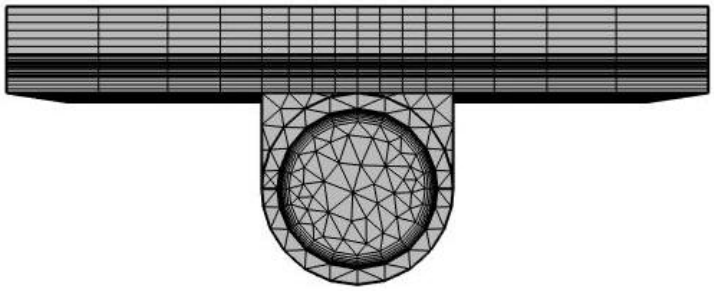


Figure 2. Mesh structure used in the numerical analysis of the PVT collector

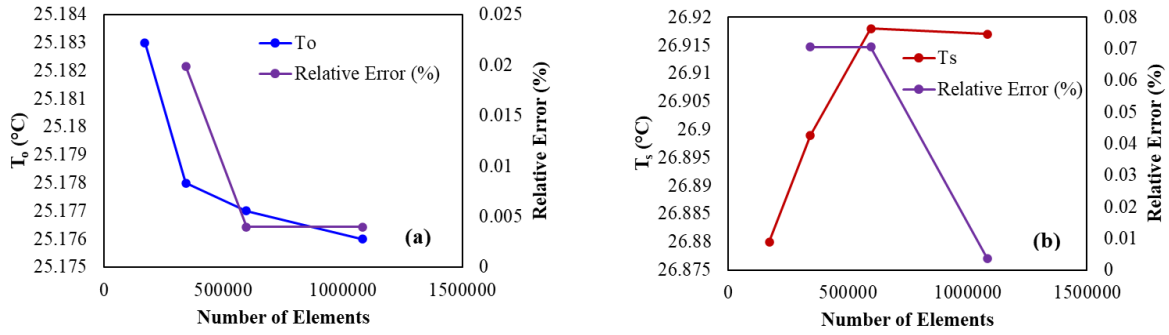


Figure 3. Variations of outlet water temperature, PV panel surface temperature, and relative error values with respect to the number of elements in the PVT collector

Validation

After achieving grid independence, validation was performed by comparing the results of this study with experimental data reported in a previous work on a PVT collector. In Figure 4, the results obtained from the numerical model developed in this study are compared with the experimental findings presented by Alous et al. [15]. For the PVT collector surface temperature and water outlet temperature, the mean and maximum relative errors were found to be 4.80% and 10.33%, and 4.17% and 6.43%, respectively. These results indicate good agreement between the numerical and experimental studies reported in the literature, with relative error values remaining within acceptable limits.

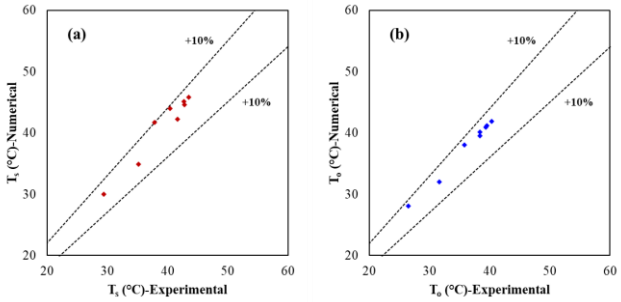


Figure 4. Comparison of numerical results with experimental data from the literature Alous et al., [15]

Machine Learning Models

In this study, three machine learning-based regression methods were employed to predict the photovoltaic panel surface temperature and outlet temperature: Random Forest (RF), Light Gradient Boosting Machine (LightGBM), and Extreme Gradient Boosting (XGBoost). These three methods are ensemble learning approaches based on decision trees, capable of making highly accurate predictions through different modeling strategies and possessing the potential to reduce generalization error. In the development of the machine learning models, 80% of the dataset was allocated for training and the remaining 20% was reserved for testing.

Random Forest (RF)

RF is an ensemble learning method based on the bagging principle, formed by combining a large number of decision trees [16]. Using bootstrap sampling, multiple subsets are generated from the training data. Independent decision trees are then trained on each subset to produce a prediction for regression problems. The final prediction is obtained by averaging the predictions from all trees. This approach reduces the risk of overfitting associated with a single decision tree and improves generalization capability. Mathematically, the prediction of an RF model consisting of *B* trees is expressed by Equation 1.

$$\hat{y} = \frac{1}{B}\sum_{b=1}^{B} h_b(x) \tag{1}$$

where, $\hat{y}$ denotes the final prediction, $h_b(x)$ represents the prediction of the *b*-th tree, and *B* is the total number of trees.

Light Gradient Boosting Machine (LightGBM)

LightGBM is an optimized version of the Gradient Boosting Decision Tree (GBDT) algorithm [17]. Unlike GBDT, LightGBM uses histogram-based feature binning and a leaf-wise growth strategy. This approach enables lower memory usage and shorter training times. Mathematically, in the t-th iteration, the prediction function of the model is updated using Equation (2):

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + \eta f_t(x_i) \tag{2}$$

where, $\hat{y}_i^{(t)}$ represents the prediction for the *i*-th instance at iteration *t*, η denotes the learning rate, and $f_t(x_i)$ represents the *t*-th weak learner.
In each iteration, the loss function is approximately obtained using a second-order Taylor expansion, as shown in Equation (3):

$$\mathcal{L}^{(t)} \approx \sum_{i=1}^{n}\left[g_i f_t(x_i) + \frac{1}{2}h_i f_t(x_i)^2\right] + \Omega(f_t) \tag{3}$$

where, g_i is the first derivative (gradient), h_i is the second derivative (hessian), and $\Omega(f_t)$ denotes the regularization term. The regularization term is calculated using Equation (4):

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} \omega_j^2 \quad (4)$$

where, T is the number of leaves, ω_j is the weight of the j-th leaf, and γ and λ are regularization parameters.

Extreme Gradient Boosting (XGBoost)

XGBoost is an enhanced version of GBDT, incorporating regularization terms and improved parallel computing capabilities [18]. To prevent overfitting, it employs L1 (Lasso) and L2 (Ridge) regularization terms. The tree structure is optimized by using both the first and second derivatives (gradient and hessian) of the loss function.
Mathematically, in the t-th iteration, the objective function is defined as in Equation (5):

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \quad (5)$$

After applying the second-order Taylor expansion, the objective function is obtained similarly to Equation (3) in LightGBM. The hyperparameters used for the three algorithms in the experimental studies are presented in Table 2.

Table 2. Hyperparameters of the machine learning models used in the experimental studies

Algorithm	Parameter	Value
Random Forest	Number of trees	100
	Fixed seed for reproducibility	42
	Learning rate	-
LightGBM	Number of trees	100
	Minimum number of samples in a leaf	1
	Minimum number of data in a leaf	1
	Fixed seed for reproducibility	42
	Learning rate	0.1
XGBoost	Number of trees	100
	Maximum depth	3
	Learning rate	0.1
	Fixed seed for reproducibility	42

Regression Model

Regression is a method used to model the relationship between an output variable and multiple input variables, and it is preferred due to its ease of use [19]. In this study, a first-order regression model was developed using the dataset to predict T_s and T_o values. This model was considered as a baseline condition for comparison with the machine learning models. In its simplest form, the first-order regression model can be represented by Equation (6).

$$\hat{y} = c_0 + \sum_{i=1}^{m} c_i x_i \quad (6)$$

3. Results and Discussion

The predictive performance of the developed models for surface temperature (T_s) and outlet temperature (T_o) was evaluated by direct comparison with numerical simulation results, as illustrated in Figure 5**.** In all models, the predicted and simulated values were closely aligned along the 1:1 reference line, indicating extremely low deviations. The error distributions

presented in Figure 6 show that the LightGBM and XGBoost models concentrated the errors near zero, thus achieving the lowest residual levels.

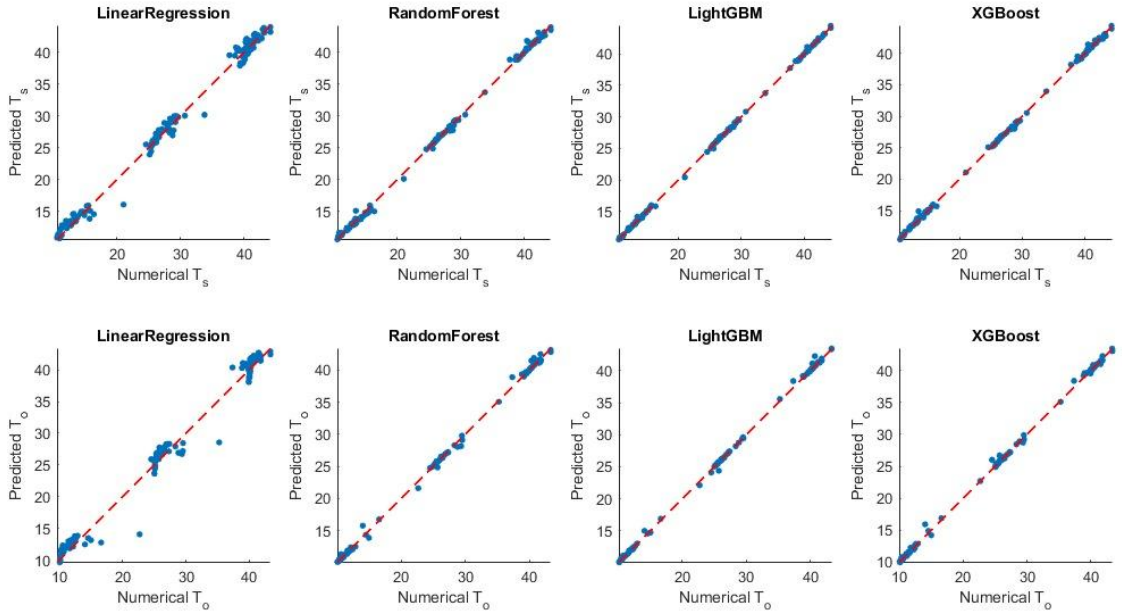


Figure 5. Comparison of predicted and numerical results for all models

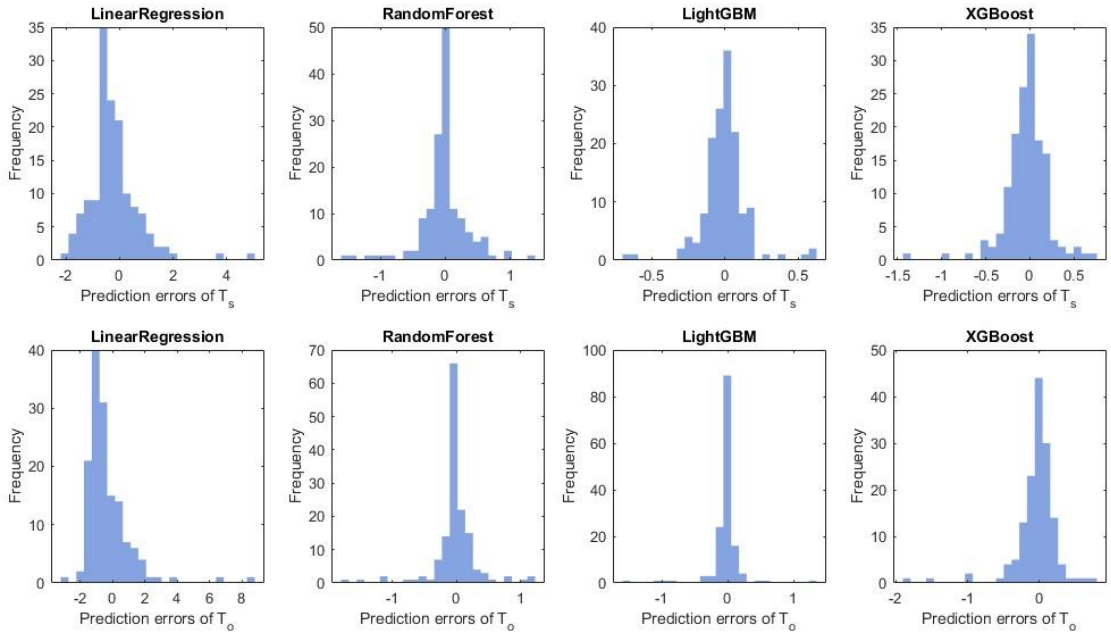


Figure 6. Error histogram for all models

As summarized in Table 3, the baseline Linear Regression model achieved $R^2 = 0.9930$ for both outputs, with Root Mean Square Error (RMSE) values of 1.010 °C for T_s and 1.000 °C for T_o. Although these results indicate a generally high predictive capability, the higher RMSE and Mean Absolute Error (MAE) values compared to ensemble methods reveal the limitations of linear regression in representing nonlinear relationships among the input parameters. For T_s prediction, LightGBM exhibited the highest performance (R2 = 0.9998, RMSE = 0.155 °C, MAE = 0.098 °C), closely followed by

XGBoost (R^2 = 0.9998, RMSE = 0.160 °C, MAE = 0.120 °C). In To prediction, XGBoost slightly outperformed LightGBM with R^2 = 0.9997, RMSE = 0.190 °C, and MAE = 0.120 °C, while LightGBM achieved R^2 = 0.9996, RMSE = 0.242 °C, and MAE = 0.106 °C. The Random Forest model also maintained high accuracy ($R^2 \geq 0.9990$) but exhibited slightly higher RMSE and MAE values compared to boosting-based methods. This outcome aligns with the known performance differences between bagging-based and boosting-based ensemble methods when modeling complex feature interactions.

Table 3. Statistical indicators for all models

Model	Output	R^2	RMSE	MAE
Linear Regression	T_s	0.9930	1.010	0.730
	T_o	0.9930	1.000	0.700
Random Forest	T_s	0.9990	0.370	0.229
	T_o	0.9991	0.347	0.182
LightGBM	T_s	0.9998	0.155	0.098
	T_o	0.9996	0.242	0.106
XGBoost	T_s	0.9998	0.160	0.120
	T_o	0.9997	0.190	0.120

Table 4 presents the constant coefficients obtained for each input parameter in the regression model. The first-order linear structure of the model simplifies the calculation process, providing a practical tool for use. This structure makes the model not only suitable for research applications but also directly applicable in engineering practice. Furthermore, it can be easily implemented in various software environments such as MATLAB, Python, and Excel. The statistical indicators obtained are within acceptable limits, highlighting the model's potential to deliver fast and reliable results as an alternative to more complex prediction methods.

Table 4. The coefficients for regression models

Output parameter	T_i	m_w	v_w	T_a	I_r (W/m^2)
T_s	0.935862	-5.383880	-0.037546	0.061926	0.003419
T_o	0.965510	-7.319854	-0.011000	0.034183	0.002027

Overall, the results confirm that LightGBM and XGBoost deliver outstanding predictive accuracy for both output parameters. Maintaining R^2 values above 0.9996 for all outputs, combined with low RMSE and MAE values and the narrow error distributions observed in Figure 6, demonstrate their reliability for real-time prediction and operational optimization of PVT systems. These results suggest that such models can effectively replace computationally expensive numerical simulations, particularly in design optimization and adaptive control applications.

4. Conclusion

In this study, four different modeling approaches were compared for predicting the surface temperature (T_s) and outlet temperature (T_o) of a PVT collector: first-order linear regression (baseline model), Random Forest, LightGBM, and XGBoost. Based on 3,645 scenarios generated in COMSOL Multiphysics, 729 data points were obtained for each output, with 80% of the data used for training and 20% for testing.

Statistical analyses showed that all models achieved high accuracy; however, the boosting-based methods, LightGBM and XGBoost, demonstrated the best performance with $R^2 > 0.9996$ and low error values. While Random Forest produced reliable results, it was slightly outperformed by the boosting algorithms. The linear regression model, although less accurate, offers practical advantages due to its computational simplicity, interpretability, and ease of implementation across various software environments.

The results confirm that LightGBM and XGBoost have strong potential as efficient alternatives to complex and time-consuming numerical simulations for real-time monitoring, adaptive control, and design optimization of PVT systems.

Future work should focus on incorporating additional operational and environmental variables, testing applicability to different PVT configurations, and exploring hybrid modeling approaches.

5. References

[1] Saidur R, Rahim NA, Islam MR, Solangi KH. "Environmental impact of wind energy". Renewable and Sustainable Energy Reviews, 16(1), 103–112, 2012.

[2] Kalogirou SA. "Solar thermal collectors and applications". Progress in Energy and Combustion Science, 30(3), 231–295, 2004.

[3] Tiwari A, Mishra RK. "Performance evaluation of photovoltaic/thermal solar air collector". Energy Conversion and Management, 52(1), 98–105, 2011.

[4] Chow TT. "A review on photovoltaic/thermal hybrid solar technology". Applied Energy, 87(2), 365–379, 2010.

[5] Tonui JK, Tripanagnostopoulos Y. "Improved PV/T solar collectors with heat extraction by forced or natural air circulation". Renewable Energy, 32(4), 623–637, 2007.

[6] Chaibi Y, Malvoni M, El Rhafiki T, Kousksou T, Zeraouli Y. "Artificial neural-network based model to forecast the electrical and thermal efficiencies of PVT air collector systems". Cleaner Engineering and Technology, 4, 100132, 2021.

[7] Khatib T, Mohamed A, Sopian K. "A review of solar energy modeling techniques". Renewable and Sustainable Energy Reviews, 45, 530–543, 2019.

[8] Abedin AH, Alim MA, Islam A, Rahman MM. "Machine learning approaches for solar radiation prediction: A review". Renewable and Sustainable Energy Reviews, 121, 109711, 2020.

[9] Mellit A, Pavan AM. "A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy".*Solar Energy, 84(5), 807–821, 2010.

[10] Gharaee H, Erfaniantti M, Bahman AM. "Machine learning development to predict the electrical efficiency of photovoltaic-thermal (PVT) collector systems". Energy Conversion and Management, 315, 118808, 2024.

[11] Wang Y, Yu L, Ali M, Khan IA, Maqsood T, Gao H, Wang Q, Guo X. "A hybrid CFD and machine learning study of energy performance of photovoltaic systems with a porous collector: Model development and validation". Case Studies in Thermal Engineering, 69, 105998, 2025.

[12] Tripathi AK, Aruna M, Elumalai PV, Karthik K, Khan SA, Asif M, Rao KS. "Advancing solar PV panel power prediction: A comparative machine learning approach in fluctuating environmental conditions". Case Studies in Thermal Engineering, 59, 104459, 2024.

[13] Diwania S, Kumar M, Kumar R, Kumar A, Gupta V, Khetrapal P. "Machine learning-based thermo-electrical performance improvement of nanofluid-cooled photovoltaic–thermal system". Energy & Environment, 35(9), 1799–1817, 2024.

[14] Ahmadi MH, Baghban A, Salwana E, Sadeghzadeh M, Zamen M, Shamshirband S, Kumar R. "Machine learning prediction models of electrical efficiency of photovoltaic-thermal collectors". Preprints, 2019.

[15] Çerçi KN. "Energy, exergy, economic and enviro-economic analysis and artificial neural network modeling of an air-cooled PVT collector with NACA 8412 airfoils". Applied Thermal Engineering, 277, 126955, 2025.

[16] Alous S, Kayfeci M, Uysal A. "Experimental investigations of using MWCNTs and graphene nanoplatelets water-based nanofluids as coolants in PVT systems". Applied Thermal Engineering, 162, 114231, 2019.

[17] Breiman L. "Random forests". Machine Learning, 45(1), 5–32, 2012.

[18] Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. "Lightgbm: A highly efficient gradient boosting decision tree". Advances in Neural Information Processing Systems, 30, 2017.

[19] Chen T, Guestrin C. "Xgboost: A scalable tree boosting system". Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794, 2016.

[20] Wang Z, Srinivasan RS. "A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models". Renewable and Sustainable Energy Reviews, 75, 796–808, 2017.

The Production of Fibres for Mineral Wool by Utilization of Olivine Extracted as A by-product of Chromite Ore

Gökhan BAŞMAN[1], Akın ODABAŞI[2], Tuğçe ERGÜLl[3], Cengiz YAŞİN[4], Hülya Kaftelen ODABAŞI[5]

[3, 4] Etikrom INC., Bingol Highway 55.Km, Kovancilar, Elâzığ/Turkey
[2] Engineering Faculty, Metallurgical and Materials Engineering, Fırat University, Elazığ/Turkey
[5]School of Civil Aviation, Department of Aircraft Maintenance and Repair, Fırat University, Elazığ/Turkey
[1]gokhan.basman@etikrom.com, [2]odabasia@firat.edu.tr, [3]Cengiz.YASIN@etikrom.com, [4]Tugce.ERGUL@etikrom.com, [5*]hkodabasi@firat.edu.tr
[1](ORCID: 0000-0001-8835-3641), [2](ORCID: 0000-0002-2556-8896), [3](ORCID: 0000-0002-6846-8166), [4](ORCID: 0000-0005-7964-3408), [5](ORCID: 0000-0003-1111-5420)

Abstract
The amount of olivine mineral released as a by-product together with chromite ore in the magnetic separation plant of Etikrom Inc. is almost the same as the amount of ferrochrome chromite ore. Due to its excellent physical, mechanical, and thermal properties, olivine is generally used as a refractory material, slag conditioner, blasting sand, and foundry sand. However, due to its high MgO and SiO_2 content (86-89%) and low iron content, olivine also has the potential to be used in mineral wool applications. This study aims to investigate the use of olivine, a ferrochrome by-product, in the production of mineral wool in order to protect natural resources and to ensure the sustainable development of the ferrochrome industry. For this purpose, briquettes with acidity values ranging between 1.81-1.97 were obtained by mixing 10-42% olivine with basalt for fibre formation for mineral wool. Limestone and dolomite were used as calcium oxide sources in the compositions. The morphology, semi-quantitative analysis, thermal properties, and phase analysis of the fibers obtained from the briquettes by fibre drawing method were evaluated. Fibre formation from briquettes containing different percentages of olivine was maximum at 20% olivine-basalt composition, while fibre morphology was obtained uniformly in the composition with 15% olivine and an acidity value of 1.97.

Keywords: olivine, mineral wool, fiber morphology, thermal propertie

1. Introduction

Rock wool, also known as stone wool, is a type of mineral wool produced by spinning fine, intertwined fibers of rock at high speed. It is produced by melting natural rocks such as basalt or industrial by-products such as slag [1], or olivine at temperatures between 1300 and 1500 °C. In the manufacture of rock wool, the flow of the molten material is of the utmost significance for fibre fabrication [2]. Fluidity behaviour generally changes according to the chemical composition of the raw material, i.e. the mineral content. The quality of stone wool can be determined by two important properties that must be present in the raw material: the acidity modulus (Ma) and the CaO/MgO ratio (≥0.5). Furthermore, the conditions for fibre preparation are only met by acidic type basalts with a SiO2 content of over 46%

Olivine, $(Mg^{2+}, Fe^{2+})_2SiO_4$, is a major component of most basic rocks and is one of the most abundant minerals in the Earth's crust (i.e., more than 50% of the upper mantle) [1]. Olivine has many different applications, such as refractory material, foundry sand, underwater pipe and cable coating material, and sandblasting in the steel industry, due to its chemical composition (MgO and SiO_2 content), low cost, high refractoriness, low thermal expansion, thermal shock resistance, and high chemical stability [3]. Olivine products are used in a wide variety of ways in the steel industry, which is one of the major consumers of these products. Some of the ways in which olivine is used in this industry include in slag conditioners, molding sand, tap hole fillers and coating materials. In addition, olivine's high melting point makes it a highly heat-resistant insulating material, which is superior to asbestos wool in the production of environmentally friendly mineral wool. At the Etikrom Inc. magnetic separation plant, approximately 500 tons of olivine mineral are released as a by-product of chromite ore processing each month. From 50 tons of ore, 25 tons of olivine mineral and the same amount of chromite ore are extracted. Olivine, which is extracted alongside chromite ore, is a material with the potential to be used in mineral wool applications due to its high MgO and SiO_2 content (86–89%) and low iron content. The utilization of large quantities of olivine mineral, which is produced as a

[1]Corresponding author

by-product in large quantities along with chromite ore, in the production of mineral wool is important in terms of the protection of natural resources within the scope of the European Green Deal.

In this study, the proportions of basalt/dolomite, basalt/limestone used in mineral wool were reduced, and ferrochrome slag was used as a substitute. In this study, mineral wool compositions were prepared by adding olivine, basalt/dolomite, and basalt/limestone in amounts ranging from 10 to 42% by weight. The fibers were evaluated by morphological, thermal, and phase analyses depending on the acidity and CaO/MgO ratios of the mixtures. In this regard, the use of industrial by-products such as olivine instead of natural resources in rock wool production will yield significant results both economically and in terms of the utilization of by-products.

2. Experimental

In this study, olivine-containing particles of size smaller than 300 μm were used as a starting material extracted as a by-product with chromite ore in Etikrom Inc. (Elazığ, Turkey). Basalt dolomite and/or calcium oxide (CaO) raw materials were used together with olivine to prepare the recipes. All raw materials were reduced to sizes smaller than 10 mm. Recipes containing different proportions of basalt, dolomite, CaO, and olivine were prepared. In addition to basalt and olivine, dolomite or CaO was added to the recipes. The basalt ratio started at 85 wt.% and was reduced to 48 wt.%.

In this study, we prepared three sets of rock wool samples. The first set of samples contained olivine with percentages between 10-42% basalt and lime (CaO), while the second set of samples was obtained from 10-15% olivine with basalt and dolomite. The last sets only contain 15 and 20 wt. % olivine with basalt. Table 1 presents the compositions of samples fabricated from olivine with varying weight percentages between 15% and 42%.

Table 1. Sample codes and compositions of mineral wools obtained from olivine

Sample code	Olivine (%)	Basalt(%)	Lime+ Basalt (%)	Dolomite+ Basalt (%)
OK-1	25	-	75	-
OK-2	30	-	70	-
OK-3	40	-	60	-
OK-4	42	-	58	-
OK-6	15	85		-
OK-7	20	80		-
OD-1	15	-		85
OD-2	10	-		90

As mentioned in the introduction section, the modulus of acidity (a_i) is important for producing high-quality mineral wool fibres. a_i is calculated using the following equation [4]:

$$a_i = \frac{mSiO + mAl_2O_3}{mCaO + mMgO} \quad (1)$$

where m: mass content of oxides (wt.%).

The modulus of acidity (a_i) should be higher than 1.8 [4] in order to obtain the high-quality mineral wool fibre, i.e., the higher the a_i of the raw basalt, the better quality in terms of strength, corrosion, temperature resistance and thermal insulation is the wool fibre as a final product. The recipes were prepared with an acidity index between 1.81 and 1.97 on average.

The mixtures prepared according to the recipes were formed into briquettes. The briquettes contained 5% cement as a binder. The prepared briquettes were then dried at 105°C for 24 hours. The middle part of the completely dry briquettes was flamed with a torch to initiate the melting process. The molten part formed was pulled upwards with

a metal rod to create a thin wire. This step is called the fiberization process. Recipes in which this process occurs are evaluated as having the fibre formation feature. Figure 1 gives the fiber formation process from the recipes given in Table 1.



Figure 1. Smelted briquette and fiber drawing

3. Results and Discussions

3.1 Powder Characterization

The chemical analysis of the olivine, basalt, and lime was taken from the XRF study and is given in Table 2. Table 2 reveals that olivine and basalt consist of three major components mainly silicon dioxide (SiO_2), magnesium oxide (MgO), and aluminum oxide (Al_2O_3). The lime we use as an alkali oxide source in our recipes is notable for its high CaO content (97%).

Table 2. Chemical analysis of olivine, basalt and lime.

	MgO	SiO_2	Al_2O_3	Fe_2O_3	Cr_2O_3	CaO	Fe	S
Olivine	50.7	37.4	1.48	6.45	1.73	0.94	4.5	<0.001
Basalt	8.96	42.8	21.3	7.3	0.44	17.95	-	<0.001
Lime	0.76	1.01	0.5	0.39	0.22	96.83	-	<0.001

From our previous work [5] XRD/Rietveld analysis results from ferrochrome olivine mineral indicated that it contains the forsterite ($Fe_{0.24}Mg_{1.76}O_4Si$, orthorombic) as a main phase (94.4%), and 1.7% lizardite ($H_4Mg_3O_9Si_2$) phase. The chemical analysis of the olivine given in Table 2 shows that the silica content is close to that of the alumina content in basalt.

3.2 Fiber Characterization

The presence of phases in the samples was identified through XRD analysis, which was performed to gain knowledge about the phases present after the fiberization process. Fiber formation occurred in the OK-6 (15 wt.% olivine, 85% basalt), OK-7(20 wt.% olivine, 80% basalt) and OD-1(10 wt.% olivine, 90% basalt + dolomite) recipes listed in Table 1; no fiber formation was observed in the other recipes. Therefore, only XRD analyses of fibres obtained from these samples are presented. Figure 2 gives the XRD patterns of the OK-6, OK-7 and OD-1 samples.

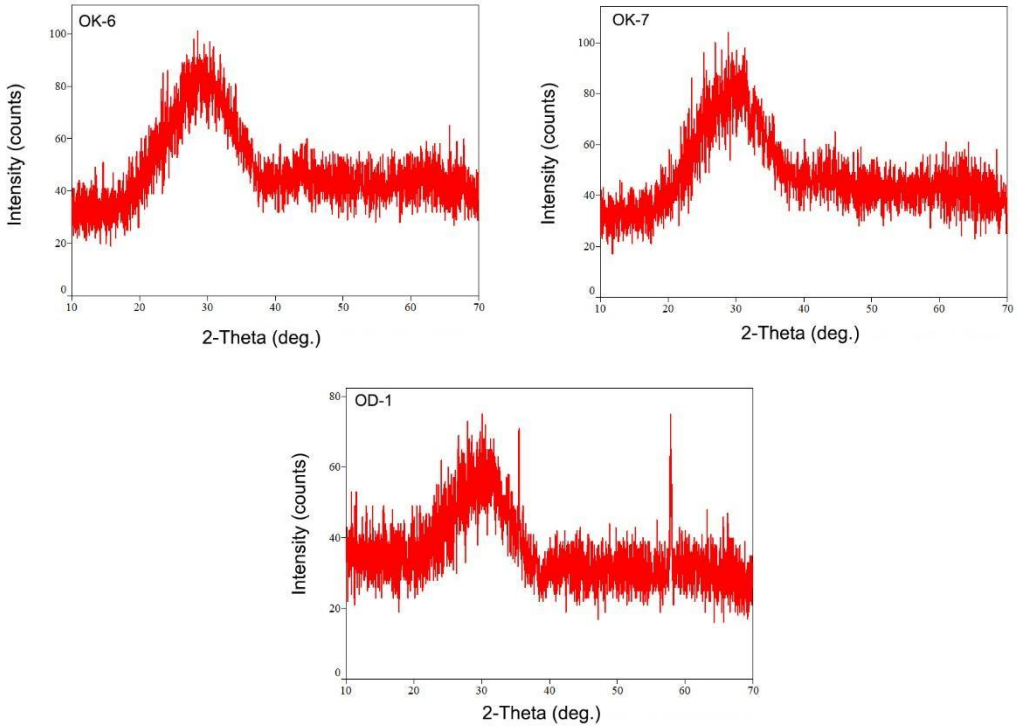


Figure 2. XRD analysis of OK-6, OK-7 and OD-1 samples.

As can be seen from the XRD analysis results, the broad amorphous peak between 2θ = 2–33° is typical of mineral wool XRD analysis, which is consistent with the literature [6, 7]. The peak observed at 2θ=44° in the XRD pattern of samples OK-6 and OK-7 is most likely a peak belonging to the $MgAl_2O_4$ spinel phase with reference code (PDF#21-1152) corresponding to (400) plane [6]. Unlike OK-6 and OK-7, the peaks observed at 2θ = 35° and 57.8° in the XRD analysis of sample OD-2, which contains 10 wt.% olivine and 90 wt.% basalt-dolomite, correspond to the (311) and (511) planes, respectively, of the $MgAl_2O_4$ spinel phase. Similar XRD patterns have also been observed in dolomite-basalt and lime-basalt samples containing ferrochrome slag [7].

Figure 3 shows the microstructural evaluation of the fibres obtained using a field emission scanning electron microscope. This illustrates the morphological characteristics of the mineral wool fibres in the OK-6, OK-7 and OD-2 samples. Figures 3a and 3b show representative fiber morphologies of OK-6, and OK-7 samples, respectively, in the first set with olivine and basalt composition. Figure 3c shows the enlarged SEM morphology of the OK-7 sample. The irregular shaped fiber morphologies belonging to OD-2 sample containing 20 wt.% dolomite, 70 wt.% basalt and 10 wt.% olivine compositions are given in Figure 3d. This inhomogeneity in fiber morphology in samples OK-7 and OD-2 is probably due to their composition. However, OK-6 sample containing 15 wt.% olivine showed better homogeneous fiber morphology. It can be noticed, OK-6 sample has the highest acidity index with a value of 1.97 compared to OK-7(1.86) and OD-2(1.84) samples. This also contributed to obtaining a more homogeneous and evenly edged fiber. EDS analysis taken from OK-7 and OD-2 fiber samples shows that the fibres are composed of oxide fomites of main elements such as Mg, Al, Si, Ca, Fe. In addition, it is noteworthy that approximately 14.3% and 6.7% carbon is found on the surfaces of OK-7 and OD-2 fibers, respectively, and this is thought to be due to the process used during the fiber drawing process.

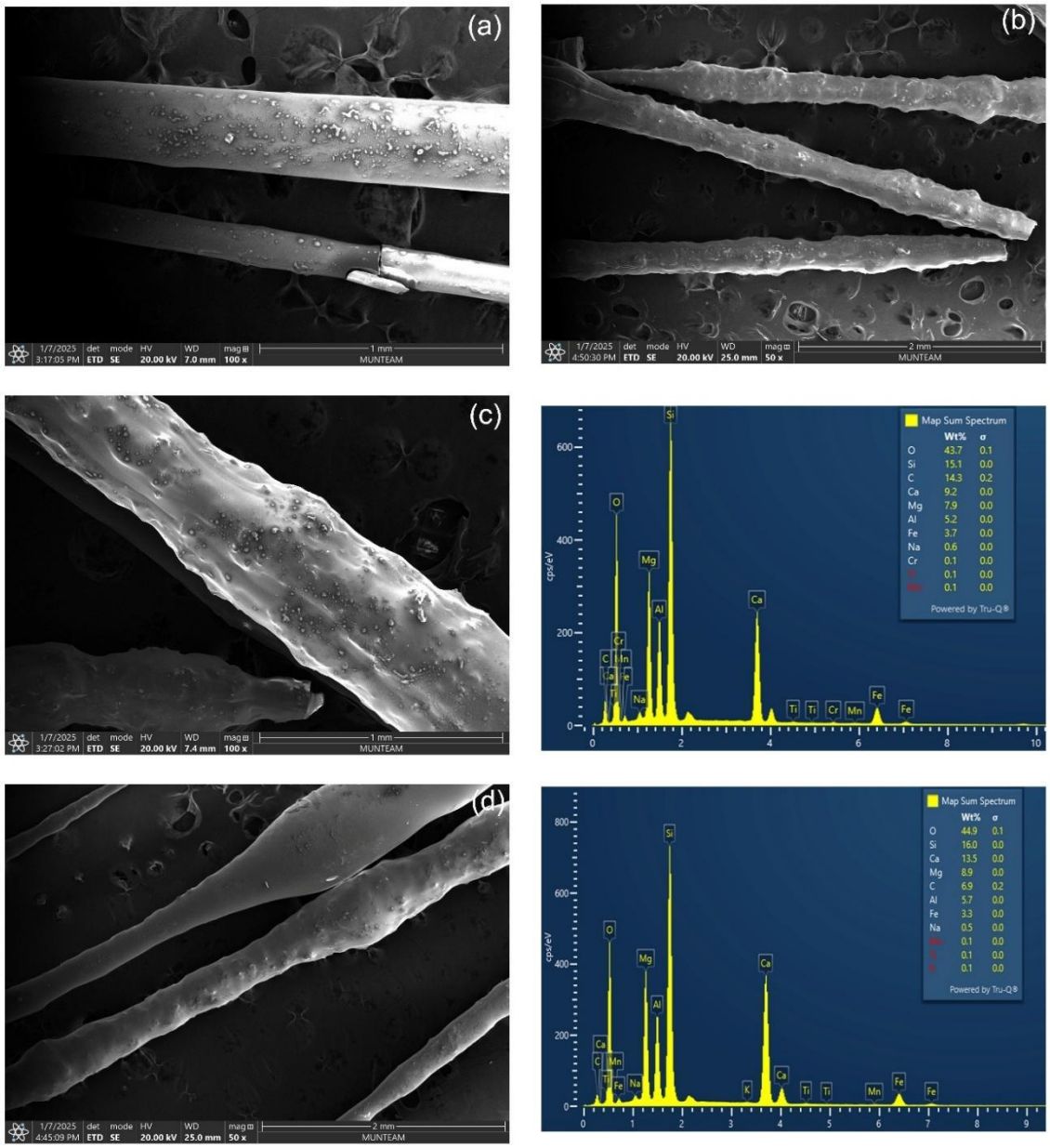


Figure 3. SEM analysis of fiber samples (a) OK-6 (b) OK-7, (c) enlarged images of OK-7 and corresponding EDS analysis (d) OD-2 and corresponding EDS analysis.

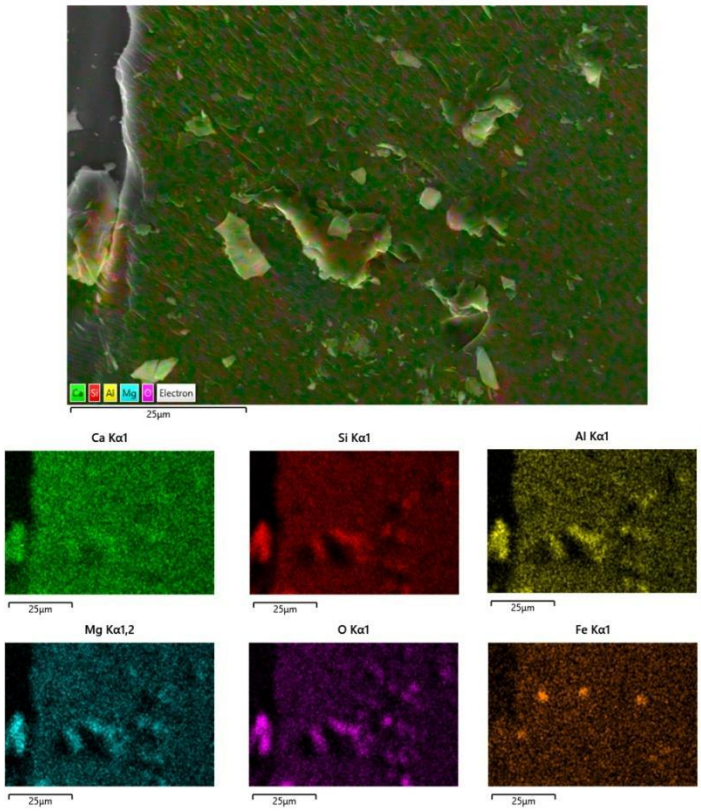


Figure 4. EDS mapping of OK-6 fiber sample and the corresponding elemental distribution of Ca, Si, Al, Mg, O, and Fe.

Energy-dispersive X-ray (EDX) spectroscopic maps provide useful information on the distribution of elements within the overall microstructure. To determine the elemental distribution of the OK-6 sample, map data were collected from the fibre surface shown in Figure 4. The elemental distribution maps of Ca, Si, Al, Mg, O, and Fe show pixels indexed from the superimposed phase distribution map. The EDS mapping results indicated that the elemental distributions of Si, Mg, Al, Ca are uniformly distributed over most of the fibre surface. In addition, iron elements measuring 4 microns in size were observed to be irregularly distributed on the fiber surfaces. EDS mapping is not provided here for OK-7 and OD-2 fibers due to their irregular structure and the presence of numerous bead-shaped structures on the fiber surfaces resulting from the fiber drawing process. Elemental mapping distribution is provided for the OK-6 sample due to its more regular fiber form compared to other samples.

Figure 5 shows the representative DTA curve of the OD-2 fiber sample containing 20 wt.% dolomite, 70 wt.% basalt and 10 wt.% olivine sample under the nitrogen atmosphere. The broad peak on DTA curve located around 230-320 °C corresponds to the release of water and organic substances. A sharp peak is observed in the DTA analyses at around 776 °C indicating the decomposition temperature of the dolomite in fiber samples. This crystallization temperature is consistent with the literature [8]. Note that the other fibres only containing basalts and olivine in their compositions therefore their DTA graphs did not shown here. Since only basalt and olivine are present in the compositions of the other fibers (OK-6 and OK-7), it was not necessary to show their DTA graphs here.

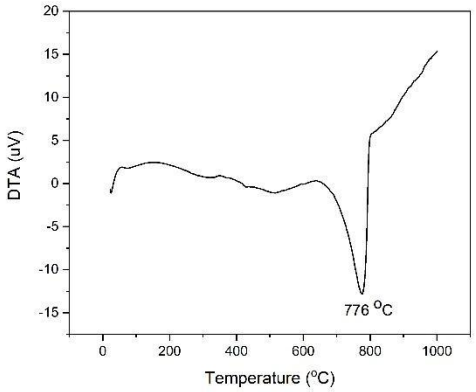


Figure 5. DTA curve of the OK-6 fiber sample

4. Conclusions

This work aimed to evaluate the possible use of olivine as a filler for the production of mineral wool. For this purpose, three sets of samples with different compositions were prepared from basalt/dolomite/olivine, basalt/lime/olivine and basalt/olivine mixture. The effects of compositional differences on the morphology and thermal properties of fibre formation were investigated. This study involved preparing briquettes containing between 10% and 42% olivine in eight different compositions, and observing fibre formation in only three of these compositions. These compositions are basalt with 15% to 20% olivine and compositions containing 10% olivine basalt and dolomite. SEM analysis showed that the morphology of fibres containing 15% olivine basalt was more homogeneous and uniform than that of the other samples. EDS mapping analyses of the fiber samples show that Ca, Al, Mg, and Si, elements are homogeneously distributed in the fiber composition. In addition, the presence of a carbon element originating from the production process in the form of droplets on the fiber surfaces of OK-7 and OD-2 fiber samples is noteworthy. It was also observed that the form of droplets was less in OK-6 sample containing basalt/dolomite. Large-scale recovery of the by-product olivine from the ferrochrome industry for use in mineral wool production can be achieved by optimising recipes with different chemical compositions, taking into account the relationship between fibre structure and properties in fibre production from briquettes, as demonstrated in this study.

Acknowledgements

This study was carried out within the scope of university-industry cooperation with the project named ETI-TAŞ-01; Investigation of Agricultural Rock Wool Production and Yield from Ferrochrome Waste and Products of ETI KROM/R&D Center. We would like to thank the fellows at the ETI KROM Inc.

5. References

[1] Isnugroho, K, Hendronursito Y, Birawidha, DC. Characterization and utilization potential of basalt rock from East-Lampung district. In IOP Conference Series: Materials Science and Engineering 285(1) 012014, IOP Publishing, 2018.

[2] Yap, Z S, Khalid, NHA, Haron, Z, Mohamed, A, Tahir, MM, Hasyim, S, Saggaff, A. Waste Mineral Wool and Its Opportunities—A Review. Materials14, 5777, 2021.

[3] Sarkar, R. Olivine—the potential industrial mineral: an overview. Transactions of the Indian Ceramic Society, 61(2), 80-82, 2002.

[4] Farouk M, Soltan A, Farrag A, El Kammar A, Hamzawy E. "Utilization of some Egyptian raw materials in rock wool industry: thermal and acoustic insulation". Egyptian Journal of Geology, 64, 155-165, 2020.

[5] Odabaşı A, Başman G, Kaftelen Odabaşı H, Ardıçoğlu H, "The Use of Olivine in the Production of Mineral Wool", 21. Uluslararası Metalurji ve Malzeme Kongresi(IMMC), İstanbul, IMMC 2022 kongre bildiriler kitabı, 481-484, 6-8 Ekim 2022.
[6] Arletti, R., Conte, S., Zanelli, C., De Iuliis, M., Di Giuseppe, D., Scognamiglio, V., ... & Gualtieri, A. F. "Recycling the product of thermal inertization of man-made vitreous fibres for the manufacture of stoneware tiles". Waste and Biomass Valorization, 14(5), 1721-1736, 2023.
[7] Odabaşı A, Başman G, Ergül T, Yaşin C, Kaftelen Odabaşı H. "Effects of Ferrochrome Slag on the Formation of Fibers for Mineral Wool Production"15th Scientific/Research Symposium with International Participation Metallic and Nonmetallic Materials, Zenica, 51, 24-25 April 2025.
[8] Gonidanga, B S, Njoya, D, Lecomte-Nana, G, & Njopwouo, D. "Phase transformation, technological properties and microstructure of fired products based on clay-dolomite mixtures". Journal of Materials Science and Chemical Engineering, 7(11), 1-14, 2019.

Slope Stability Analysis Using the Limit Equilibrium Method

Kübra BULUT[1], Zekai ANGIN[2], Hakan ERSOY[3]

[1]Civil Engineering, Faculty of Engineering, Fırat University, Elazığ, Türkiye.
[2]Civil Engineering, Faculty of Engineering, Karadeniz Technical University, Trabzon, Türkiye
[3]Geological Engineering, Faculty of Engineering, Karadeniz Technical University, Trabzon, Türkiye.
[1]k.bulut@firat.edu.tr, [2]angin@ktu.edu.tr, [3]ersoy@ktu.edu.tr
[1](ORCID: 0009-0003-5613-0218), [2](ORCID: 0000-0002-6981-6396),[3](ORCID: 0000-0001-5556-547X)

Abstract
This study presents the results of geotechnical analyses conducted for a landslide that occurred in the Esenyurt neighborhood of the Ortahisar district, Trabzon, Türkiye. Subsurface conditions, including soil parameters and stratigraphic profiles, were characterized through comprehensive field investigations and laboratory testing. In order to mitigate potential damage to a buried pipeline located beneath the landslide-affected area, a bored pile system was proposed as a slope stabilization measure. Limit equilibrium analyses, employing both the Bishop simplified method and the Janbu method, were performed using Slide software under static and dynamic loading conditions for the natural slope configuration. The critical slip surfaces were identified, and initial stabilization trials involved the implementation of a single-row bored pile system. However, analysis results indicated that this configuration did not achieve the required factor of safety in accordance with TS 8853 standards. Consequently, a double-row bored pile system was designed, with pile length and center-to-center spacing optimized through parametric studies. The final configuration achieved adequate stability under both static and dynamic conditions, demonstrating its effectiveness in reducing landslide susceptibility and safeguarding critical buried infrastructure within the project area..

Keywords: landslide, slope stability, bored pile system, limit equilibrium analysis, dynamic analysis.

1. Introduction

Slope stability issues are a critical concern in mountainous and hilly regions, where the risk of mass movements, such as landslides, is significant. These phenomena are influenced by various factors, including geographic, topographic, climatic, and lithological conditions. In such terrains, the stability of the slopes is often compromised, leading to the formation of conditions conducive to soil sliding. Natural disasters, such as earthquakes and heavy rainfall, can exacerbate these risks, further threatening the safety of surrounding infrastructure and communities [1]. In Turkey, which is geographically and geologically prone to natural disasters, landslides are a frequent occurrence, particularly in the Black Sea and Eastern Anatolia regions. The increasing frequency of landslides in these areas presents a significant challenge to urbanization and infrastructure development.

Studies have compared two- and three-dimensional modeling techniques for slope stability analysis, demonstrating that three-dimensional models provided a higher safety factor by 4% compared to two-dimensional models and emphasized the reliability of finite element method (FEM) analysis when used in combination with limit equilibrium methods [2].

Other research has addressed slope stability problems by investigating various engineering methods, employing a back analysis approach to identify failure mechanisms and determine shear resistance parameters. Comparisons of results obtained from Plaxis 2D and Slide V6.0 software confirmed the consistency of the analyses, enhancing the reliability of the findings. Additionally, pile reinforcement techniques have been suggested as an effective solution to improve slope stability, particularly when combined with drainage measures [3]. The effects of rising water levels in road embankment slopes have also been investigated, revealing that increasing water levels positively influenced the safety factor, while the properties of soil and fill materials played a crucial role in stabilizing slopes [5]. The use of pile systems for slope stabilization has been found effective in mitigating slope failure in case studies [4]. Furthermore, excessive rainfall has been identified as a significant factor contributing to slope instability, with the critical role of water pressure in destabilizing slopes being highlighted in previous studies [6,7]. These findings underline the need for comprehensive engineering solutions, including drainage systems, to counter the impact of climatic factors on slope stability.

[1]Corresponding author

Overall, the application of advanced engineering methods such as the finite element method and limit equilibrium analysis provides valuable insights into slope stability. The integration of these methods with reinforcement techniques, including pile systems, is essential for developing effective solutions to mitigate landslide risks and ensure the safety of infrastructure in vulnerable areas.

The primary objective of this study is to analyze a previously occurred landslide event using limit equilibrium methods via the Slide software and to determine the most appropriate engineering solution based on the obtained results. Landslides pose significant threats to life and property, particularly in regions characterized by steep topography and active tectonic structures. Considering its geological and topographical characteristics, along with climatic conditions, Turkey is among the countries highly susceptible to mass movements [1,9]. The prevalence of steep slopes, the irregular precipitation regime associated with a semi-arid climate, and uncontrolled land use are the main factors triggering landslide occurrences. According to AFAD [1], landslides and rockfalls are among the most frequently encountered natural hazards in Turkey, with approximately 21% of all recorded landslides occurring in the provinces of Trabzon, Rize, Erzurum, and Giresun. The Eastern Black Sea Region, in particular, exhibits the highest landslide density in the country due to high precipitation levels, steep slopes, and weak lithological units. Trabzon ranks first with 1,517 recorded cases since 1950, corresponding to an average of 23 landslide events per year. Historical records reveal that landslides in Turkey have caused severe loss of life and property. For example, the landslide that occurred on 23 June 1988 in Çatak Village of Maçka District, Trabzon, claimed 64 lives, while the mass movement triggered by heavy rainfall in July 1995 in Senirkent District, Isparta, resulted in 74 fatalities [11,8]. In Turkey, landslides typically occur in the form of rockfalls, slides, and flows, or as combinations of these types [1]. Landslide density maps [10] indicate that such events are particularly concentrated in the Black Sea region. In addition to geological weaknesses, improper land use and insufficient engineering measures further increase the level of risk.

2. Methodology

In the study area, particularly around Esiroğlu neighborhood, the dominant lithological unit generally consists of sandstone, marl, clayey limestone interbedded with basalt, andesite, and pyroclastic rocks. This unit, typically dark-colored, locally exhibits fractured and jointed textures. Chloritization and calcitization are widespread within the rocks, and well-developed pillow structures are observed in certain areas. Based on fossil evidence, the age of the unit has been determined as Upper Cretaceous [12]. Only volcanic units outcrop in the project site and its surroundings, with no sedimentary formations observed.



Figure 1. Satellite images of the region. (a) Satellite image from 2010, (b) Satellite image from 2014, (c) Satellite image from 2018, (d) Satellite image from 2023.

Figure 1 presents the satellite images of the landslide area under investigation from previous years. In Figure 1a, the 2010 satellite image clearly shows the pipeline route, which is distinctly traceable due to excavations carried out during construction activities. This route served as the primary reference dataset for subsequent analyses.

Figure 1b displays the 2014 satellite image, in which a landslide event is observed approximately 10 m west of the pipeline. Slope failures at the upper elevations were identified as a result of excavation works conducted during this period.

Figure 1c illustrates the 2018 satellite image, revealing that the landslide occurring west of the pipeline had progressed towards the road. This movement caused surface deformations on the concrete pavement in the area.

Figure 1d, dated May 17, 2023, indicates the development of a landslide immediately upslope of the pipeline. The leakage from a potable water pipeline passing beneath the concrete road at the upper elevation infiltrated into the subsurface, increasing the pore water pressure. This condition led to the formation of tensile cracks along a north–south orientation, followed by circular sliding towards the west along the crack alignment.

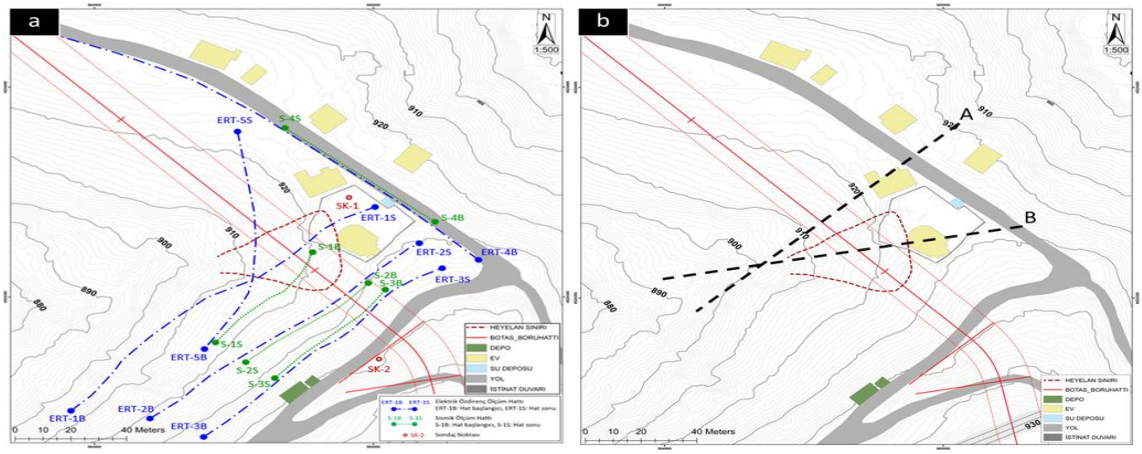


Figure 2. (a) Site layout of the landslide area, showing borehole locations and geophysical survey lines; (b) cross-sections selected for the analyses.

The locations of the boreholes and geophysical survey points conducted within the landslide area are shown in Figure 2a. These investigations were planned to accurately determine the soil profile and subsurface conditions.
On the site plan, two cross-sections were identified considering the locations of the boreholes, the boundaries of the landslide mass, and the slope direction (Figure 2b). These sections were selected to represent the area affected by the landslide, and slope stability analyses were performed on these sections.

Table 1. Detailed information associated drilling works carried out within and around the landslide area

Drilling no	Elevation (m)	Depth (m)	lithology	geotechnical behavior
SK-1	925	0.0-14.0	Silt, clay	Soil
		14.0-21.0	highly weathered rock	Soil-rock
		21.0-30.0	moderately weathered rock	Rock
SK-2	939	0.0-4.0	highly weathered rock	Soil-rock
		4.0-36.0	moderately weathered rock	Rock

Table 1 presents the geological information obtained from the boreholes conducted in the study area.

Table 2. laboratory results of soil layers

Depth(m)	Lithology	Unit Weight (kN/m3)	Cohesion (kPa)	shear strength angle
0-14	Soil	18	28	10
14-21	highly weathered rock	20	700	27
21-30	moderately weathered rock	22	2600	33

Table 2 presents the laboratory test results of the soil layers. In the stability analyses, a seismic acceleration of 0.13 g was considered, and the Mohr–Coulomb failure criterion was adopted. The groundwater level was assumed to be at a depth of 3 m.

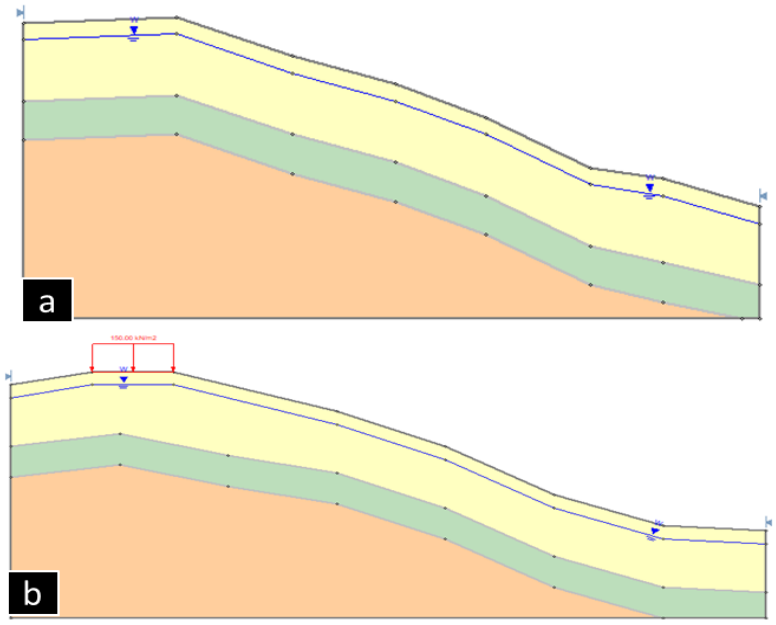


Figure 3. Topography of the soil used in slope stability analyses: (a) Topography of Section A–A, (b) Topography of Section B–B.

Figure 3 illustrates the lithological variations with depth for (a) Section A–A and (b) Section B–B. The sections were developed using borehole data and laboratory results (Table 2), revealing the distribution and characteristics of soil layers at different depths. Without any improvement or stability-enhancing intervention, the safety factors of the slope in its natural state were evaluated under both dynamic and static conditions. The analyses were performed using the Slide software, employing the Bishop and Janbu methods, and the results are presented in Table 3.

Table 3. Limit equilibrium analyses conducted on the existing condition

Situation	Method	A-A section	B-B section
Static	Bishop	0.873	0.760
analysis	Janbn	0.839	0.730
Dynamic	Bishop	0.623	0.568
abalysis	Janbu	0.598	0.545

Initially, the benching method was considered for slope stabilization (Figure 4). However, the presence of a pipeline located approximately 1.5 m below the existing topography rendered the implementation of this method infeasible.

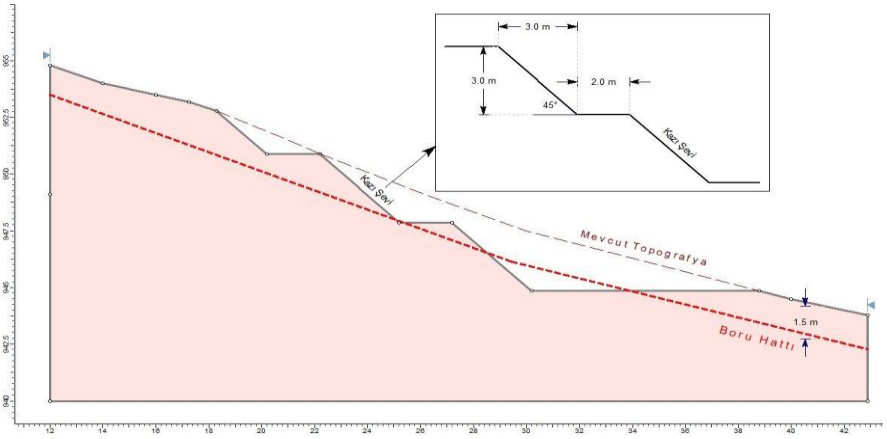


Figure 4. Representation of benching application on slope geometry

A bored pile system was selected as the stabilization method. In the initial stage, a single row of bored piles was installed at the upper elevation of the pipeline. The lengths and spacings of the piles were varied to examine their effects on potential slip circles and the factor of safety.

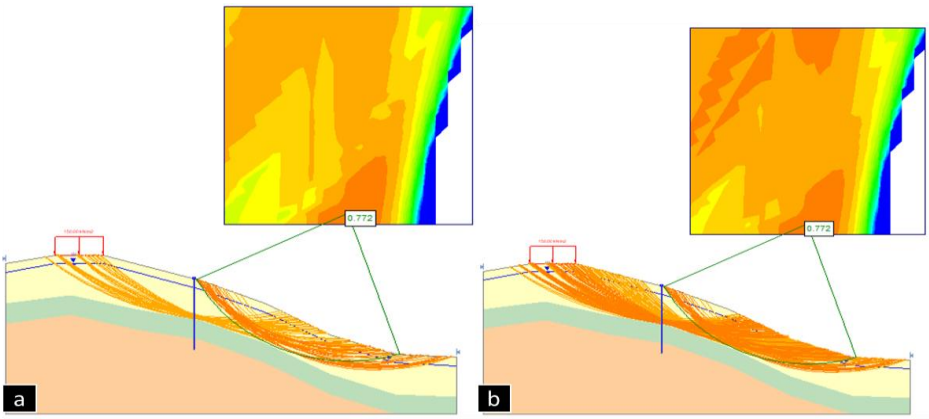


Figure 5. Stability analysis results for the B-B cross-section using the Janbu method with a single-row bored pile system, showing the factor of safety and slip circles. (a) Bored pile system with 25 m pile length and 0.80 m spacing, (b) Bored pile system with 25 m pile length and 0.50 m spacing.

As shown in Figure 5, reducing the spacing between piles did not result in a significant change in the factor of safety. However, it was observed that the frequency and location of potential slip circles changed considerably. This indicates that while pile spacing has a limited impact on the overall slope stability, it significantly influences the geometry of the failure mechanism and the distribution of potential slip surfaces.

The single-row bored pile application was found to be insufficient in increasing the factor of safety, particularly at the lower pile elevations. Therefore, it was proposed to place a second row of bored piles at the lower elevation of the pipeline. This approach is expected to not only increase the factor of safety but also alter the depth and position of potential slip circles.

In Figure 6, the thick blue lines represent the bored pile rows. The pile alignments were designed to be parallel to the pipeline route and positioned to cover the landslide area. Furthermore, to ensure effective slope stability, the piles

were carefully placed perpendicular to the slope direction. In this way, bored pile rows were positioned at both the upper and lower elevations of the pipeline, aiming to achieve optimal stability.

Figure 7 illustrates the effect of the double-row bored pile system on the factor of safety and potential slip circles for the B-B cross-section using the Janbu method [15]. The findings indicate that changes in pile spacing in a double-row system create notable differences in both the factor of safety and the distribution of slip circles. Although pile length is known to be a critical factor in engineering practice, the Slide program analyses showed that variations in pile length did not provide a significant improvement in the factor of safety when the piles remained below the lower bounds of the slip circles. However, pile spacing was found to have a considerable influence on the density and positions of the slip circles.

Static analyses revealed that the double-row bored pile system with 25 m pile length and 0.50 m spacing did not meet the minimum factor of safety value of 1.2 specified in the Turkish Standard TS 8853 [13]. Therefore, both static and dynamic analyses were repeated with variations in pile length and spacing.

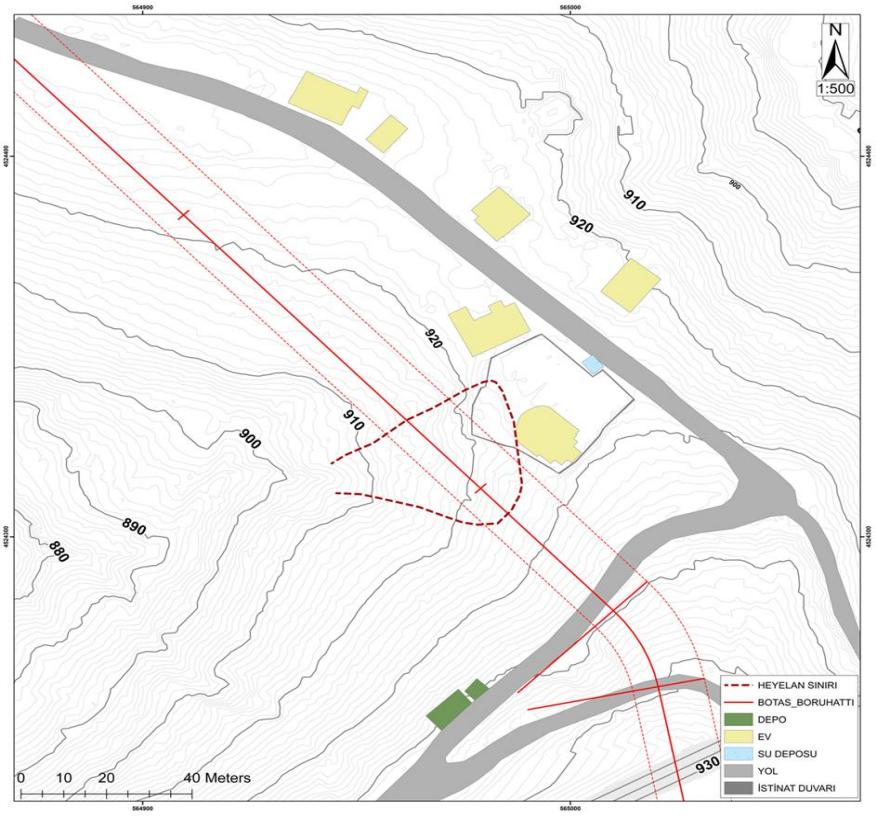


Figure 6. Locations of bored piles on the site plan of the landslide area

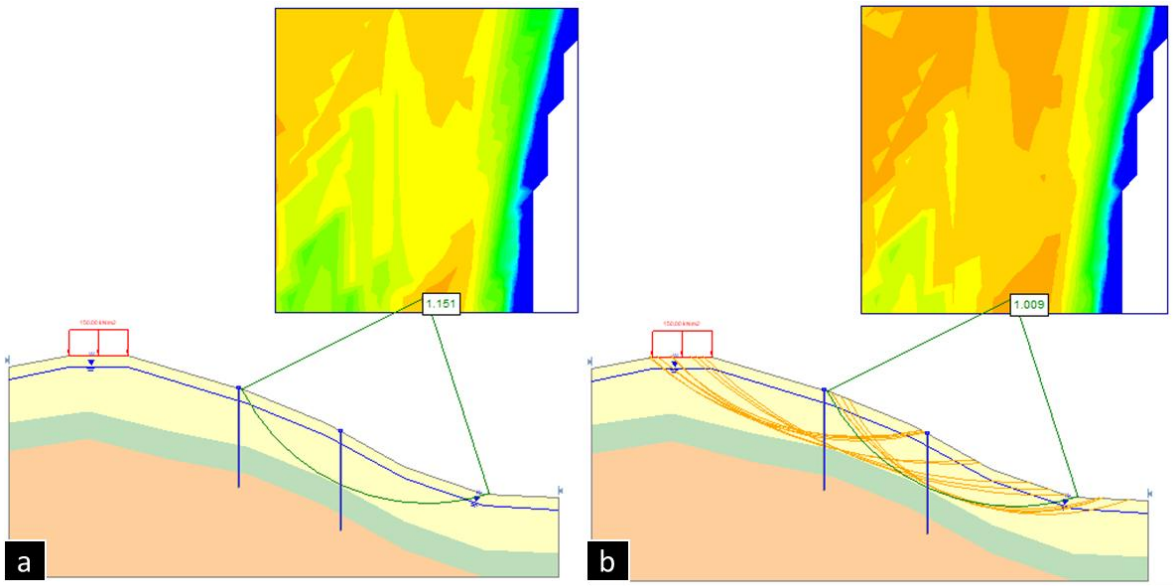


Figure 7. Double-row bored pile stability analyses for the B-B cross-section using the Janbu method, showing the factor of safety and slip circles: (a) bored pile system with 25 m length and 0.80 m spacing, (b) bored pile system with 25 m length and 0.50 m spacing.

3. Results

The influence of pile length on the factor of safety in single-row bored pile applications was examined under both static and dynamic conditions. In this study, Bishop's [14] method of slices and Janbu's [15] limit equilibrium method were selected, as the Bishop method is known for producing reliable results, particularly for circular slip surfaces, while the Janbu method offers a broader application range by accommodating both circular and non-linear complex slip surfaces. This approach enabled a comparison of the safety factors obtained from different analytical methods, allowing an assessment of method dependency in the results.

Initially, the pile length was set at 20 m, and analyses performed without socketing into the bedrock showed no significant change in the factor of safety. Subsequently, the length was increased to 25 m; however, no noticeable differences were observed in the slip circles or the factor of safety (Tables 3 and 4).

At this stage, the pile length was kept constant at 25 m, and the pile spacing was varied between 0.80 m and 0.50 m for repeated analyses. The results indicated that the slip circles with the lowest safety factors remained concentrated below the influence zone of the piles, while a reduction in slip circles occurred within the pile influence area. This configuration (25 m length, 0.50 m spacing) was also evaluated under dynamic loading conditions (Table 4).

For all static analyses of the single-row bored pile configuration, the factor of safety was found to be 0.953 (Bishop) and 0.839 (Janbu) for section A-A, and 0.760 and 0.730, respectively, for section B-B (Table 3). Under dynamic conditions, the factors of safety were 0.697 (Bishop) and 0.598 (Janbu) for section A-A, and 0.627 and 0.581, respectively, for section B-B (Table 4). These results fall below the minimum factor of safety of 1.20 for static conditions stipulated in TS 8853 [13]. Moreover, since the slip circles were concentrated at lower elevations outside the pile influence zone, a second row of piles was recommended at the lower elevation of the pipeline.

In the double-row pile application, the initial configuration consisted of piles 25 m in length and spaced at 0.80 m. Static analysis results showed factors of safety of 1.238 (Bishop) and 1.148 (Janbu) for section A-A, and 1.097 and 1.009, respectively, for section B-B (Table 3).

To assess the effect of pile spacing on the factor of safety, the spacing was reduced to 0.50 m and subsequently to 0.30 m, with both static and dynamic analyses repeated. For 0.50 m spacing, the static factors of safety were 1.411 (Bishop) and 1.304 (Janbu) for section A-A, and 1.261 and 1.151, respectively, for section B-B. The dynamic results for the same configuration were 1.037 and 0.958 for section A-A, and 0.959 and 0.879 for section B-B (Table 4). These dynamic values remain below the 1.00 limit specified in TS 8853 [13] for dynamic conditions.

In the 0.30 m spacing configuration, static analyses yielded safety factors of 1.932 (Bishop) and 1.849 (Janbu) for section A-A, and 1.540 and 1.408 for section B-B. Dynamic analyses showed 1.373 and 1.321 for section A-A, and 1.159 and 1.065 for section B-B. These results exceed the TS 8853 [13] minimum limits for both static (1.20) and dynamic (1.00) conditions, indicating adequate stability.

In practice, to further enhance the safety margin, it was targeted that pile lengths be socketed into bedrock by at least 5 m. Accordingly, the pile length was increased to 28 m while maintaining a 0.30 m spacing. The analyses revealed no significant changes in the factors of safety or the positions of the slip circles; however, considering the additional security provided by bedrock socketing, the 28 m length, 0.30 m spacing double-row bored pile system was concluded to be the most suitable engineering solution.

Table 4. Safety factores obtained as a result of static analyses performed on the sections

system	bored pile length (m)	distance between bored piles (m)	Safety Factor			
			A-A section		B-B section	
			bishop	janbu	bishop	janbu
Natural state	-	-	0.873	0.839	0.760	0.730
single row bored piles	20	0.50	0.953	0.888	0.826	0.772
single row bored piles	25	0.80	0.953	0.888	0.826	0.772
single row bored piles	25	0.50	0.953	0.888	0.826	0.772
double row bored piles	25	0.80	1.238	1.148	1.097	1.009
double row bored piles	25	0.50	1.411	1.304	1.261	1.151
double row bored piles	25	0.30	1.932	1.849	1.540	1.408
double row bored piles	28	0.30	1.932	1.849	1.540	1.408

Table 5. Safety factores obtained as a result of dynamic analyses performed on the sections

system	bored pile length (m)	distance between bored piles (m)	Safety Factor			
			A-A section		B-B section	
			bishop	janbu	bishop	janbu
Natural state	-	-	0.623	0.598	0.568	0.545
single row bored piles	25	0.50	0.697	0.646	0.627	0.581
double row bored piles	25	0.50	1.037	0.958	0.959	0.879
double row bored piles	25	0.30	1.373	1.321	1.159	1.065
double row bored piles	28	0.30	1.373	1.321	1.159	1.065

4. References

[1] AFAD. "Türkiye Afet Müdahale Planı". Afet ve Acil Durum Yönetimi Başkanlığı, Ankara, Türkiye, 2018.
[2] Taşkıran T, vd. "Şev stabilitesinin iki ve üç boyutlu modeller ile incelenmesi". Dicle Üniversitesi Mühendislik Fakültesi Dergisi, 6(1), 1-8, 2015.
[3] Ün B, Yıldız A. "Şev Stabilitesi Probleminin Geri Analizle Çözümü: Örnek Bir Vaka". Academic Platform-Journal of Engineering and Science, 9(1), 174-181, 2021.
[4] Karadağ MB, vd. "A Case Study of the Landslide in Çiftlikkoy District Genclik Street". Düzce Üniversitesi Bilim ve Teknoloji Dergisi, 7, 640-649, 2019.
[5] Yamak S. Karayolu şevlerinde boşluk suyu basınç değişikliklerinin şev stabilitesine etkisinin incelenmesi. MSc Thesis, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara, Türkiye, 2017.
[6] Suradi M, vd. "An experimental and numerical study of a landslide triggered by an extreme rainfall event in northern Australia". Landslides, 13, 1125-1138, 2016.
[7] Zhang LL, vd. "Stability analysis of rainfall-induced slope failure: a review". Proceedings of the Institution of Civil Engineers-Geotechnical Engineering, 164(5), 299-316, 2011.
[8] Fidan H, Görüm T. "Türkiye'de ölümcül heyelan olaylarının mekânsal ve zamansal analizi". Türk Coğrafya Dergisi, 75, 1-16, 2020
[9] Öztürk T. "Türkiye'de heyelanlar ve coğrafi dağılış özellikleri". Doğu Coğrafya Dergisi, 7(8), 15-34, 2002.
[10] T.C. Bayındırlık ve İskan Bakanlığı Afet İşleri Genel Müdürlüğü. "Türkiye Heyelan Yoğunluk Haritası". Ankara, Türkiye, 2008.
[11] TRT Haber. "Çatak Köyü Heyelanı". https://www.trthaber.com (Accessed 01.08.2021).
[12] Güven İH, vd. MTA Genel Müdürlüğü 1/100 000 ölçekli açınsama nitelikli Türkiye Jeolojisi haritalan serisi. MTA Genel Müdürlüğü, Ankara, Türkiye, 1993.
[13] TS 8853. "Yamaç ve Şevlerin Dengesi ve Hesap Metotları – Zeminde". Türk Standardları Enstitüsü, Ankara, Türkiye, 1991.
[14] Bishop AW. "The Use of the Slip Circle in the Stability Analysis of Slopes". Geotechnique, 5(1), 7–17, 1955.
[15] Janbu N. "Application of Composite Slip Surface for Stability Analysis". Proceedings of the European Conference on Stability of Earth Slopes, Stockholm, Sweden, 43–49, 1954.

Explainable Diabetes Prediction in a Large-Scale Clinical Dataset Using TabNet and SHAP

Keremcan KARAKAYA[1] , Mustafa ULAŞ[2]

[1,2]Department of Artificial Intelligence and Data Engineering, Engineering Faculty, Firat University, Elazig, Turkey.
[1]k.karakaya@firat.edu.tr, [2]mustafaulas@firat.edu.tr
[1](ORCID: 0009-0007-0910-4198), [2](ORCID: 0000-0002-0096-9693)

Abstract
In this study, the Comprehensive Diabetes Clinical Dataset, which contains more than 100.000 examples shared on Kaggle, was analyzed using a TabNet-based deep learning architecture. The SHAP method was integrated to explain the model's decision-making mechanism, thereby revealing the features on which the decision was based at both the individual and overall levels. In the post-training analysis, the model's performance was examined in detail by visually evaluating the confusion matrix, ROC curve and precision-recall curves. According to the results obtained, the model demeonstrated a performance level of approximately 97.3% accuracy, an F1 score of 0.81 and an AUC value of 0.98. SHAP analysis revealed that HbA1c and blood sugar levels contributed the most to the prediction. Considering that TabNet is typically applied to small datasets in the literature, this study is considered one of the first examples to systematically use the TabNet+SHAP combination on a large-scale clinical dataset.

Keywords: diabetes prediction, tabnet, shap, explainable ai, clinical data

1. Introduction

Diabetes is a chronic, lifelong metabolic disease that results from the pancreas not producing enough insulin or the body not being able to effectively use the insülin it produces [1]. It is estimated that approximately 463 million adults worldwide live with diabetes, and this number is expected to increase in the coming years [2]. Early diagnosis is critical for preventing complications and improving quality of life [3].

The digitization of electronic health records (EHR) has made health data accesible for big data analysis [4]; machine learning and deep learning-based models, integrated into clinical decision support systems, have created significant opportunities in diagnosis, risk prediction and patient-specific treatment processes [5]. For models that achieve successful results on structured (tabular) data to be used safely in clinical applications, both accuracy and interpretability are required [6].

Developed by Arık and Pfister in 2019, the TabNet model is a deep learning architecture that achieves high performance and decision interpretability on tabular data by using attention mechanism and sparse feature selection strategies [7]. Studies report that TabNet outperforms powerful tree-based methods such as XGBoost and LighGBM in comparative tests and thanks to its attention-based structure, makes it possible to understand which features are included in the decision-making process.

The use of TabNet for diabetes prediction is generally limited to smaller-scale datasets such as the Pima Indians In a study examining the combination of Bayes optimization-based Tabnet + SHapley Additive exPlanations (SHAP), the TabNet model achieved an accuracy rate of 92.2% on the Pima dataset and 99.4% on the ESDRPD dataset [8]. In onether study, TabNet was used after applying PCA-based preprocessing steps and achieved an accuracy rate of approximately 98% [9].

On the other hand, there is no study yet on the systematic use of TabNet on a Kaggle dataset known as the Comprehensive Diabetes Clinical Dataset, which contains over 100.000 observations supported by clinical variables such as age, gender, BMI, HbA1c, blood sugar and smoking history [10]. In this context, this research represents an important innovation in the literature, as TabNet models have been evaluated on a larger, heterogeneous dataset in terms of predictive model performance, scale and clinical diversity.

[1]Corresponding author

The SHAP method was used to clearly reveal the features contributing to each prediction of the model, both numerically and graphically. With this analysis, it became possible to clearly see the strength and direction of the inclusion of critical clinical variables such as HbA1c and blood sugar in the model's decision-making process [11]. In this regard, it was concluded that this study made a meaningful contribution to the literature in terms of scale, model accuracy and clinical explainability.

2. Literature Review

Machine learning-based models are increasingly being applied to health data and have shown significant success, particularly in the early diagnosis of chronic diseases such as diabetes. TabNet models developed on the Pima Indians Diabetes Dataset, one of the most frequently used datasets in the literature, are notable for their accuracy rates of up to 92.2% [8]. The same model has been reported to achieve a high accuracy. The same model has been reported to achieve a high accuracy rate of 99.4% in lesser-known datasets such as the End-Stage Diabetic Retinopathy Progression Dateset (ESDRPD). In these studies, the SHapley Additive Explanations (SHAP) approach was used for explainability, and the underlying reasons for the model's decisions were presented visually.

Deep learning architectures specific to tabular data, such as TabNet and TabTransformer, have achieved remarkable success in healthcare applications in recent years. For example, in a study conducted in 2025, TabNet and TabTransformer outperformed traditional ML methods in chemical toxicity prediction tasks and provided SHAP-supported interpretability [12]. Similarly, in another study, a TabNet-based model was developed to predict chronic kidney disease stages in diabetic patients; the model achieved 94% accuracy based on features such as serum creatinine, age, gender, and made decision processes clinically understandable through SHAP-based explanations [13]. SHAP-based explanation methods can provide more stable and individual example-sensitive predictions compared to other popular explainable artifical intelligence (XAI) techniques such as LIME and Anchors [14]. This enables the development of interpretable and reliable models that can be used by both healthcare professionals and researchers, especially in high-risk areas such as healthcare.

The majority of the studies mentioned above have been conducted on relatively small datasets such as those of the Pima Indians. In this context, the analyses conducted using the 100.000-row "Comprehensive Diabetes Clinical Dataset" used in this study contribute significantly to the literature in terms of both sample size and the advanced deep learning architecture TabNet used. Additionally, ensuring the interpretability of the model using SHAP represents an important originality in terms of both clinical decision support systems and the research field.

3. Materials and Methods

3.1. Dataset

In this study, the Comprehensive Diabetes Clinical Dataset, which is publicly available on Kaggle and contains more than 100.000 individual observations, was used [10]. The dataset includes numerous clinical and demographic variables that could aid in the prediction of diabetes. These variables include age, gender, race, smoking history, body mass index (BMI), HbA1c level, blood sugar level, presence of hypertension and history of cardiovascular disease. The target variable, diabetes, indicates whether an individual has diabetes in a binary format (0: No, 1: Yes). This large and diverse dataset is highly suitable for evaluating the generalizability of the model.

3.2. Data Preprocessing

The dataset underwent a comprehensive preprocessing phase prior to model training. First, missing values in numerical continuous variables were imputed using the SimpleImputer class from Scikit-learn, filling in the missing values with the mean value of the corresponding column [15]. This approach was chosen to prevent incorrect results caused by missing values without losing any data. Next, categorical variables such as gender, smoking_history and location were converted to numerical form using the Label Encoding method to enable model to process these columns internally [16]. After digitization, all features were rescaled to have a mean of 0 and a standart deviation of 1 using Scikit-learn's StandartScaler class [17]. This normalization contributes to the stable and fast learning of

gradient-based learning algorithms, as all features are on a similar scale and extremely large values do not disrupt the model's optimization.

Finally, after separating the feature (X) and target (y) variables, the dataset was split into two parts with an 80% training – 20% testing ratio, using the stratify parameter to maintain the class distribution. This ensured that both the training and testing sets contained a balanced number of diabetes-positive and diabetes-negative examples, thereby improving the model's generalization ability.

3.3. Model: TabNet

In this study, the TabNetClassifier architecture, which has been particularly successful on structured data (tabular data), was chosen as the classification model [7]. TabNet combines attention and sparse feature selection mechanisms to provide both high accuracy and interpretability. In this regard, it offers a more powerful and flexible alternative to traditional machine learning algorithms.

The hyperparameter configuration used during model training is summarized in Table 1. The training process was performed on GPU hardware, and an early stopping mechanism was enabled to prevent the model from overfitting.

Table 1. TabNet model hyperparameter settings

Parameter	Value
Optimization Algorithm	Adam
Learning Rate	0.02
Maximum Epoch	30
Early Stopping	10 epochs
Batch Size	1024
Virtual Batch Size	128
Training Hardware	GPU

The classification performance metrics obtained as a result of model training are presented in Table 2. High accuracy and AUC values indicate that the model is effective in both overall performance and recognizing positive class diabetes cases.

Table 2. TabNet model performance results

Metric	Value
Accuracy	0.9727
F1 Score	0.81
ROC AUC Score	0.9767

These results show that the model is not only accurate overall, but also demonstrates resilience to class imbalance. The F1 score proves that the model can successfully identify cases of diabetes, which is the positive class. The ROC AUC score of 0.97 indicates that the model's ability to distinguish between classes is quite high.

3.4. Model Explainability: SHAP

In order to make the decision mechanism of the developed model more understandable and interpretable, SHAP (Shapley Additive exPlanations), one of the explainable artificial intelligence (XAI) methods, was used. SHAP is a method based on game theory that can quantitatively determine the contribution of each feature of the prediction

result. Especially in high-risk areas such as healthcare, SHAP is important in that it transparently reveals the extent to which the model relies on certain inputs.

In this study, SHAP analysis was applied to 100 randomly selected examples from the test data. According to the results obtained, features such as HbA1c, blood_glucose_level and BMI were found to be the most effective variables in predicting whether an individual has diabetes. The effects of these variables with high SHAP values on both the positive (1) and negative (0) classes were visually presented using a beeswarm plot Fig. 5. This allowed not only the order of importance of the variables to be analyzed, but also the positive or negative direction of each variable's effect.

The explanations obtained with SHAP allow for an intuitive examination of the model's "decision structure" and produce reliable decision support outputs for doctors in clinical applications. In addition, SHAP's ability to provide sample-based explanations offers significant advantages for patient-based evaluation and comparisons [18]. In the literature, various studies have emphasized that SHAP is more stable and sensitive compared to other XAI methods such as LIME and Anchors [19].

Thanks to this explainability approach, not only high accuracy but also a reliable and explainable model has been achieved. This strengthens the model's acceptability in terms of ethics, legality and security for applications in the healthcare field.

3.5. Performance Visualizations

The model's success level was evaluated using both numerical metrics and visual analysis. As a result of the classification performed on the test data, it was observed that the model accurately identified the negative class (non-diabetic individuals) but made a certain number of misclassifications for the positive class (diabetic patients). The confusion matrix clearly illustrating this situation if shown in "Fig. 1".

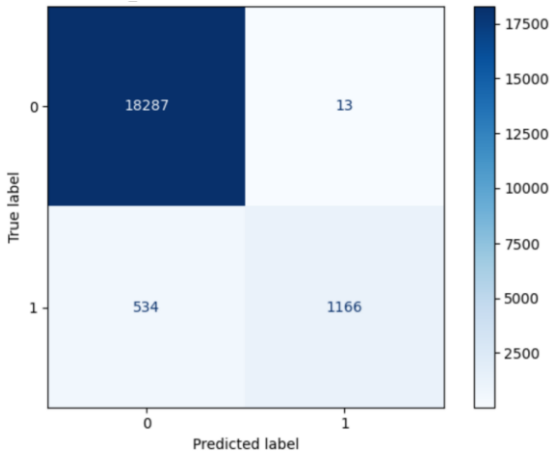


Figure 1. Confusion matrix

The overall classification performance of the model was evaluated using the ROC curve and an AUC score of approximately 0.98 was obtained. This high value indicates that the model can make a strong distinction between positive and negative classes, as shown in Fig. 2. However, considering the class imbalance in the dataset, the precision-recall curve more accurately reflects the model's performance on the positive class. This curve demonstrates that the model provides high precision and recall, as shown in Fig. 3.

The internal importance analysis was conducted to understand which characteristic of the model are given more consideration revealed that variables such as HbA1c level, blood clucose level and age stand out, as shown in Fig. 4.

In addition, SHAP analysis was performed on 100 test samples to increase the explainability of individual predictions. The results of this analysis also revealed that variables such as HbA1c and blood_glucose_level have a decisive influence on individual decisions. The distribution and interpretability of SHAP values are presented in Fig. 5.

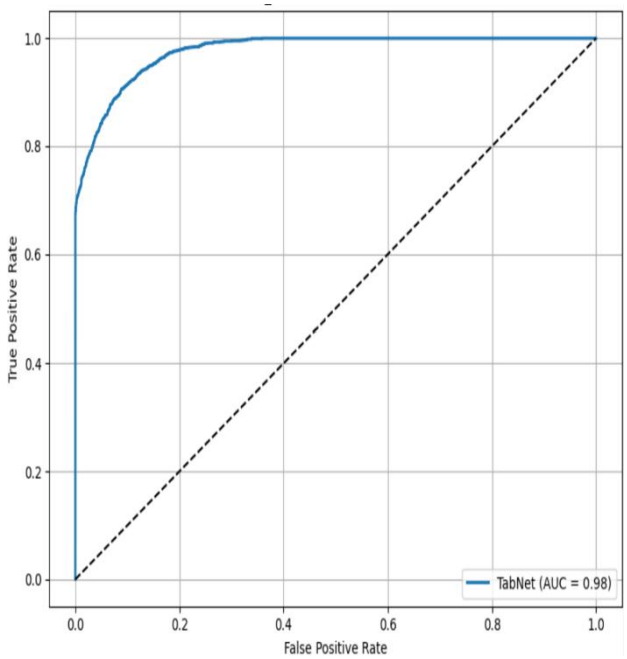


Figure 2. ROC Curve

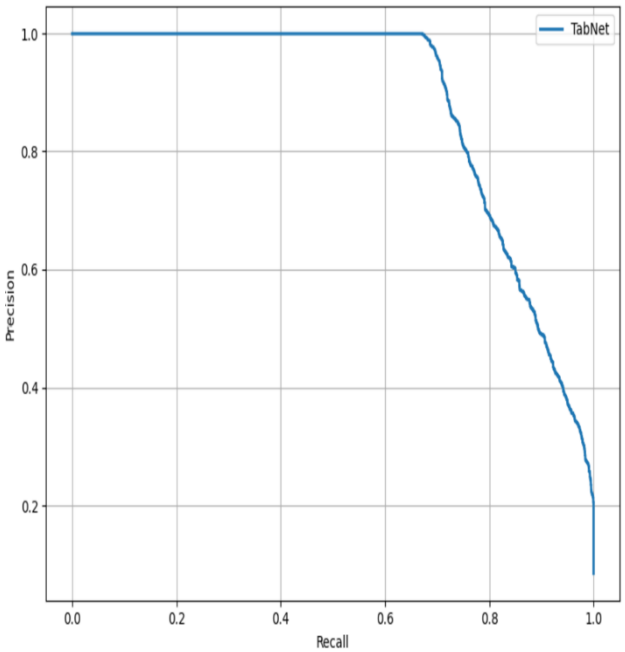


Figure 3. Precision-Recall Curve

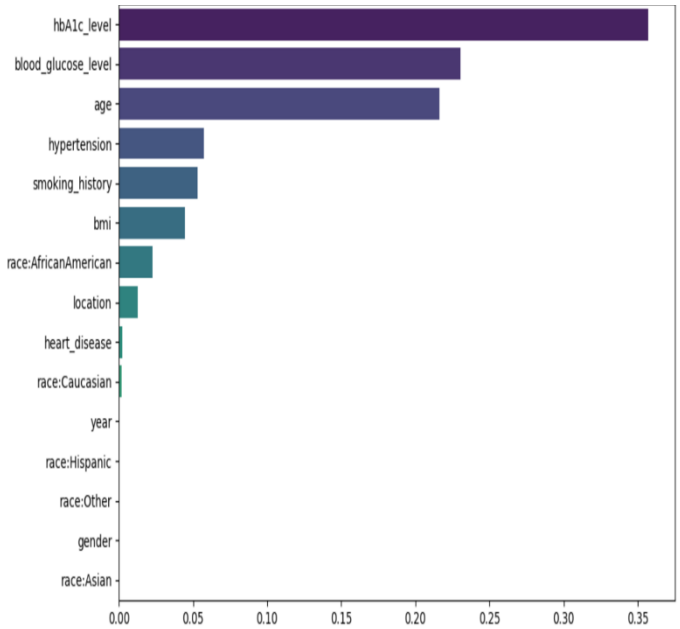


Figure 4. Feature Importance

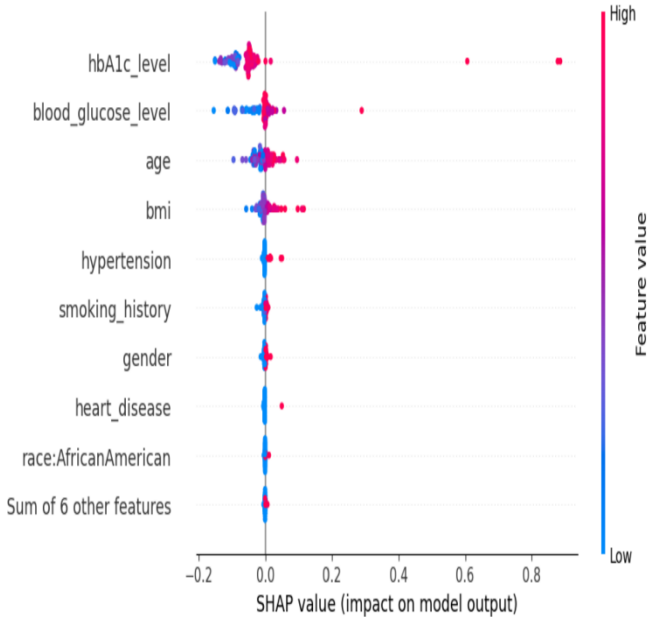


Figure 5. SHAP Summary Plot

4. Experimental Results and Discussion

As a result of the experiments conducted, the TabNet model has yielded highly successful results on the large-scale "Comprehensive Diabetes Clinical Dataset". With 97.27% accuracy and a ROC AUC score of 0.9767, the model has demonstrated strong performance in terms of both overall classification performance and recognition of the positive class (diabetes patients). The F1 score reaching a high value of 0.81 indicates that the model offers a balanced performance, particularly for the positive class.
Visual analyses also support this success. The confusion matrix shown in "Fig. 1" indicates that the model correctly indetified the negative class but made a certain number of false negative predictions in the positive class. This highlights the need for more precise modeling of the positive class, which is of critical importance in sensitive areas such as healthcare.

The ROC and Precision-Recall curves shown is "Fig. 2" and "Fig. 3" demonstrate the model's high discriminatory power and performance on imbalanced datasets. In particular, high precision and recall values indicate the reliability of decisions on the positive class.

The feature importance assessment in Fig. 4 reveals that variables such as HbA1c, blood_glucose_level and age are decisive in model decisions. The explainability analysis performed using the SHAP method and shown in "Fig. 5" supports these findings and enables the model's decision structure to be understood at the individual level. The interactive explanations provided by SHAP ensure that the model is more transparent and acceptable for clinical decision support systems.

The accuracy, F1 and ROC AUC scores achieved in this study are higher than those in previous TabNet-based studies (e.g., those conducted on small-scale datasets such as Pima Indians or ESDRPD). Another important contribution of this study to the literature is the first detailed evaluation of the TabNet model using a large clinical dataset of 100.000 lines. In this respect, our study offers a unique approach not only in terms of performance but also in terms of scalability and explainability.

5. Conclusion and Future Works

In this paper, diabetes prediction was performed using the TabNet architecture on a large-scale clinical dataset and the model's explainability was examined using the SHAP method. The applied model demonstrated successful results in terms of both performance and transparency, with high accuracy, strong discriminative ability and interpretable decision structures. As a contribution to the literature, this study is pioneering in that it is the fist systematic application of the TabNet architecture on a large clinical dataset of 100.000 row. Additionally, in addition to the high performance achieved, the detailed analysis of SHAP-based explainability has demonstrated that the model's decision-making mechanism has been made more understandable and reliable for healthcare professionals. In this regard, the study offers not only a prediction engine but also a reliability-based artificial intelligence solution in the context of clinical decision support systems.

In future studies, it is planned to address performance losses caused by class imbalance using advanced sampling techniques such as SMOTE and ADASYN. In addition, a more comprehensive model analysis process can be carried out by comparing the TabNet architecture with other modern deep learning-based structures such as TabTransformer. Furthermore, exploring long-term prediction capabilities using patient tracking data containing time series and integrating these structures with TabNet are also potential research topics. Finally, integrating the developed model into clinical decision support systems and subjecting it to user experience evaluations with clinicians through a prototype application will be an important step in terms of the applicability of artificial intelligence-based healthcare solutions in the field.

6. References

[1] N. A. ElSayed *vd.*, "Introduction and Methodology: *Standards of Care in Diabetes—2023*", *Diabetes Care*, c. 46, sy Supplement_1, ss. S1-S4, Oca. 2023, doi: 10.2337/dc23-sint.

[2] "Home", Diabetes Atlas. Erişim: 21 Temmuz 2025. [Çevrimiçi]. Erişim adresi: https://diabetesatlas.org/
[3] N. Noorbakhsh-Sabet, R. Zand, Y. Zhang, ve V. Abedi, "Artificial Intelligence Transforms the Future of Health Care", *The American Journal of Medicine*, c. 132, sy 7, ss. 795-801, Tem. 2019, doi: 10.1016/j.amjmed.2019.01.017.
[4] W. Raghupathi ve V. Raghupathi, "Big data analytics in healthcare: promise and potential", *Health Inf Sci Syst*, c. 2, sy 1, Ara. 2014, doi: 10.1186/2047-2501-2-3.
[5] R. Miotto, F. Wang, S. Wang, X. Jiang, ve J. T. Dudley, "Deep learning for healthcare: review, opportunities and challenges", *Briefings in Bioinformatics*, c. 19, sy 6, ss. 1236-1246, Kas. 2018, doi: 10.1093/bib/bbx044.
[6] S. Lundberg ve S.-I. Lee, "A Unified Approach to Interpreting Model Predictions", 2017, *arXiv*. doi: 10.48550/ARXIV.1705.07874.
[7] S. O. Arik ve T. Pfister, "TabNet: Attentive Interpretable Tabular Learning", 2019, *arXiv*. doi: 10.48550/ARXIV.1908.07442.
[8] L. P. Joseph, E. A. Joseph, ve R. Prasad, "Explainable diabetes classification using hybrid Bayesian-optimized TabNet architecture", *Computers in Biology and Medicine*, c. 151, s. 106178, Ara. 2022, doi: 10.1016/j.compbiomed.2022.106178.
[9] M. Ibrahim, "43 PUBLICATIONS 112 CITATIONS SEE PROFILE", *Technology Vol.*, c. 11, sy 1.
[10] "Comprehensive Diabetes Clinical Dataset(100k rows)". Erişim: 21 Temmuz 2025. [Çevrimiçi]. Erişim adresi: https://www.kaggle.com/datasets/priyamchoksi/100000-diabetes-clinical-dataset
[11] K. Pang, "A comparative study of explainable machine learning models with Shapley values for diabetes prediction", *Healthcare Analytics*, c. 7, s. 100390, Haz. 2025, doi: 10.1016/j.health.2025.100390.
[12] F. M. Mustafa *vd.*, "TabNet and TabTransformer: Novel Deep Learning Models for Chemical Toxicity Prediction in Comparison With Machine Learning", *J of Applied Toxicology*, May. 2025, doi: 10.1002/jat.4803.
[13] M. N. H. Chowdhury *vd.*, "Deep learning for early detection of chronic kidney disease stages in diabetes patients: A TabNet approach", *Artificial Intelligence in Medicine*, c. 166, s. 103153, Ağu. 2025, doi: 10.1016/j.artmed.2025.103153.
[14] "A Comparative Analysis of SHAP, LIME, ANCHORS, and DICE for Interpreting a Dense Neural Network in Credit Card Fraud Detection", içinde *Communications in Computer and Information Science*, Cham: Springer Nature Switzerland, 2024, ss. 365-383. doi: 10.1007/978-3-031-63803-9_20.
[15] "SimpleImputer", scikit-learn. Erişim: 21 Temmuz 2025. [Çevrimiçi]. Erişim adresi: https://scikit-learn/stable/modules/generated/sklearn.impute.SimpleImputer.html
[16] "LabelEncoder", scikit-learn. Erişim: 21 Temmuz 2025. [Çevrimiçi]. Erişim adresi: https://scikit-learn/stable/modules/generated/sklearn.preprocessing.LabelEncoder.html
[17] "StandardScaler", scikit-learn. Erişim: 21 Temmuz 2025. [Çevrimiçi]. Erişim adresi: https://scikit-learn/stable/modules/generated/sklearn.preprocessing.StandardScaler.html
[18] A. Salih *vd.*, "A Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME", 2023, doi: 10.48550/ARXIV.2305.02012.
[19] A. Nayebi, S. Tipirneni, B. Foreman, C. K. Reddy, ve V. Subbian, "An Empirical Comparison of Explainable Artificial Intelligence Methods for Clinical Data: A Case Study on Traumatic Brain Injury", *AMIA Annu Symp Proc*, c. 2022, ss. 815-824, 2022.

Modernization of The Refueling Probe With an Electromagnetic Magnet

Ebru KARAKÖSE[1], Gülşah AYTEMİZ[2], Mustafa BAŞTÜRK[3], Ömer Faruk OĞUZ[4]

[1,2,3,4] Aviation Electrical and Electronics Department, School of Civil Aviation, Fırat University, Elazığ, Türkiye
1-eozbay@firat.edu.tr, 2-230576012@fira.edu.tr, 3-230576058@firat.edu.tr, 4-220576026@firat.edu.tr
[1]0000-0003-1191-6375, [2]0009-0009-7206-7929, [3]0009-0001-5091-5503, [4]0009-0002-8097-5424

Abstract
There are problems caused by the weight of the refueling probe during refueling in aircraft. Lifting and docking this probe by technicians is a highly demanding process that requires physical force. Uncontrolled falls, bumps and related accidents may occur due to the weight of the probe. In this study, after the fuel requirement of modern passenger aircraft, one of the stages to be performed is the refueling probe [the part that allows the two components to be refueled to each other at the prop] to the fuel inlet flange [the flange is the hydrant or the sealing element between the two connections.] Problems with placement are examined. It is aimed to minimize the post-alignment docking problems caused by the weight of the probe, which will be lifted dozens of meters above the ground, and the annular structure of the flange. The current problem is that by adding an electromagnetic magnet to the end of the fuel filling flange and using a controller circuit, where we can adjust the attraction and electrical power of this magnet through the potentiometer, the refueling hydrant [hydrant will be used to transport fuel underground in modern airports in aviation and the fuel from underground will be replenished dispenser(The dispenser is called the vehicle that connects the fuel from underground with the hydrant located in the aircraft wing) On the side of the structure connecting the pipelines with], it was tried to be solved by modifying the refueling probe with an iron-based ring material that could interact with electromagnetic energy. The added electromagnetic magnet will strengthen the connection between the aircraft fuel inlet flange and the probe, facilitating the synchronization and docking of the connection between the aircraft fuel filler flange and the refueling probe as a result of the weight of the probe during the connection, and will greatly benefit the technicians undertaking the refueling. The electromagnetic magnet used will be placed in the fuel filling flange so that the current weight of the refueling probe is left stable and the fuel filling processes will be both faster and less strenuous. Therefore, the reduction of aircraft downtime by reducing fuel filling times will contribute to the operational process both physically and temporally. It is expected to open new doors for future research.

Keywords: refueling, probe, electromagnetic magnet, flight safety

1. Introduction

Civil aviation is one of the most important and reliable sectors for both international and domestic passenger transportation. The primary factor contributing to this significance is the ability to cover thousands of kilometers much faster than land or sea vehicles. As in every sector, safety is a fundamental component of the civil aviation industry. Since the inception of civil aviation, approximately 99% success has been achieved in flights, demonstrating the sector's high level of reliability. This success is the result of an intensive work and operational process, which includes all procedures carried out from the moment the aircraft approaches the passenger boarding bridge until it takes off again. These procedures cover the inspection of hydraulic systems, mechanical systems, pressurization and oxygen systems, as well as aircraft refueling operations. Special regulations have been issued to ensure that operations are carried out flawlessly [1]. These regulations also address the requirements for the safe refueling of modern aircraft, which plays a vital role in successful take-off.

In the existing literature, aircraft refueling is carried out in two main ways. In small aircraft, due to the low fuel demand, fuel transported by tankers is directly supplied to the aircraft. However, at large airports, this method is insufficient because of the high fuel demand of the aircraft. In such cases, fuel is transported through underground pipelines and transferred to the aircraft with the help of vehicles called “dispensers.” In both methods, in order to perform refueling, the probe coming from the fuel tanker or dispenser must be connected to the hydrant inlet located on the aircraft wing. Currently, this connection is made manually, using human power. During this process, the probe must be positioned by lifting it from the ground to the aircraft wing, thus gaining potential energy through human effort. As the height increases, the hose attached to the probe adds weight, increasing the potential energy of the

[1]Corresponding author

probe, which makes stabilizing the probe more difficult and causes time loss due to the difficulty of the connection [2-3]. Under normal conditions, the fuel truck approaches the aircraft at the gate. The fueling hose on the truck is extended and aligned with the fuel flange located beneath the aircraft's wing. After reaching the appropriate height, the probe is locked in place using a rotary mechanism on the probe. Only after completing this secure attachment can the actual fueling process begin. Once the refueling is completed, the connections are disengaged, finalizing the procedure. This entire process takes more than 30 minutes and involves lifting the probe and hose assembly approximately 50 kg in total weight to a height required for aircraft like the Boeing 747-400, which stands 5.6 meters above ground level. Lifting such a load presents a significant physical burden on the operator and poses potential safety risks [4-8].

One significant incident highlighting the importance of proper fuel handling occurred on September 5, 2001, during British Airways Flight 2019. After landing at Denver International Airport (USA) following a flight from Gatwick International Airport (UK), a fuel operator initiated the refueling process using a lifting platform to access the aircraft's fuel flange. However, due to the angled attachment of the fuel probe to the flange, fuel sprayed uncontrollably, resulting in an immediate fire [9].

The use of electromagnetic clamping systems in industrial applications is quite limited. CNC machines have a wide range of applications in the production processes of nearly all industries. Magnetic holders used in CNC machines utilize magnetic force from permanent magnets to clamp workpieces securely for safe operation. Magnetic chucks are alternatives to traditional chucks attached to different types of machines. These chucks provide consistent clamping pressure without causing any deformation to the workpiece. Therefore, magnetic holders ensure operator safety, precision in machining, and accuracy in the final product. A magnetic chuck typically consists of a plate, an electromagnet, and a control unit. When a ferromagnetic material approaches, the magnetic field helps with clamping. Additionally, the Dragon spacecraft developed by SpaceX performs the docking approach to the International Space Station automatically. This system operates under electromechanical control. In the field of aviation, studies on aircraft refueling have mainly focused on aerodynamic design and mechanical durability during the process of supplying fuel through hoses and connecting the fuel probe to the tanker [10-14].

This study aims to prevent possible mechanical and system failures that may threaten flight safety before or during flights. To mitigate these challenges, a novel solution has been proposed: integrating stainless steel rings on the fuel probe and installing an electromagnetic magnet on the aircraft's fuel flange. This configuration allows the probe to be attracted and aligned by the magnet as it reaches the necessary height. An electromagnetic magnet has been placed on the aircraft's fuel hydrant. A ferromagnetic ring has been attached to the fuel supply probe. When this ring enters the magnetic field, the fuel probe is pulled toward the fuel flange. This pulling force acts against the downward weight caused by the probe's potential energy. As a result, the maneuverability of the probe during transport and connection increases, facilitating the work of the refueling operator. Compared to current methods in the literature, this approach offers a safer connection and saves time by reducing physical labor [15,17].

2. Proposed Method

Most of the materials used in the aviation industry are aluminum-based materials. Since the probe to be used is also made of an aluminum alloy, the electromagnet cannot attract the probe. Modifying the aluminum alloy of the probe would reduce its corrosion resistance and require the replacement of all fuel probes. Furthermore, if the probe corrodes, the evaporation of Jet A1 fuel used in civil aviation (a fuel produced for gas turbine aircraft engines) may lead to hazardous consequences. Jet A1 fuel is obtained by distilling crude oil within a temperature range of 150°C to 270°C. Its freezing point is –47°C and its flash point is 38°C. This fuel, which is straw-yellow in color, has a high evaporation rate. Therefore, in the event of spillage or contact with air, it can ignite easily [18-22]. As an alternative solution to this issue, it is planned to install a ferrous alloy (cast iron) ring onto the aluminum alloy probes without making any changes to the probes themselves. Cast iron, which contains between 2.06% and 4% carbon, will preserve the corrosion resistance of the probe due to its ferromagnetic property and strength, while allowing it to be attracted by the magnet. The electromagnetic magnet circuit was designed using an Arduino board. The circuit includes an Arduino microcontroller, transistor, diode, 220Ω resistor, LCD screen (16x2), breadboard, jumper cables, and a potentiometer. The potentiometer controls the power of the magnet to prevent possible faulty

probe connections, and this power control value is displayed to the fuel supply operator via the LCD screen. The flowchart of the code used to program the Arduino is provided in Figure 1. Since these circuit components are not corrosion-resistant, they are coated with petroleum-derived plastic, which has high corrosion resistance.

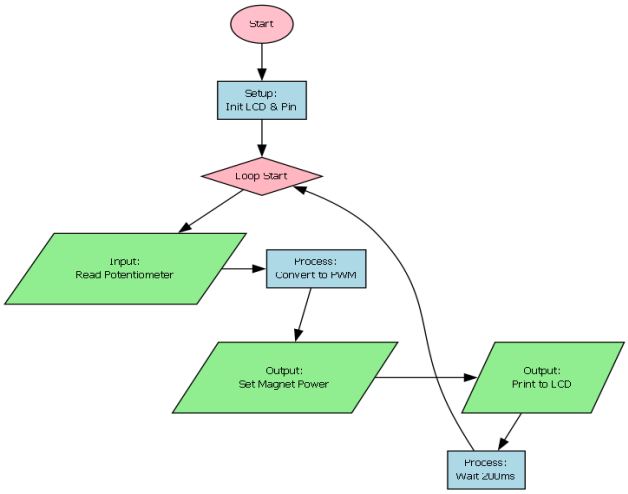


Figure 1. Flowchart of codes used in Arduino

As a result, the fuel supply probe an indispensable part of routine fuel supply operations—has been redesigned and modified in a way that makes it easier and safer to use, contrary to the conventional design. The nose gasket located in the latching mechanism of the 347GF (fuel supply probe) needs to be replaced over time due to wear. As a solution, it was deemed appropriate to modify the 347GF by adding a ferrous alloy ring to its neck and attaching an electromagnetic magnet to the hydrant flange. An electromagnetic magnet is a type of magnet that temporarily gains magnetic properties through the magnetic field produced when an electric current passes through a wound wire coil. Electromagnetic magnets have been widely used in real-life applications to facilitate work by utilizing their attractive force [22-23].

The following formulas are used to calculate the attractive force of the electromagnetic magnet [24].

$$B = \frac{(\mu \,.\, N \,.\, I)}{L} \tag{1}$$

B = Symbol of the intensity of the magnetic field
μ = Symbol of magnetic permeability of the material ($\frac{H}{m}$)
N = Refers to the winding number of the coil.
I = Current through the coil (Amper)
L = Coil length

$$F = \frac{(B^2 . A)}{2 \,.\, \mu_0} \tag{2}$$

F = Force of attraction (Newton)
B = Magnetic field intensity (Tesla)
A = Active area (m^2)
μ_0 = Magnetic permeability of vacuum ($4 \,.\, \square \,.\, 10^{-7}\ \frac{H}{m}$)

$$W = F \,/\, G \tag{3}$$

W = Load that can be carried (kg)
g = Acceleration of gravity ($9{,}81\ m/s^2$)

Since the load to be carried by the electromagnetic magnet is 22 kg, the force to be exerted by the electromagnet is approximately 215.82 N (Newtons). When the 24V DC voltage supplied from the aircraft and the magnetic field strength are calculated, the magnetic flux density is found to be 0.233 Tesla. By winding approximately 1000 turns on a 10 cm coil and calculating the current passing through it, the current value is approximately 18.54 amperes. These data are sufficient to lift the gross load of the fueling probe. Using iron alloy rings, the head of the 347-GF probe was removed and integrated onto the neck section (Figure 2) [25]. To give a brief idea about the research, the initial design of the probe head is shown in Figure 2. A ferromagnetic metal ring has been positioned in the red-colored area labeled as number 1 (representing the neck of the probe).

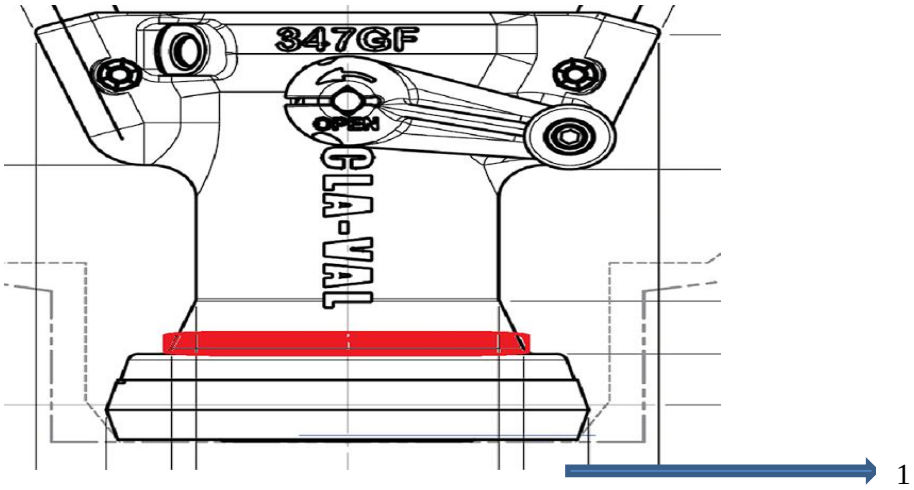


Figure 2. 347GF refueling probe [16]

A cylindrical electromagnetic magnet has been added onto the hydrant flange. Thus, a magnetic attractive force is provided between the iron alloy rings added to the 347GF and the hydrant. The control of the electromagnetic magnet is managed through the necessary circuit connections using a microcontroller board. A cylindrical electromagnetic magnet was added onto the flange of the hydrant. This created a magnetic attractive force between the iron alloy rings on the 347GF and the hydrant. The control of the electromagnetic magnet was designed using a microcontroller board with the necessary circuit connections.

In the existing system, the fuel supply probe's head must be manually aligned and fixed during refueling. This method requires considerable attention and labor. It is especially challenging under bad weather conditions and during night shifts. The stages of the system are described as follows:

- An electromagnetic magnet is integrated into the probe head.
- A ferromagnetic ring compatible with the magnet is mounted on the surface of the probe head.
- The 115-volt AC voltage from the aircraft is converted to 24 V DC voltage using a regulator to power the circuit, and the system is activated.
- The technician performing the refueling brings the supply hose close to the fuel probe.
- When the electromagnet is activated, the probe head automatically aligns and locks into place.
- After securing, the fuel coupling is attache.
- The electromagnet power is turned off, and fuel supply begins.
- Once fueling is completed, the electromagnet is powered again at full strength, and the locking mechanism is released.
- The electromagnets power is then reduced, and the probe is detached from the aircraft.
- The refueling process is completed safely and with time efficiency.

The flow diagram of the refueling coupling performed by the proposed method is given in Figure 3.

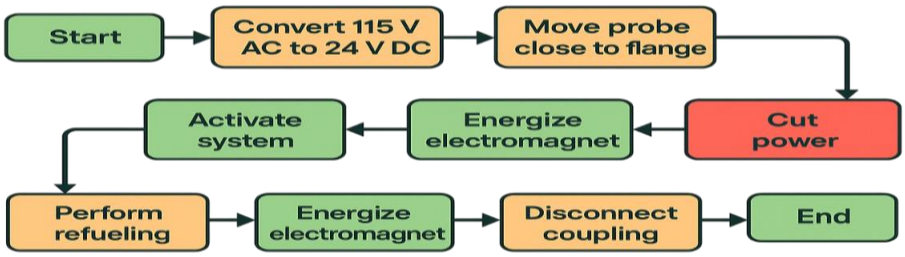


Figure 3. Flow chart for coupling connection

The drawing that technically specifies the positioning of the connection coupler made with the proposed method and the changes made on the aircraft and the prop. The visual obtained with the Python language and the Visual Studio Code application is given in figure 4.

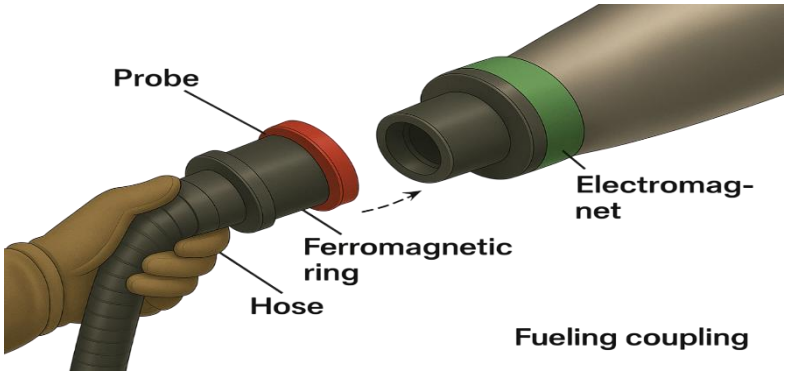


Figure 4. Connection coupling image obtained with Visual Studio Code application.

3. Experimental Results

The electromagnetic magnet to be mounted on the aircraft's fuel hydrant flange must have specific characteristics. The pulling force of the electromagnet depends on its core, voltage source, and diameter. The specifications of the electromagnet are presented in Table 1, calculated using formulas (1), (2), and (3). The PWM values, PWM ratios, magnetic flux density (Gauss), force, and effective range of the electromagnet at 0–24 V DC voltage are given in Table 2. The values presented in Table 2 are illustrated in Figure 3 using Python programming language in Visual Studio Code.

Table 1. Characteristics of the magnet used

Property	Value
External Diameter	151-200 mm (≈ 175 mm avg.)
Inner Diameter	101 mm+
Height	40 mm
The Power of Attraction	50-75 kg
Value of Gauss	7000-8000 Gauss
Type of	Ring type, electric

Source of Tension	24V DC
The max. Temperature of Operation	100 °C

Table 2. Findings based on the properties of the magnet used

PWM Value (0-255)	PWM Ratio (%)	Source (V)	Gauss (G)	The Force (kg)	The Effective Distance (mm)
0	%0	0V	0G	0 kg	0 mm
64	%25	6V	~2000G	~15 kg	~5-10 mm
128	%50	12V	~4000G	~30 kg	~10-15mm
192	%75	18V	~6000G	~50 kg	~15-20 mm
255	%100	24V	8000G	75 kg (maks.)	20+ mm

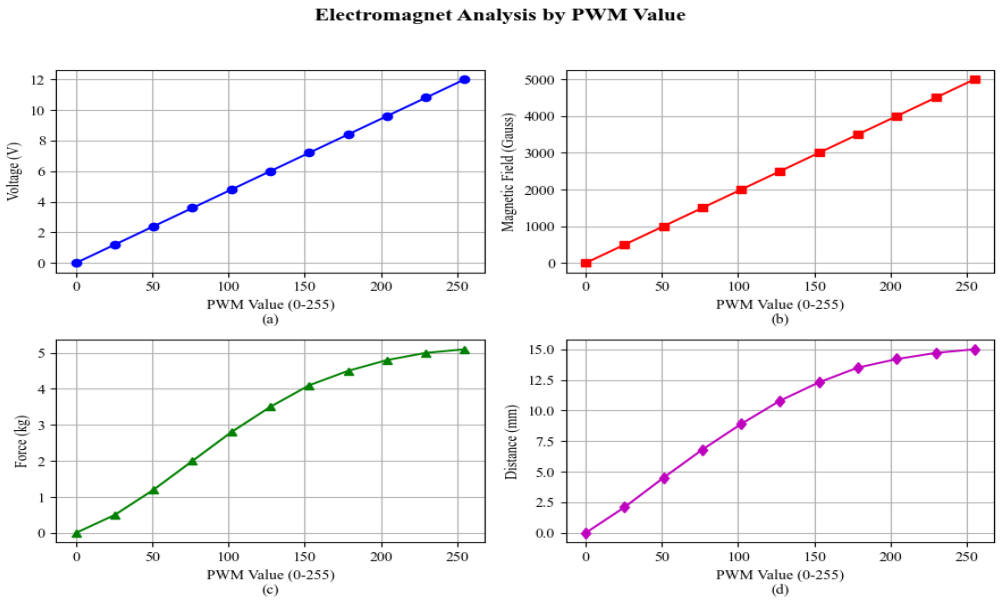


Figure 5. Electromagnet analysis by PWM Values (a) increase in PWM value by voltage quantity (b) the magnetic field relative to the increase in the value of PWM (c) the force to be applied by the magnet relative to the value of PWM (d) the distance that the electro magnet acts according to the value of PWM

If a comparison is made in terms of assembly time, a 38% reduction in assembly time is shown in Table 3.

Table 3. Comparison of connection times between the existing and designed

Type of System	Average Connectin Time (min)
Traditional	9.8 min
Electromagnet Assisted	6.1 min

Under normal conditions, the weight of the fuel supply probe and the connected system was measured using a spring-type dynamometer. A 42% reduction in the required physical force before and after adding the electromagnetic magnet attachment is shown in Table 4.

Table 4. Weights of the system used under normal conditions and the configured system

Type of System	Average Force (N)
Traditional	27.3 N
Electromagnet Assisted	15.8 N

The performance aspects of the system used are provided in Table 5.

Table 5. Characteristics of the electromagnet used

Property	Value
Power of attraction	50-75 kg
Gauss Value	7000-8000 Gauss
Source	24 V DC
Power Consumption (Avg.)	36W
Max. Tampreture	100°C

4. Conclusion and Discussion

In this study, the workload and time loss issues experienced by employees performing the "fueling procedures," which is one of the ground operations in aviation, were identified. According to the literature, the coupling connection between the fueling probe and the fueling flange must be carried out in accordance with established procedures. The difficulty during the connection process and the workload caused by potential energy during the connection lead to accidents and safety vulnerabilities. The inevitable problems were detected. As a solution to these problems, a circuit was designed using a microcontroller system and an electromagnetic magnet. An algorithm and flowchart to be uploaded to the Arduino microcontroller board for the electromagnetic magnet-supported design were developed [As shown in Figure 1]. The connection steps made with the proposed method were listed in items and a flow chart was created [As shown in Figure 3]. The connection coupling image obtained by coding the Python language and using the Visual Studio Code application was obtained. [As shown in Figure 4]. The properties of the magnet used were demonstrated using the equations (1), (2), and (3). The functionality of the electromagnetic magnet with the specified characteristics is presented in the findings [It is shown in Table 2], including the electromagnet analysis based on PWM values: (a) increase of PWM value according to voltage, (b) magnetic field, (c) force applied by the magnet, and (d) effective range of the electromagnet [As shown in Figure 5], as indicated in the tables and figures. The refueling time with the traditional system used at airports was determined to be approximately 9.8 minutes. With the modified probe, this time was reduced to about 6.1 minutes, showing an approximate 38% decrease in refueling time. The comparison of connection times between the system and the traditional system is provided in [As shown in Table 3]. The comparison of force requirements between the traditional system and the electromagnetic magnet-supported system is recorded in [As shown in Table 4]. The performance values of the electromagnetic magnet-supported system are shown in [As shown in Table 5]. All calculated values were based on the 24V DC voltage obtained by converting the internally available 115V AC voltage on the aircraft using a regulator.

Acknowledgements

This study has been supported by the TÜBİTAK 2209-A project numbered 1919B012473527.

5. References

[1] Turkish civil aviation authority. "Directive on procedures and principles for the preparation and approval of the master minimum equipment list (MMEL) and minimum equipment list (MEL)". Ankara, Turkey, Directive, 2011.

[2] ENERGY INSTITUTE. “EI 1585: GUIDANCE ON THE USE OF HYDRANT SYSTEM EMERGENCY STOP CONTROLS”. London, UK, 2018.
[3] Drury CG. *Human Factors in Aviation Maintenance*. In: Salvendy G. *Handbook of Human Factors and Ergonomics*, 3rd ed., pp. 1013-1033, USA, John Wiley & Sons, 2009.
[4] Aviation learnings team. “Aircraft refueling safety procedures & precautions | Safety in aircraft fueling on ground”. https://aviationlearnings.com/aircraft-refueling-safety-procedures-precautions-safety-in-aircraft-fueling-on-ground/, (28.5.2025).
[5] Ayson E, Dhanani R, Parker G. “THE 400FUEL SYSTEM”. *SAE Technical Paper*, 700276, 1970
[6] NATA AERO. “FUEL SYSTEM SERVICING”. 2003.
[7] And L, Whyte R. “JET FUEL İN CANADİAN OPERATİONS”. *SAE Technical Paper*, 1976.
[8] BOEING. “747-400 AIRPLANE CHARACTERISTICS FOR AIRPORT PLANNING”. USA, 2023.
[9] SKYbrary Aviation Safety. “B772, DENVER CO USA, 2001”. Online Report, SKYbrary, 2001.
[10] Haug H, Kubler D. *Industrial Robotics: Theory, Modelling and Control*. Springer, 2017.
[11] Patle BK, Sahu S. “A REVIEW ON MAGNETIC WORK HOLDING SYSTEMS”. *International Journal of Engineering Research & Technology (IJERT)*, 3(7), 1162-1166, 2014.
[12] Wegener K, Kyburz P. “INTELLIGENT CLAMPING SYSTEMS: A REVIEW”. *CIRP Journal of Manufacturing Science and Technology*, 5(2), 77-88, 2012.
[13] Stanley RK. *Advances in Magnetic Materials: Processing, Properties, and Performance*. ASM International, 2011.
[14] Fehse W. *Automated Rendezvous and Docking of Spacecraft*. Cambridge, United Kingdom, Cambridge University Press, 2003.
[15] McFatter J, Keiser K, Rupp T. “NASA DOCKING SYSTEM BLOCK 1: NASA'S NEW DIRECT ELECTRIC DOCKING SYSTEM SUPPORTING ISS AND FUTURE HUMAN SPACE EXPLORATION”. *44th Aerospace Mechanisms Symposium*, Cleveland, OH, USA, May 16-18, 2018.
[16] CLA-VAL COMPANY. “347GF AIRCRAFT UNDERWING REFUELING VALVE”. Costa Mesa, California, USA, Report, 347GF, 2020.
[17] Fitzgerald AE, Kingsley C Jr, Umans SD. *Magnetic Circuits and Magnetic Materials*. In: *Electric Machinery*, 6th ed., pp. 2-42, New York, NY, USA, McGraw-Hill, 2003.
[18] TURKISH CIVIL AVIATION AUTHORITY (SHGM). *Refueling Procedures and Safety*. Ankara, Turkey, SHGM Publications, 2024.
[19] Rosales Calderon O, Tao L, Abdullah Z, Moriarty K, Smolinski S, Milbrandt A, Talmadge M, Bhatt A, Zhang Y, Ravi V, Skangos C, Tan E, Payne C. “SUSTAINABLE AVIATION FUEL (SAF) STATE-OF-INDUSTRY REPORT: STATE OF SAF PRODUCTION PROCESS”. National Renewable Energy Laboratory, Golden, CO, USA, Technical Report, NREL/TP-5100-87802, 2024.
[20] TURKISH CIVIL AVIATION AUTHORITY (SHGM). “SUSTAINABLE AVIATION FUEL DIRECTIVE (SHT-SAF)”. Ankara, Turkey, Directive, SHT-SAF, 2022.
[21] Wojtas P, et al. “MAGNETIC PROPERTIES OF CAST IRONS”. *Journal of Magnetism and Magnetic Materials*, 2018.
[22] Callister WD Jr, Rethwisch DG. *Magnetic Properties*. In: *Materials Science and Engineering: An Introduction*, 10th ed., 714-722, Hoboken, NJ, USA, Wiley, 2018.
[23] Fontana MG. *Corrosion Engineering*. USA, McGraw-Hill, 1986.
[24] Serway RA, Jewett JW Jr. *Physics for Scientists and Engineers*. 9th ed. Boston, MA, USA, Brooks/Cole, 2014.
[25] Griffiths DJ. *Electrodynamics*. In: *Introduction to Electrodynamics*, 3rd ed., 285-301, Upper Saddle River, NJ, USA, Prentice Hall, 1999.

Effects Of Sodium Hydroxide On The Unconfined Compressive Strength Of Low Plasticity Soils

Tacettin GEÇKİL[1] , Talha SARICI [2], Özge Nur ÇETKİN [3]

[1,2] Inonu University, Civil Engineering Department, Malatya, Turkey.
[3] Bitlis Eren University, Vocational School of Technical Sciences, Bitlis, Turkey.
[1] tacettin.geckil@inonu.edu.tr, [2] talha.sarici@inonu.edu.tr, [3]oncetkin@beu.edu.tr
[1](ORCID: 0000-0001-8070-6836), [2](ORCID: 0000-0001-8488-5851),[3](ORCID: 0000-0002-0515-0337)

Abstract
As a result of the rapidly increasing environmental problems today, various alternative and environmentally friendly solutions are being produced in the construction industry as in many other areas. This study investigated the improvement of a low plasticity clayey soil with sodium hydroxide (NaOH). To this end, the soil's geotechnical properties were determined by applying sieve analysis, hydrometer, pycnometer, and consistency limit (Atterberg) tests. Then, NaOH was added to improve the soil, and pure and additive soil samples were prepared. The prepared pure and additive soil samples were subjected to the standard Proctor test in accordance with ASTM D698 and their optimum water content and maximum dry unit weight were determined. Based on the data obtained from the standard Proctor test, unconfined compression test samples were prepared in accordance with ASTM D2166. After 1 and 7 days of curing, the pure and additive soil samples were subjected to unconfined pressure tests. The tests revealed that after 7 days of curing, the unconfined compressive strength of the NaOH added soil sample increased by 4.38 times compared to the pure soil sample. As a result, in the improvement achieved by the addition of NaOH to the low plasticity clay soil, it was observed that the NaOH addition had a positive effect on the soil cohesion and internal friction angle.

Keywords: low plasticity, clay soil, sodium hydroxide, unconfined compressive strength

1. Introduction

With the recent increase in population, there has also been a rise in the demand for various needs. Among these needs, construction structures such as housing and transportation come first. One of the most important factors for the construction of these structures under desired conditions is the ground on which they are built. However, soils do not always provide the desired conditions. For this reason, various methods have been developed to improve the engineering properties of soils. Among these developed methods, soil stabilization with various additives is the most preferred.

Kulanthavel et al. (2023) investigated the changes that occur when clay soil is stabilized with Portland cement, sodium silicate, and sodium hydroxide. The optimum additive content was determined as 10% Portland cement + 4% sodium silicate + 8% sodium hydroxide. Following the unconfined compressive strength (UCS) test, the maximum UCS value was measured as 212.22 kPa. Hydration and the formation of pozzolanic components were confirmed through Scanning Electron Microscope (SEM) analyses [1]. Ali et al. (2020) investigated the effects of three sodium-based additive solutions — sodium hydroxide (NaOH), sodium carbonate (Na_2CO_3), and sodium chloride (NaCl) — on low-plasticity clay soil mixtures. For this purpose, unconfined compressive strength, California Bearing Ratio (CBR), and swelling potential tests were conducted. It was observed that the addition of NaOH and Na_2CO_3 increased strength, while NaCl had a limited effect [2]. Estabrgh et al. (2022) used ground granulated blast furnace slag (GGBFS), GGBFS + magnesium oxide (MgO), and GGBFS + MgO + cement mixtures for clay soil stabilization. The mixtures were subjected to Atterberg limit and standard compaction tests. The results indicated that all additives increased strength, with the improvement being dependent on additive content and curing time [3]. Soğancı et al. (2023) studied the effects of marble dust and GGBFS on improving the geotechnical properties of low- and high-plasticity clays. Mixture ratios ranged between 5% and 20%, and the optimum additive content was determined as 15% marble dust + 10% GGBFS for low-plasticity clay soils, and 10% marble dust + 15% GGBFS for high-plasticity clay soils [4]. Çınar and Erbaşı (2022) investigated the usability of geopolymers. It was determined that the mixing ratio of alkali activators provided sufficient diversity for alumina and silica gelation. The geopolymer mixture was observed to provide greater strength compared to cement-based mixtures [5].

[1]Corresponding author

In this study, the changes in unconfined compressive strength resulting from the improvement of low-plasticity clay soil with low bearing capacity using sodium hydroxide were investigated. For this purpose, sieve analysis, hydrometer, pycnometer, and consistency limit tests were conducted to determine the geotechnical properties of the soil. Then, stabilized soil samples were prepared by adding sodium hydroxide to the soil. Both untreated and stabilized soil samples were subjected to standard Proctor and unconfined pressure tests.

2. Materials and Methods

2.1 Materials

The soil used in this study was obtained from the campus of İnönü University, Faculty of Engineering. According to the Unified Soil Classification System (USCS), the soil was classified as low-plasticity clay (CL). The properties of the soil used are presented in Table 1.

Table 1. Physical properties of the soil

Variable Name	Symbol	Unit	Value
Particle unit weight	γs	g/cm³	2.566
Maximum dry unit weight	γkmax	g/cm³	1.869
Optimum water content	wopt	%	14.50
Liquid limit	wL	%	24.44
Plastic limit	wp	%	19.50
Plasticity index	PI	%	11.31
Classification (USCS)	–	–	CL
AASHTO Classification	–	–	A-6

The sodium hydroxide (NaOH) solution used in the study was prepared by mixing crystalline NaOH with distilled water. The molarity of the solution was chosen as 8 M (320 g/L), determined with reference to the literature [6,7]. The properties of NaOH are given in Table 2 [8].

Table 2. Properties of sodium hydroxide (NaOH)

Physical and Chemical Properties	Sodium Hydroxide
Molecular formula	NaOH
Molecular weight (g/mol)	40.00
Color	White
pH	13–14
Relative density (g/cm³)	2.13

2.2 Methods

In order to determine the geotechnical properties of the soil, sieve analysis and hydrometer tests (ASTM D422-63), pycnometer test (ASTM D854), and consistency limit tests (ASTM D4318) were performed. Images related to the tests are shown in Figure 1.

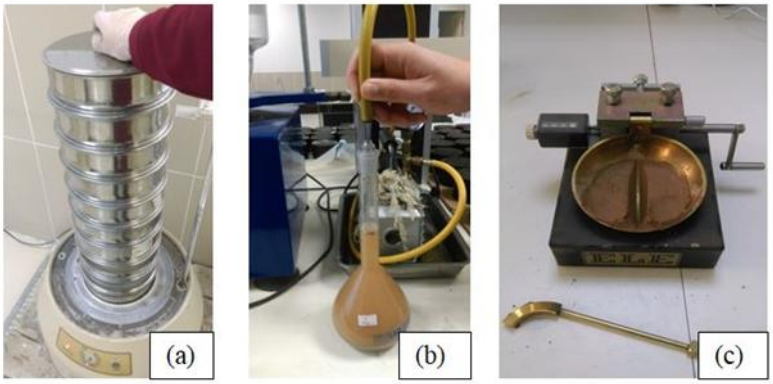


Figure 1. Sieve (a), pycnometer (b), and consistency limit (c) tests

After determining the geotechnical properties of the soil, an 8 M NaOH solution was added, and both untreated and stabilized soil samples were prepared. The prepared samples were tested using the Standard Proctor test (ASTM D698) and the Unconfined Compressive Strength (UCS) test (ASTM D2166).

The Standard Proctor test was used to determine the maximum dry unit weight and optimum water content of the samples. The dry unit weight (γk) of the soil indicates its degree of compaction; the higher the γk value, the more compacted the soil is [9]. Based on the maximum dry unit weight (γk) and optimum water content (wopt) values obtained from the Standard Proctor test, UCS test samples were prepared.

The UCS test is conducted to determine the shear strength of cohesive soils [10]. UCS test samples were prepared in accordance with the standard and were subjected to testing after curing periods of 1 and 7 days. The UCS testing apparatus is shown in Figure 2.



Figure 2. Unconfined compressive strength testing apparatus

3. Results and Discussion

In this study, which investigates the effects of sodium hydroxide (NaOH) on the unconfined compressive strength of low-plasticity clay, the geotechnical properties of the soil were first determined through sieve analysis, hydrometer, pycnometer, and consistency limit tests. Based on the data obtained from these tests, the grain size distribution curve of the soil is presented in Figure 3.

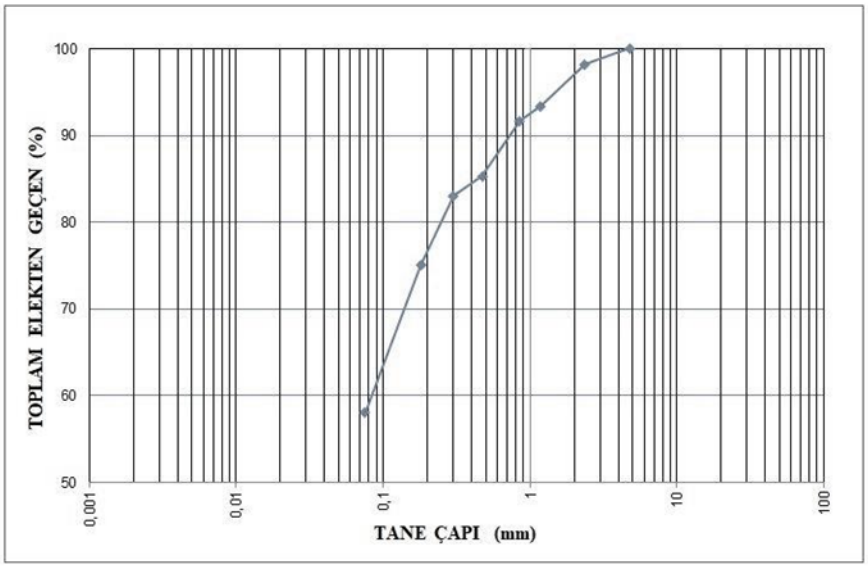


Figure 3. Grain size distribution curve of the soil

After determining the geotechnical properties of the soil, an 8 M NaOH solution was added, and both untreated and stabilized soil samples were prepared for the Standard Proctor test. At the end of the Standard Proctor test, the optimum water content (wopt) and maximum dry unit weight (γk) values were determined, and the compaction curves were plotted (Figure 4). The γk and wopt values are presented in Table 3.

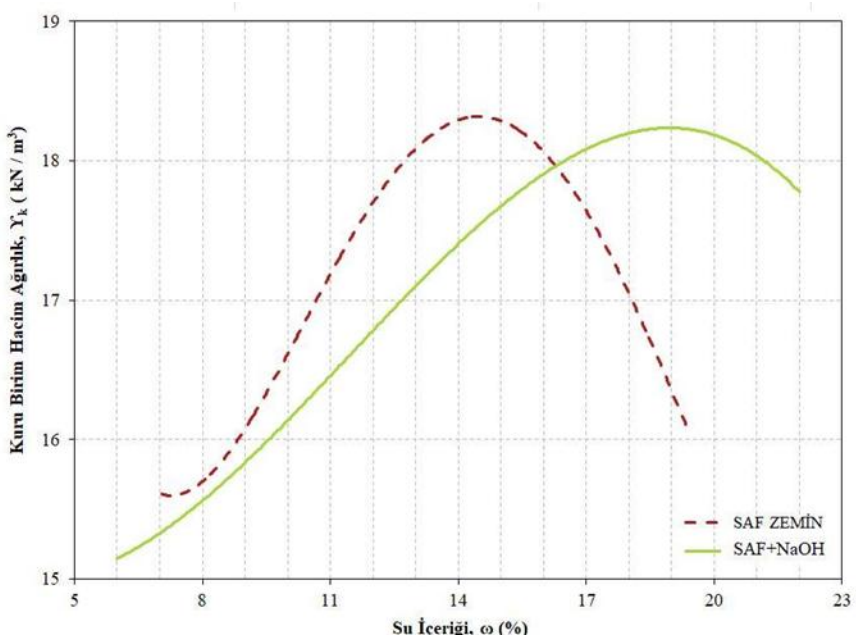


Figure 4. Compaction curve

Table 3. Results of the Standard Proctor test

Test Sample	Maximum Dry Unit Weight γk (kN/m³)	Optimum Water Content wopt (%)
Untreated Soil	18.33	14.50
Untreated Soil + NaOH	18.15	14.30

From the Standard Proctor test results, it was observed that the maximum dry unit weight decreased, while the optimum water content increased. Based on the γk and wopt values obtained from the Standard Proctor test, UCS test specimens were prepared, and after curing periods of 1 and 7 days, the soil samples were subjected to the unconfined compressive strength (UCS) test. In the UCS test, the stress (q) and strain (Ɛ) values of the samples were determined. The stress–strain curve obtained at the end of the test is presented in Figure 5, while the corresponding stress–strain values are given in Table 4.

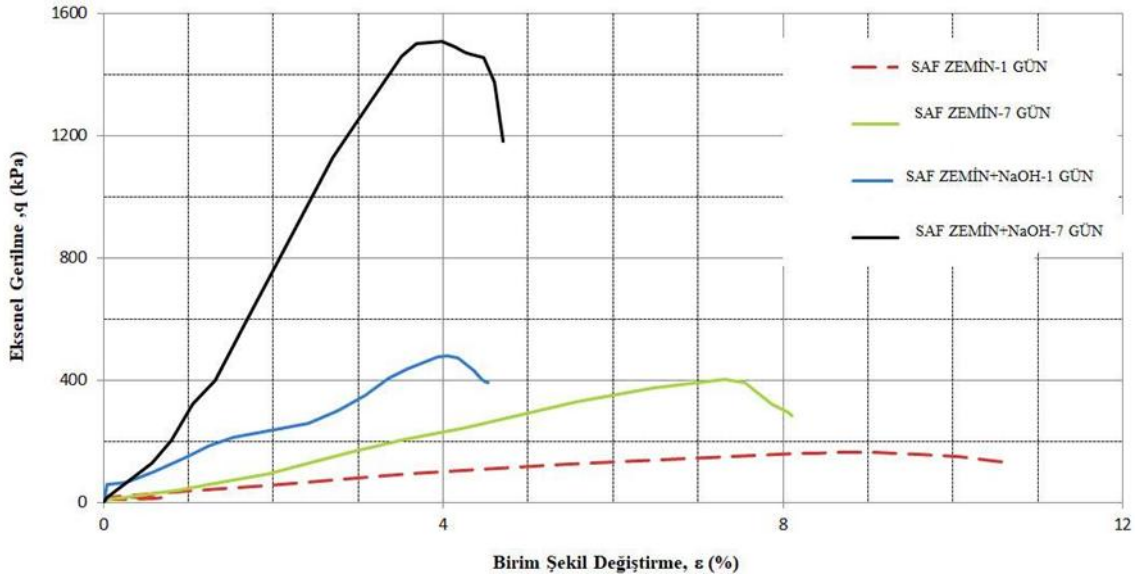


Figure 5. Stress–strain curve obtained from UCS test

Table 4. UCS and strain values obtained from the UCS test

Test Sample	UCS, q (kPa)	Strain at UCS, Ɛ (%)
Untreated Soil – 1 Day	163.628	9.047
Untreated Soil – 7 Days	323.747	7.308
Untreated Soil + NaOH – 1 Day	478.047	4.044
Untreated Soil + NaOH – 7 Days	1419.921	3.378

From the results of the unconfined compressive strength (UCS) test, it was observed that the UCS value of the soil increased with the addition of NaOH. Examination of Table 4 indicates that stabilization with NaOH gains strength over time, highlighting the necessity of evaluating the long-term strength values of the samples [7].

4. Conclusions

The findings obtained in this study, which investigated the effects of sodium hydroxide on the unconfined compressive strength of low-plasticity clay soils, can be summarized as follows:

From the Standard Proctor test results, it was observed that the addition of NaOH increased the optimum water content of the soil while decreasing the maximum dry unit weight.

From the UCS test results, it was determined that the UCS value increased with curing time when NaOH was added to the soil. This indicates that stabilization with NaOH gains strength over time. The highest UCS value was obtained in NaOH-stabilized soil after 7 days of curing, and the strength was found to be 4.38 times greater compared to untreated soil.

In conclusion, it can be stated that curing time has a significant effect on the unconfined compressive strength of soils stabilized with sodium hydroxide. Furthermore, NaOH can be considered as an alternative to traditional stabilizing agents such as cement, potentially contributing positively to environmental sustainability.

5. References

[1] Kulanthaivel, P., Selvakumar, S., Soundara, B., & Krishnaraja, A. R. (2023). Strength enhancement of clay soil stabilized with ordinary portland cement, sodium silicate and sodium hydroxide. *International Journal of Pavement Research and Technology*, *16*(5), 1297-1310.

[2] Ali, H. E., Asmel, N. K., Ganiyu, A. A., & Tijani, H. (2020). Effect of sodium compounds additives on the strength of cement-stabilized soils. *Engineering and Applied Science Research*, *47*(3), 287-296.

[3] Estabragh, A. R., Jahani, A., Javadi, A. A., & Babalar, M. (2022). Assessment of different agents for stabilisation of a clay soil. *International Journal of Pavement Engineering*, *23*(2), 160-170.

[4] Soğancı, AS, Yenginar, Y. ve Orman, A. (2023). Mermer tozu ve granüle yüksek fırın cürufu ile stabilize edilmiş killi zeminlerin jeoteknik özellikleri. *KSCE İnşaat Mühendisliği Dergisi* , *27* (11), 4622-4634.

[5] Çınar, M., & Erbaşı, B. (2022). GEOTEKNİK UYGULAMALARDA GEOPOLİMERLERİN KULLANILABİLİRLİĞİNİN İNCELENMESİ, LİTERATÜR ÇALIŞMASI. *Kahramanmaraş Sütçü İmam Üniversitesi Mühendislik Bilimleri Dergisi*, *25*(4), 774-789.

[6] Somna K. , Jaturapitakkul C. , Kajitvichyanukul P. , Chindaprasit P. , (2011) Fuel Volume 90, Issue 6, Pages 2118-2124

[7] Eskisar, T., & Gizem, A. K. S. U. (2020). Zeminlerde Tek Fazli Geopolimerizasyon Uygulamasi Ve Geopolimerizasyonun Serbest Basinç Mukavemeti Üzerindeki Etkisi. Konya Mühendislik Bilimleri Dergisi, 8(3), 466-478.

[8] Ekinci E. (2017). Volkanik tüf kullanılarak üretilen geopolimer betonların bazı mekanik ve fiziksel özelliklerinin araştırılması (Master's thesis, İnönü Üniversitesi Fen Bilimleri Enstitüsü).

[9] Geçkil T., Sarıcı T., Yıldıran E.S., 'Kireç Katkısı ile Kil Bir Zeminin Dayanımının İyileştirilmesi', Çukurova Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, 34(4), ss. 171-179, Aralık 2019.

[10] Yılmaz I. Vd.(2017), Zemin Mekaniği Labaratuvar Deneyleri ve Çözümlü Problemler, Seçkin Kitabevi Ankara

Examination Of The Behavior Of Soil Stabilized With Silica Fume

Tacettin GEÇKİL[1] , Talha SARICI [2], Özge Nur ÇETKİN [3]

[1,2] Inonu University, Civil Engineering Department, Malatya, Turkey.
[3] Bitlis Eren University, Vocational School of Technical Sciences, Bitlis, Turkey.
[1] tacettin.geckil@inonu.edu.tr, [2] talha.sarici@inonu.edu.tr, [3]oncetkin@beu.edu.tr
[1](ORCID: 0000-0001-8070-6836), [2](ORCID: 0000-0001-8488-5851),[3](ORCID: 0000-0002-0515-0337)

Abstract
This study investigated the changes in the behavior of a low-plasticity clay soil stabilized with silica fume (SF), an industrial waste product. For this purpose, the soil was first subjected to sieve analysis, hydrometer, pycnometer, and consistency limit (Atterberg) tests to determine its geotechnical properties. Then, in accordance with the purpose of the study, SF was added to the soil at 2.5% and 30% by weight to prepare modified soil samples. The prepared pure and modified soil samples were subjected to standard Proctor and unconfined compression tests in accordance with ASTM D698 and ASTM D2166 standards, respectively. The tests revealed that the unconfined compressive strength of the SF added soil increased by 2.90 times compared to the pure soil after 7 days of curing. Consequently, it can be stated that soil stabilization with SF has a positive effect on the curing time, cohesion, and internal friction angle. At the same time, it is thought that the use of silica fume, which is an industrial waste, will contribute to the country's economy and the environment by carrying out an environmentally friendly process.

Keywords: silica fume, stabilization, unconfined pressure test

1. Introduction

Soils used as building foundations or construction materials do not always meet the desired engineering properties. When soils fail to meet these engineering requirements, they can cause various related problems. Therefore, soils are improved using various methods. Among these improvement techniques, one of the most preferred is soil stabilization using various additives [1,2]. Aksoy (2010) conducted a study in which silica fume (SF) was added to clay soil to prepare test specimens. Unconfined compressive strength tests were carried out on the prepared specimens after curing periods of 1, 7, and 30 days. The results showed that while the plasticity of the soil remained unchanged, there was a decrease in the maximum dry unit weight [3]. Fındık (2024) investigated the stabilization of clay soils using soft clay and iron slag. For this purpose, slag–clay mixtures were prepared in proportions of 10%, 15%, and 20% of the soil's dry weight. The prepared specimens were examined through durability and compaction tests. The results indicated that moisture content, swelling percentage, and swelling pressure values decreased in the stabilized soil samples [4]. Uysal et al. (2020) examined the engineering properties of fine-grained soils stabilized with different waste materials. In this context, fly ash, silica fume, blast furnace slag, and glass powder were added to the fine-grained soil for stabilization. The study concluded that blast furnace slag was the most effective additive in reducing the plasticity index and increasing undrained shear strength [5]. Phanikumar et al. (2020) studied the plasticity and compaction properties of clay soil improved with silica fume. The Atterberg limit values of the clay decreased with increasing silica fume content. In addition, the California Bearing Ratio (CBR) of silica fume–treated clay improved significantly [6]. Ramazeni et al. (2023) used glass powder, silica fume, and sugarcane ash for geopolymerization. The study found that increasing the sugarcane ash content significantly improved the unconfined compressive strength and geopolymer gel formation. Furthermore, the addition of glass powder and silica fume was also observed to increase unconfined compressive strength [7].

2. Materials And Methods

2.1. Materials

The soil used in this study was obtained from the campus of the Faculty of Engineering, İnönü University. Based on laboratory test results, the soil was classified as low-plasticity clay (CL) according to the Unified Soil Classification System (USCS). The properties of the soil used in the study are presented in Table 1.

[1]Corresponding author

Table 1. Physical properties of the soil

Variable Name	Symbol	Unit	Value
Specific gravity of soil particles	γs	g/cm³	2.566
Maximum dry unit weight	γdmax	g/cm³	1.869
Optimum moisture content	wopt	%	14.50
Liquid limit	wL	%	24.44
Plastic limit	wp	%	19.50
Plasticity index	PI	%	11.31
Classification (Unified Soil Classification System – USCS)	–	–	CL
AASHTO classification	–	–	A-6

The silica fume used in this study was obtained from Eti Electrometallurgy Company in Antalya, Türkiye. The properties of the silica fume (SF) used in the study are presented in Table 5.2.

Table 2. Properties of silica fume (SF)

Property	Composition (%)
Al_2O_3	1.2
CaO	0.4
Fe_2O_3	0.6
K_2O	1.4
MgO	8.8
MnO	0.1
Na_2O	1.0
SO_3	0.6
SiO_2	83.4

2.2. Methods

In order to determine the soil classification used in this study, sieve analysis (ASTM C136/C136M), hydrometer test (ASTM D422-63), pycnometer test (ASTM D824), and Atterberg limits tests (ASTM D4318) were conducted. The testing equipment used for soil classification is presented in Figure 1.

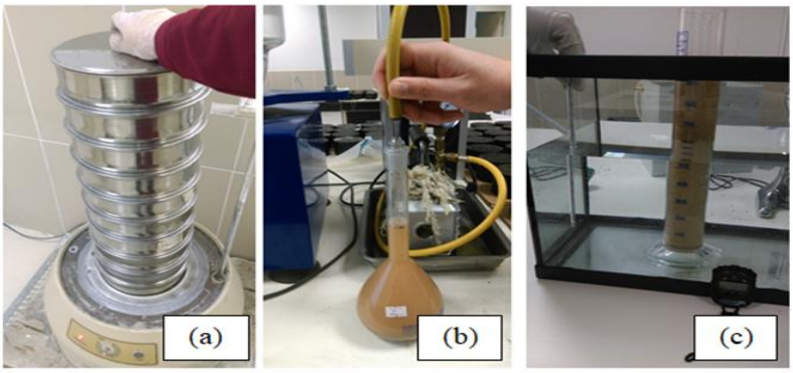


Figure 1. Experimental devices: Sieve Shaker (a), Pycnometer (b), Hydrometer (c)

After determining the soil classification, stabilized mixture samples were prepared by adding SD to the soil at weight percentages of 2.5%, 5%, 10%, 15%, 20%, 25%, and 30%. The prepared samples were subjected to the Standard Proctor (ASTM D698) and Unconfined Compressive Strength (ASTM D2166) tests.

In the Standard Proctor (compaction) test, the degree of compaction is measured by the dry unit weight of the soil (γk) The higher the γk value, the more compacted the soil is. The water content corresponding to the maximum γk value is called the optimum moisture content (w_{opt}) [8]. For each additive ratio determined for the Standard Proctor test, mixture samples were prepared. Based on the data obtained from the Standard Proctor test, samples for the Unconfined Compressive Strength test were prepared.

The Unconfined Compressive Strength test is conducted on cohesive soils to determine their shear strength. The samples were tested after 1 and 7 days of curing, and their undrained shear strength and cohesion values were measured. Figure 2 shows the unconfined compressive strength testing apparatus.



Figure 2. Unconfined compressive strength testing apparatus

3. Results and Discussion

In this study, the effects of SF stabilization on the unconfined compressive strength of a low-plasticity clay soil with poor bearing capacity were investigated. The geotechnical properties of the soil were determined using sieve analysis, hydrometer, pycnometer, and consistency (Atterberg limits) tests. Based on the data obtained from these experiments, the soil gradation curve was created and is presented in Figure 3.

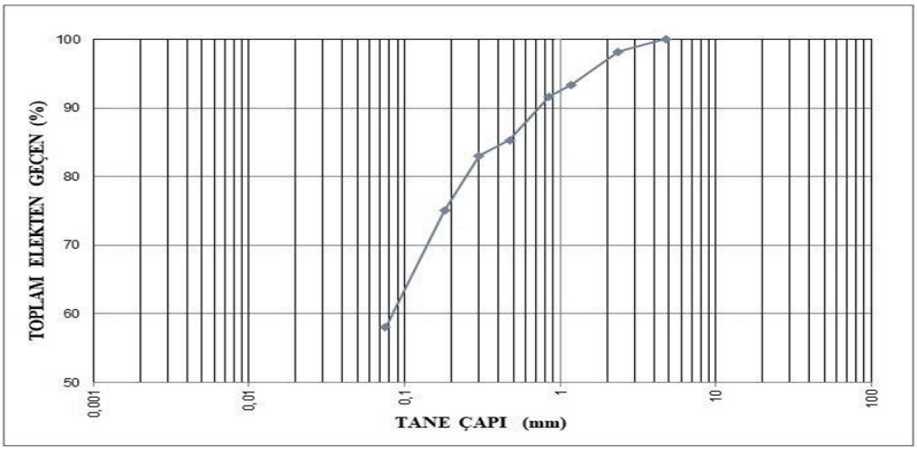


Figure 3. Soil gradation curve

After determining the geotechnical properties of the soil, stabilized soil samples were prepared by adding SD at weight percentages of 2.5%, 5%, 10%, 15%, 20%, 25%, and 30%. The prepared samples were subjected to the Standard Proctor test to determine the maximum dry unit weights (γk) and optimum moisture contents (w_{opt}) and the

compaction curve shown in Figure 4 was plotted. The determined maximum dry unit weights and optimum moisture contents are presented in Table 3**.**

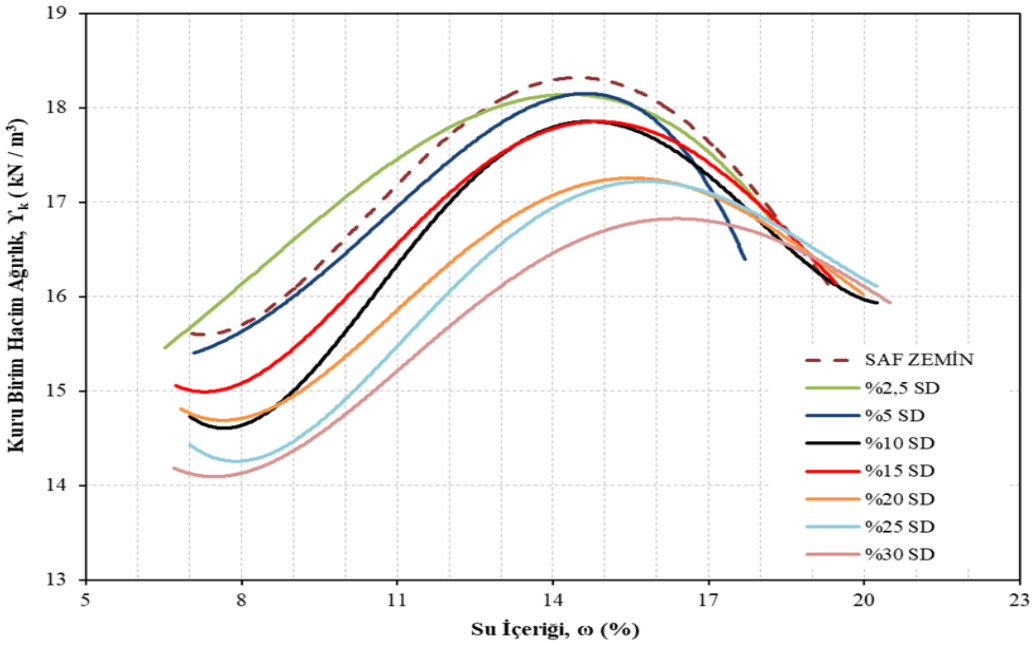


Figure 4. Standard proctor test compaction curves

Table 3. Standard proctor test results

Test Sample	Maximum Dry Unit Weight, γk (kg/cm³)	Optimum Moisture Content, w (%)
Pure Soil	18.33	14.50
Pure Soil + 2.5% SD	18.15	14.30
Pure Soil + 5% SD	18.15	14.60
Pure Soil + 10% SD	17.85	14.70
Pure Soil + 15% SD	17.85	14.90
Pure Soil + 20% SD	17.27	15.50
Pure Soil + 25% SD	17.22	15.80
Pure Soil + 30% SD	16.82	16.40

When the Standard Proctor test results were examined, it was observed that the addition of SF increased the optimum moisture content and decreased the maximum dry unit weight. This change was found to be consistent with the literature. This effect is thought to be due to SF's water absorption capacity and its fine-grained nature [9,10].
Based on the γk_{max} and w_{opt} values obtained from the Standard Proctor test, samples for the Unconfined Compressive Strength test were prepared. The prepared samples were tested after 1 and 7 days of curing, and the stress (q) and strain (ε) values were measured.

The stress–strain graph after 1 day of curing is shown in Figure 5, and the corresponding stress–strain values are given in Table 4**.** The stress–strain graph after 7 days of curing is shown in Figure 6**,** and the corresponding stress–strain values are presented in Table 5**.**

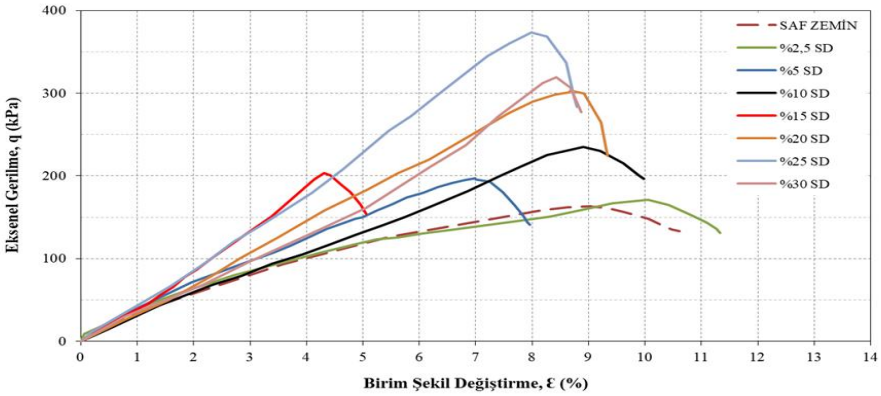


Figure 5. Stress–Strain (q–ε) Curve obtained from the unconfined compressive strength test (1 day curing)

Table 4. Stress–Strain (q–ε) Values obtained from the unconfined compressive strength test (1 day curing)

Test Sample	Unconfined Compressive Strength (UCS), q (kPa)	Unit Strain, ε (%)
Pure Soil	163.63	9.05
Pure Soil + 2.5% SD	170.83	10.05
Pure Soil + 5% SD	197.08	6.98
Pure Soil + 10% SD	234.73	8.91
Pure Soil + 15% SD	203.19	4.31
Pure Soil + 20% SD	303.31	8.68
Pure Soil + 25% SD	373.21	7.99
Pure Soil + 30% SD	319.43	8.43

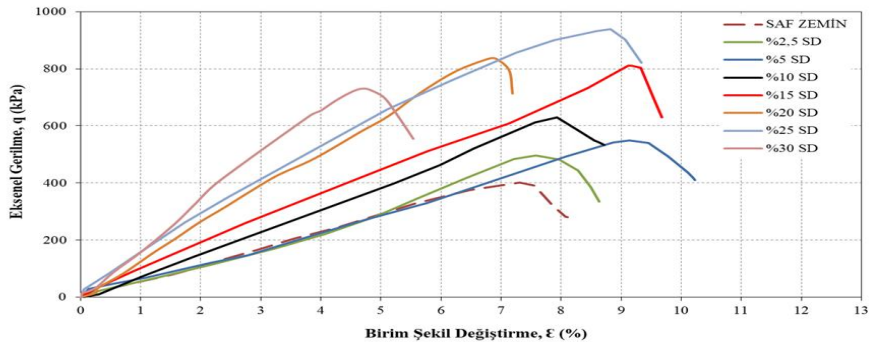


Figure 6. Stress–Strain (q–ε) Curve obtained from the unconfined compressive strength test (7 days curing)

Table 5. Stress–Strain (q–ε) values obtained from the unconfined compressive strength test (7 Days Curing)

Test Sample	Unconfined Compressive Strength (UCS), q (kPa)	Unit Strain, ε (%)
Pure Soil	323.75	7.31
Pure Soil + 2.5% SD	495.72	7.58
Pure Soil + 5% SD	549.38	9.13
Pure Soil + 10% SD	629.19	7.93
Pure Soil + 15% SD	811.36	9.17
Pure Soil + 20% SD	835.10	6.92
Pure Soil + 25% SD	939.45	8.82
Pure Soil + 30% SD	730.71	4.71

In the SF (silica fume) stabilized samples, it was observed that the strength values increased after 1 and 7 days of curing. This increase can be attributed to the fact that SF is a pozzolanic material, which gains strength over time [9,11].

4. Conclusions

This study investigated the effects of SF stabilization on the unconfined compressive strength of a low-bearing-capacity subgrade soil. Based on the obtained data, the following conclusions were drawn:

- Examination of the Standard Proctor test results showed that as the additive content increased, the water content values increased while the dry unit weight decreased. This change is thought to be due to SF's fine-grained nature.
- Examination of the Unconfined Compressive Strength test results showed that as the additive content increased, the strength values also increased, with the highest strength observed at 25% SF content. This increase indicates that SF has a significant effect on the soil's cohesion.
- In conclusion, soil stabilization with SF has a positive effect on unconfined compressive strength. Additionally, using SF, an industrial by-product, is expected to contribute to both the national economy and the environment.

5. References

[1] Çetin, A. Y. (2011). Yüksek plastisiteli kil zeminlerin alternatif malzemeler ile yüzeysel zemin stabilizasyonu (Doctoral dissertation, Fen Bilimleri Enstitüsü).

[2] Geçkil, T., Tanyıldızı, M. M., & Yıldıran, E. S. (2020). Yüksek Fırın Cürufu ile Stabilize Edilmiş Killi Bir Zeminin Yol Esnek Üstyapı Tabaka Kalınlıklarına ve Maliyetine Etkileri. Firat University Journal of Engineering, 32(2).

[3] Aksoy, H. S. (2010). SİLİS DUMANININ KİLLİ BİR ZEMİN ÜZERİNDEKİ STABİLİZASYON ETKİSİ. *Fırat Üniversitesi Doğu Araştırmaları Dergisi*, *9*(1), 74-78.

[4] Fındık, M. (2024). Kilin Çelik Cürufu ile Stabilizasyonu. *Teknik Bilimler Dergisi*, *14*(2), 19-28.

[5] Uysal, F., Yılmaz, V., & Topçu, H. M. (2020). Farklı atık malzemeler ile stabilize edilmiş ince daneli zeminin mühendislik özellikleri. *Çukurova Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, *35*(1), 19-26.

[6] Phanikumar, B. R., & e, R. R. (2020). Silica fume stabilization of an expansive clay subgrade and the effect of silica fume-stabilised soil cushion on its CBR. *Geomechanics and Geoengineering*, *15*(1), 64-77.

[7] Ramezani, SJ, Toufigh, MM ve Toufigh, V. (2023). Toprak stabilizasyonu için şeker kamışı bagası külü bazlı jeopolimerde cam tozu ve silika dumanının kullanımı. *İnşaat Mühendisliğinde Malzeme Dergisi* , *35* (4), 04023042.

[8] Geçkil T., Sarıcı T., Yıldıran E.S., 'Kireç Katkısı ile Kil Bir Zeminin Dayanımının İyileştirilmesi', Çukurova Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, 34(4), ss. 171-179, Aralık 2019.

[9] Sarıcı T, (2019). Puzolan Ile Güçlendirilmiş İnşaat Ve Yıkıntı Atıklarının Granüler Dolgu Olarak Kullanılabilirliğinin Değerlendirilmesi, İnönü Üniversitesi, FBE, Doktora Tezi.

[10] Türköz, M., Umu, S. U., & Öztürk, O. (2021). Effect of Silica Fume as a Waste Material for Sustainable Environment on the Stabilization and Dynamic Behavior of Dispersive Soil. Sustainability, 13(8), 4321.

[11] Geçkil, T., Tanyildizi, M. M., & Yildiran, E. S. (2020). Yol Altındaki Killi Bir Zeminin Uçucu Kül Ile İyileştirilmesinin Karayolu Üstyapısına Ve Maliyetine Etkileri. Savunma Bilimleri Dergisi, (38), 1-21.

Determining the Optimal Stator Slot–Rotor Pole Combination for Brushless DC and Permanent Magnet Synchronous Motor

Hayatullah NORY[1] , Ahmet ORHAN [2], Cansu AKSOY[3]

[1,2]Electrical-Electronics Engineering, Engineering Faculty, Firat University, Elaziğ, Türkiye.
3Electrical-Electronics Engineering, Engineering Faculty, Ege University, Izmir, Türkiye.
[1]hayatullahnory@gmail.com, [2]aorhan@firat.edu.tr, [3]aksoycansu94@gmail.com
[1](ORCID: 0000-0002-3375-4744), [2](ORCID: 0000-0003-1994-4661),[3](ORCID: 0000-0001-5854-9061)

Abstract
This study investigates the optimal stator slot–rotor pole combination for Brushless Direct Current and Permanent Magnet Synchronous Motors to enhance electromagnetic performance and manufacturability. A comprehensive 13 × 20 slot–pole matrix was developed, evaluating key design parameters including winding factor, winding layout, number of slots per pole per phase, symmetry number, cogging torque characteristics via least common multiple analysis, and magneto-motive force harmonics. The methodology integrates analytical formulations and tabulated results to identify configurations that minimize torque ripple, suppress undesirable harmonics, and maximize efficiency. Both concentrated and distributed winding arrangements were examined, highlighting trade-offs in performance and complexity. High winding factor values were associated with improved torque production and reduced losses, while optimal least common multiple values contributed to shifting torque ripple to less detrimental frequency ranges. Harmonic analysis emphasized the mitigation of lower-order non-triplen harmonics to prevent efficiency loss and vibration. The findings provide a structured approach for selecting slot–pole combinations that balance electrical, magnetic, and mechanical requirements, offering practical guidance for the design of high-performance Permanent Magnet Synchronous Motor and Brushless Direct Current motors in industrial, automotive, and aerospace applications.

Keywords: slot–pole combination, winding factor, cogging torque, harmonics, PMSM, BLDC, Symmetry, Winding schematic, Distributed winding, Concentrated winding, LCM, GCD, MMF

1. Introduction

Permanent Magnet Synchronous Motors (PMSMs) and Brushless Direct Current (BLDC) motors are among the most widely adopted electric machines due to their compact design, high torque per rated volume (TRV), and superior electrical efficiency. These characteristics have enabled their extensive use in a variety of modern applications, including electric vehicles, unmanned aerial vehicles (UAVs) such as drones, computer numerical control (CNC) machines, industrial automation systems, and aerospace technologies.

A key factor contributing to the performance and application versatility of PMSM and BLDC motors is their torque–speed characteristic. This relationship can be divided into two distinct operating regions as shown in Figure, the constant torque region and the constant power region. From standstill up to the base speed, the motor operates in the constant torque region, delivering a steady torque output. Beyond the base speed, operation transitions into the constant power region, where torque decreases as speed increases while power output remains constant.

PMSM and BLDC motors are available in numerous structural and magnetic topologies. Structural configurations include outer rotor and inner rotor designs. Magnet arrangements vary across surface-mounted (SM), V-type, interior permanent magnet (IPM), and inset magnet (IM) configurations. Magnetization orientations may be parallel or radial, while flux path arrangements can be classified as axial flux or radial flux. Furthermore, magnet materials differ in composition, with common choices including neodymium–iron–boron (NdFeB), ferrite, and samarium–cobalt (SmCo) [1] [2] [3].

Irrespective of these classifications, PMSMs and BLDC motors generally achieve the highest efficiencies, high TRV, and superior power density compared to other electrical motor types. High efficiency results in reduced electrical losses and lower temperature rise during operation. However, these advantages are accompanied by certain drawbacks, primarily related to permanent magnet dependency. Approximately 90% of global rare-earth permanent magnet production particularly NdFeB is concentrated in China, creating a supply chain vulnerability. Additionally,

[1]Corresponding author

NdFeB magnets rely on rare earth elements, which are finite in availability. While recycling technologies for NdFeB magnets exist, the process is limited to specialized facilities, and the performance of recycled magnets is generally inferior to that of new ones [4]. Despite these limitations, PMSMs and BLDC motors continue to be the preferred choice in many applications, as other motor types present their own constraints.

Among the PMSM and BLDC motor design considerations, one of the most critical is the stator slot–pole combination. The selection of an optimal combination forms the foundation of motor design, influencing winding layout, magnetic performance, and mechanical characteristics. In this study, a 13 × 20 slot-pole matrix is developed to evaluate possible configurations. The methodology for determining the optimal combination considers winding factor (k_w), winding schematic, the number of slots per pole per phase (q_1), winding schematic symmetry number, cogging torque ripple as evaluated via least common multiple (LCM), magneto-motive force harmonics, and the stator outer diameter.

The Materials and Methods section details the formulation, calculation, and example-based explanation of each criterion, while the Implementation section presents four comprehensive tables covering winding factor, number of slots per pole per phase, symmetry number, and LCM values for the complete 13 × 20 slot-pole combination matrix [1].

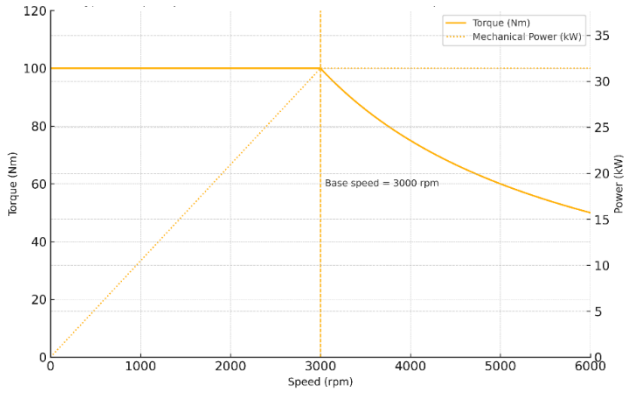


Figure 1. Typical Torque-Speed characteristic of a PMSM and BLDC motor (Constant torque- Constant power)

2. Material and Methods

In this section, the relationships aof number of slot-pole combination with key design parameters including the winding factor, winding schematic, winding schematic symmetry number, number of slots per pole per phase, cogging torque ripple as evaluated through the least common multiple (LCM) and magneto-motive force (MMF) harmonics relation are explained in sequence.

2.1. Winding Factor (k_w): is a dimensionless parameter that quantifies the effectiveness of a stator winding in converting electrical input into a sinusoidal MMF distribution within the air gap. It indicates how closely the spatial arrangement of coils approaches the ideal condition for generating the desired fundamental magnetic field while suppressing undesirable harmonic components. The value of the winding factor is always less than or equal to unity, with k_w =1 representing a theoretically perfect winding configuration for the fundamental harmonic. This parameter has a significant impact on the electromagnetic performance of the machine, as it directly affects the developed torque, induced electromotive force (EMF), and overall efficiency. A higher winding factor results in greater induced EMF and torque for a given current, whereas a lower winding factor leads to diminished performance and increased harmonic distortion. Since harmonic content contributes to additional copper and iron losses, the winding factor also influences machine heating and operational efficiency. Consequently, optimizing $k_w$is a critical objective in the design of slot–pole combinations.

The winding factor ( ) is the product of two components: the distribution factor (k_d) and the pitch factor (k_p):

$$k_w = k_d . k_p \quad (1)$$

Pitch Factor (Coil Span Factor): Accounts for the effect of short-pitch (fractional-pitch) windings:

$$k_p = cos(\frac{\epsilon}{2}) \quad (2)$$

where: ϵ = coil span shortfall in electrical degrees from full pitch.

Distribution Factor (Breadth Factor): Accounts for the distribution of coils in multiple slots per pole per phase:

$$k_d = \frac{sin(\frac{q1.\beta}{2})}{m.sin(\frac{\beta}{2})} \quad (3)$$

Where, q_1 = number of slots per pole per phase, β = slot pitch in electrical degrees
Multiplying k_d with k_p,

$$k_d = \frac{sinsin\left(\frac{q1.\beta}{2}\right)}{m.sinsin\left(\frac{\square}{2}\right)} . cos(\frac{\epsilon}{2}) \quad (4)$$

2.2. Winding Schematic: As shown in Table [1-4], the slot–pole combinations are classified into four distinct categories, each represented by a specific color. The black cells correspond to unbalanced windings, which should be avoided in all cases as they represent non-viable configurations. The green cells indicate combinations that yield a concentrated winding arrangement, also referred to in some literature as "needle winding" or "tooth winding." The red cells denote integer-slot windings, while the orange cells represent fractional-slot windings. Both red and orange categories are collectively referred to as distributed windings.

2.3. The number of slots per pole per phase (q1): Usually denoted as q_1, is a key design parameter in electrical machines that specifies how many stator slots are allocated to each pole of each phase winding. It is calculated using the expression:

$$q_1 = \frac{N_{slot}}{N_{pole} N_{phase}} \quad (5)$$

This parameter directly affects the winding layout, as it determines how coils are distributed in the slots and consequently influences the winding factor. Certain values of q_1 can reduce undesirable space harmonics, thereby lowering torque ripple and acoustic noise. When q_1 is an integer, the machine employs an integral-slot winding, which generally offers a higher winding factor but may result in increased cogging torque. In contrast, a fractional value of q_1 produces a fractional-slot winding, which can suppress specific harmonics and reduce cogging torque, although it makes the winding design more complex. In the context of slot-pole combination optimization, q_1 also has a strong influence on other critical design parameters, including the symmetry number, the least common multiple (LCM), and the winding factor.

2.4. Number of Winding Symmetry: The symmetry number in a winding schematic refers to how many times the winding pattern repeats around the stator circumference and is determined by the greatest common divisor (GCD) of the number of stator slots and poles as shown in Figure 2. A higher symmetry number simplifies manufacturing by allowing repetitive coil arrangements, ensures balanced electromagnetic performance, reduces harmonic distortion, improves fault tolerance in multiphase machines, and facilitates design optimization by making electromagnetic analysis and comparisons between slot-pole combinations more straightforward.

2.5. Cogging Torque Ripple: is an undesired torque ripple in permanent magnet machines that occurs due to the interaction between the rotor's permanent magnets and the stator's teeth. It is position-dependent and present even

when no current flows through the stator windings. Least Common Multiple (LCM) of the number of stator slots (N_{slot}) and the number of poles (N_{pole}) determines the least angular displacement over which the relative position of stator teeth and rotor poles repeats. The cogging torque period in mechanical degrees is:

$$\theta_{cog} = \frac{360\,^{\circ}}{LCM(N_{slot}, N_{pole})} \tag{6}$$

The cogging torque frequency is proportional to the LCM:

$$f_{cog} = LCM\left(N_{slot}, N_{pole}\right).f_{mech} \tag{7}$$

Where, f_{mech} is the rotor's mechanical rotational frequency.

A higher least common multiple (LCM) results in a shorter repetition period and a higher cogging torque frequency, shifting the torque ripple to a range that is less noticeable mechanically and easier to filter through the system's inertia. Conversely, a lower LCM leads to a longer repetition period and lower cogging torque frequency, which can make the torque ripple more pronounced and potentially problematic at low speeds. Therefore, when selecting a slot-pole combination, designers often aim for a high LCM to minimize low-frequency torque ripple while also optimizing other parameters such as winding factor and harmonic content.

2.6. Winding Harmonics: This section emphasizes the significance of MMF harmonics when selecting slot-pole combinations. Harmonics in electrical machines affect performance in various ways. Lower-order odd harmonics, such as the 3rd, 5th, and 7th, can distort the rotating magnetic field, resulting in torque ripple, vibrations, and additional losses. In a balanced three-phase system, triplen harmonics (multiples of the 3rd) sum in the neutral conductor and may cause overheating; However, in the MMF they cancel out, which explains their absence in the harmonic charts. Non-triple odd harmonics, such as the 5th and 7th, generate rotating fields in reverse or forward directions, reducing efficiency and introducing pulsating torque. Even harmonics, generally absent in symmetrical three-phase windings, are indicative of waveform asymmetry; when present, they can introduce a DC component in the magnetic flux, leading to increased core saturation, heating, and vibration. Figure 3. shows the winding harmonics for 18 slot- 14 pole combination.

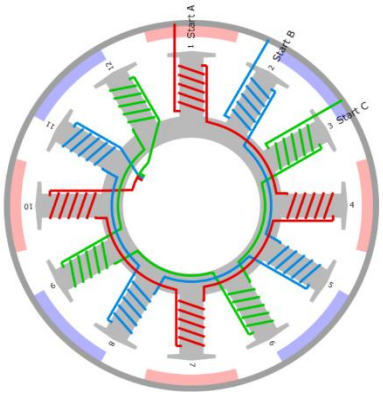


Figure 2. Winding symmetry for 12 slot-8 pole combination, the symmetry number is 4, [5]

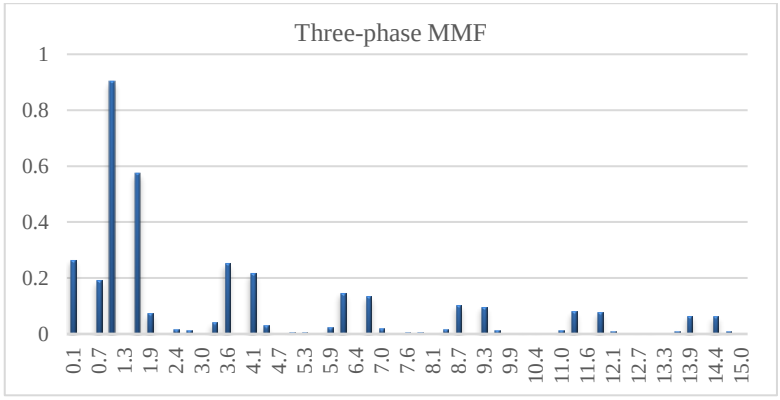


Figure 3. Three-phase winding harmonics for 18 slot – 14 poles combination [6]

3. Implementation and Calculation Results

In this section, a 13 × 20 matrix of stator-slot–rotor-pole combinations is presented. Table 1 provides the winding factor values for the matrix, while Table 2 lists the calculated values of slots per pole per phase. Table 3 reports the symmetry numbers for the stator winding configurations, and Table 4 presents the least common multiple (LCM) values used for cogging-torque assessment. Together, these tables provide the key numerical parameters required to

evaluate and compare each slot–pole combination in terms of electromagnetic performance and manufacturability [6].

Table 1. Calculated winding factor values for 13 x 20 slot-pole matrix

Pole\Slot	3	6	9	12	15	18	21	24	27	30	33	36	39
2	0.866	1	0.945	0.966	0.951	0.96	0.953	0.958	0.954	0.957	0.954	0.956	0.954
4	0.866	0.866	0.617	1	0.951	0.945	0.953	0.966	0.954	0.951	0.954	0.96	0.954
6			0.866			1			0.945			0.966	
8	0.866	0.866	0.945	0.866	0.711	0.617	0.538	1	0.941	0.951	0.954	0.945	0.954
10	0.866	0.5	0.945	0.966	0.866	0.735	0.65	0.588	0.525	1	0.946	0.942	0.954
12			0.866	Slo		0.866			0.617			1	
14	0.866	0.5	0.617	0.966	0.951	0.902	0.866	0.766	0.695	0.64	0.591	0.551	0.511
16	0.866	0.866	0.328	0.866	0.951	0.945	0.89	0.866	0.766	0.711	0.659	0.617	0.574
18									0.866				
20	0.866	0.866	0.328	0.5	0.866	0.945	0.953	0.966	0.877	0.866	0.778	0.735	0.689
22	0.866	0.5	0.617	0.259	0.711	0.902	0.953	0.958	0.915	0.874	0.866	0.786	0.74
24			0.866			0.866			0.945			0.866	
26	0.866	0.5	0.945	0.259	0.389	0.735	0.89	0.958	0.954	0.936	0.903	0.87	0.866
28	0.866	0.866	0.945	0.5	0.199	0.617	0.866	0.966	0.954	0.951	0.928	0.902	0.863
30			0.866			0.5			0.945			0.966	
32	0.866	0.866	0.617	0.866	0.199	0.328	0.65	0.866	0.915	0.951	0.954	0.945	0.917
34	0.866	0.5	0.328	0.966	0.389	0.167	0.538	0.766	0.877	0.936	0.954	0.956	0.936
36									0.866				
38	0.866	0.5	0.328	0.966	0.711	0.167	0.282	0.588	0.766	0.874	0.928	0.956	0.954
40	0.866	0.866	0.617	0.866	0.866	0.328	0.142	0.5	0.695	0.866	0.903	0.945	0.954

Table 2. Calculated number of slot per pole per phase values for 13 x 20 slot-pole matrix

Pole\Slot	3	6	9	12	15	18	21	24	27	30	33	36	39
2	0,500	1,000	1,500	2,000	2,500	3,000	3,500	4,000	4,500	5,000	5,500	6,000	6,500
4	0,250	0,500	0,750	1,000	1,250	1,500	1,750	2,000	2,250	2,500	2,750	3,000	3,250
6	0,167	0,333	0,500	0,667	0,833	1,000	1,167	1,333	1,500	1,667	1,833	2,000	2,167
8	0,125	0,250	0,375	0,500	0,625	0,750	0,875	1,000	1,125	1,250	1,375	1,500	1,625
10	0,100	0,200	0,300	0,400	0,500	0,600	0,700	0,800	0,900	1,000	1,100	1,200	1,300
12	0,083	0,167	0,250	0,333	0,417	0,500	0,583	0,667	0,750	0,833	0,917	1,000	1,083
14	0,071	0,143	0,214	0,286	0,357	0,429	0,500	0,571	0,643	0,714	0,786	0,857	0,929
16	0,063	0,125	0,188	0,250	0,313	0,375	0,438	0,500	0,563	0,625	0,688	0,750	0,813
18	0,056	0,111	0,167	0,222	0,278	0,333	0,389	0,444	0,500	0,556	0,611	0,667	0,722
20	0,050	0,100	0,150	0,200	0,250	0,300	0,350	0,400	0,450	0,500	0,550	0,600	0,650
22	0,045	0,091	0,136	0,182	0,227	0,273	0,318	0,364	0,409	0,455	0,500	0,545	0,591
24	0,042	0,083	0,125	0,167	0,208	0,250	0,292	0,333	0,375	0,417	0,458	0,500	0,542
26	0,038	0,077	0,115	0,154	0,192	0,231	0,269	0,308	0,346	0,385	0,423	0,462	0,500
28	0,036	0,071	0,107	0,143	0,179	0,214	0,250	0,286	0,321	0,357	0,393	0,429	0,464
30	0,033	0,067	0,100	0,133	0,167	0,200	0,233	0,267	0,300	0,333	0,367	0,400	0,433
32	0,031	0,063	0,094	0,125	0,156	0,188	0,219	0,250	0,281	0,313	0,344	0,375	0,406

34	0,029	0,059	0,088	0,118	0,147	0,176	0,206	0,235	0,265	0,294	0,324	0,353	0,382
36	0,028	0,056	0,083	0,111	0,139	0,167	0,194	0,222	0,250	0,278	0,306	0,333	0,361
38	0,026	0,053	0,079	0,105	0,132	0,158	0,184	0,211	0,237	0,263	0,289	0,316	0,342
40	0,025	0,050	0,075	0,100	0,125	0,150	0,175	0,200	0,225	0,250	0,275	0,300	0,325

Table 3. Calculated winding symmetry [GCD] values for 13 x 20 slot-pole matrix for cogging-torque assessment

Pole\Slot	3	6	9	12	15	18	21	24	27	30	33	36	39
2	1	1	1	1	1	1	1	1	1	1	1	1	1
4	1	2	1	2	1	2	1	2	1	2	1	2	1
6			3			3			3			3	
8	1	2	1	4	1	2	1	4	1	2	1	4	1
10	1	1	1	1	5	1	1	1	1	5	1	1	1
12			3			6			3			6	
14	1	1	1	1	1	1	7	1	1	1	1	1	1
16	1	2	1	4	1	2	1	8	1	2	1	4	1
18									9				
20	1	2	1	2	5	2	1	2	1	10	1	2	1
22	1	1	1	1	1	1	1	1	1	1	11	1	1
24			3			6			3			12	
26	1	1	1	1	1	1	1	1	1	1	1	1	13
28	1	2	1	2	1	2	7	2	1	2	1	2	1
30			3			3			3			3	
32	1	2	1	4	1	2	1	8	1	2	1	4	1
34	1	1	1	1	1	1	1	1	1	1	1	1	1
36									9				
38	1	1	1	1	1	1	1	1	1	1	1	1	1
40	1	2	1	4	5	2	1	4	1	10	1	4	1

Table 4. Calculated LCM values for 13 x 20 slot-pole matrix for cogging-torque assessment

Pole\Slot	3	6	9	12	15	18	21	24	27	30	33	36	39
2	6	6	18	12	30	18	42	24	54	30	66	36	78
4	12	12	36	12	60	36	84	24	108	60	132	36	156
6	6	6	18	12	30	18	42	24	54	30	66	36	78
8	24	24	72	24	120	72	168	24	216	120	264	72	312
10	30	30	90	60	30	90	210	120	270	30	330	180	390
12	12	12	36	12	60	36	84	24	108	60	132	36	156
14	42	42	126	84	210	126	42	168	378	210	462	252	546
16	48	48	144	48	240	144	336	48	432	240	528	144	624
18	18	18	18	36	90	18	126	72	54	90	198	36	234
20	60	60	180	60	60	180	420	120	540	60	660	180	780
22	66	66	198	132	330	198	462	264	594	330	66	396	858
24	24	24	72	24	120	72	168	24	216	120	264	72	312
26	78	78	234	156	390	234	546	312	702	390	858	468	78
28	84	84	252	84	420	252	84	168	756	420	924	252	1092
30	30	30	90	60	30	90	210	120	270	30	330	180	390

32	96	96	288	96	480	288	672	96	864	480	1056	288	1248
34	102	102	306	204	510	306	714	408	918	510	1122	612	1326
36	36	36	36	36	180	36	252	72	108	180	396	36	468
38	114	114	342	228	570	342	798	456	1026	570	1254	684	1482
40	120	120	360	120	120	360	840	120	1080	120	1320	360	1560

4. References

[1] J.R. Handershot, T.J.E Miller, Design of Brushless Permanent-Magnet Machines, Florida: Motor Design Books LLC 740 Egret Walk Ln,Venice, Florida 34292, USA, 2010.

[2] T. Miller, SPEED's Electric Machines, Simcenter SPEED, 2002-2014.

[3] J.R. HendershotJr. T.J.E. Miller, Design Studies in Electric Machines., Florida: Motor Design Books LLC 740 Egret Walk Ln,Venice, Florida 34292, USA, 2020.

[4] "hypromag.com," [Online]. Available: https://hypromag.com/rare-earth-magnet-recycling/. [Accessed 10 08 2025].

[5] "www.bavaria-direct.co.za," [Online]. Available: https://www.bavaria-direct.co.za/scheme/calculator/. [Accessed 12 08 2025].

[6] Emetor, "Emetor," Emetor AB, 10 07 2025. [Online]. Available: https://www.emetor.com/. [Accessed 110 07 2025].

International Conference on Advances and Innovations in Engineering

3.ICAIE

KS-Net: Multi-layer network model for determining the rotor type from motor parameters in interior PMSMs

Kıvanç DOĞAN[1] , Ahmet ORHAN [2]

[1 a)] Graduate School of Natural and Applied Sciences, Electrical and Electronic Engineering, Firat University, Elazığ, Türkiye
[1 b)] Department of Electrical and Electronic Engineering, Faculty of Engineering, Firat University, Elazığ, Türkiye.
[2]Department of Electrical and Electronic Engineering, Faculty of Engineering, Firat University, Elazığ, Türkiye.
[1]kivanc.dogan@firat.edu.tr, [2]aorhan@firat.edu.tr
[1](ORCID: 0000-0003-4832-1412), [2](ORCID: 0000-0003-1994-4661)

Abstract
The demand for high efficiency and precise control in electric drive systems has led to the widespread adoption of Interior Permanent Magnet Synchronous Motors (IPMSMs). The performance of these motors is significantly influenced by rotor geometry. Traditionally, rotor shape analysis has been conducted using the finite element method (FEM), which involves high computational costs. This study aims to classify the rotor shape (2D type, V type, Nabla type) of IPMSMs using electromagnetic parameters (L_d, L_q, φ_a) through machine learning-based methods and to demonstrate the applicability of this approach as an alternative to classical methods. In this context, a custom deep learning model, KS-Net, developed by the user, was comparatively evaluated against Cubic SVM, Quadratic SVM, Fine KNN, Cosine KNN, and Fine Tree algorithms. The balanced dataset, consisting of 9,000 samples, was tested using 10-fold cross-validation, and performance metrics such as accuracy, precision, recall, and F1-score were employed. The results indicate that the Cubic SVM and Quadratic SVM algorithms classified all samples flawlessly, achieving 100% accuracy, while the KS-Net model achieved 99.98% accuracy with only two misclassifications, demonstrating competitiveness with classical methods. This study shows that the rotor shape of IPMSMs can be predicted with high accuracy using data-driven approaches, offering a fast and cost-effective alternative to FEM-based analyses. The findings provide a solid foundation for accelerating motor design processes, developing automated rotor identification systems, and enabling data-driven fault diagnosis in engineering applications.

Keywords: Machine Learning, IPMSM, Rotor Shape Estimation, Parameter Identification, Motor Design Optimization, Rotor Geometry.

1. Introduction

In modern electric drive systems, the demand for high efficiency, compact structure, and precise control has led to the widespread adoption of advanced motor types, particularly Interior Permanent Magnet Synchronous Motors (Interior PMSMs). One of the most critical factors determining the performance of Interior PMSMs is the geometric structure of the rotor. Rotor shape directly affects key parameters such as torque production, efficiency, magnetic saturation, thermal behavior, and fault tolerance. Therefore, predicting the rotor shape based on motor parameters is of great importance for both design optimization and the monitoring and fault diagnosis of existing systems. In recent years, machine learning algorithms have been effectively employed in various applications in the field of electric machines, including parameter estimation, fault diagnosis, and performance analysis. The preference for data-driven approaches over complex physical models shortens the modeling process and reduces the need for expert knowledge. In this context, developing machine learning-based methods to predict rotor shape from Interior PMSM parameters has emerged as an innovative research area in engineering. Rotor shape encompasses numerous physical characteristics, such as magnet placement, the position of magnetic barriers, and the structure of bridges and ribs. These geometric features directly influence many performance parameters, including torque production, efficiency, magnetic saturation behavior, temperature distribution, and fault tolerance. [1]. The successful outcomes achieved in applications such as parameter estimation, fault diagnosis, and torque and position prediction in PMSMs indicate that a similar methodology can be applied to rotor shape prediction. In particular, the flexibility and speed offered by machine learning algorithms in modeling multidimensional and complex relationships provide significant added value to engineering design processes. Furthermore, the decisive role of rotor topology in motor performance encourages the widespread adoption of such prediction and optimization studies in industrial applications [2], [3], [4], [5], [6]. In traditional methods, analyses related to rotor shape are generally performed using the Finite Element Method (FEM), through which various design scenarios are examined. However, FEM analyses are computationally expensive, and since they must be rerun for each new design, they slow down engineering processes and increase costs [7]. At this point, machine learning (ML)-based methods emerge as an innovative paradigm in motor

[1]Corresponding author

design. ML algorithms have the potential to achieve highly accurate predictions by learning complex, multivariate relationships from data that are difficult to model physically. Indeed, the literature includes numerous studies demonstrating that various PMSM parameters—such as torque, temperature, and rotor position—can be successfully predicted using algorithms like artificial neural networks and support vector machines (SVMs) [8], [9]. Traditionally, the methods used to determine rotor shape and parameters have largely relied on engineering expertise, experience, and extensive simulations. However, these processes are constrained by high computational costs and long development times. In recent years, the application of machine learning algorithms in engineering disciplines has introduced a new paradigm for solving such complex problems. In particular, regression and classification algorithms can learn from large datasets to model and predict the complex relationships between motor parameters and rotor shape [10]. In one study, three different machine learning algorithms—Support Vector Regression (SVR), Random Forest Regressor, and Polynomial Regression—were compared for the prediction of Interior PMSM parameters. Random Forest Regressor demonstrated the highest R^2 values, indicating that it best explained the variation of the dependent variables in the model [10]. This finding demonstrates that machine learning algorithms offer higher accuracy and generalization capability compared to traditional methods. In the specific context of rotor shape prediction, however, studies directly based on machine learning classification remain limited. Nevertheless, it is well established that a strong causal relationship exists between rotor shape and electromagnetic parameters (L_d, L_q, φ_a) [11]. Therefore, accurately predicting and analyzing rotor shape is of great importance in motor design and manufacturing. The success of machine learning algorithms used for rotor shape prediction also depends on the quality and diversity of the dataset on which the model is trained. Simulation data obtained through the finite element method, experimental measurements, and parametric design studies provide a rich source of data for training machine learning models [1], [7]. These data enable the model to learn different rotor topologies and parameter combinations. In this study, the prediction of rotor shape (2D type, V type, Nabla type) was targeted using the electromagnetic parameters (L_d, L_q, φ_a) of IPMSM motors. For this purpose, a custom artificial neural network model, KS-Net, developed by the user, was comparatively evaluated alongside Cubic SVM, Quadratic SVM, Fine Tree, Fine KNN, and Cosine KNN algorithms. The performance of the models was tested using 10-fold cross-validation, and their effectiveness was compared through accuracy rates and confusion matrix. The originality of this study lies in introducing KS-Net as a data-driven approach for rotor shape classification, providing a fast and cost-effective alternative to traditional FEM-based analyses. The test results revealed that the proposed KS-Net algorithm exhibits strong generalization capability and demonstrates competitive, and in some cases superior, performance compared to traditional algorithms. This study demonstrates the applicability of a data-driven approach in IPMSM motor design and provides a foundation for future applications in parametric optimization, monitoring, and fault diagnosis.

2. Materials and Methods

The primary objective of this study is to evaluate the performance of machine learning algorithms and a custom-designed artificial neural network model (KS-Net) in rotor shape classification based on the electromagnetic parameters (L_d, L_q, φ_a) of IPMSM motors. A high-quality and balanced dataset was generated through systematic simulations based on parametric modeling. The dataset comprises three different rotor topologies, each consisting of 3,000 samples:

- 2D type rotor
- V type rotor
- Nabla type rotor

Since all classes contain an equal number of samples, a balanced learning environment was ensured for the classification algorithms. Each sample is characterized by three fundamental parameters (L_d, L_q, φ_a) representing the electromagnetic behavior of the rotor. These parameters were derived based on the motor's operating conditions and rotor geometry, processed and analyzed in MATLAB. They were used directly as model inputs without requiring an additional feature engineering step.

Within the classification process, the custom-designed wide neural network, KS-Net, was evaluated alongside classical algorithms such as Cubic SVM, Quadratic SVM, Fine KNN, Cosine KNN, and Fine Tree. All algorithms were trained in the MATLAB Classifier Learner interface and tested using the 10-fold cross-validation method.

The model performances were compared using accuracy rate and confusion matrix analyses. Through this comprehensive evaluation approach, the ability of the developed models to correctly distinguish between different rotor topologies was analyzed in detail. The obtained results demonstrate that electromagnetic parameters are highly determinant of IPMSM rotor shape and that high-accuracy classification is achievable through data-driven methods.

2.1. Dataset

In this study, a dataset consisting of the electromagnetic parameters of IPMSM motors—namely, the d–q axis inductances $L_d, L_q,$ and the static magnetic flux φ_a —was utilized. The dataset comprises a total of 9,000 samples corresponding to three different rotor topologies. Each rotor shape contains 3,000 samples, resulting in balanced classes. Table 1 presents the distribution of the dataset.

Table 1. Dataset

No	Class	Number of Data
1	2D	3000
2	V	3000
3	Nabla	3000

The data were obtained from a simulation framework, and the parameters L_d, L_q, φ_a were normalized for each sample. The fact that these parameters represent the electromagnetic structure of the motor and are directly dependent on rotor topology has enabled their use in training the classification algorithms in this study.

2.2. Feature Selection and Preprocessing

The parameters L_d, L_q, φ_a in the dataset were used directly for classification. No dimensionality reduction or filtering was applied, only standard scaling was performed on the input parameters. All algorithms were trained and tested using the same input vectors. This approach was specifically adopted to ensure the comparability of the algorithms.

2.3. Classification Algorithms Used

For rotor shape classification, both traditional and custom-developed machine learning algorithms were employed. KS-Net is a custom artificial neural network architecture developed by the user and designed for classification tasks in the MATLAB environment. The model features a deep feedforward neural network structure consisting of three hidden layers.

- The first hidden layer consists of 128 neurons,
- The second layer contains 64 neurons,
- The third layer consists of 32 neurons.

n each hidden layer, the ReLU (Rectified Linear Unit) activation function was used to enhance the nonlinear learning capability. The model was trained for 1000 iterations to ensure stable parameter optimization and convergence of the learning process. With this structure, KS-Net achieved high accuracy in rotor shape prediction based on electromagnetic parameters and successfully distinguished between the classes. The multi-layer perceptron (MLP) architecture of the model is illustrated in Figure 1. [12].

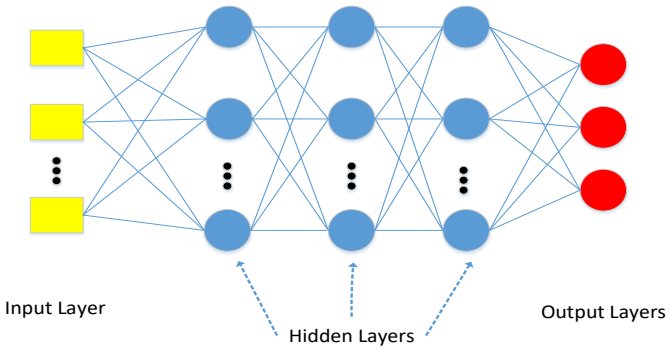


Figure 2. 3-layer MLP architecture

The other ML algorithms used for comparison are presented in Table 2.

Table 2. Configuration parameters of the proposed classification algoritms

Classifier	Configuration Parameters	
Cubic SVM	Kernel function: Kernel scale: Box constraint level:	Cubic Automatic 1
Quadratic SVM	Kernel function: Kernel scale: Box constraint level:	Quadratic Automatic 1
Fine Tree	Maximum number of splits: Split criterion:	100 Gini's diversity index
Fine KNN	Number of neighbors: Distance metric: Distance weight:	100 5 Manhattan (Cityblock) Equal
Cosine KNN	Number of neighbors: Distance metric: Distance weight:	10 Cosine Equal

These models were trained and tested using the MATLAB Classifier Learner Toolbox. A 10-fold cross-validation method was employed for each model, allowing a more reliable assessment of their generalization performance.

3. Results and Discussion

In this study, various machine learning algorithms were employed to classify the rotor types of Interior Permanent Magnet Synchronous Motors (IPMSMs). The dataset consists of samples from three different rotor geometries—2D type, V-type, and Nabla-type—where each sample includes the parameters (L_d, L_q, φ_a) representing the electromagnetic characteristics of the motor. Six different classification algorithms were evaluated: Cubic SVM, Quadratic SVM, Fine KNN, Cosine KNN, Fine Tree, and the custom deep learning model KS-Net developed by the user. The performance of the models was compared based on accuracy rates and confusion matrices analyses.

The results indicate that the Cubic SVM and Quadratic SVM algorithms classified all test samples without error, achieving 100% accuracy. The proposed KS-Net model achieved 99.98% accuracy with only two misclassifications, reaching a performance level comparable to SVM-based methods. The Fine KNN model misclassified only one Nabla sample and achieved 99.99% accuracy. Cosine KNN produced four misclassifications, resulting in 99.93%

accuracy. The Fine Tree model yielded the lowest performance, with a total of 85 misclassifications and 99.06% accuracy, with most errors occurring between the 2D and V classes.

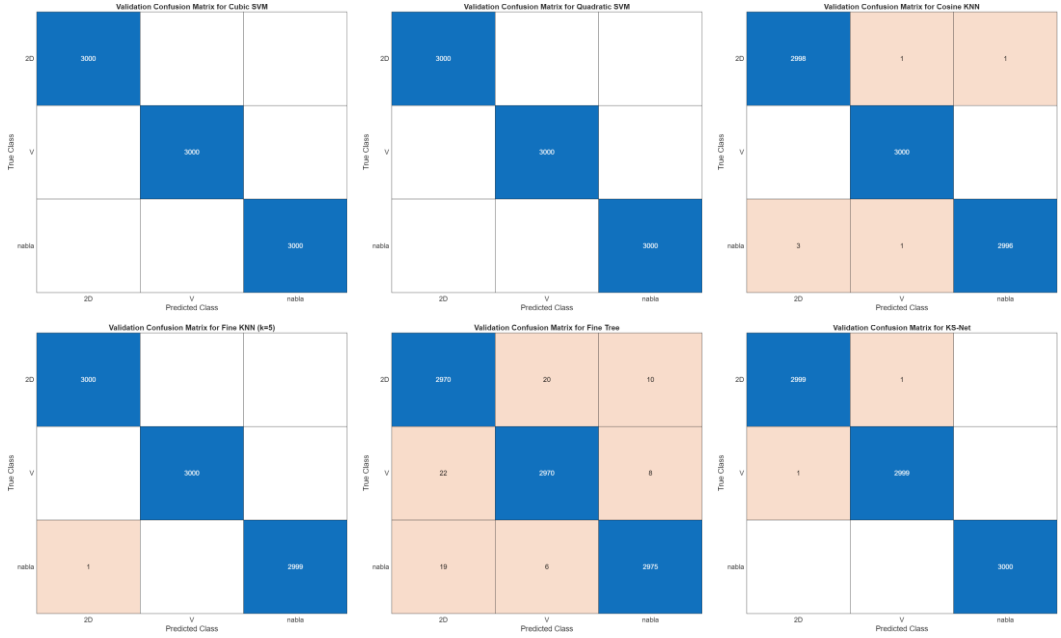


Figure 2. The confusion matrices for rotor type prediction correspond, respectively, to Cubic SVM, Quadratic SVM, Cosine KNN, Fine KNN, Fine Tree, and KS-Net.

As shown in Fig. 1, the Cubic SVM and Quadratic SVM algorithms distinguished all classes without error, while KS-Net produced only two misclassifications. The Fine KNN model made a single error in the Nabla class, whereas the Cosine KNN model exhibited four errors, most of which occurred in the Nabla class. In the Fine Tree model, the errors were distributed across classes, with notable confusion particularly between the 2D and V classes. These results demonstrate that the proposed KS-Net model achieves a level of accuracy comparable to classical machine learning algorithms and effectively distinguishes between classes.

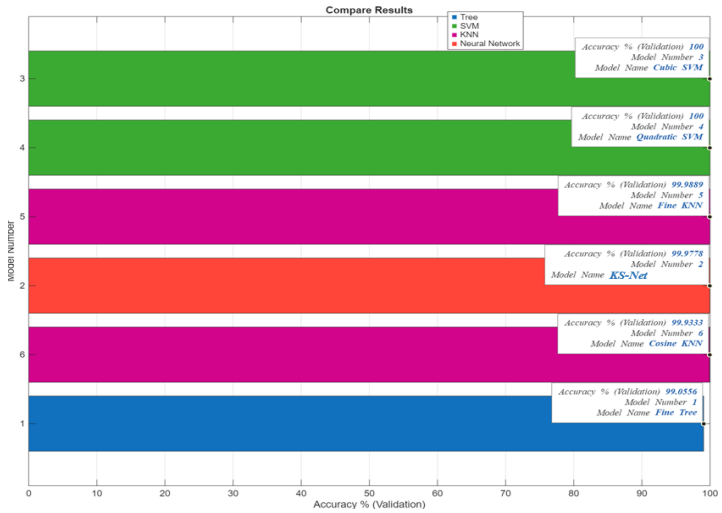


Figure 3. Comparison results of classical machine learning algorithms and the KS-Net algorithm

The validation accuracies of all classifiers used in the study are shown in Fig. 2. In the bar chart, the colors represent the algorithm families: blue corresponds to the tree-based model (Fine Tree), green to the SVM-based models (Cubic

SVM and Quadratic SVM), pink to the KNN-based models (Fine KNN and Cosine KNN), and red to the proposed neural network-based model (KS-Net). According to the results, the Cubic SVM and Quadratic SVM algorithms classified all test samples without error, achieving 100% accuracy. This indicates that the dataset can be completely separated using nonlinear SVM kernels and that SVM-based methods are highly effective for this problem. The Fine KNN model achieved 99.99% accuracy with only a single misclassification, exhibiting a performance very close to the SVM-based models. The proposed KS-Net model achieved a similar level of performance with 99.98% accuracy, demonstrating competitiveness with classical methods. Cosine KNN achieved a high accuracy of 99.93%, with a limited number of misclassifications observed. The Fine Tree algorithm yielded the lowest performance with 99.06% accuracy, which can be explained by the inability of decision trees to separate inter-class boundaries as precisely as other algorithms. Overall, the accuracy rates shown in the figure indicate that all models exhibit high success in rotor type classification. However, the SVM-based models demonstrate near-perfect classification capability, while the proposed KS-Net achieves an accuracy comparable to classical methods. These results demonstrate that rotor type classification based on motor parameters can be performed reliably and that the method can be effectively utilized in motor design and automatic rotor type identification. Following this general assessment, a comparative performance table containing the accuracy, precision, recall, F1 score and specificity values of each model is presented in Table 3 to quantitatively illustrate the classification performance of the models.

Table 3. Comparative performance metrics of rotor type classification models

Metrics	Cubic SVM	Quadratic SVM	Fine KNN	Cosine KNN	Fine Tree	KS-Net
Accuracy	1.00	1.00	0.9999	0.9993	0.9906	0.9998
Precision	1.00	1.00	0.9999	0.9993	0.9905	0.9998
Recall	1.00	1.00	0.9999	0.9993	0.9905	0.9998
F1 Score	1.00	1.00	0.9998	0.9993	0.9905	0.9998
Specificity	1.00	1.00	0.9999	0.9996	0.9953	0.9998

The obtained results clearly demonstrate that all algorithms used in this study achieved very high performance in rotor type classification. The Cubic SVM and Quadratic SVM models reached a perfect score of 1.00 (100%) across all metrics, classifying every sample in the dataset without error. This indicates that the dataset is structurally fully separable by SVM-based models. The Fine KNN model achieved performance very close to these two models, yielding values around 0.9999 for all metrics and making only a single misclassification. The Cosine KNN model achieved metric values in the range of 0.9993 to 0.9996, with a total of four misclassifications, demonstrating very high accuracy. The KS-Net model, as the proposed custom deep learning approach, achieved 0.9998 across all metrics, exhibiting performance competitive with classical machine learning algorithms. In contrast, the Fine Tree algorithm showed lower performance compared to the other models, with metrics ranging between 0.9905 and 0.9953. This result can be attributed to decision trees being less precise in defining inter-class boundaries compared to SVM- and KNN-based methods.

In general, the table indicates that the proposed KS-Net model achieved nearly equivalent performance to SVM-based methods and proved to be a strong alternative to classical approaches. The fact that all models achieved accuracy above 99% demonstrates that rotor type classification based on motor parameters can be performed with a high degree of reliability.

4. Conclusion

In this study, the rotor shape (2D type, V-type, Nabla-type) of Interior PMSM motors was classified based on their electromagnetic parameters (L_d, L_q, φ_a) using machine learning algorithms. In this context, the custom-designed neural network model KS-Net was compared with Cubic SVM, Quadratic SVM, Fine KNN, Cosine KNN, and Fine Tree algorithms, and a comprehensive performance analysis was conducted. According to the results, Cubic SVM and Quadratic SVM classified the dataset with 100% accuracy, predicting all samples without error. This indicates that the classes in the dataset are distinctly separable. KS-Net achieved 99.98% accuracy with only two misclassifications, demonstrating performance competitive with the SVM models in terms of overall success. Fine KNN and Cosine KNN also achieved very high accuracies, though some inter-class confusions occurred in certain

samples. Fine Tree was the weakest model in terms of inter-class separation, producing more than 60 misclassifications in total. This study presents an effective data-driven approach for rotor shape classification as an alternative to FEM-based analyses. The proposed method is particularly advantageous for accelerating the design process, developing diagnostic systems when rotor geometry cannot be directly measured, and providing a foundation for data-driven fault diagnosis approaches.

In essence, the originality of this study lies in introducing KS-Net as a data-driven method for high-accuracy rotor shape classification, providing a fast and cost-effective alternative to FEM-based analyses and establishing a solid foundation for automated rotor identification and data-driven fault diagnosis in future engineering applications.
In future work, the KS-Net model will be trained on larger and more diverse datasets to enhance its generalization capability, along with hyperparameter optimization and the inclusion of new parametric inputs related to rotor geometry (e.g., rotor volume, air gap). Furthermore, it is planned to move beyond classification toward optimizing rotor design using regression models and developing automated design systems.

Data available: https://www.kaggle.com/datasets/uuuuuuuuu/dataset-iron-losses-ipmsms

5. References

[1] D. K. Athanasopoulos, V. I. Kastros, and J. C. Kappatou, "Electromagnetic analysis of a PMSM with different rotor topologies," in *2016 XXII International Conference on Electrical Machines (ICEM)*, IEEE, Sep. 2016, pp. 306–312. doi: 10.1109/ICELMACH.2016.7732543.

[2] I. Vascan, F. P. Piglesan, C. Martis, and I. Barsan, "PMSM Rotor Topologies for Automotive HVAC System," in *2019 Electric Vehicles International Conference (EV)*, IEEE, Oct. 2019, pp. 1–4. doi: 10.1109/EV.2019.8892852.

[3] T. Cherriere, S. Hlioui, L. Laurent, F. Louf, H. Ben Ahmed, and M. Gabsi, "Topology Optimization of Asymmetric PMSM Rotor," in *2022 International Conference on Electrical Machines (ICEM)*, IEEE, Sep. 2022, pp. 469–475. doi: 10.1109/ICEM51905.2022.9910650.

[4] L. Papini and C. Gerada, "Sensitivity analysis of rotor parameters in solid rotor induction machine," in *2017 IEEE International Electric Machines and Drives Conference (IEMDC)*, IEEE, May 2017, pp. 1–7. doi: 10.1109/IEMDC.2017.8002345.

[5] L. Yang and A. Shami, "On hyperparameter optimization of machine learning algorithms: Theory and practice," *Neurocomputing*, vol. 415, pp. 295–316, Nov. 2020, doi: 10.1016/j.neucom.2020.07.061.

[6] M. S. Kaleli, A. R. Kaleli, and C. Yalçıner, "Torque Prediction Based Performance Analysis of Small-Scale Wind Turbines Using Data Driven Modelling Methods," *Erzincan Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, vol. 14, no. 1, pp. 260–269, Mar. 2021, doi: 10.18185/erzifbed.758924.

[7] J.-B. Park, M. Johnson, and H. A. Toliyat, "Interior PMSM double layer rotor core parametric design study," in *IECON 2014 - 40th Annual Conference of the IEEE Industrial Electronics Society*, IEEE, Oct. 2014, pp. 695–701. doi: 10.1109/IECON.2014.7048576.

[8] W. Nawae and K. Thongpull, "PMSM Torque Estimation Based on Machine Learning Techniques," in *2020 International Conference on Power, Energy and Innovations (ICPEI)*, IEEE, Oct. 2020, pp. 137–140. doi: 10.1109/ICPEI49860.2020.9431433.

[9] Y. Wang, Y. Hoole, and K. Haran, "Position Estimation of Outer Rotor PMSM Using Linear Hall Effect Sensors and Neural Networks," in *2019 IEEE International Electric Machines & Drives Conference (IEMDC)*, IEEE, May 2019, pp. 895–900. doi: 10.1109/IEMDC.2019.8785178.

[10] R. Savant, A. A. Kumar, and A. Ghatak, "Prediction and Analysis of Permanent Magnet Synchronous Motor parameters using Machine Learning Algorithms," in *2020 Third International Conference on Advances in Electronics, Computers and Communications (ICAECC)*, IEEE, Dec. 2020, pp. 1–5. doi: 10.1109/ICAECC50550.2020.9339479.

[11] X. Liang, F. Liu, W. Li, M. Wang, P. Zheng, and Z. Gu, "A Novel Rotor Re-Construction Method for Improving the Electromagnetic Performance of the Interior PMSM," in *2022 25th International Conference on Electrical Machines and Systems (ICEMS)*, IEEE, Nov. 2022, pp. 1–4. doi: 10.1109/ICEMS56177.2022.9983390.

[12] B. Çelik and A. Akbal, "NeuroParkNet: A New Neural Network Model for Classification of Parkinson's Disease," *GAZI UNIVERSITY JOURNAL OF SCIENCE*, pp. 1–1, Apr. 2025, doi: 10.35378/gujs.1602747.

Quantum-Enhanced Algorithmic Frameworks for Autonomous Vehicles: A Systematic Review and Analysis Emerging Approaches

Seda Nur GÜNGÖR[1] , Mehmet KARAKÖSE[2]

[1,2]Department of Computer Engineering, Faculty of Engineering, Fırat University, Elazığ, Turkey.
[1]sngungor@firat.edu.tr, [2]mkarakose@firat.edu.tr
[1](ORCID: 0000-0003-1849-3555), [2](ORCID: 0000-0002-3276-3788)

Abstract
In recent years, quantum computing approaches have shown great potential in solving high-dimensional and computationally complex problems that are difficult for classical methods to address. Autonomous vehicle systems require advanced capabilities in real-time decision-making, object detection, route planning, multi-sensor data processing, and secure communication. In this context, integrating quantum algorithms into autonomous vehicles has become a significant research focus. This study systematically examines the applications of quantum-assisted algorithms in autonomous vehicle systems. Three key application areas have been identified: route planning and traffic optimization, object detection and image processing, and intelligent subsystems for autonomous driving. In the domain of object detection, Quantum Convolutional Neural Networks (QCNNs) have achieved comparable accuracy rates with fewer parameters and have demonstrated faster inference capabilities in specific sub-tasks. In Quantum Reinforcement Learning (QRL), Hierarchical QRL has improved safety scores by 15.8% with 1/36 fewer parameters. Despite these promising results, quantum computing still faces challenges such as limited access to hardware, lack of real-device testing, NISQ-era constraints, barren plateau issues, and scalability limitations. In this systematic review, the transformative potential of quantum computing in autonomous vehicle technologies has been demonstrated, taking into account existing limitations, and a roadmap has been proposed for future interdisciplinary research collaborations in this emerging field.

Keywords: Quantum Computing, Autonomous Vehicles, Quantum Machine Learning, Route Planning, Object Detection

1. Introduction

In autonomous vehicle technologies, the increasing need for high-accuracy, real-time execution of fundamental functions such as decision-making, route planning, system optimization, and environmental perception has highlighted the limitations of classical computing systems due to their insufficient computational power and discontinuities. In this context, quantum computing has emerged as a powerful alternative. With its foundations in quantum mechanics—such as superposition, entanglement, and parallel processing capabilities—quantum computing offers potential advantages in solving complex and computation-intensive problems. Over the past five years, research efforts have increasingly integrated quantum computing into the decision-support mechanisms of autonomous systems in various domains, including route and trajectory planning, environmental and object detection, secure inter-vehicle communication, and fleet optimization. These studies aim to leverage the high computational power and potential benefits offered by quantum technologies to enhance the performance of autonomous vehicles and equip them with advanced capabilities.

Quantum algorithms have the potential to offer faster and more efficient solutions than classical algorithms in addressing complex problems such as route planning and traffic optimization. D-Wave, the company that developed the world's first commercial quantum computer, proposes the use of quantum annealing technology for logistics and route planning in road traffic, particularly for solving NP-hard problems related to vehicle route selection [1].

Quantum-assisted sensing and image processing hold significant potential for enhancing the speed and accuracy of sensor data processing through quantum machine learning algorithms, as well as for improving object recognition and environmental perception tasks via quantum-enhanced image processing techniques. For instance, a Quantum Deep Convolutional Neural Network (QDCNN) has been proposed to increase safety and reliability in autonomous transportation systems The proposed method demonstrates superior computational efficiency by achieving faster detection times compared to classical algorithms. Another example of studies in this domain involves the development of hybrid classical–quantum deep learning models for traffic sign classification [2]. Compared to classical deep learning models, improved accuracy has been achieved for traffic sign classification. By leveraging the fundamental principles of quantum

[1]Corresponding author

mechanics, quantum sensing techniques employing quantum sensors and quantum light sources—capable of performing measurements with higher precision than classical sensors—have enabled the development of quantum LIDAR systems that provide higher-resolution distance measurements [3].

Localization, a critical component of mobile robotics, is essential for safe navigation and obstacle avoidance. However, existing probabilistic localization methods are computationally intensive, especially when dealing with large maps or high-resolution sensor data. In this context, the application of Grover's algorithm to the localization problem in mobile robots has demonstrated notable speed-ups compared to classical methods, despite hardware resource limitations [4]. Zhuang et al. [5], investigated the potential impacts of quantum computing technologies in Intelligent Transportation Systems (ITS). Hybrid approaches such as Quantum Genetic Algorithm and Learning Vector Quantization (QGA-LVQ) have been proposed for quantum traffic management problems, addressing the issue of local minima often encountered by classical neural networks during traffic prediction tasks [6]. Quantum Particle Swarm Optimization (QPSO) has been employed to more accurately model phase transitions and breakpoints in traffic signal timing [7]. Hybrid approaches based on Quadratic Unconstrained Binary Optimization (QUBO), implemented on D-Wave devices, offer real-time optimization potential for large-scale urban planning and optimization problems [8]. In autonomous driving, quantum approaches such as Quantum Convolutional Neural Networks (QCNN) [9], [10] have started to be used in object detection and scene understanding tasks, achieving comparable accuracy with fewer resources. Advanced Driver Assistance Systems (ADAS) aim to enhance driving safety by analyzing data from multiple sensors such as radar, LIDAR, cameras, and ultrasonic sensors in real time. However, current ADAS systems face limitations including high latency, limited computational capacity, insufficient accuracy, slow processing speeds, and inadequacies in security and privacy. To leverage the advantages of quantum computing in this context, Rattan et al. [11], demonstrated the superiority of models such as Quantum Support Vector Machine (QSVM), Quantum Neural Network (QNN), and Quantum-Like Bayesian (QLB), when integrated into ADAS and autonomous vehicle (AV) systems, compared to classical approaches. Although Quantum Reinforcement Learning (QRL) has shown great potential for solving complex decision-making problems that are intractable or difficult for classical methods, practical applications still face several challenges. Limitations such as difficulties in encoding classical data into quantum circuits and the dependence on quantum hardware during inference restrict the applicability of QRL in real-world problems. To overcome these limitations, Liu et al. [12] proposed the Quantum Trainable Reinforcement Learning (QTRL) algorithm. In this architecture, data is not encoded into quantum circuits; instead, the QNN learns only the parameters of a classical policy model. As a result, the depth and width of the quantum circuit become independent of the input size. Moreover, the final output obtained from QTRL is entirely classical data, allowing inference to be performed using classical hardware. For a comprehensive overview, Table 1 summarizes several representative studies employing quantum computing in autonomous vehicle systems.

Table 5. Literature review

	Problem	Approach	Advantage
[13]	Real-time road and object detection yielding low accuracy due to shadow artifacts in image analysis	Quantum Deep Convolutional Neural Network (QDCNN)	Shadow detection was achieved in 0.0049 seconds. Demonstrated faster performance and greater robustness under noise compared to classical systems.
[2]	Traffic sign classification	Hybrid Classical-Quantum Deep Learning (HCQ-DL)	Achieved 85% accuracy with significantly higher robustness than classical models. Over 91% accuracy was obtained under GA and FGSM attacks. Quantum advantage was achieved without retraining or classical data augmentation
[14]	Target detection under high environmental noise where classical LIDAR systems fail	Dispersed non-classical target detection (DNCTD)	Achieved superior performance with up to 43.1 dB SNR improvement, 3D target detection, and high noise tolerance.
[11]	Driving style classification using IMU sensor data (aggressive/normal/slow)	Quantum Support Vector Machine (QSVM)	Achieved 53% accuracy with shorter processing time (25 ms vs. 173 ms) compared to classical SVM.
[11]	Traffic sign recognition from RGB images (e.g., stop, speed limit)	QSVM + VGG-16 + PCA	Achieved 62% accuracy with 24.5% better performance than classical SVM at lower processing time.

Table 6. (Continue) Literature review

[11]	Driver state monitoring from grayscale images (e.g., eyes open/closed, yawning)	QSVM + VGG-16 + PCA	Achieved similar accuracy with 24% faster inference compared to classical models.
[12]	Overhead of quantum data encoding and dependency on quantum hardware during inference in quantum reinforcement learning	Quantum-Train Reinforcement Learning (QTRL)	Eliminates quantum dependency at inference; achieves comparable or superior RL performance with significantly fewer parameters
[15]	Cooperative route planning for real-time multi-vehicle coordination	Optimized QCNN	Outperformed both classical CNN and QCNN with higher accuracy, lower loss, and faster convergence.
[16]	High computational cost of multi-channel data processing for object detection in autonomous driving using classical CNNs	Fast Quantum Convolution (FQC) + QCNN	Achieved 13% increase in accuracy and 50% reduction in activation time.
[17]	Safe obstacle-avoidance motion planning under limited computational resources in autonomous vehicles	Hierarchical Reinforcement Learning (HQRL)	Achieved higher average safety score than HRL with 1/36 fewer parameters. Improvements observed in pedestrian distance and lane-center alignment scores.

This study provides a comprehensive analysis of recent research on the application of quantum computing in autonomous systems. The primary objective is to evaluate the advantages of quantum computing and the outcomes obtained in enabling enhanced capabilities in autonomous systems. Widely adopted algorithms are examined in detail, with a particular focus on their integration into autonomous vehicles and the architectural role of quantum circuits within these systems. Architectural diagrams are analyzed to establish an efficient infrastructure for understanding how quantum circuits are embedded and utilized throughout the system architecture.

2. Quantum-based Autonomous Systems

In recent years, autonomous vehicle technologies have gained significant momentum, driven by the growing demand for intelligent systems capable of environmental perception, decision-making, and motion planning. Although classical artificial intelligence and deep reinforcement learning algorithms employed in these systems have achieved successful results in large state and action spaces, they still face several challenges during the optimization process, such as learning instability, low computational efficiency, and convergence to local minima. In this context, quantum computing offers substantial potential. Advances in quantum computing—particularly quantum circuit-based learning algorithms—leverage quantum principles such as superposition and entanglement, and possess the capability to outperform classical methods in terms of speed, efficiency, and accuracy.

2.1 Quantum Computing

Quantum computing, grounded in the principles of quantum mechanics, offers promising solutions for problems in which classical methods struggle in terms of resource requirements and computation time [18]. Quantum circuits are constructed using quantum gates. Unlike traditional computing, which relies on bits, quantum computing utilizes qubits. In contrast to classical bits, qubits can represent multiple states simultaneously due to quantum properties such as superposition and entanglement [19]. Quantum circuits are read from left to right, and each line in the circuit diagram corresponds to qubit Quantum gates correspond mathematically to specific matrices. The output of the circuit is obtained by multiplying the state vectors of the qubits at each stage with the matrices representing the quantum gates [20].

2.2 Common Quantum Algorithms Applied in AV Research

Variational Quantum Circuits (VQC), also referred to as parameterized quantum circuits or quantum neural networks, can be defined as the quantum machine learning equivalent of classical neural networks, in which learnable parameters (weights) are replaced with tunable parameters of quantum gates embedded in quantum circuits [21]. VQC-based models operate through hybrid architectures where classical and quantum components are jointly optimized. These models are particularly favored in studies due to their applicability on Near-Term Intermediate-Scale Quantum (NISQ) devices [22]. Quantum Machine Learning (QML) encompasses a broad class of machine learning approaches enhanced by the integration of quantum computing. The models developed within the QML framework gain learning capability through the parameterization of quantum circuits. Quantum Neural Network (QNN), implement quantum gates in place of classical activation functions and weights, and rely on measurement-based inference [23]. Designed based on VQC principles, QNNs are trained using hybrid optimization algorithms instead of traditional backpropagation. Similarly, Quantum Support Vector Machine (QSVM) is an adaptation of the classical SVM algorithm to quantum computing, offering the potential to efficiently compute similarity measures—especially inner products—that are otherwise computationally demanding using classical methods [24]. Quantum Convolutional Neural Network (QCNN) is a quantum machine learning model composed of multi-layer quantum circuits that acquire learning capacity via parameterized gates and perform operations such as convolution, pooling, and nonlinear transformations based on quantum measurement [25]. Parameterized Quantum Circuit (PQC) are circuits that learn over the angular parameters of quantum gates using classical optimization algorithms. They can be applied in both supervised and unsupervised learning models [26]. Quantum Reinforcement Learning (QRL) is a model that adapts the core components of classical reinforcement learning (agent, environment, action, reward) into the quantum computing paradigm [27]. It aims to enhance the agent's learning process by incorporating fundamental quantum properties such as superposition, entanglement, and measurement. In this direction, quantum circuit-based variants of classical Q-Learning or Deep Q-Network (DQN) algorithms have been developed, which perform learning over parameterized policies or value functions [28]. All these methods aim to integrate the capabilities of quantum computing into machine learning workflows in order to overcome challenges associated with computational complexity, inefficiency in learning, and resource limitations inherent in classical approaches. Given the presence of real-time, high-dimensional, and decision-intensive problems in autonomous vehicle systems, the applicability of quantum-based approaches in this domain is becoming increasingly important.

3. Quantum-Enhanced Methodologies in Autonomous Vehicle Subsystems

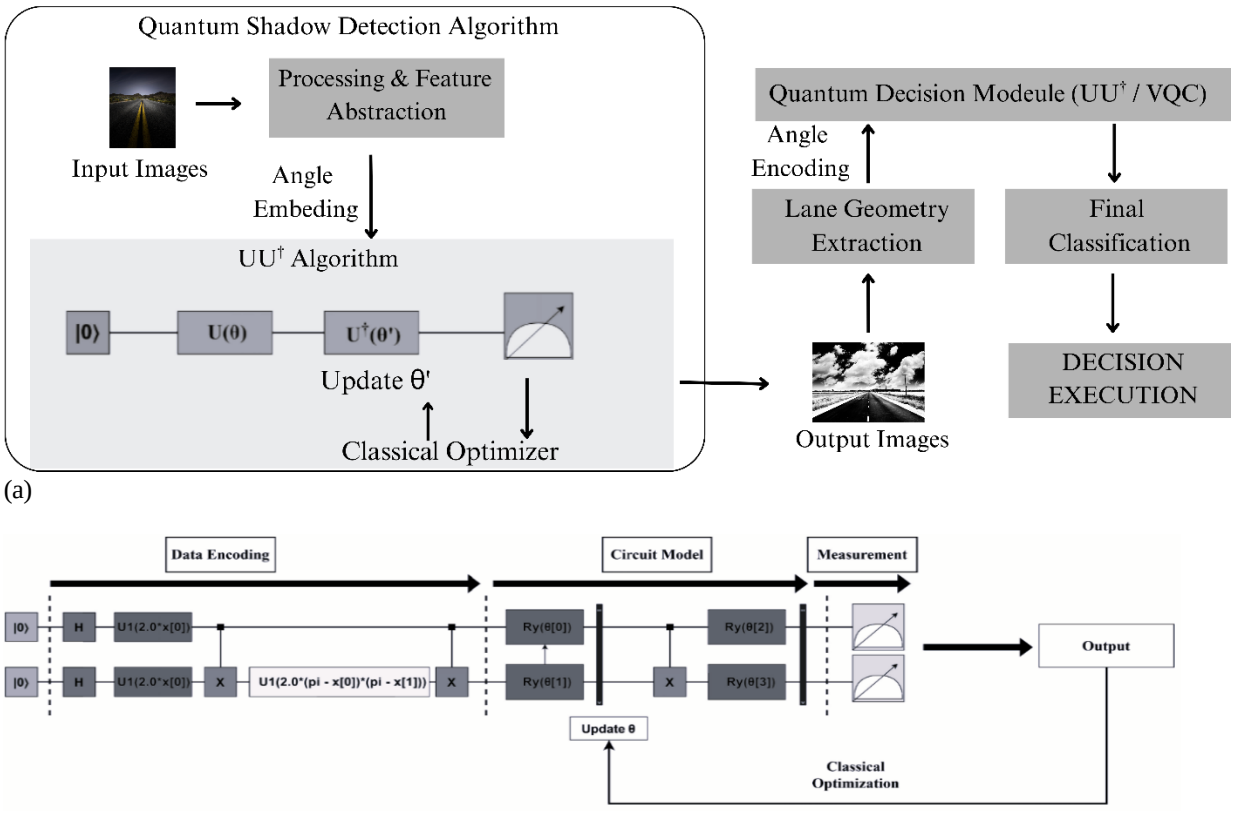
Due to the challenges faced by classical methods in meeting the requirements of autonomous vehicle systems—such as real-time data processing, high accuracy, low latency, and multi-sensor data integration in object detection, decision-making, and planning processes—quantum computing has begun to offer significant contributions to the subsystems of autonomous vehicles. In this context, analyzing the overall structures of the conducted studies and examining the underlying logic of the QML algorithms developed for specific problems will provide a strong foundation. The reviewed studies offer a basis for understanding both the integration of quantum computing and its architectural contributions within autonomous systems.

Efforts to enhance the safety and reliability of autonomous vehicles are primarily aimed at overcoming challenges such as uncertainty and computational complexity, particularly in real-time decision-making processes involving tasks like lane detection, shadow detection, and object recognition. In this context, the study presented in Figure 1 illustrates the process of enhancing lane and shadow detection by integrating quantum computing and deep learning techniques [13]. The QDCNN (Quantum Deep Convolutional Neural Network) architecture consists of two main stages: the quantum shadow detection algorithm and the lane decision and execution phase. Initially, the input image undergoes a preprocessing phase to detect shadows. This is followed by grayscale conversion using Gaussian blurring and the Canny edge detection function. Subsequently, Hough transform is applied to identify lane lines. After these steps, the number of lines, intersection points, and slope parameters are processed using the UU† and VQC [29] algorithms to support the decision-making and final prediction processes. In the QDCNN architecture proposed in [13], the VQC circuit starts with data encoding, which converts classical data into quantum representations. This is followed by the learning phase of the VQC, which consists of parameterized transformations that enable the formation of non-linear decision boundaries between data samples. Measurement gates and a classical optimization loop follow this stage. The optimization loop provides a classical

feedback mechanism to align the quantum circuit output with the desired targets. As a result, quantum computing is directly integrated into both the preprocessing phase and the final decision-making stage.

Deep learning-based image processing modules used in the perception systems of autonomous vehicles exhibit significant vulnerabilities against adversarial attacks [30], [31]. These attacks can cause severe misclassifications—such as interpreting a stop sign as a speed limit sign—by introducing minor perturbations to the input images. Figure 2 presents a hybrid classical–quantum deep learning model developed to address this issue [2]. For the transfer learning stage, pretrained AlexNet and VGG-16 models are employed. Subsequently, a Quantum Neural Network (QNN) replaces the classical neural network structure. The QNN layer comprises components similar to the previous architecture, including a data encoding layer and Variational Quantum Circuit (VQC) layers. Before the measurement layer, single-qubit gates (U1, U2) are used as a pre-measurement layer. These gates serve to transform the qubit states one final time to ensure accurate projection during measurement. In the proposed architecture [2], the VQC and entanglement structures demonstrate greater robustness compared to classical models in suppressing perturbations caused by adversarial attacks.

Although classical CNN architectures [32] offer high accuracy in object detection systems, computational complexity increases rapidly in applications involving multi-channel visual data. When 3D data, RGB images, and LiDAR maps are used together, the resulting increase in model size, memory requirements, and processing latency in real-time scenarios creates significant bottlenecks. To overcome these challenges, the FQCNN (Fast Quantum Convolutional Neural Network) architecture [16] Integrates quantum computing by loading multi-channel data into a single quantum system. The data are encoded onto the same qubit using rotation gates Rx, Ry, and Rz with different angles. Figure 3(a) illustrates this encoding scheme. Similar to other architectures, entanglement and parameterized quantum circuits (PQC) are used to implement the convolutional operations.
This system enables the processing of multi-channel data using the limited resources of quantum hardware by significantly reducing computational complexity. As a structured version of other QCNN architectures, this design employs a sequential encoding method to prevent unnecessary circuit growth within the quantum circuit [16]. Additionally, the training of the Quantum Region Proposal Network (QRPN) is optimized through a heterogeneous knowledge distillation mechanism, which integrates knowledge transfer techniques from classical neural networks into the quantum environment. This knowledge transfer process is depicted in Figure 3(b).



(b)

Figure 1. QDCNN architecture [13]. (a) System model, (b) VQC

The obstacle avoidance problem in autonomous vehicles is a complex optimization challenge that must be solved under limited computational resources. To address this problem—which requires real-time and safe route planning—the Hierarchical Quantum Reinforcement Learning (HQRL) framework [17] integrates Quantum Reinforcement Learning (QRL) with Hierarchical Reinforcement Learning. The architecture is designed in two levels: a policy over options and an intra-option policy. In the first level, the available options for each state are selected by a VQC-based Quantum Processing Unit (QPU). At the second level, the chosen sub-policy is responsible for executing specific tasks, which include lane following and obstacle avoidance. The training algorithm is based on the classical option-critic architecture, but in this case, only the option policy is trained. Policy updates are performed through gradients calculated over the total reward function [17]. The VQC used in the architecture is illustrated in Figure 4.

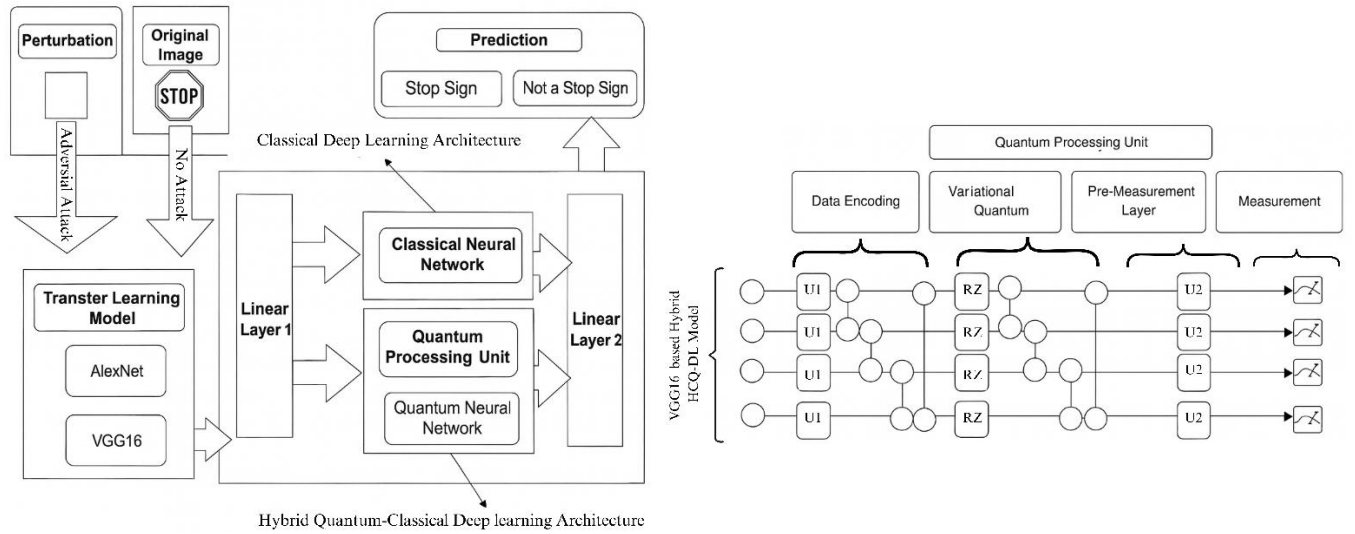


Figure 2. C-DL and HCQ-DL architecture [2]. (a) System model, (b) VGG16 based HCQ-DL model

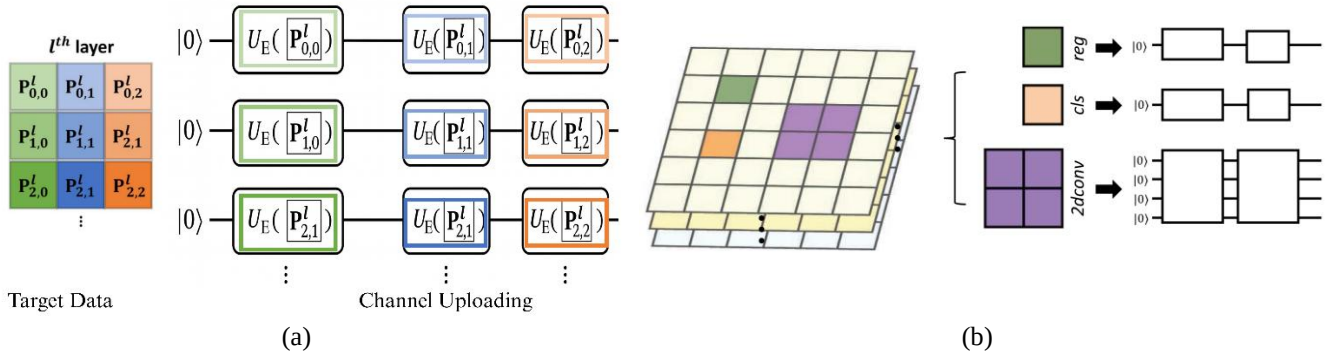


Figure 3. FQCNN architecture [16]. (a) Channel uploading, (b) Implementation details of out proposed QRPN

Although QRL holds great potential to solve problems that are intractable or difficult for classical methods, its practical implementation is limited due to challenges in encoding input data into quantum circuits and the persistent need for quantum hardware during the inference phase. To overcome these limitations, the Quantum-Train Reinforcement Learning (QTRL) architecture was proposed [12] In this approach, the parameters of the classical policy network are generated by a Quantum Neural Network (QNN) and a mapping model. This eliminates the need to encode input data into quantum

circuits, enabling the QNN to produce policy parameters independently of the input size, thereby decoupling circuit depth and width from the input dimensionality. Since the output of QTRL is a classical model, the inference process can be executed on classical hardware. This presents a significant advantage in latency-sensitive autonomous driving scenarios.

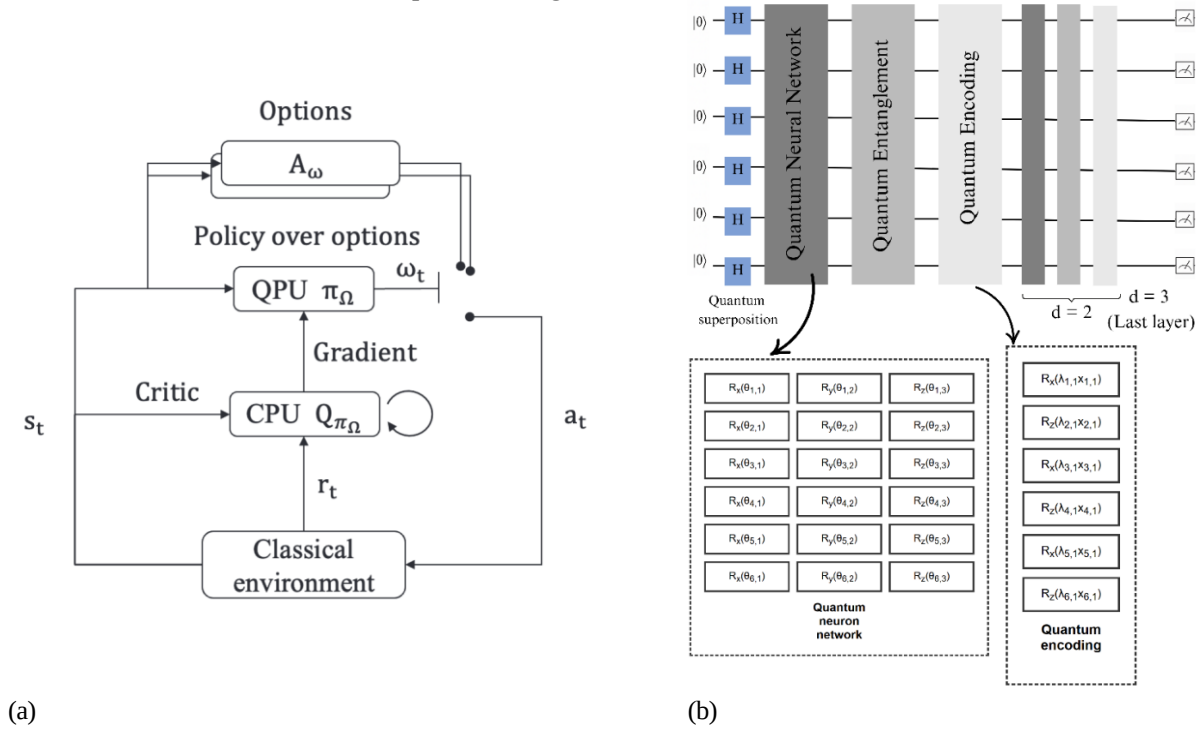


Figure 4. Option-critic with QRL architecture [17]. (a) architecture diagram, (b) VQC structure

Address the NP-hard problem of collision-free navigation in autonomous driving, a quantum-assisted deep reinforcement learning (DRL) algorithm [33], is proposed and illustrated in Figure 6. In this architecture [33] , unlike many other QRL approaches, quantum computation is employed only during the training phase, while the testing phase does not require quantum hardware. The actor–critic-based architecture implements the critic component using a parameterized quantum circuit (PQC). The overall framework consists of a classical feature encoder and a representation passed through an LSTM module, followed by a classical actor and a quantum-assisted critic component. An embedding method is utilized to flexibly encode high-dimensional classical data into a limited number of qubits. Within the quantum critic network, a hybrid structure combining the PQC and fully connected classical layers is used to evaluate the Q-value.

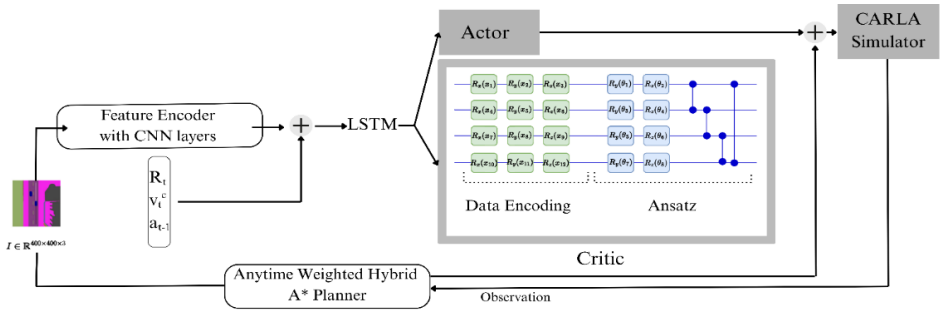


Figure 5. Quantum supported deep reinforcement learning [33]

4. Experimental Results - Comparative Analysis and Cross-Domain Evaluation

The performance advantages achieved through the integration of quantum algorithms into autonomous vehicle systems are comprehensively presented in Table 2. Across a wide range of domains—from shadow detection and adversarially robust traffic sign recognition to multi-channel object detection and policy learning in reinforcement learning scenarios—quantum-assisted architectures consistently demonstrate improvements in accuracy, processing speed, computational efficiency, and parameter reduction. Table 3 categorizes the quantum advantages obtained in the reviewed studies based on their respective problem domains. According to the findings of this systematic review, the HQRL and QTRL frameworks achieved significant parameter reductions of 47.7% and 96%, respectively, while maintaining or improving learning performance and safety metrics. Similarly, the Nav-Q architecture demonstrated a 10-fold improvement in effective model complexity and significantly lower variance during training, enabled by quantum computation, which led to enhanced learning stability. The FQCNN architecture provided a 13% increase in accuracy while halving activation latency, thereby emphasizing the practical applicability of quantum convolutional methods for real-time, high-dimensional input processing. These results confirm that quantum methodologies not only reduce resource consumption but also improve robustness, convergence speed, and scalability in autonomous systems.

Table 7. Experimental results

Reference	Algorithm	Accuracy	Speed	Parameters
[13]	QDCNN	- Track-1: $UU^{\dagger}$ %70, VQC %40 - Track-2: VQC &60 - CARLA: VQC %60	0.0049 s (shadow detection)	1 qubit ($UU^{\dagger}$), 2 qubit (VQC), 4 total parameters (VQC), 1 parameter ($UU^{\dagger}$ centroid)
[2]	HCQ-DL (VQC + QNN)	No attack: %98 (VGG), %96 (AlexNet) FGSA: %92-94 GA: %96-98 PGD: %85 (AlexNet)	Training Time: 28-47 min (HCQ-DL) (AlexNet faster than VGG)	2.5M-14.8M (depends on backbone) 3-6 qubits used in QNN
[16]	FQCNN	Mean average precision (IoU ≥ 0.5) = 58.1%	51.7% faster than CNN	4-qubit (2x2 QCNN), 1-qubit (1x1), heterogeneous KD (γ = 0.3), C = 64
[17]	HQRL	15.8% ↑ safety score improvement over HRL	Real-time simulation on 100 m track with moving pedestrians	78 (VQC-based HQRL) vs. 2790 (NN-based HRL);1/13 params
[12]	QTRL	CartPole main reward: 4493.8 / 500 MiniGrid main reward: 0.900 / 1.0	Classcial inference (real-time suitable)	CartPole: 651 MiniGrid: 2529 (vs. 4835 classical) 47.7% parameter reduction
[33]	Nav-Q	Safety Index = 3 (scenarios with <20% crashes/near-misses); Mean return ≈ 0.10	Average time to goal: 11.47 s; Training time ≈ 10.87 days	53 (critic only, 4 qubits, 2 layes) 98% parameter reduction, better stability (IQR: 20.02 vs 230.42), 10x effective dimension

Table 3. Problem domains of algorithms and the quantum advantages they achieve

Reference	Algorithm	Problem Domain	Quantum Advantage
[13]	QDCNN	Shadow detection	Real-time processing
[2]	HCQ-DL (VQC + QNN)	Traffic sign classification adversarial resilience	Adversarial robustness
[16]	QCNN/QODE	Multi-channel Detection	Parameter efficiency
[17]	HQRL	Obstacle avoidance	Parameter efficency
[12]	QTRL	Policy gradient learning	Parameter generation efficiency + hardware independence
[33]	Nav-Q	Collision free navigation	Model Complexity reduction + training stability

5. Conclusion

This study systematically examines the integration of quantum computing techniques into various subsystems of autonomous vehicle technologies. It evaluates the quantum-assisted architectures developed in the recent literature along with their experimental results. The evaluation clearly indicates that significant advantages have been achieved over classical methods in terms of computational efficiency, robustness, and architectural scalability. Foundational approaches such as Variational Quantum Circuits (VQC), Quantum Neural Networks (QNN), and Quantum Reinforcement Learning (QRL), along with their advanced variants and hybrid quantum-classical architectures, exhibit the potential to overcome limitations such as resource constraints and the need for real-time inference. Despite the constraints of NISQ-era devices, architectural frameworks like QTRL and HQRL represent practical examples of how quantum computing can be applied in real-world autonomous systems. The findings of this review not only highlight the potential of quantum computing to enhance the capabilities of intelligent transportation systems, but also provide a foundational perspective by jointly presenting architectural strategies for applying both QCNN and QRL models to different subsystems in autonomous vehicles.

6. References

[1] R. Harikrishnakumar, S. Nannapaneni, N. H. Nguyen, J. E. Steck, and E. C. Behrman, "A Quantum Annealing Approach for Dynamic Multi-Depot Capacitated Vehicle Routing Problem," May 27, 2020, *arXiv*: arXiv:2005.12478. doi: 10.48550/arXiv.2005.12478.

[2] R. Majumder *et al.*, "Quantum Computing Supported Adversarial Attack-Resilient Autonomous Vehicle Perception Module for Traffic Sign Classification," Apr. 17, 2025, *arXiv*: arXiv:2504.12644. doi: 10.48550/arXiv.2504.12644.

[3] R. G. Torrome and S. Barzanjeh, "Advances in Quantum Radar and Quantum LiDAR," *Progress in Quantum Electronics*, vol. 93, p. 100497, Jan. 2024, doi: 10.1016/j.pquantelec.2023.100497.

[4] U. Antero, B. Sierra, J. Oñativia, A. Ruiz, and E. Osaba, "Robot localization aided by quantum algorithms," May 01, 2025, *arXiv*: arXiv:2502.00077. doi: 10.48550/arXiv.2502.00077.

[5] Y. Zhuang *et al.*, "Quantum Computing in Intelligent Transportation Systems: A Survey," *CHAIN*, vol. 1, no. 2, pp. 138–149, Jun. 2024, doi: 10.23919/CHAIN.2024.000007.

[6] "Application of Quantum Genetic Optimization of LVQ Neural Network in Smart City Traffic Network Prediction | IEEE Journals & Magazine | IEEE Xplore." Accessed: May 29, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9107118

[7] "New method of traffic flow forecasting based on quantum particle swarm optimization strategy for intelligent transportation system - Zhang - 2021 - International Journal of Communication Systems - Wiley Online Library." Accessed: May 29, 2025. [Online]. Available: https://onlinelibrary.wiley.com/doi/full/10.1002/dac.4647

[8] F. Neukart, G. Compostella, C. Seidel, D. von Dollen, S. Yarkoni, and B. Parney, "Traffic Flow Optimization Using a Quantum Annealer," *Front. ICT*, vol. 4, Dec. 2017, doi: 10.3389/fict.2017.00029.

[9] S. Oh, J. Choi, and J. Kim, "A Tutorial on Quantum Convolutional Neural Networks (QCNN)," in *2020 International Conference on Information and Communication Technology Convergence (ICTC)*, Oct. 2020, pp. 236–239. doi: 10.1109/ICTC49870.2020.9289439.

[10] I. Kerenidis, J. Landman, and A. Prakash, "Quantum Algorithms for Deep Convolutional Neural Networks," Nov. 04, 2019, *arXiv*: arXiv:1911.01117. doi: 10.48550/arXiv.1911.01117.

[11] A. Rattan, A. Rudra Pal, and M. Gurusamy, "Quantum Computing for Advanced Driver Assistance Systems and Autonomous Vehicles: A Review," *IEEE Access*, vol. 13, pp. 17554–17582, 2025, doi: 10.1109/ACCESS.2025.3532958.

[12] C.-Y. Liu, C.-H. A. Lin, C.-H. H. Yang, K.-C. Chen, and M.-H. Hsieh, "QTRL: Toward Practical Quantum Reinforcement Learning via Quantum-Train," in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Sep. 2024, pp. 317–322. doi: 10.1109/QCE60285.2024.10299.

[13] A. Meghanath, S. Das, B. K. Behera, M. A. Khan, S. Al-Kuwari, and A. Farouk, "QDCNN: Quantum Deep Learning for Enhancing Safety and Reliability in Autonomous Transportation Systems," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–11, 2025, doi: 10.1109/TITS.2025.3546353.

[14] P. S. Blakey *et al.*, "Quantum and non-local effects offer over 40 dB noise resilience advantage towards quantum lidar," *Nat Commun*, vol. 13, no. 1, p. 5633, Sep. 2022, doi: 10.1038/s41467-022-33376-9.

[15] M. Seyfipoor, M. J. S. Zafarqandi, and S. Mohammadi, "Real-Time Collaborative Route Planning in Autonomous Vehicles by Leveraging Optimized Quantum Neural Networks," in *2025 29th International Computer Conference, Computer Society of Iran (CSICC)*, Feb. 2025, pp. 1–7. doi: 10.1109/CSICC65765.2025.10967421.
[16] E. J. Roh, H. Baek, D. Kim, and J. Kim, "Fast Quantum Convolutional Neural Networks for Low-Complexity Object Detection in Autonomous Driving Applications," *IEEE Transactions on Mobile Computing*, vol. 24, no. 2, pp. 1031–1042, Feb. 2025, doi: 10.1109/TMC.2024.3470328.
[17] "Motion Planning of Vehicle Obstacle Avoidance by Quantum Reinforcement Learning | IEEE Conference Publication | IEEE Xplore." Accessed: May 30, 2025. [Online]. Available: https://ieeexplore.ieee.org/document/10450836
[18] S. N. Gungor and M. Karakose, "A Quantum Computing Based Solution Approach for NP-Complete Problems," in *2025 29th International Conference on Information Technology (IT)*, Feb. 2025, pp. 1–4. doi: 10.1109/IT64745.2025.10930248.
[19] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*, 10th anniversary edition. Cambridge: Cambridge university press, 2010.
[20] N. F. Bar and M. Karakose, "Quantum Circuit Dataset Generator Approach for Deep Learning based Solutions," in *2024 28th International Conference on Information Technology (IT)*, Feb. 2024, pp. 1–4. doi: 10.1109/IT61232.2024.10475754.
[21] N. F. Bar, M. Yenilmez, S. Aksu, and M. Karakose, "A Quantum Computing based Approach for Sentiment Analysis in Bilateral Conversations," in *2024 28th International Conference on Information Technology (IT)*, Feb. 2024, pp. 1–4. doi: 10.1109/IT61232.2024.10475718.
[22] "Variational quantum algorithms | Nature Reviews Physics." Accessed: Jul. 29, 2025. [Online]. Available: https://www.nature.com/articles/s42254-021-00348-9
[23] "The quest for a Quantum Neural Network | Quantum Information Processing." Accessed: Jul. 29, 2025. [Online]. Available: https://link.springer.com/article/10.1007/s11128-014-0809-8
[24] P. Rebentrost, M. Mohseni, and S. Lloyd, "Quantum Support Vector Machine for Big Data Classification," *Phys. Rev. Lett.*, vol. 113, no. 13, p. 130503, Sep. 2014, doi: 10.1103/PhysRevLett.113.130503.
[25] "Quantum convolutional neural networks | Nature Physics." Accessed: Jul. 29, 2025. [Online]. Available: https://www.nature.com/articles/s41567-019-0648-8
[26] "A generative modeling approach for benchmarking and training shallow quantum circuits | npj Quantum Information." Accessed: Jul. 29, 2025. [Online]. Available: https://www.nature.com/articles/s41534-019-0157-8
[27] V. Dunjko and H. J. Briegel, "Machine learning & artificial intelligence in the quantum domain: a review of recent progress," *Rep. Prog. Phys.*, vol. 81, no. 7, p. 074001, Jun. 2018, doi: 10.1088/1361-6633/aab406.
[28] "Variational Quantum Circuits for Deep Reinforcement Learning | IEEE Journals & Magazine | IEEE Xplore." Accessed: Apr. 14, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9144562
[29] U. Azad, B. K. Behera, E. A. Ahmed, P. K. Panigrahi, and A. Farouk, "Solving Vehicle Routing Problem Using Quantum Approximate Optimization Algorithm," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 7, pp. 7564–7573, Jul. 2023, doi: 10.1109/TITS.2022.3172241.
[30] S.-M. Moosavi-Dezfooli, A. Fawzi, and P. Frossard, "DeepFool: A Simple and Accurate Method to Fool Deep Neural Networks," presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2574–2582. Accessed: Jul. 31, 2025. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2016/html/Moosavi-Dezfooli_DeepFool_A_Simple_CVPR_2016_paper.html
[31] N. Akhtar and A. Mian, "Threat of Adversarial Attacks on Deep Learning in Computer Vision: A Survey," *IEEE Access*, vol. 6, pp. 14410–14430, 2018, doi: 10.1109/ACCESS.2018.2807385.
[32] F. Sultana, A. Sufian, and P. Dutta, "A Review of Object Detection Models Based on Convolutional Neural Network," in *Intelligent Computing: Image Processing Based Applications*, J. K. Mandal and S. Banerjee, Eds., Singapore: Springer, 2020, pp. 1–16. doi: 10.1007/978-981-15-4288-6_1.
[33] A. Sinha, A. Macaluso, and M. Klusch, "Nav-Q: quantum deep reinforcement learning for collision-free navigation of self-driving cars," *Quantum Mach. Intell.*, vol. 7, no. 1, p. 19, Feb. 2025, doi: 10.1007/s42484-024-00226-4.

Turkish Regulatory QA: A Hybrid BM25 and Pretrained Transformer Retrieval Framework with Label-Guided Question-Answering

İlhan AYDIN[1], Zeynep KARACA[2]

[1,2]Department of Computer Engineering, Faculty of Engineering, Fırat University, Elazığ, Turkey.
[1]iaydin@firat.edu.tr, [2]zkaraca@firat.edu.tr
[1](ORCID: 0000-0001-6880-4935), [2](ORCID: 0009-0004-4335-5853)

Abstract
This study presents a user-oriented, interactive question–answering (QA) system developed for Turkish regulatory texts. The proposed system identifies relevant regulatory provisions by hybridizing lexical (BM25) and semantic (transformer-based) similarity measures. First, candidate provisions are ranked by BM25; subsequently, semantic similarity is computed using embeddings extracted from pretrained transformers such as BERTurk and XLM-RoBERTa. The system also includes a classifier component that assigns user queries to predefined regulatory headings, thereby increasing the relevance of the returned answers. Experiments on a manually annotated dataset derived from the regulations of Fırat University show that the proposed hybrid approaches improve retrieval and classification metrics compared to single-method baselines. The developed interactive interface enhances the user experience from query input to answer presentation.

Keywords: pretrained transformers, hybrid retrieval, semantic search, regulatory text, Turkish NLP, question-answering

1. Introduction

Regulatory texts form the basis of decision-making processes across a variety of domains, including public institutions, universities and the private sector. However, these documents are often long, technical and complex, which makes it difficult for users to access the information they need quickly and accurately [1]. This problem becomes particularly pronounced in contexts where regulations change frequently and timely access to information is critical. The distinctive terminology and structural complexity of legal documents further exacerbate these difficulties. The increasing volume of laws, articles and regulations over time necessitates the development of automated information-access solutions [2]. Natural Language Processing (NLP) based QA systems can enhance the accessibility of regulatory texts by extracting contextually relevant information from documents. In the field of information retrieval, both lexical (sparse) and semantic (dense) approaches are widely used. Lexical ranking algorithms such as BM25 provide fast and effective term matching, whereas transformer-based models represent texts in dense vector spaces and are better able to capture semantic proximity [3]. Hybrid information-retrieval methods aim to combine the advantages of lexical and semantic approaches to achieve higher accuracy and coverage [1]. In this study, we compare hybrid retrieval systems—which integrate lexical and semantic signals—with QA systems based solely on large language models, using a manually annotated dataset derived from the regulations of Fırat University. An independent classifier component for query labeling is also included to support the pipeline.

The principal contributions of this work are twofold: (1) the creation of a manually annotated dataset derived from a Turkish university regulation, and (2) a comparative evaluation of single-method and hybrid configurations that combine BM25 lexical ranking with transformer-based models such as BERTurk and XLM-RoBERTa. This comparison provides a practical reference for retrieval and classification tasks in the Turkish regulatory domain.

Retrieval-based QA systems have been intensively studied across multiple domains (e.g., legal, biomedical, and technical documentation). Surveying the literature in this area reveals several notable works. Rayo et al. [4] developed a hybrid retrieval system that combines BM25 with a fine-tuned sentence transformer for extracting information from regulatory texts; the retrieved passages were then integrated with a Retrieval-Augmented Generation (RAG) pipeline based on large language models. Their dataset consisted of large-scale regulatory documents, and their results demonstrated that the hybrid system significantly improved Recall@10 and MAP@10 relative to purely lexical or purely semantic baselines.

[1]Corresponding author

Wiratunga et al. [5] proposed a CBR-RAG approach that integrates Case-Based Reasoning (CBR) with RAG for legal QA; retrieval and similarity measures from the CBR cycle were used to enrich LLM prompts, and the results showed that the contextual enrichment provided by CBR notably improved the quality and accuracy of generated answers. Esposito et al. [6] developed a hybrid query-expansion method combining MultiWordNet-derived terms with Word2Vec embeddings and demonstrated its effectiveness on Italian cultural-heritage documents, reporting improvements in sentence retrieval and correct-answer rates. Agarwal et al. [7] proposed a hybrid graph-based approach that leverages LLMs for complex QA over tabular and textual data and evaluated it on Hybrid-QA and OTT-QA datasets using GPT-3.5, GPT-4 and LLaMA-3; their method increased Exact Match in zero-shot scenarios while reducing token usage. Min et al. [8] examined the integration of table–text hybrid data into LLM-based QA systems and compared various table-to-text transformation methods using industrial datasets; they found that LLM-based and Markdown conversion approaches often achieved superior retrieval and QA performance by improving semantic alignment. Zhang et al. [9] introduced HD-RAG and the DocRAGLib dataset for hybrid documents that combine text and hierarchical tables; HD-RAG outperformed contemporary methods on retrieval accuracy and QA performance. Verma et al. [10] proposed HybridContextQA, a hybrid method that combines knowledge graphs with vector-based retrieval to improve LLM performance on complex conditional questions. Doan et al. [11] combined text and knowledge-graph embeddings in a hybrid retrieval method to enhance the accuracy of RAG systems in health-technology applications, reporting significant improvements in retrieval accuracy and the quality of LLM-generated answers. Tachioka [12] combined BM25-based similarity scoring with LLM-based summarization to produce QA summaries from unstructured texts; BM25 was used to select relevant answer sentences which were then summarized by an LLM, and the hybrid approach yielded significantly better results than LLM-only summarization.

2. Materials and Methods

In this section, the dataset, the models and retrieval architecture used, the hybrid scoring method, and the evaluation metrics are examined in detail. The experimental results and comparative analyses are presented in Section 3, while Section 4 discusses the findings, limitations, and directions for future work.

2.1 Dataset

The dataset employed in this study was extracted from the regulations of Fırat University and manually annotated and structured into an Excel (.xlsx) file. For each provision, the dataset contains the following columns: *Article Number*, *Text*, *Label* and *Answer*. The *Article Number* field denotes the regulation clause identifier; *Text* contains the question or query associated with the provision; *Label* records the provision's category or heading; and *Answer* contains the passage selected as the response. The structure of the dataset is shown in Table 1.

Table 9. Example records from the annotated regulatory dataset

Article Number	Text	Label	Answer
Madde 5	Üniversitenin öğretim birimlerine öğrenci kabulü nasıl olur?	Kayıt ve Kabul	Üniversitenin öğretim birimlerine öğrenci kabulü, ÖSYM yolu ile ve gerekli birimlerde Rektörlükçe düzenlenecek özel yetenek sınavına dayanılarak yapılır.
Madde 5	Üniversiteye girmeye hak kazanan ve başka bir öğretim kurumunda veya Üniversitenin başka bir biriminde kayıtlı olmayan öğrenciler Kayıtlarını ne zaman yaptırır?	Kayıt ve Kabul	Üniversiteye girmeye hak kazanan ve başka bir öğretim kurumunda veya Üniversitenin başka bir biriminde kayıtlı olmayan öğrenciler, ÖSYM'nce belirlenen ve ilan edilen günlerde kayıtlarını yaptırırlar.
Madde 5	Zamanında bizzat başvurmayan ve istenilen belgeleri tamamlamayan öğrencilere ne olur?	Kayıt ve Kabul	Zamanında bizzat başvurmayan ve istenilen belgeleri tamamlamayan öğrenciler, bu hakkını kaybeder.

Madde 13	Mazeret sınavı hangi öğrenciler için yapılır?	Mazeret Sınavı	Mazeret sınavı, ara sınav gününü kapsayan tarihlerde Rektörlük tarafından görevlendirilen veya ilgili birimin yönetim kurulunca kabul edilebilir bir mazereti nedeniyle ara sınava katılamayan öğrenciler için yapılır.
Madde 13	Mazeret sınavı hangi sınavlar için ve yarıyıl / yılda kaç defa yapılır?	Mazeret Sınavı	Mazeret sınav hakkı sadece ara sınavlar için verilir. Mazeret ara sınavı, ara sınav sayısına bakılmaksızın her yarıyıl / yılda, her ders için bir defa yapılır.
Madde 13	Öğrenci mazeret sınavına girebilmek için mazeretini nereye ve en geç ne zamana kadar bildirmelidir?	Mazeret Sınavı	Öğrenci, mazeretini Dekanlığa veya Müdürlüğe, ara sınavların bitiminden sonra, en geç bir hafta içerisinde bildirmek zorundadır.
Madde 13	Mazeret sınavı ne zaman yapılır?	Mazeret Sınavı	Mazeret sınavı, yarıyıl / yılın son haftasında ilgili birimler tarafından yapılır.
Madde 13	Mazeret sınavında hangi konulardan sorumludur?	Mazeret Sınavı	Öğrenci mazeret sınavında, o tarihe kadar olan müfredat programından sorumludur.
Madde 13	Mazeret sınav notu hangi sınavın notu yerine geçer?	Mazeret Sınavı	Mazeret sınav notu, bu öğrenciler için katılmadıkları ara sınav notlarının yerine geçer.

As shown in Table 1, long texts corresponding to a single article are split into smaller fragments that can answer different user questions; therefore, the dataset may contain multiple examples sharing the same Article Number. Dataset statistics are presented in Table 2.

Table 10. Dataset statistics

Description	Value
Total number of examples	87
Number of unique articles	46
Number of unique labels	43
Average answer length (words)	40.38

The label distribution in the dataset is relatively imbalanced; certain labels are overrepresented (e.g., “Definition” = 8, “Make-up Exam” = 6, “Registration Renewal” = 6, “Exam Format” = 4, and “Graduation” = 1). This class imbalance may affect classifier training and evaluation.

2.2. Models and Architecture

In this study, we compare four principal system configurations: (1) BERTurk (`dbmdz/bert-base-turkish-cased`), (2) BM25 + BERTurk (hybrid), (3) XLM-RoBERTa (`xlm-roberta-base`), and (4) BM25 + XLM-RoBERTa (hybrid). BERTurk is a BERT variant pretrained specifically for Turkish and is widely used in the community; XLM-RoBERTa is a multilingual model pretrained at scale and has demonstrated strong performance across both low- and high-resource languages. The selection of these models is motivated by the need for robust Turkish language support as well as the desire to evaluate multilingual baselines for comparison. The overall system architecture is illustrated in Figure 1.

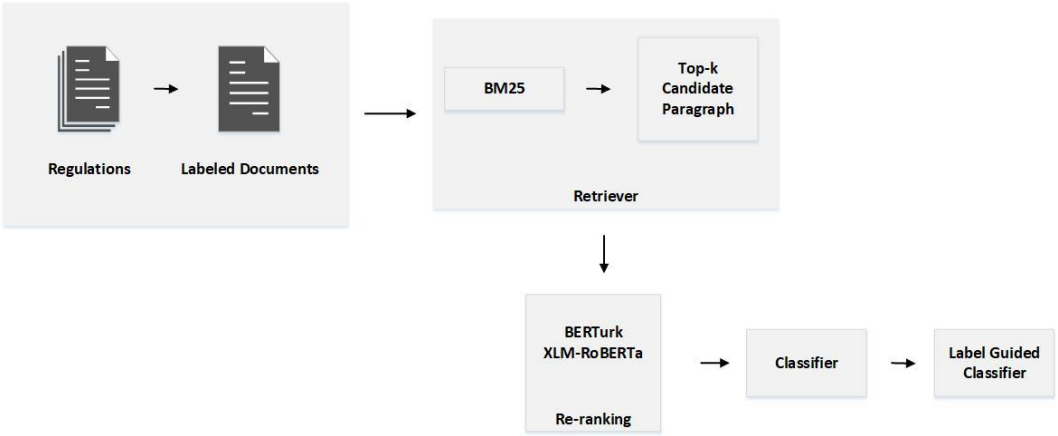


Figure 8. System architecture for the regulatory QA pipeline.

Figure 1 depicts the two-stage retrieval pipeline used throughout the experiments: (1) a fast lexical retriever (BM25) that produces a compact candidate set from the regulation corpus after minimal preprocessing, and (2) a transformer-based reranker that computes semantic similarity scores for those candidates using pretrained encoders. In our implementation the reranking component is instantiated separately with BERTurk and XLM-RoBERTa (i.e., experiments run each encoder independently, and hybrid variants combine BM25 + the chosen encoder). For each encoder, query and candidate texts are tokenized, encoded, pooled into fixed-length sentence vectors (attention-masked mean pooling), L2-normalized and compared by dot product to obtain dense similarity scores. Finally, the system fuses lexical and dense signals (score fusion) and optionally consults a label-guided classifier to filter or re-rank results before presenting the top-K answers to the user.

2.2.1. Retriever Component – BM25

For candidate generation we employ the widely used BM25 ranking algorithm, originating from the Okapi information retrieval framework. BM25 provides an effective retrieval backbone for scoring the relevance between documents and user queries [4]. Intuitively, BM25 gives higher weight to terms that are frequent in a given document but rare across the collection, thereby identifying terms that are more informative for that document. Because BM25 produces efficient and accurate term-matching ranks, it remains a strong baseline for initial candidate retrieval in the retrieval pipeline [13], [14]. Equation 1 shows the BM25 calculation.

$$score_{BM25}(d,Q) = \sum_{i=1}^{|Q|} idf(q_i).\frac{tf(q_i,d).(k_1+1)}{tf(q_i,d)+k_1.B} \quad (1)$$

In Equation (1) [14], $tf(q_i,d)$ denotes the term-frequency weight, $df(q_i)$ denotes the inverse document frequency weight for the query term q_i and k_1 denotes a parameter that typically takes a value between 1 and 3. In this study, BM25 was applied to tokenized *Answer* texts (preserving Turkish characters) to generate top-K candidates for candidate generation.

2.2.2. Dense Components – Transformer-based Embeddings

The dense (semantic) similarity component is computed from sentence embeddings extracted using pretrained transformer models. In this study, we employ BERTurk, a BERT variant pretrained on Turkish, and XLM-RoBERTa as a multilingual baseline. After tokenizing queries and candidate texts with the models' subword tokenizers, the token-level hidden states are aggregated into a single sentence representation via mean pooling, with the attention mask applied as weights. The resulting representations are L2-normalized, and query–document similarity is computed by dot-product. This procedure provides an efficient and reproducible solution for domain-specific semantic matching. The dense similarity scores are computed for the candidate set produced by BM25. The effectiveness of transformer-based dense embedding approaches for retrieval tasks has been demonstrated in prior work such as DPR and related studies [16].

2.2.3. BM25 + Dense Fusion Hybrid Sequencing

The retrieval pipeline is designed as a two-stage approach. In the first stage, a fast lexical filter generates candidate documents; in the second stage, these candidates are re-examined using transformer-based semantic similarity. The two signals (lexical and semantic) are scaled at the query level and combined into a single fused score, with the fusion weights adjusted according to the system's preferences for recall versus semantic precision. This design leverages BM25's broad coverage together with the semantic discriminative power of transformer embeddings to produce the final ranking.

2.2.4 Classifier Component

A transformer-based classifier was employed to assign user queries to predefined regulatory headings. The classifier was fine-tuned using a standard train–validation split and common optimization procedures; a detailed hyperparameter table for the training and evaluation protocol is provided in Table 3.

Table 11. Key hyperparameters used in the experiments

Parameter	Value	Description
Max Length	128	Maximum token length for the tokenizer/mode
Train Batch Size	16	Batch size for classifier training
Embed Batch	16	Batch size for embedding extraction
Learning Rate	2e-5	Learning rate for classifier optimization
Epoch	40	Number of training epochs for the classifier
BM25 Top-K	50	Number of candidates retrieved by BM25 (candidate generation)
Final Top-K	5	Number of final top-k results shown to the user

3. Results

In this section we present the experimental findings of the proposed question–answering system. The evaluation is based on both quantitative metrics and qualitative examples. Experiments were conducted on four model configurations: BERTurk, XLM-RoBERTa, BM25 + BERTurk, and BM25 + XLM-RoBERTa. The retrieval-focused evaluation compares the performance achieved by combining classical lexical signals (BM25) with transformer-based embedding signals.

3.1 Retrieval Results

In our study, Top-*k* metrics were employed to evaluate retrieval performance. These metrics measure the system's ability to present the correct answer within the top *k* returned results. The obtained results are presented in Table 4**.**

Table 12. Comparative performance metrics for Transformer-Based and BM25-enhanced models

Model	Top-1	Top-3	Top-5	MRR
BERTurk	50.0	83.33	83.33	62.96
XLM-RoBERTa	48.5	76.66	76.66	59.67
BM25+BERTurk	88.88	94.44	94.44	91.66
BM25+XLM-RoBERTa	72.22	94.4	94.4	82.40

As shown in Table 4, the purely dense embedding-based models BERTurk and XLM-RoBERTa achieved Top-1 scores of 50.0% and 48.5%, respectively. In contrast, the BM25+BERTurk combination demonstrated the highest performance, achieving 88.88% in Top-1 and 94.44% in both Top-3 and Top-5. The BM25+XLM-RoBERTa model exhibited a lower Top-1 performance of 72.22% but still reached high scores of 94.4% in Top-3 and Top-5. In terms of Mean Reciprocal Rank (MRR), the BM25+BERTurk model achieved the best performance with 91.66%. These findings indicate that hybrid BM25 and dense embedding models significantly outperform purely dense models.

3.2. Qualitative Examples

To complement the quantitative evaluation, representative qualitative examples are presented below. The Table 5 reports, for each model and selected user query: (i) the ground-truth article number with its true answer excerpt, and (ii) the model's top predicted passages (article number, short excerpt) together with the model score. These cases illustrate typical successes and failure modes of dense-only and hybrid retrieval configurations in the regulatory domain.

Table 13. Representative qualitative retrieval example

Model	Question	Article Number & True Answer	Predict Article Number & Predict Answer	Score
BERTurk	Mazeret sınavı hangi öğrenciler için yapılır?	Madde 13 - Ders kayıt işlemleri, öğrencilere internet ortamında duyurulur. Öğrenciler derslere kayıtlarını, Üniversite Senatosunun tespit ettiği esaslar ve Rektörlük tarafından belirlenen usullere göre, bilgisayar ortamında interaktif olarak yenilerler. Kayıt yenileme işlemlerinin tümünden öğrenci sorumludur. Kaydın son şekli, danışman öğretim elemanı tarafından dönemin ilk iki haftası içerisinde onaylanır.	Madde 13 - Mazeret sınav notu, bu öğrenciler için katılmadıkları ara sınav notlarının yerine geçer.	86.89
			Madde 13 - Mazeret sınav hakkı sadece ara sınavlar için verilir. Mazeret ara sınavı, ara sınav sayısına bakılmaksızın her yarıyıl / yılda, her ders için bir defa yapılır.	84.39
			Madde 17 - Tek ders sınavı, bu Yönetmeliğin 14 üncü maddesindeki şartları daha önce yerine getirmek kaydıyla mezuniyet aşamasına gelip, sadece tek dersten FF notu olan öğrenciler için yapılan sınavdır.	83.22
			Madde 19 - Muafiyet sınavı, Üniversiteye yeni kayıt yaptıran öğrencilere, ortak zorunlu yabancı dil dersinden muaf olmaları için, öğretim yılı başında açılan bir sınavdır.	82.70
			Madde 13 - Mazeret sınavı, ara sınav gününü kapsayan tarihlerde Rektörlük tarafından görevlendirilen veya ilgili birimin yönetim kurulunca kabul edilebilir bir mazereti nedeniyle ara sınava katılamayan öğrenciler için yapılır.	82.28
BERTurk	Dersi geçebilmek için genel sınav notu en az kaç olmalıdır?	Madde 15 - Yarıyıl / yılsonu sınavında 100 üzerinden Meslek Yüksekokulları için en az 35, diğer tüm akademik birimler için 50 puan alma zorunluluğu vardır. Yarıyıl/yıl sonu sınavına girmeyen veya bu sınavdan en az yukarıda belirtilmiş olan puanı alamayan öğrenciler FF başarı notu ile değerlendirilir.	Madde 14 - Yarıyıl / yılsonu sınavında 100 üzerinden Meslek Yüksekokulları için en az 35, diğer tüm akademik birimler için 50 puan alma zorunluluğu vardır. Yarıyıl/yıl sonu sınavına girmeyen veya bu sınavdan en az yukarıda belirtilmiş olan puanı alamayan öğrenciler FF başarı notu ile değerlendirilir.	90.04

Table 14. (Continue) Representative qualitative retrieval example

			Madde 15 - Bütünleme sınavında meslek yüksekokulları için en az 35, diğer tüm akademik birimler için 50 puan alma zorunluluğu vardır. Bütünleme sınavında alınan not, o dersin yarıyıl/yılsonu notu yerine geçer.	89.71
			Madde 12 - b) Ara sınavın başarı notuna katkısı %40'dır.	88.9
			Madde 15 - Öğrenciler, akademik takvimde belirtilen sürede enteraktif olarak kayıt yaptırmak şartıyla, bütünleme sınavına, DD ve DC aldığı derslerden de girebilir. Genel sınavda FF almış olan öğrencilerin bütünleme sınavına girebilmeleri için kayıt yaptırmaları gerekmez.Son ara sınav notu bütünleme sınavında da geçerlidir.	88.26
			Madde 30 - Bir öğrencinin mezun olabilmesi için, bu Yönetmeliğin 8 inci maddesinde belirtilen yasal süre içinde, izlemekte olduğu programda belirtilen tüm derslerden en az (CC), stajlardan da (B) almak zorundadır. GNO'su 2.00 ve üzerinde olan öğrenciler, (DC) ve (DD) aldıkları derslerden de başarılı kabul edilirler. GNO, aynı zamanda mezuniyet not ortalamasıdır.	88.09
BM25 + BERTurk	Mazeret sınavı hangi öğrenciler için yapılır?	Madde 13 - Ders kayıt işlemleri, öğrencilere internet ortamında duyurulur. Öğrenciler derslere kayıtlarını, Üniversite Senatosunun tespit ettiği esaslar ve Rektörlük tarafından belirlenen usullere göre, bilgisayar ortamında interaktif olarak yenilerler. Kayıt yenileme işlemlerinin tümünden öğrenci sorumludur. Kaydın son şekli, danışman öğretim elemanı tarafından dönemin ilk iki haftası içerisinde onaylanır.	Madde 13 - Mazeret sınav hakkı sadece ara sınavlar için verilir. Mazeret ara sınavı, ara sınav sayısına bakılmaksızın her yarıyıl / yılda, her ders için bir defa yapılır.	95.25
			Madde 13 - Mazeret sınavı, ara sınav gününü kapsayan tarihlerde Rektörlük tarafından görevlendirilen veya ilgili birimin yönetim kurulunca kabul edilebilir bir mazereti nedeniyle ara sınava katılamayan öğrenciler için yapılır.	90.57
			Madde 13 - Mazeret sınavı, yarıyıl / yılın son haftasında ilgili birimler tarafından yapılır.	87.25
			Madde 13 - Mazeret sınav notu, bu öğrenciler için katılmadıkları ara sınav notlarının yerine geçer.	79.15
			Madde 18 - Son yarıyıl / yılın sonunda izledikleri programdaki bütün derslerden en az (DD) notu aldıkları halde, Genel Not Ortalamaları 2.00'ın altına düşen öğrenciler son 2 (iki) yılda; (DD), (DC) veya (CC) notu aldıkları en çok üç dersten mezuniyet "Genel Not Ortalaması Baraj Sınavı"na girerler. Bu sınav haklarını, takip eden yıl başlangıcındaki ders kayıt haftasında kullanırlar. Bu sınava giremeyenler için mazeret sınavı açılmaz.	66.56

Table 15. (Continue) Representative qualitative retrieval example

BM25 + BERTurk	Dersi geçebilmek için genel sınav notu en az kaç olmalıdır?	Madde 15 - Yarıyıl / yılsonu sınavında 100 üzerinden Meslek Yüksekokulları için en az 35, diğer tüm akademik birimler için 50 puan alma zorunluluğu vardır. Yarıyıl/yıl sonu sınavına girmeyen veya bu sınavdan en az yukarıda belirtilmiş olan puanı alamayan öğrenciler FF başarı notu ile değerlendirilir.	Madde 18 - Son yarıyıl / yılın sonunda izledikleri programdaki bütün derslerden en az (DD) notu aldıkları halde, Genel Not Ortalamaları 2.00'ın altına düşen öğrenciler son 2 (iki) yılda; (DD), (DC) veya (CC) notu aldıkları en çok üç dersten mezuniyet "Genel Not Ortalaması Baraj Sınavı"na girerler. Bu sınav haklarını, takip eden yıl başlangıcındaki ders kayıt haftasında kullanırlar. Bu sınava giremeyenler için mazeret sınavı açılmaz.	93.01
			Madde 14 - Yarıyıl / yılsonu sınavında 100 üzerinden Meslek Yüksekokulları için en az 35, diğer tüm akademik birimler için 50 puan alma zorunluluğu vardır. Yarıyıl/yıl sonu sınavına girmeyen veya bu sınavdan en az yukarıda belirtilmiş olan puanı alamayan öğrenciler FF başarı notu ile değerlendirilir.	88.58
			Madde 15 - Bütünleme sınavında meslek yüksekokulları için en az 35, diğer tüm akademik birimler için 50 puan alma zorunluluğu vardır. Bütünleme sınavında alınan not, o dersin yarıyıl/yılsonu notu yerine geçer.	85.44
			Madde 12 - Her ders için, bir yarıyılda/yılda en az bir ara sınav yapılır.	74.78
			Madde 15 - Öğrenciler, akademik takvimde belirtilen sürede enteraktif olarak kayıt yaptırmak şartıyla, bütünleme sınavına, DD ve DC aldığı derslerden de girebilir. Genel sınavda FF almış olan öğrencilerin bütünleme sınavına girebilmeleri için kayıt yaptırmaları gerekmez.Son ara sınav notu bütünleme sınavında da geçerlidir.	70.55
XLM-RoBERTa	Mazeret sınavı hangi öğrenciler için yapılır?	Madde 13 - Ders kayıt işlemleri, öğrencilere internet ortamında duyurulur. Öğrenciler derslere kayıtlarını, Üniversite Senatosunun tespit ettiği esaslar ve Rektörlük tarafından belirlenen usullere göre, bilgisayar ortamında interaktif olarak yenilerler. Kayıt yenileme işlemlerinin tümünden öğrenci sorumludur. Kaydın son şekli, danışman öğretim elemanı tarafından dönemin ilk iki haftası içerisinde onaylanır.	Madde 11 - Öğrencinin bir üst yarıyıl / yıldan ders alabilmesi için, önceki yarıyıl / yıllardan dersinin kalmaması ve GNO'sunun en az 3.00 olması gerekir. Bu durumda öğrenci, bulunduğu yarıyıl / yılın toplam kredisinin % 20'sini aşmayacak ölçüde üst yarıyıl / yıldan ders alabilir.	49.2
			Madde 12 - c) Ayrı ders niteliğindeki proje, laboratuvar, klinik, seminer ve benzeri derslerin yarıyıl / yıl içi değerlendirmeleri, ara sınav notlarının yerine geçebilir.	44.3
			Madde 22 - Bir dersin sınavı, o dersi vermekle sorumlu öğretim elemanlarınca düzenlenir ve yürütülür. Sorumlu öğretim elemanlarının sınav günü üniversitede bulunmaması halinde sınavı kimin yapacağı, ilgili birim başkanı tarafından belirlenerek dekanlığa/müdürlüğe bildirilir. Sınavların düzenli biçimde yürütülebilmesi için ilgili birimdeki bütün öğretim elemanları görevlendirilebilir.	40.0

Table 16. (Continue) Representative qualitative retrieval example

			Madde 6 - İkinci öğretimde yatay geçişler, 19.11.1992 tarihli ve 3843 sayılı "Yükseköğretim Kurumlarında İkili Öğretim Yapılması, 2547 Sayılı Yükseköğretim Kanununun Bazı Maddelerinin Değiştirilmesi ve Bu Kanuna Bir Madde Eklenmesi Hakkında Kanun" hükümlerine ve Yükseköğretim Kurulu tarafından belirlenen ilkelere göre yapılır	39.7
			Madde 8 - Yarıyıl/yılsonu sınavına girebilme şartını yerine getirmediği ders ve diğer eğitim-öğretim çalışmalarının toplam sayısı beşten fazla olan öğrenciler bu ek sınav hakkından yararlanamaz. Bu öğrencilerin Üniversite ile ilişikleri kesilir.	37.6
XLM-RoBERTa	Dersi geçebilmek için genel sınav notu en az kaç olmalıdır?	Madde 15 - Yarıyıl / yılsonu sınavında 100 üzerinden Meslek Yüksekokulları için en az 35, diğer tüm akademik birimler için 50 puan alma zorunluluğu vardır. Yarıyıl/yıl sonu sınavına girmeyen veya bu sınavdan en az yukarıda belirtilmiş olan puanı alamayan öğrenciler FF başarı notu ile değerlendirilir.	Madde 11 - Öğrencinin bir üst yarıyıl / yıldan ders alabilmesi için, önceki yarıyıl / yıllardan dersinin kalmaması ve GNO'sunun en az 3.00 olması gerekir. Bu durumda öğrenci, bulunduğu yarıyıl / yılın toplam kredisinin % 20'sini aşmayacak ölçüde üst yarıyıl / yıldan ders alabilir.	52.4
			Madde 12 -.c) Ayrı ders niteliğindeki proje, laboratuvar, klinik, seminer ve benzeri derslerin yarıyıl / yıl içi değerlendirmeleri, ara sınav notlarının yerine geçebilir.	47.1
			Madde 22 - Bir dersin sınavı, o dersi vermekle sorumlu öğretim elemanlarınca düzenlenir ve yürütülür. Sorumlu öğretim elemanlarının sınav günü üniversitede bulunmaması halinde sınavı kimin yapacağı, ilgili birim başkanı tarafından belirlenerek dekanlığa/müdürlüğe bildirilir. Sınavların düzenli biçimde yürütülebilmesi için ilgili birimdeki bütün öğretim elemanları görevlendirilebilir.	43.2
			Madde 8 - Yarıyıl/yılsonu sınavına girebilme şartını yerine getirmediği ders ve diğer eğitim-öğretim çalışmalarının toplam sayısı beşten fazla olan öğrenciler bu ek sınav hakkından yararlanamaz. Bu öğrencilerin Üniversite ile ilişikleri kesilir.	41.7
			Madde 6 - İkinci öğretimde yatay geçişler, 19.11.1992 tarihli ve 3843 sayılı "Yükseköğretim Kurumlarında İkili Öğretim Yapılması, 2547 Sayılı Yükseköğretim Kanununun Bazı Maddelerinin Değiştirilmesi ve Bu Kanuna Bir Madde Eklenmesi Hakkında Kanun" hükümlerine ve Yükseköğretim Kurulu tarafından belirlenen ilkelere göre yapılır.	41.6

Table 17. (Continue) Representative qualitative retrieval example

BM25+XLM-RoBERTa	Mazeret sınavı hangi öğrenciler için yapılır?	Madde 13 - Ders kayıt işlemleri, öğrencilere internet ortamında duyurulur. Öğrenciler derslere kayıtlarını, Üniversite Senatosunun tespit ettiği esaslar ve Rektörlük tarafından belirlenen usullere göre, bilgisayar ortamında interaktif olarak yenilerler. Kayıt yenileme işlemlerinin tümünden öğrenci sorumludur. Kaydın son şekli, danışman öğretim elemanı tarafından dönemin ilk iki haftası içerisinde onaylanır.	Madde 13 - Mazeret sınavı, yarıyıl / yılın son haftasında ilgili birimler tarafından yapılır.	79.51
			Madde 13 - Mazeret sınavı, ara sınav gününü kapsayan tarihlerde Rektörlük tarafından görevlendirilen veya ilgili birimin yönetim kurulunca kabul edilebilir bir mazereti nedeniyle ara sınava katılamayan öğrenciler için yapılır.	77.37
			Madde 13 - Mazeret sınav hakkı sadece ara sınavlar için verilir. Mazeret ara sınavı, ara sınav sayısına bakılmaksızın her yarıyıl / yılda, her ders için bir defa yapılır.	75.85
			Madde 18 - Son yarıyıl / yılın sonunda izledikleri programdaki bütün derslerden en az (DD) notu aldıkları halde, Genel Not Ortalamaları 2.00'ın altına düşen öğrenciler son 2 (iki) yılda; (DD), (DC) veya (CC) notu aldıkları en çok üç dersten mezuniyet "Genel Not Ortalaması Baraj Sınavı"na girerler. Bu sınav haklarını, takip eden yıl başlangıcındaki ders kayıt haftasında kullanırlar. Bu sınava giremeyenler için mazeret sınavı açılmaz	64.0
			Madde 22 - Bir dersin sınavı, o dersi vermekle sorumlu öğretim elemanlarınca düzenlenir ve yürütülür. Sorumlu öğretim elemanlarının sınav günü üniversitede bulunmaması halinde sınavı kimin yapacağı, ilgili birim başkanı tarafından belirlenerek dekanlığa/müdürlüğe bildirilir. Sınavların düzenli biçimde yürütülebilmesi için ilgili birimdeki bütün öğretim elemanları görevlendirilebilir.	65.49
BM25+XLM-RoBERTa	Dersi geçebilmek için genel sınav notu en az kaç olmalıdır?	Madde 15 - Yarıyıl / yılsonu sınavında 100 üzerinden Meslek Yüksekokulları için en az 35, diğer tüm akademik birimler için 50 puan alma zorunluluğu vardır. Yarıyıl/yıl sonu sınavına girmeyen veya bu sınavdan en az yukarıda belirtilmiş olan puanı alamayan öğrenciler FF başarı notu ile değerlendirilir.	Madde 18 - Son yarıyıl / yılın sonunda izledikleri programdaki bütün derslerden en az (DD) notu aldıkları halde, Genel Not Ortalamaları 2.00'ın altına düşen öğrenciler son 2 (iki) yılda; (DD), (DC) veya (CC) notu aldıkları en çok üç dersten mezuniyet "Genel Not Ortalaması Baraj Sınavı"na girerler. Bu sınav haklarını, takip eden yıl başlangıcındaki ders kayıt haftasında kullanırlar. Bu sınava giremeyenler için mazeret sınavı açılmaz.	87.54
			Madde 15 - Bütünleme sınavında meslek yüksekokulları için en az 35, diğer tüm akademik birimler için 50 puan alma zorunluluğu vardır. Bütünleme sınavında alınan not, o dersin yarıyıl/yılsonu notu yerine geçer.	75.37

Table 18. (Continue) Representative qualitative retrieval example

			Madde 25 - Harf başarı notlarına göre öğrencinin durumu aşağıdaki şekilde değerlendirilir: a) Bir dersten (AA), (BA), (BB), (CB) ve (CC) notlarından birini alan öğrenci o dersi başarmış sayılır. b) Bir dersten alınan (DC) ve (DD) notları, bu dersin "koşullu" olarak başarıldığını gösterir. c) (FF) notunu alan öğrenci, o dersi başaramamış sayılır ve aynı dersi verildiği ilk yarıyıl / yılda tekrarlar. d) Muaf Sayma (M); bir yükseköğretim kurumundan yatay geçişle kabul edilen veya Öğrenci Seçme Sınavına girerek yeniden kayıt hakkı kazanan öğrencilerin ilgili programa intibaklarında, önceden izledikleri yükseköğretim programında almış oldukları ve ilgili Yönetim Kurulu Kararı ile başarılı sayıldıkları dersler için kullanılır. e) Başarı notu belirtilmeyip (M) ve (B) olarak gösterilen dersler, genel not ortalamasının hesabında değerlendirmeye katılmaz. f) (K) ve (D) notu, genel not ortalaması hesabında (FF) notu işlemi görür. (K) ve (D) ile değerlendirilen dersler, verildiği ilk yarıyıl / yılda tekrar alınır.	72.39
			Madde 12 - Her ders için, bir yarıyılda/yılda en az bir ara sınav yapılır.	68.07
			Madde 14 - Yarıyıl / yılsonu sınavında 100 üzerinden Meslek Yüksekokulları için en az 35, diğer tüm akademik birimler için 50 puan alma zorunluluğu vardır. Yarıyıl/yıl sonu sınavına girmeyen veya bu sınavdan en az yukarıda belirtilmiş olan puanı alamayan öğrenciler FF başarı notu ile değerlendirilir.	66.91

The hybrid configurations, particularly BM25 + BERTurk, consistently rank the most relevant provisions at the top of the list, exhibiting both high Top-1 accuracy and high MRR. For the examined queries, the hybrid model assigns very high scores to the correct Article 13/15 variants (e.g., 95.25%, 93.01%), thereby corroborating the quantitative advantage reported in Section 3.1. In contrast, the dense-only models (BERTurk and XLM-RoBERTa) often retrieve passages that are thematically or procedurally related—for example, registration procedures or general exam regulations—but they sometimes fail to prioritize the exact statutory clause that directly answers the query. This behaviour is reflected in lower scores and more variable article predictions for XLM-RoBERTa, whose top candidates occasionally include unrelated articles with substantially lower scores. The hybrid pipelines benefit from BM25's robust lexical anchoring, which surfaces candidates containing salient keywords (e.g., "mazeret", "ara sınav"); transformer-based dense re-scoring then discriminates among these candidates and improves top-rank precision and MRR. Typical error modes include paraphrase/granularity mismatch—where semantic similarity matches a related but differently scoped clause (e.g., registration versus exam detail)—and topical drift in the multilingual model, possibly due to subword tokenization mismatches or weaker Turkish-specific modeling. Finally, many top predictions consist of near-duplicate phrasings of the same article (e.g., variants of Article 13); applying post-processing such as deduplication or article-level aggregation (for example, merging passage scores by Article Number before final ranking) could further improve the user-facing output.

4. References

[1] D. Sultania *et al.*, "Domain-specific Question Answering with Hybrid Search," Dec. 21, 2024, *arXiv*: arXiv:2412.03736. doi: 10.48550/arXiv.2412.03736.

[2] M. Çetiner, A. Yıldırım, C. Öksüz, and B. Onay, "Mevzuat Verisetinde Soru Cevaplama Uygulamasi Question Answering Application on Legalisation Dataset," in *2021 6th International Conference on Computer Science and Engineering (UBMK)*, Sept. 2021, pp. 603–607. doi: 10.1109/UBMK52708.2021.9558981.

[3] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," May 24, 2019, *arXiv*: arXiv:1810.04805. doi: 10.48550/arXiv.1810.04805.

[4] J. Rayo, R. de la Rosa, and M. Garrido, "A Hybrid Approach to Information Retrieval and Answer Generation for Regulatory Texts," Feb. 24, 2025, *arXiv*: arXiv:2502.16767. doi: 10.48550/arXiv.2502.16767.

[5] N. Wiratunga *et al.*, "CBR-RAG: Case-Based Reasoning for Retrieval Augmented Generation in LLMs for Legal Question Answering," Apr. 04, 2024, *arXiv*: arXiv:2404.04302. doi: 10.48550/arXiv.2404.04302.

[6] M. Esposito, E. Damiano, A. Minutolo, G. De Pietro, and H. Fujita, "Hybrid query expansion using lexical resources and word embeddings for sentence retrieval in question answering," *Information Sciences*, vol. 514, pp. 88–105, Apr. 2020, doi: 10.1016/j.ins.2019.12.002.

[7] A. Agarwal, G. S, and C. Devaguptapu, "Hybrid Graphs for Table-and-Text based Question Answering using LLMs," Jan. 29, 2025, *arXiv*: arXiv:2501.17767. doi: 10.48550/arXiv.2501.17767.

[8] D. Min *et al.*, "Exploring the Impact of Table-to-Text Methods on Augmenting LLM-based Question Answering with Domain Hybrid Data," Apr. 09, 2024, *arXiv*: arXiv:2402.12869. doi: 10.48550/arXiv.2402.12869.

[9] C. Zhang and Q. Chen, "HD-RAG: Retrieval-Augmented Generation for Hybrid Documents Containing Text and Hierarchical Tables," Apr. 13, 2025, *arXiv*: arXiv:2504.09554. doi: 10.48550/arXiv.2504.09554.

[10] G. Verma *et al.*, "HybridContextQA: A Hybrid Approach for Complex Question Answering using Knowledge Graph Construction and Context Retrieval with LLMs".

[11] N. N. Doan, A. Härmä, R. Celebi, and V. Gottardo, "A Hybrid Retrieval Approach for Advancing Retrieval-Augmented Generation Systems," in *Proceedings of the 7th International Conference on Natural Language and Speech Processing (ICNLSP 2024)*, M. Abbas and A. A. Freihat, Eds., Trento: Association for Computational Linguistics, Oct. 2024, pp. 397–409. Accessed: Aug. 12, 2025. [Online]. Available: https://aclanthology.org/2024.icnlsp-1.41/

[12] Y. Tachioka, "Question Answer Summary Generation from Unstructured Texts by Using LLMs," pp. 261–268, 2024, doi: 10.1007/978-981-96-0914-7_20.

[13] S.-H. Xiong, Z.-H. Wang, Z.-S. Chen, G. Li, and H. Zhang, "Text classification of public online messages in civil aviation: A N-BM25 weighted word vectors method," *Information Sciences*, vol. 704, p. 121956, June 2025, doi: 10.1016/j.ins.2025.121956.

[14] C. R. Pitoyo, S. Sulistyo, and I. Hidayah, "Comparison of Cosine Similarity and BM25 in Book Procurement Prioritization Based on Multi-Source Data Integration," in *2025 International Conference on Computer Sciences, Engineering, and Technology Innovation (ICoCSETI)*, Jan. 2025, pp. 90–94. doi: 10.1109/ICoCSETI63724.2025.11019421.

[15] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," Aug. 27, 2019, *arXiv*: arXiv:1908.10084. doi: 10.48550/arXiv.1908.10084.

[16] V. Karpukhin *et al.*, "Dense Passage Retrieval for Open-Domain Question Answering," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, B. Webber, T. Cohn, Y. He, and Y. Liu, Eds., Online: Association for Computational Linguistics, Nov. 2020, pp. 6769–6781. doi: 10.18653/v1/2020.emnlp-main.550.

Mechanical Stirrer-Assisted Production of $FeCl_3$ from by Using HCl and Mill Scale

Hasan NİZAMOĞLU[1] , M. Deniz TURAN[2] , Zeynel Abidin SARI[3]

[1,2]Department of Metallurgy and Materials Engineering, Fırat University, 23119 Elaziğ, Turkey
[3]Department of Metallurgy-Iskenderun Vocational School of Higher Education, Iskenderun Technical University, Hatay 31200, Turkey
[1]hsnizamoglu@gmail.com, [2]mdturan@firat.edu.tr, [3]zeynel.sari@iste.edu.tr
[1](ORCID: 0000-0002-1801-7156), [2](ORCID: 0000-0002-2136-1425), [3](ORCID: 0000-0001-5932-2141)

Abstract
During the production process in foundries and rolling mills, which hold a significant position in the metallurgical industry, the high temperature and oxidizing environment cause the formation of mill scale on the surfaces of hot slabs and billets. This mill scale, primarily composed of iron oxides, is considered a substantial metallurgical waste, yet it cannot be economically utilized due to its low metal content. The evaluation of mill scale in the iron and steel sector is a common issue worldwide, engaging researchers frequently. Mill scale, which is considered as waste, was used to produce high-value and oxidizing iron (III) chloride solution. In this study a mechanical stirrer which is a powerful mixing system, was used to dissolve iron from mill scale and produce $FeCl_3$ chloride. The type of propeller used here is a double-bladed propeller made of teflon, allowing the magnetic scale to be mixed more effectively. The optimum leaching conditions were determined: 60 min, leaching time, 105 °C, leaching temperature, 7 M, HCl concentration, 1/10, solid-liquid ratio, 400 rpm mechanical stirrer, -74 µm particle size of mill scale. Under these leaching conditions, 100 % Fe dissolution, 55.05 g/L of Fe^{3+} concentration and 5.03 of Fe^{3+}/Fe^{2+} ratio was determined.

Keywords: hydrometallurgy, mill scale, hydrochloric acid, metallurgical waste

1. Introduction

In foundries and continuous casting facilities, during the annealing process of steel slabs and billets coming out of the reheating furnaces, an iron oxide layer forms due to oxidation, known as mill scale. Mill scale is considered a significant waste due to its high iron content in the casting and rolling sections of steel production plants. Although the exact amount is not known in our country's foundries, it is estimated that mill scale accounts for about 4-5% of the produced steel. Layers formed as oxidation products on the metal surface at high temperatures are referred to as mill scale. There are three types of iron oxide layers present on the steel surface at high temperatures [1,2].

The simplest form of mill scale is iron oxide (FeO). When the surface temperature exceeds 900 °C during heating, mill scale formation accelerates. The mill scale that forms is predominantly 80% wustite, 18% magnetite, and 2% α-hematite. It has been observed that mill scale thickness is thinner with rapid heating, whereas it increases with longer holding times [3,4]. Considering the ability of iron within mill scale to be converted into forms highly capable of oxidation in solution environments (Fe^{3+}= +0.771V / Fe^{2+}= -0.440V), it suggests potential use as an oxidizing leaching reagent in hydrometallurgical extraction from sulfur-bearing ores and concentrates. Therefore, starting from the premise of this study, an iron (III) chloride ($FeCl_3$) solution was produced from mill scale, which has high oxidation potential.
On the other hand, hydrometallurgical metal extraction techniques come into prominence because of it has some advantages. Particularly, there is a need for an oxidative leaching reactant during leaching of sulfurous ore/concentrates. Iron ions are known one of economical oxidants [5]. Both Fe^{3+} and Fe^{2+} are easily soluble in chlorinated solution. Iron (III) chloride is widely used in the leaching of sulphide ores and concentrates, so that it has been applied to chalcopyrite, sphalerite and galena concentrates by researchers [6.7].

In this study, it was aimed to produce $FeCl_3$ chloride solution that can be used as an oxidizing leaching reactant from scale by using HCl and a mechanical stirrer.

2. Material and Methods

Mill scale, obtained approximately 15 kg from the stock area and stored, is a mill scale obtained by removing the layer accumulated on the surface of billets during slab casting and transporting it to the waste area with pressurized water at İskenderun Iron and Steel factory. Samples of layered mill scale were first washed several times in the laboratory with

[1]Corresponding author

water to remove the oil on the surface. Subsequently, the mill scale samples obtained for obtaining material in the -200 mesh fraction for experiments were subjected to size reduction in a ring mill. The ground mill scale was passed through suitable porous sieves and kept in sealed containers for use in experiments

In the characterization studies performed on the samples, the it was weighed in sufficient quantities for the determination of chemical content were analyzed in ACME laboratories by AQ-274 method using ICP (Inductively Coupled Plasma) method (Table 1).

Table 1. Chemical compounds of mill scale.

Element	**Fe**	**Ca**	Na	**Cu**	**Ni**	**Mn**	**Cr**
%	68	0.13	0.02	0.05	0.03	0.64	0.03

As a result of the dissolution process, it is observed that the chemical composition of mill scale is predominantly composed of iron (approximately 68% Fe). On the other hand, the structural analysis of mill scale samples was conducted using the powder diffraction method based on XRD (Bruker / D8 Advance) at the Central Laboratory of Firat University (Figure 1). According to the XRD analysis of mill scale, the structure is predominantly composed of wustite (FeO), hematite (Fe_2O_3), and magnetite (Fe_3O_4). Based on the differences obtained in structural analyses, it is observed that the iron in mill scale inherently possesses different oxidation states, which could serve as a source for iron chloride formation in the presence of HCl acid.

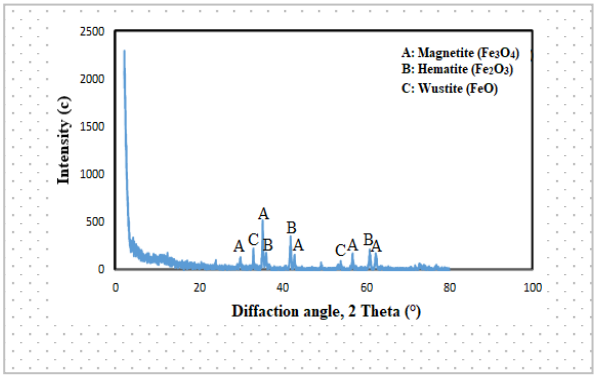


Figure 1. XRD pattern of Mill Scale.

The effect of various parameters was investigated in presence of HCl with mechanical stirrer. It is clear that it is a necessity to work with the mixing system that will not be affected by the magnetic properties of iron in the works to be carried out at this stage since the scale structure is mostly composed of iron [8]. All experiments were conducted using a double-bladed mechanical stirrer made from teflon. The mechanical stirring speed was maintained at 400 rpm. The experiments were conducted using a mantle heater and 250 ml Erlenmeyer flasks. A reflux cooling system was used for experiments at high temperatures. In all experiments, mechanical stirrer speed was kept constant as 400 rpm.

3. Results and Discussion

In the leaching studies conducted, the solubility of iron from mill scale was examined under the influence of various leaching parameters. Treatment of mill scale with HCl was conducted using a mechanical stirring system. According to the experimental results, it was determined that the mechanical stirring system yielded the best results at high temperature and high HCl concentration. For the conditions required to obtain the leaching solution in the second stage of leaching studies, it is clear that determining conditions with high Fe^{3+} concentration and a high Fe^{3+}/Fe^{2+} ratio in the solution is crucial.

The effectiveness of Fe^{3+} concentration at different temperatures was examined in Figure 2. It is observed that the concentration of Fe^{3+} increased into the solution with increasing temperature. Also, Fe^{3+} concentration increased with increasing leaching time, however, it is not very effective at low temperatures. Iron extraction efficiency increased in the presence of different leaching temperature and leaching time (Figure 3).

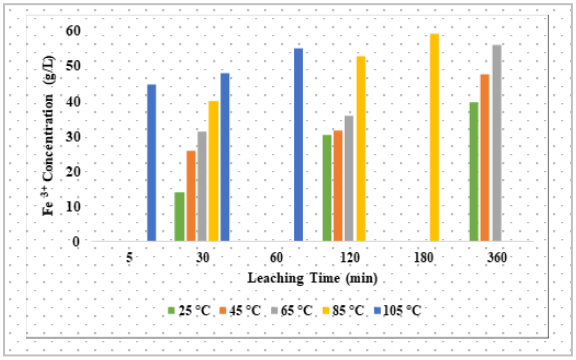


Figure 2. Effect of leaching temperature and leaching time on iron concentration under optimum conditions in the presence HCl of mill scale.

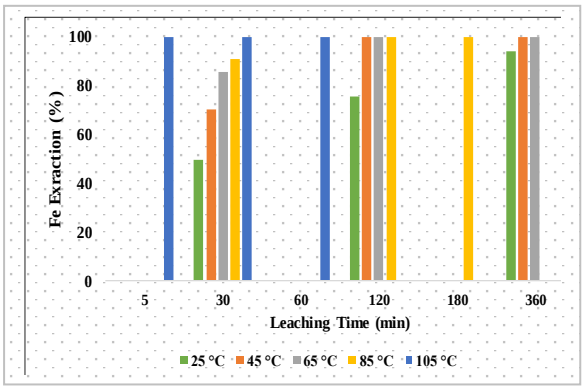


Figure 3. Effect of leaching temperature and leaching time on iron extraction under optimum conditions in the presence HCl of mill scale.

According to the results obtained from experimental studies, limitations on leaching times were imposed in some experimental groups due to excessive evaporation in leach solutions at high temperatures (85°C and 105°C). However, the reduction in leaching times with increasing temperature was not perceived as detrimental. This is because it was observed that even at low durations, all mill scale dissolved completely, and the Fe^{3+} concentration and Fe^{3+}/Fe^{2+} ratio increased.
In Figure 4, it is seen that Fe^{3+}/ Fe^{2+} ratio is increasing with leaching temperature and leaching time. This study demonstrates that effective production of $FeCl_3$ from mill scale can be achieved at high temperatures and reduced durations. The experiments were conducted as follows: at 105°C, experiments were conducted for 5, 30, and 60 minutes. At 85°C, experiments were conducted for 30, 120, and 180 minutes. The study was conducted in this manner due to the reasons mentioned above.

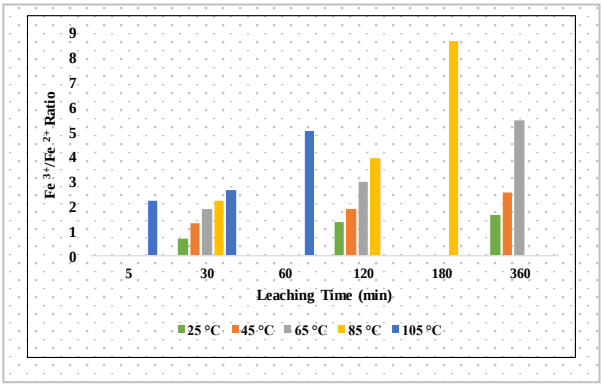


Figure 4. Effect of leaching temperature and leaching time on Fe^{3+}/ Fe^{2+} ratio under optimum conditions in the presence HCl of mill scale.

4. Conclusions

In this study, the conditions of iron chloride formation that can be used as an oxidizing leaching reactant were investigated in presence of HCl by mechanical stirring. Various experimental parameters were optimized for purpose of obtain to maximum iron dissolution and maximum Fe^{3+}/Fe^{2+} ratio in leaching solution. Range of investigating parameters is leaching temperature (25-105 °C), leaching time (5-360 min), HCl concentration (0.5-7 M), solid-liquid ratio (1/3-1/50). The optimum leaching conditions were determined: 60 min, leaching time, 105 °C, leaching temperature, 7 M, HCl concentration, 1/10, solid-liquid ratio, 400 rpm mechanical stirrer, -74 µm particle size of mill scale. Under these leaching conditions, 100 % Fe dissolution, 55.05 g/L of Fe^{3+} concentration and 5.03 of Fe^{3+}/Fe^{2+} ratio was determined.

Acknowledgement

The authors would like to acknowledge the Research Foundation of Fırat University under project no FUBAP-MF.18.32 for supporting the study.

5. References

[1] The Manufacture of Hot Strip Mill Products “Oxside removal”. Chapter 33, page 1083-1096.
[2] Legodi M. A, De Waal D. “The preparation of magnetite, goethite, hematite and maghemite of pigment quality from mill scale iron waste”. Days and Pıgments, 78, 177, 2008.
[3] Ismayadi I, Abdul Halim A, Azdiya Suhada Abdul A, Idza Riati I, Farah Nabilah S, Raba’ah Syahidah A. “Phase, morphological, and magnetic properties of iron oxide nanoparticles extracted from mill scale waste and its surface modification with CTAB surfactant”. Journal of the Australian Ceramic Society, 56:729–743, 2020.
[4] Nur Asyıkın Ahmad N, Raba’ah Syahıdah A, Muhammad Syazwan M, Abdul Halım S, Ismayadı I, Hasfalına Che M, Norlaıly Mohd S, Nor Hapıshah A. “Magnetite Nanoparticles (MNPs) Used as Cadmium Metal Removal from the Aqueous Solution from Mill Scales Waste Sources”. Sains Malaysiana, 49(4), 847-858, 2020.
[5] Dutrizac J. E. “The leaching of sulphide minerals in chloride media”. Hydrometallurgy, 29 (1-3), 1-45, 1992.
[6] Havlik T, Scrobian M, Balaz P, and Kammel R. “Leaching of chalcopyrite concentrate with ferric chloride”. International Journal of Mineral Processing, 43, 61-72, 1995.
[7] Mulak W, and Wawrzak D. “Ferric chloride leaching of lead sulphide concentrate”. 5th International Mineral Processing Symposium, Progress in Mineral Processing Technology, 1994.
[8] Turan M. D, Altundogan S. H. “Leaching of copper from chalcopyrite concentrate by using ammonium persulphate in an autoclave: Determination of most suitable impeller type by using response surface methodology”. Central South University Press and Springer-Verlag Berlin Heidelberg, 20, 622−628, 2013.

A Hybrid Quantum Federated Learning Approach for Image Classification

Niyazi Furkan BAR[1] , Mehmet KARAKÖSE [2]

[1,2]Department of Computer Engineering, Firat University, Elazig, Turkey.
[1]nfbar@firat.edu.tr, [2]mkarakose@firat.edu.tr
[1](ORCID: 0000-0002-3393-004X), [2](ORCID: 0000-0002-3276-3788)

Abstract
Data security and privacy have become critical issues in deep learning applications. Federated learning is a widely used approach to address this critical issue. However, federated learning has limitations such as computational complexity and the processing of high-dimensional data. Therefore, a quantum federated learning approach, which combines the potential computational capacity and power offered by quantum computing with federated learning, is introduced. In this study, a hybrid quantum federated learning approach is proposed for image classification. The proposed approach includes a classical server and classical-quantum hybrid clients. The hybrid clients contain a pretrained model and perform feature extraction from this model. Furthermore, the performance of the approach is maximized by optimizing the parameters of the variational quantum circuit on the quantum side, as well as the parameters of the layer where the pretrained model performs feature extraction. The proposed approach was tested in a real-world simulation. These test results are presented graphically and compared with similar studies in the literature. The results validate the performance of the proposed approach.

Keywords: quantum computing, quantum federated learning, quantum neural network, image classification

1. Introduction

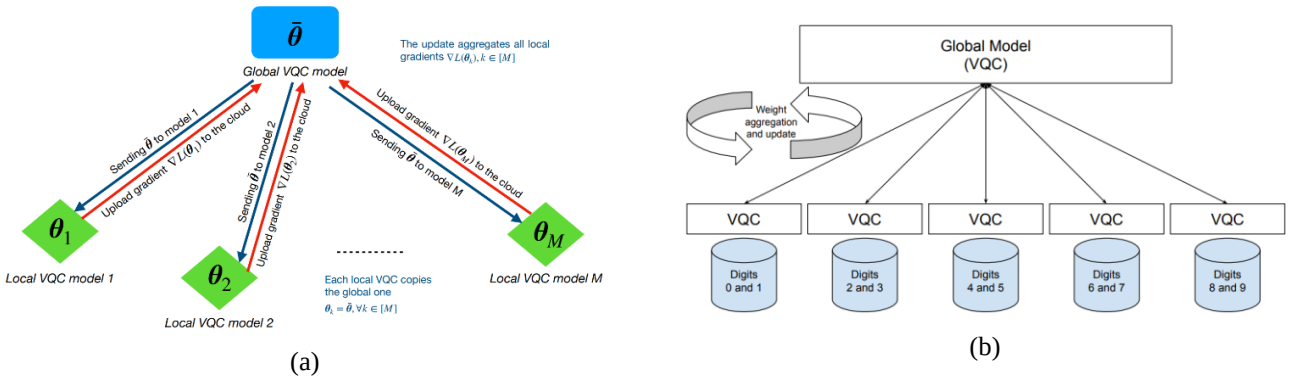
Today, with the increase in data sources and computational power, image classification in deep learning has become quite common. However, the increasing size of data requires high processing power to process these data. Therefore, using the central processing unit (CPU) alone is insufficient. In addition to the CPU, other classical computing units such as graphics processing units (GPU), tensor processing units (TPU), and neural processing units (NPU) are widely used [1]. The concept of quantum computing, first introduced in the 1980s [2,3], has become increasingly exploitable over the past decade. With this potential, many applications on classical computers have begun to be implemented on quantum computing. In fact, topics such as machine learning, artificial neural networks, and deep learning are being studied on quantum computing. Quantum neural networks (QNN), a neural network application implemented with quantum computing, are used in many areas such as classification [4], navigation [5], and healthcare [6].

In addition to these developments, security and privacy issues regarding data processing in the increasingly digital world continue to gain importance. This focus on privacy and security has emerged due to user demands, the enactment of legal regulations in many countries, such as the General Data Protection Regulation (GDPR) in the European Union and the Personal Data Protection Law (KVKK) in Turkey, and the reluctance of companies to share their data [7,8]. In response to this demand, a method called federated learning (FL) was introduced by Google in 2016 [9]. Federated learning, unlike classical learning, allows training by acquiring data from distributed edge devices instead of a single data source [10]. This architecture offers advantages in privacy and security, while its distributed structure offers advantages in hardware and time costs. While currently applied in many areas, it has limitations in complex models, such as model security, computational complexity, and processing high-dimensional data. Therefore, quantum computing, a combination of quantum computing and quantum federated learning (QFL), has emerged.

In this study, a hybrid QFL approach is proposed that ensures data security and privacy. The proposed approach features a classical server and quantum-classical hybrid clients. The proposed approach features a classical server and quantum-classical hybrid clients. The contribution of this approach is that feature extraction is performed on hybrid clients using a pre-trained model and the layer from which the features are extracted is optimized using variational quantum circuit (VQC) parameters. The optimized parameters are combined using a weighted average method on the classical server and distributed to the other clients. The proposed approach was tested in a simulation environment to ensure its suitability for real-world conditions.

[1]Corresponding author

There are many studies in the literature using QFL. Among these, studies on image classification show that Sünkel et al. [11] used QNN on classical federated learning neural network data and tested this QFL method on the MNIST dataset. Qi et al. [12] proposed a federated quantum natural gradient descent (FQNGD) algorithm and tested it on the MNIST dataset. Song et al. [13] proposed an architecture that uses classical clients instead of quantum clients and tested their proposed approach using the MNIST dataset. The basic block diagrams of the approaches proposed by Qi et al. [12] and Sünkel et al. [11] are given in Figure 1. It is seen that these methods similarly consist of a central quantum server and quantum clients.



(a) (b)

Figure 9. Block diagrams of similar QFL studies in the literature, (a) Block diagram of the method in Qi et al. [12], (b) Block diagram of the method in Sünkel et al. [11].

The contributions of this study are as follows:

- Significant performance improvement compared to standard QFL.
- A classical server, quantum-classical hybrid client architecture is proposed.
- The layer where features are taken from the pretrained model has been optimized.
- A federation strategy based on random client selection was implemented, where only a limited number of clients participated in each training round.

2. Quantum Federated Learning

Quantum computing is a computing paradigm implemented using quantum mechanics. Quantum mechanics offers low computational cost, high computational power, and capacity. The superposition, entanglement, and parallelism properties of quantum mechanics highlight these advantages. The fundamental processing unit in quantum computing is called a quantum bit (qubit). Unlike bits in classical computing, qubits can have values not only of 0 or 1 but also of any value within this range. This is as given in (1).

$$|\psi\rangle = \begin{bmatrix} c_0 \\ c_1 \end{bmatrix}, \ |0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
$$|\psi\rangle = c_0.|0\rangle + c_1.|1\rangle \ , \ where \ {c_0}^2 + {c_1}^2 = 1 \quad (1)$$

Here ${c_0}^2$ represents the probability of the qubit being 0, and ${c_1}^2$ represents the probability of it being 1.

Quantum Neural Networks (QNNs) are artificial neural network applications implemented using quantum computation. Variational quantum circuits (VQCs), also known as parameterized quantum circuits (PQCs), are commonly used in quantum neural network applications. These circuits are a hybrid classical-quantum approach, where the parameters of the quantum gates in the circuit are optimized by a classical computer.

Federated Learning (FL) is a secure distributed learning framework where distributed clients collaborate to create a virtual model to solve a problem without sharing raw data. The virtual model is a global model used to collect data from all clients and serve other clients. FL is similar to the classical centralized training model in that it receives parameters from clients and sends them back to them. Quantum federated learning (QFL) uses QNN to ensure data privacy. QFL can be a system

comprised entirely of quantum servers and clients, while classical systems comprised of servers and quantum clients are also available. In both systems, each client performs its own training, combining its parameters with global parameters as in classical FL.

3. The Proposed Approach

In this study, we propose a classical-quantum hybrid quantum federated learning approach that uses a pretrained model to extract features for images requiring data privacy and security. The proposed approach employs a classical server and multiple classical-quantum clients, in accordance with QFL. The classical server side essentially consists of three operations: sending parameters to local clients, collecting parameters from local clients, and combining incoming parameters. The classical-quantum client side consists of the classical side: receiving the raw input image, preprocessing it, extracting features from the pretrained model, and optimizing VQC parameters. The quantum side includes a VQC circuit, which includes data encoding, parameterized quantum gates, and measurement operations. The architecture of the proposed approach is shown in Figure 2, and the internal structure of a quantum-classical hybrid client is shown in Figure 3.

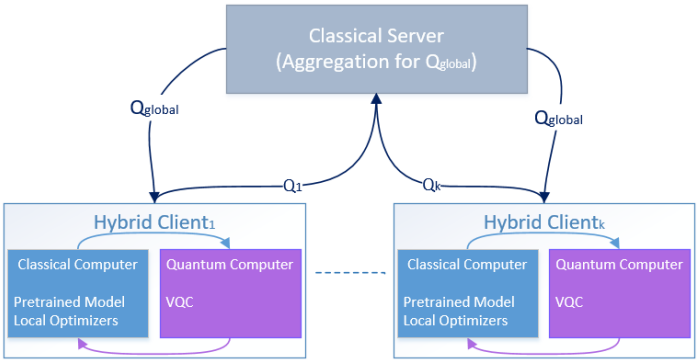


Figure 10. Block diagram of the architecture in the proposed approach

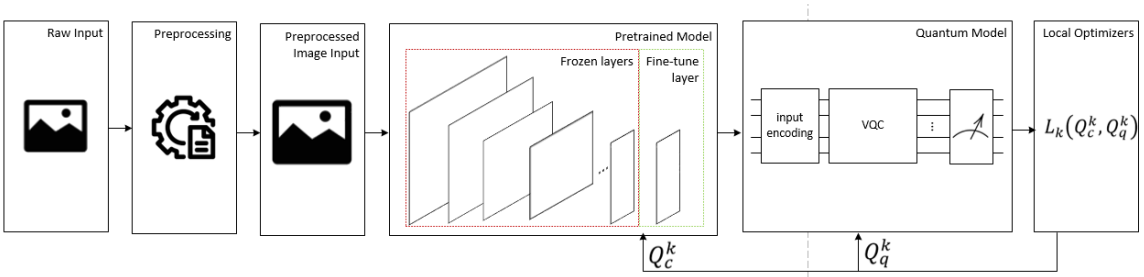


Figure 11. Internal structure of the hybrid client in the proposed approach

The general steps of the proposed approach are as follows:

- The server distributes the parameters of the global models to the clients.
- Hybrid clients send raw data to the image pre-processing step.
- Hybrid clients perform feature extraction by feeding the processed image to the pretrained model.
- It sends the features extracted from the pretrained model to the quantum computing model, where it performs the classification process.
- The results from the quantum computation model are optimized in the local optimizer, where both the parameters for the fine-tuned layer of the pretrained model and the parameters of the VQC are optimized.
- In the specified number of rounds, local parameters are sent, combined, global parameters are updated and distributed back to the clients.

In classical-quantum hybrid architectures, pretrained models are commonly kept fixed and only VQC is trained. In the proposed approach, the final layers of the pretrained model are simultaneously optimized along with VQC. To this end, the initial layers of the pretrained model are fixed; only the last n fully connected layers are updated with VQC parameters on the client side using gradient-based methods. This aims to increase the compatibility between classical feature extraction and quantum classification and improve the overall performance of the hybrid architecture. The total loss function optimized on the client k is as given in (2).

$$L_k(Q_c^k, Q_q^k) = L_{CE}\left(y_k, f_q\left(f_c(x_k, Q_c^k), Q_q^k\right)\right) \quad (2)$$

4. Simulation Results

To verify its performance, the proposed approach was simulated using Python and Tensorflow on a computer with Windows 11, an AMD Ryzen 9 7940HS processor, and 64 GB of RAM. The proposed approach was simulated in three rounds on a system with four clients using the MNIST dataset, which has a total of 10 classes. To ensure heterogeneity in real-world scenarios, clients were randomly selected at 50% in each round. The parameters used in the simulation are given in Table 1.

Table 19. Parameters

Parameter Name	Parameter Value
Pre-trained model (for feature extraction)	VGG16
Accuarcy Metrik	Sparse Categorical Accuracy
Loss Function	Sparse Categorical Crossentropy
Optimizer	Adam
Learning rate	0.001
Number of epochs	10
Number of clients	4
Number of rounds	3
The ratio of the number of clients to be trained in each round	50%

The simulation was performed using a classical federated learning algorithm, standard quantum federated learning, and the quantum-classical hybrid federated learning algorithm in the proposed approach. The results obtained from the proposed approach are presented in the graphs. The accuracy rate and loss value of the clients in each round are shown in the graphs.

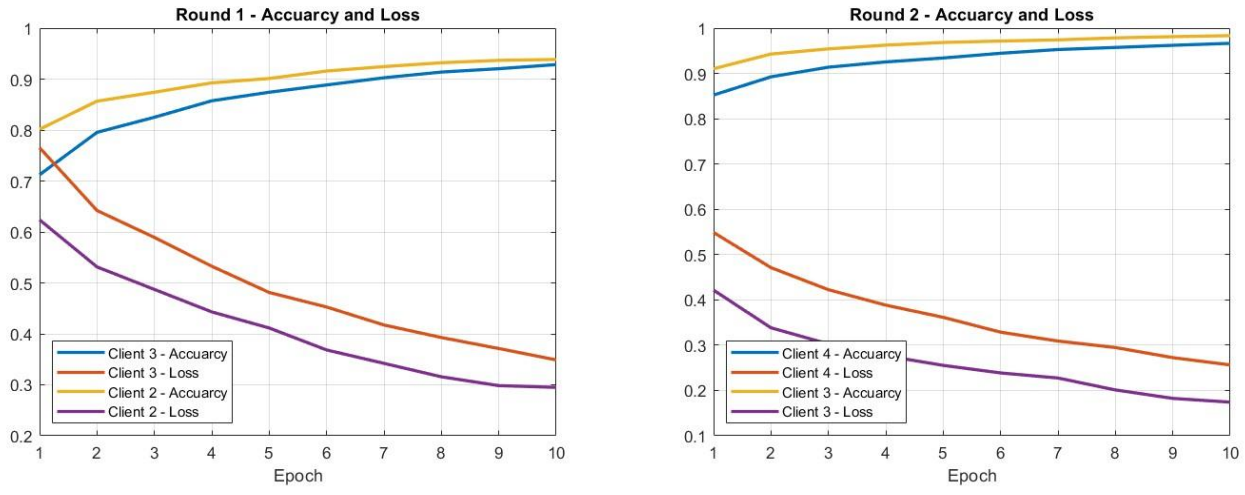
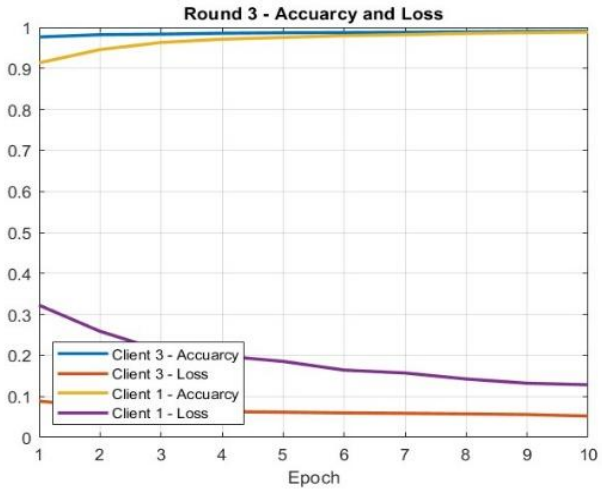


Figure 12. Accuracy and loss values obtained during training



(c)

Figure 13. (Continue) Accuracy and loss values obtained during training

The selected clients in each round during training and the highest accuracy rates achieved by these clients are given in Table 2.

Table 20. Accuracy rates achieved by clients during training

		Standart Quantum FL	The Proposed Approach
Round 1	**Client 3**	66.03%	92.85%
	Client 2	66.90%	93.85%
Round 2	**Client 4**	68.11%	96.65%
	Client 3	70.24%	98.34%
Round 3	**Client 3**	69.57%	99.12%
	Client 1	68.29%	98.85%
Average Accuarcy Rate		68.19%	96.61%

As seen in Table 2, the accuracy rate increases significantly as the number of rounds increases for both the standard QFL and the proposed approach. Specifically, Client 3, selected in each round, consistently achieved the highest accuracy rates in both the classical approach, the standard QFL, and the proposed approach. Comparing the accuracy rates and average accuracy rates between the methods, the performance achieved by our proposed approach is highly acceptable. This finding clearly demonstrates the potential of a quantum-classical hybrid system to solve complex problems with high accuracy despite the limitations of existing quantum hardware, thanks to the adaptive optimization of the classical feature extraction layers.

The accuracy rates of quantum federated learning approaches implemented with the same data sets found in the literature are given in Table 3.

Table 3. Comparison of QFL studies in the literature in terms of dataset, number of classes, and accuracy rate

Reference	Dataset	Number of Classes	Accuarcy Rate
Sünkel et al. [11]	MNIST	10	51.8%
Yun et al.[14]	Mini-MNIST	4	77%
Zhang et al.[15]	MNIST-3	3	70%
Zhao et al.[16]	MNIST-2	2	92.7%
	Fashion-MNIST	10	75.4%
Standart QFL	MNIST	10	68.19%
The Proposed Approach	MNIST	10	96.61%

5. Conclusion

This work proposes a quantum computing-based hybrid quantum federated learning approach for image classification applications requiring data privacy and security. This proposed approach integrates a classical pretrained model into a quantum-classical hybrid architecture to significantly increase the performance and speed of the QFL process. The proposed approach is designed to capture the heterogeneity of real-world scenarios. In these simulation results, the standard QFL approach achieves an average accuracy of 68.19%, while our proposed approach achieves an average accuracy of 96.61%. This finding clearly demonstrates the potential applicability of our proposed approach to similar problems and provides a robust quantum solution for federated learning in real-world scenarios, despite the current limitations of today's NISQ quantum computers.

6. References

[1] A. Chalumuri, R. Kune, and B. S. Manoj, “A hybrid classical-quantum approach for multi-class classification,” Quantum Information Processing, vol. 20, no. 3. Springer Science and Business Media LLC, Mar. 2021. doi: 10.1007/s11128-021-03029-9.

[2] G.-L. Chen, X.-H. Song, S. E. Venegas-Andraca, and A. A. Abd ElLatif, “QIRHSI: novel quantum image representation based on HSI color space model,” Quantum Information Processing, vol. 21, no. 1. Springer Science and Business Media LLC, Dec. 23, 2021. doi: 10.1007/s11128-021-03337-0.

[3] N. F. Bar and M. Karakose, "Quantum Circuit Dataset Generator Approach for Deep Learning based Solutions," 2024 28th International Conference on Information Technology (IT), Zabljak, Montenegro, 2024, pp. 1-4, doi: 10.1109/IT61232.2024.10475754.

[4] N. F. Bar, H. Yetis and M. Karakose, "A Quantum-Classical Hybrid Classifier Using Multi-Encoding Method for Images," 2023 27th International Conference on Information Technology (IT), Zabljak, Montenegro, 2023, pp. 1-4, doi: 10.1109/IT57431.2023.10078617.

[5] N. F. Bar, H. Yetis, and M. Karakose, “An efficient and scalable variational quantum circuits approach for deep reinforcement learning,” Quantum Inf Process, vol. 22, no. 8, Aug. 2023, doi: 10.1007/s11128-023-04051-9.

[6] N. F. Bar, M. Yenilmez, S. Aksu and M. Karakose, "A Quantum Computing based Approach for Sentiment Analysis in Bilateral Conversations," 2024 28th International Conference on Information Technology (IT), Zabljak, Montenegro, 2024, pp. 1-4, doi: 10.1109/IT61232.2024.10475718.

[7] C. Qiao, M. Li, Y. Liu, and Z. Tian, “Transitioning From Federated Learning to Quantum Federated Learning in Internet of Things: A Comprehensive Survey,” IEEE Commun. Surv. Tutorials, vol. 27, no. 1, pp. 509–545, Feb. 2025, doi: 10.1109/comst.2024.3399612.

[8] A. Gupta, M. Kumar Maurya, K. Dhere, and V. Kumar Chaurasiya, “Privacy-Preserving Hybrid Federated Learning Framework for Mental Healthcare Applications: Clustered and Quantum Approaches,” IEEE Access, vol. 12, pp. 145054–145068, 2024, doi: 10.1109/access.2024.3464240.

[9] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, “Communication-Efficient Learning of Deep Networks from Decentralized Data,” arXiv, 2016, doi: 10.48550/ARXIV.1602.05629.

[10] J. Wen, Z. Zhang, Y. Lan, Z. Cui, J. Cai, and W. Zhang, “A survey on federated learning: challenges and applications,” Int. J. Mach. Learn. & Cyber., vol. 14, no. 2, pp. 513–535, Nov. 2022, doi: 10.1007/s13042-022-01647-y.

[11] L. Sünkel, P. Altmann, M. Kölle, and T. Gabor, “Quantum Federated Learning for Image Classification,” Proceedings of the 16th International Conference on Agents and Artificial Intelligence. SCITEPRESS - Science and Technology Publications, pp. 936–942, 2024. doi: 10.5220/0012421200003636.

[12] J. Qi and M.-H. Hsieh, “Federated quantum natural gradient descent for quantum federated learning,” Federated Learning. Elsevier, pp. 329–341, 2024. doi: 10.1016/b978-0-44-319037-7.00028-4.

[13] Y. Song et al., “A quantum federated learning framework for classical clients,” Sci. China Phys. Mech. Astron., vol. 67, no. 5, Mar. 2024, doi: 10.1007/s11433-023-2337-2.

[14] W. J. Yun, J. P. Kim, S. Jung, J. Park, M. Bennis, and J. Kim, “Slimmable Quantum Federated Learning,” 2022, arXiv. doi: 10.48550/ARXIV.2207.10221.

[15] Y. Zhang, C. Zhang, C. Zhang, L. Fan, B. Zeng, and Q. Yang, “Federated Learning with Quantum Secure Aggregation,” 2022, arXiv. doi: 10.48550/ARXIV.2207.07444.

[16] H. Zhao, “Non-IID quantum federated learning with one-shot communication complexity,” Quantum Mach. Intell., vol. 5, no. 1, Jan. 2023, doi: 10.1007/s42484-022-00091-z.

Electric Field Estimation in High Voltage Ceramic Insulator Using Deep Learning and Random Forest

İrem GÖRGÖZ[1] , Mehmet CEBECİ [2]

[1,2]Electrical and Electronics Engineering Department, Engineering Faculty, Fırat University, Elazığ, Türkiye.
[1]igorgoz@firat.edu.tr, [2]mcebeci@firat.edu.tr
[1](ORCID: 0000-0003-2803-1119), [2](ORCID: 0000-0002-2971-6788)

Abstract
In this study, the electric field distribution of a three-dimensional U70 BLP fog-type insulator under various contamination and voltage conditions was predicted. Data obtained from COMSOL simulations were analyzed using Deep Neural Network (DNN) and Random Forest (RF) regression techniques. Ablation studies were conducted by varying the layer depths in the DNN and adjusting the hyperparameters in the RF to identify the best-performing models within the tested parameter ranges.The results demonstrate that the RF model outperforms the DNN, achieving lower MAE and MSE values. These approaches provide a reliable prediction tool for insulator design and safety assessments.

Keywords: ceramic insulator, deep learning, electric field, pollution, random forest

1. Introduction

Insulators are widely used in the power industry, from generation to transmission and distribution lines, to maintain electrical insulation and to mechanically support the load between the conductor and the ground in power equipment systems [1]. In recent years, the rapid development of industry and the economy has led to a gradual deterioration in air quality. Similarly, electricity demand and voltage levels have increased significantly [2]. As a result, the safety of high and extra-high voltage transmission lines has been significantly affected by contamination-induced flashovers under normal operating conditions. Contamination on insulators can lead to chemical interactions among hardened deposits, causing thermal aging and wear of conductors and line hardware, the formation of crusted layers on conductors and equipment, and a reduction in insulation performance. Therefore, contamination flashover observed in insulators used in high-voltage transmission lines has become one of the most critical issues to consider in the operation and design of high-voltage insulation systems [3].

The electric field intensity in an insulator increases with the level of contamination. The surface under polluted conditions contains both soluble and insoluble contaminants. The soluble layer becomes conductive in the presence of water. Consequently, surface leakage current flows along the conductive layer under the influence of the electric field, altering the electric field distribution within the insulator [4]. Uneven electric field distribution on the insulator surface affects its flashover characteristics and can compromise the insulator over time due to prolonged corona discharges. Given the significant impact of electric field distribution on insulator performance, numerous studies have focused on analyzing electric field patterns in insulators [5], [6].

Artificial intelligence and machine learning techniques are increasingly employed for fault detection in insulators and the classification of contamination conditions. One such method, Random Forest (RF), is an ensemble learning technique that constructs multiple decision trees during training and outputs the mode of the classes as the final prediction. By aggregating the predictions of many trees, RF reduces overfitting and enhances model robustness. Each tree is trained on a randomly selected subset of features and data, improving the model's generalization capability. Random Forest has been effectively used for contamination condition detection [7], for condition assessment using leakage current data [8] and for predicting lightning strike criteria. Decision tree-based models have also demonstrated significantly faster training and optimization times compared to neural network-based models.[9].
Another widely used approach for prediction and classification tasks is deep learning. Recent studies have shown that artificial neural networks (ANNs) perform exceptionally well in complex machine vision problems. ANNs achieve high accuracy in tasks such as classification, segmentation, and object detection. They have been applied for early fault detection

[1]Corresponding author

through arc identification [10], for fault detection and classification [11], for predicting insulator condition using leakage current data [12] and for estimating electric field and temperature distributions in insulators[13].

In this study, the electric field distribution of a three-dimensional U70 BLP silicone insulator under various contamination and voltage conditions was investigated. Data obtained from finite element method (FEM) simulations were used to predict the electric field using Deep Neural Network (DNN) and Random Forest (RF) models. Model performance was evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE), and R^2 metrics, and the effects of different hyperparameters were analyzed through ablation studies. The results demonstrate the potential of machine learning approaches for predicting the electrical performance of insulators, providing a valuable reference for engineering applications.

2. Method

In this study, a single-unit U70-BLP fog-type insulator with a surface leakage distance of 470 mm was modeled using the finite element method (FEM) in COMSOL Multiphysics. The insulator geometry was constructed in detailed three-dimensional form, and all relevant material properties were defined under the corresponding physics modules (e.g., porcelain, air, contaminated surface regions, cement, cap, and pin components). To simulate environmental conditions, the insulator was placed inside an air block measuring 500 mm × 500 mm. This block was designed to minimize the influence of boundary conditions on the electric field distribution and to obtain a more realistic solution. For the electric field analysis, a fine mesh was applied, with increased resolution at the insulator surface and other critical regions. The single-unit insulator geometry considered two distinct contamination regions (upper and lower) on the surface, and eight points were identified where the electric field intensity was expected to be critical. The fog-type insulator geometry and mesh structure are shown in Figure 1.

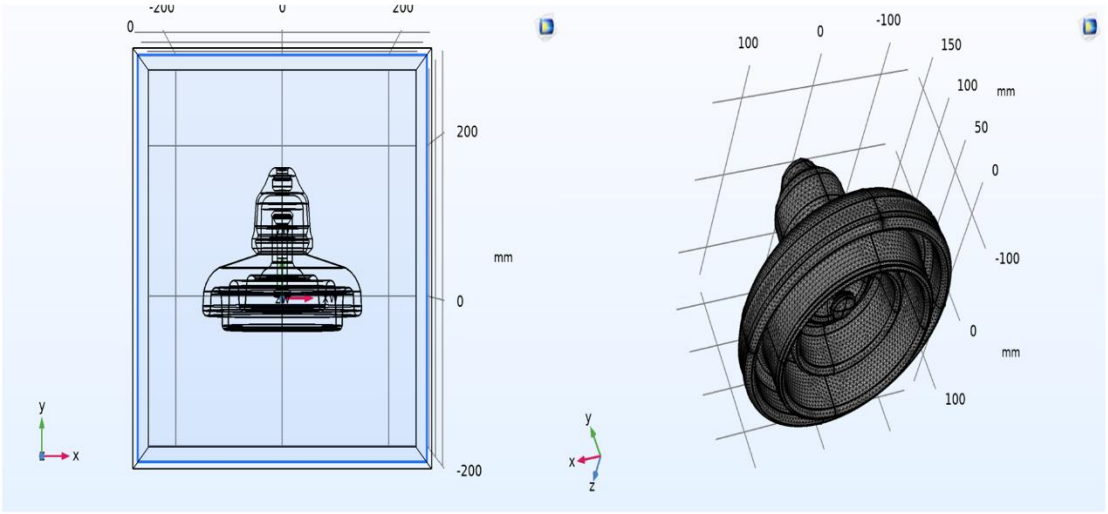


Figure 1. Geometry and mesh structure of the fog-type insulator.

The voltage applied to the insulator was investigated at five different levels between the nominal operating voltages of 11 kV and 15 kV. The contamination conductivity values were varied across five levels, ranging from 10^{-6} S/m (lightest contamination) to 10^{-2} S/m (heaviest contamination). The electrical material properties used for the electric field analysis of the single-unit U70-BLP insulator are listed in Table 1.

Table 8. Electrical material properties

Property	Porcelain/glass	cement	iron
Electrical conductivity: σ (S/m)	10^{-14}	10^{-14}	$1.12\times10^{+7}$
Relative permittivity: ε_r	4.2	2.09	1

The dataset for prediction was generated using COMSOL simulations. The applied voltage was varied between 11, 11.25, 12.5, 13.75, and 15 kV, while the upper and lower contamination conductivity values were set at 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, and 10^{-2} S/m. The input features consisted of four variables: two surface regions (upper and lower), one critical point (one of eight points), and the applied voltage. The output variable was the electric field. The upper and bottom contamination regions of the insulator, along with the eight critical points, are shown in Figure 2.

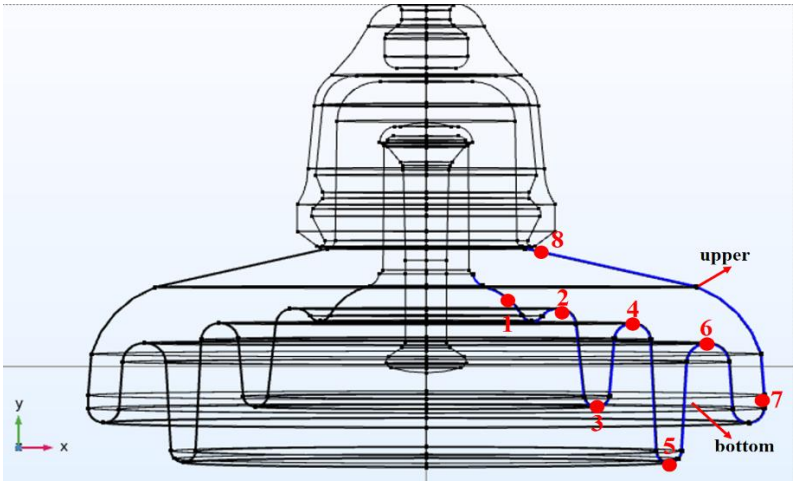


Figure 2. Contamination regions and critical points of the insulator

For each critical point, the electric field was calculated under $5^3 = 125$ different conditions, resulting in a total of $8 \times 125 = 1000$ data for the eight critical points. These data were used to train both Deep Neural Network (DNN) and Random Forest (RF) models. In both DNN and RF approaches, the dataset was divided into training, validation, and test sets, with 64% of the data used for training, 16% for validation, and 20% for testing.

Training was conducted within the framework of ablation studies, testing different combinations of hyperparameters. Model performance was evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE), and the coefficient of determination (R^2) to assess prediction accuracy. These metrics allow for a comparative analysis of the effects of different hyperparameter settings on model error and generalization capability. The metrics are defined as follows:
MAE (Mean Absolute Error) represents the average absolute difference between the model predictions and the true values. A smaller MAE indicates that the model's predictions are closer to the actual values. MAE is calculated using Equation 1.

$$\text{MAE}=\frac{1}{n}\sum_{i=1}^{n}|y_i - y_{ii}| \quad (1)$$

MSE (Mean Squared Error) is the average of the squared errors, giving greater weight to larger errors. A smaller MSE indicates that the model generally makes accurate predictions. MSE is calculated using Equation 2.

$$MSE=\frac{1}{n}\sum_{i=1}^{n}(y_i - y_{ii})^2 \quad (2)$$

R^2 (R-Squared) indicates how well the model explains the variance in the data. An R^2 value close to 1 implies that the model predicts the actual values very accurately, whereas an R^2 value near 0 or negative indicates poor predictive performance. R^2 is calculated using Equation 3.

$$R^2=1-\frac{\sum_{i=1}^{n}(y_i-y_{ii})^2}{\sum_{i=1}^{n}(y_i-y_{iii})^2} \quad (3)$$

n = Number of data points, y_i=Actual value, y_{ii} =Estimated value, y_{iii}=The mean of the actual values

2.1. Comsol Multiphysics Simulations

COMSOL Multiphysics is an interactive and powerful environment for solving and modeling a wide range of scientific and engineering problems based on partial differential equations. In this software, instead of explicitly defining the governing equations, modeling can be performed by specifying physical quantities such as material properties, loads, and boundary conditions.[2]. Using COMSOL, it is possible to incorporate different physical conditions into the model. COMSOL is a finite element solver and analysis software with applications across multiple physics domains, particularly for coupled or multiphysics problems [3].

The Electric Currents interface can be used to calculate electric field, current, and potential distributions. Conductivity is important for current calculations, as it depends on the surface resistance. Therefore, in addition to the dielectric constant (ε_r), the conductivity (σ) must also be specified.

In the COMSOL software, the set of equations given by Equations (4), (5), and (6) is solved in the frequency domain to compute the electric field and potential distribution.

$$\nabla . J = Q_j \quad (4)$$

$$J = (\sigma + jw\varepsilon_0\varepsilon_r)\mathrm{E} + J_e \quad (5)$$

$$\mathrm{E}=-\nabla V \quad (6)$$

In the given equations, ε is electrical permittivity, J is current density, σ denotes the conductivity of the contamination layer, Q_j is the current source (A/m³), and J_e is the external surface current density.
The variation of leakage distance and electric field for a specific condition, based on data obtained from COMSOL, is shown in Figure 3.

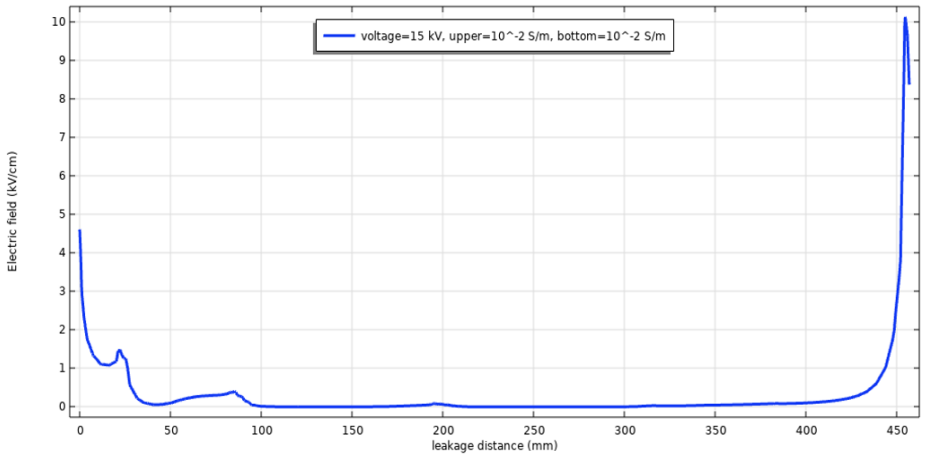


Figure 3. Electric field variation with leakage distance.

2.2. Ablation Tests

Ablation studies are conducted to systematically investigate the effects of model structure and hyperparameters on performance. In this study, both Deep Neural Network (DNN) and Random Forest (RF) models were employed. For the DNN, prediction accuracy was evaluated by varying the number of layers, number of neurons, epochs, batch size, and learning rate. For the Random Forest model, the number of trees (n_estimators), maximum depth (max_depth), minimum samples for leaves and splits (min_samples_leaf / min_samples_split), feature selection method (max_features), and bootstrap parameter were tested. This approach allowed for a comparative analysis of the influence of each hyperparameter on model performance.

2.2.1. DNN Ablation Tests

In the electric field prediction study using Deep Neural Networks (DNN), an ablation study was conducted by varying network depths and hyperparameters. Network structures ranged from a single layer to four layers, with the number of neurons per layer varying from 64 to 256. Although the ablation study was conducted on Deep Neural Networks (DNNs), a single-layer configuration was also evaluated. Evaluating a single-layer network allows for understanding the impact of network depth on model performance and highlights the necessity of deeper architectures for capturing complex relationships in the data. During training, different combinations of epochs (the number of times the entire training dataset is passed through the model), batch size (the number of samples processed in each step), and learning rate (the step size for updating model weights) were tested. This approach allowed for the assessment of the effects of network depth and training parameters on prediction accuracy, enabling the selection of the optimal configuration. The tested layers and hyperparameter settings are listed in Table 2.

Table 2. DNN ablation study –layer structures and hyperparameters

Layer Configuration	Neurons Per Layer	Epochs	Batch Size	Learning Rate
1 Layer	64	100,150	32,50	0.001, 0.0005
2 Layer	128-64	100,150	32,50	0.001, 0.0005
3 Layer	128-64-32	100,150	32,50	0.001, 0.0005
4 Layer	256-128-64-32	100,150	32,50	0.001, 0.0005

A total of 32 experiments were conducted by combining four different network depths with two choices each of epoch (100, 150), batch size (32, 50), and learning rate (0.001, 0.0005) to evaluate the effect of these hyperparameters on the model performance.

2.2.2. Random Forest Ablation Tests

The Random Forest (RF) model was evaluated using different combinations of hyperparameters. The number of trees (n_estimators) determines how many decision trees are used in the model; the maximum depth (max_depth) limits the depth of each tree; the minimum number of samples per leaf (min_samples_leaf) specifies the minimum number of data points required at a leaf node; the minimum number of samples required to split a node (min_samples_split) indicates the minimum data points needed for a node to be divided; the feature selection method (max_features) defines how many features are randomly selected at each split; and bootstrap (bootstrap) controls whether sampling is done with replacement. Using these hyperparameter combinations, a total of 96 different experiments were conducted with the Random Forest model. The hyperparameters used for testing are listed in Table 3.

Table 3. Random Forest ablation study – hyperparameters

Hyperparameter	Values
n_estimators	200,500
max_depth	None,8,16
min_samples_leaf	1,3
min_samples_split	2,6
max_features	Sqrt,0.5
bootstrap	True,False

If bootstrap is set to True, bootstrap sampling is performed for each tree; that is, random samples with replacement are drawn from the training data. This approach increases the diversity of the trees in the Random Forest and generally improves generalization.

If bootstrap is set to False, each tree is built using the entire training dataset, meaning no sampling is performed. In this case, the trees may be more similar, reducing model diversity and potentially increasing the risk of overfitting. If max_depth is set to None, the depth of the trees is unlimited.

3. Findings and Discussions

3.1. Deep Neural Network

The ablation study revealed that the worst-performing model was a single-layer Neural Network. This model consisted of a single layer with 64 neurons, was trained for 100 epochs, with a batch size of 50 and a learning rate of 0.0005. The performance metrics are provided in Table 4. These results indicate that a single-layer network with a low number of parameters is insufficient to capture the complex relationships within the dataset.

Table 4. Performance Metrics of the Worst-Performing Model

Metric	Train	Validation	Test
MAE	1.4386	1.4242	1.4671
MSE	2.9187	3.0144	3.1316
R^2	0.2550	0.2906	0.2412

The model performance graphs are shown in Figure 4.

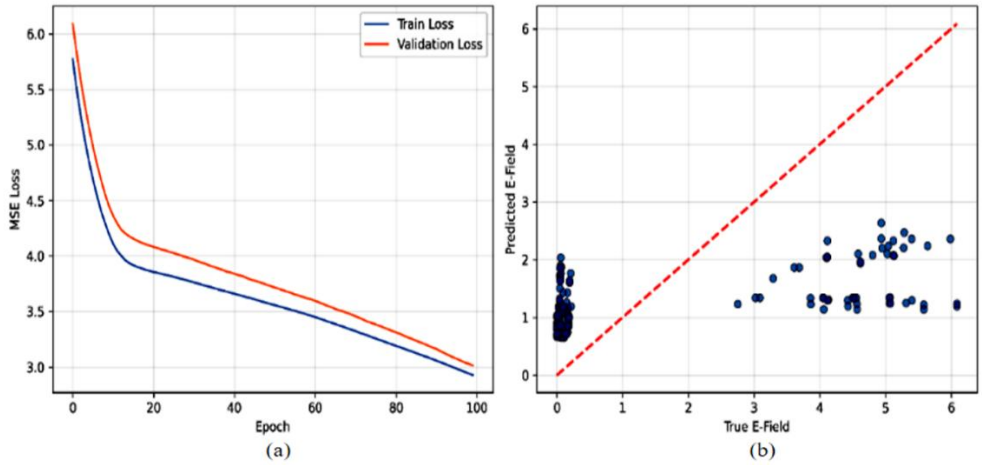


Figure 4. The worst performing model. (a) Training and Validation Loss vs. Epochs, (b) Predicted vs. True E-Field

The best performance was achieved with a four-layer DNN model (256–128–64–32 neurons). The model was trained for 100 epochs with a batch size of 50 and a learning rate of 0.001. While the previous worst-performing model (single layer, 64 neurons, learning rate 0.0005) exhibited notably high errors, increasing the number of layers and using a higher learning rate significantly improved the model's generalization capability. The performance metrics are provided in Table 5.

Table 5. Performance Metrics of the Best-Performing Model

Metric	Train	Validation	Test
MAE	0.1102	0.1386	0.1247
MSE	0.0587	0.0779	0.0672
R^2	0.9850	0.9817	0.9837

The model performance graphs are shown in Figure 5.

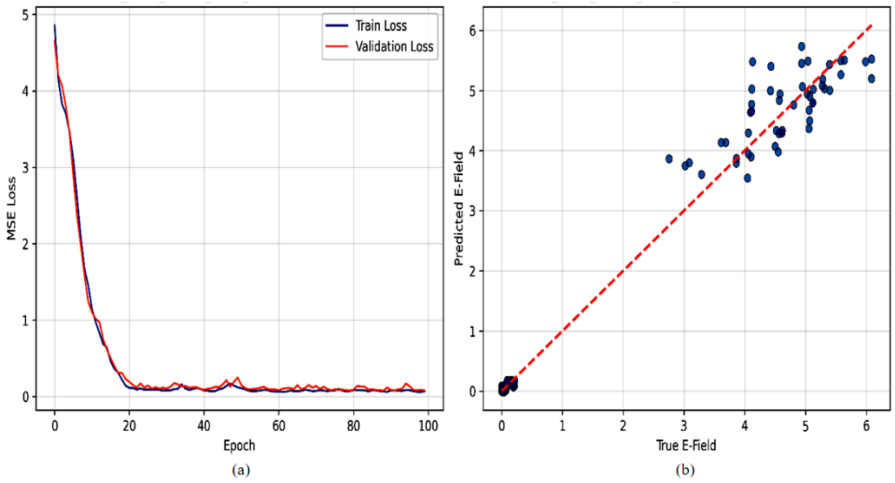


Figure 5. The best performing model. (a) Training and Validation Loss vs. Epochs, (b) Predicted vs. True E-Field

3.2. Random Forest

The ablation study showed that the lowest-performing Random Forest model was trained with 200 trees (n_estimators), no maximum depth limitation (max_depth=None), a minimum of 3 samples per leaf (min_samples_leaf=3), a minimum of 2 samples required to split a node (min_samples_split=2), square root feature selection (max_features='sqrt'), and bootstrap sampling enabled (bootstrap=True). The performance metrics are presented in Table 5. Although the training results were quite good, the increase in error on the validation and test sets indicates that the model's generalization capability is limited and suggests a tendency toward overfitting. The performance metrics are presented in Table 6.

Table 6. Performance metrics of the worst-performing Random Forest model

Metric	Train	Validation	Test
MAE	0.1139	0.1748	0.1707
MSE	0.0353	0.1130	0.0895
R^2	0.9910	0.9734	0.9783

The scatter plot of the actual versus predicted electric field values, illustrating model performance, is shown in the Figure 6.

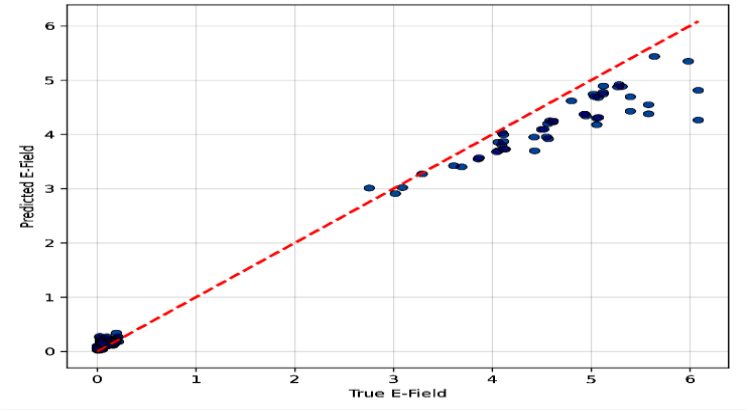


Figure 6. Predicted vs. True E-Field (The worst performing model-Random Forest)

The highest-performing Random Forest model was trained with 500 trees (n_estimators), no maximum depth limitation (max_depth=None), a minimum of 1 sample per leaf (min_samples_leaf=1), a minimum of 2 samples required to split a node (min_samples_split=2), square root feature selection (max_features='sqrt'), and without bootstrap sampling (bootstrap=False). The performance metrics are presented in Table 7.

Table 7. Performance metrics of the best-performing Random Forest model

Metric	Train	Validation	Test
MAE	0.0005	0.0707	0.0666
MSE	0.0005	0.0279	0.0217
R^2	0.999	0.9934	0.9948

The scatter plot of the actual versus predicted electric field values, illustrating model performance, is shown in the Figure 7.

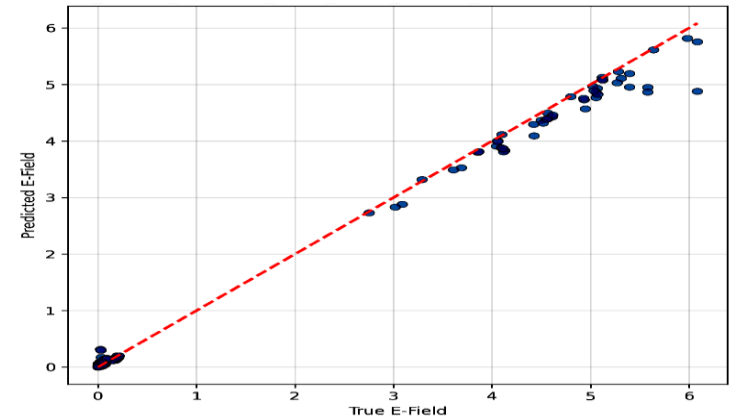


Figure 7. Predicted vs. True E-Field (The best performing model-Random Forest)

4. Results

In this study, electric field prediction was performed for a U70 BLP-type silicone insulator under different contamination and voltage conditions. The electrical behavior of the insulator was modeled by considering field intensities under high voltage and the effects of contamination. Predictions made using Deep Learning (DNN) and Random Forest (RF) methods provide valuable insights for understanding the insulator's performance and safe

operating limits. This approach emphasizes not only model performance but also the ability to accurately reflect the insulator's electrical behavior.
The test MAE was 0.1247 for the DNN and 0.0666 for the RF model, with RF reducing the error by approximately half. Increasing the number of layers and selecting an appropriate learning rate in the DNN reduced errors and improved the model's generalization capability. For the Random Forest model, both test and validation errors remained very low, indicating strong predictive capability.

While the DNN's performance is highly sensitive to hyperparameters such as network depth and learning rate, proper configuration is critical. In contrast, Random Forest exhibited a more stable structure.

Overall, Random Forest outperformed the DNN in terms of both accuracy and error metrics. This demonstrates that, even with a limited dataset (1000 data), RF can deliver strong performance. Although the DNN can achieve high accuracy with sufficient data and optimal hyperparameter selection, RF generally provides more reliable results across various ablation tests, making it a method that achieves robust performance with less effort.

5. References

[1] M. A. Douar, A. Mekhaldi, and M. C. Bouzidi, "Flashover process and frequency analysis of the leakage current on insulator model under non-uniform pollution conditions," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 17, no. 4, pp. 1284–1297, 2010, doi: 10.1109/TDEI.2010.5539701.

[2] H. Benguesmia, N. M'Ziou, and A. Boubakeur, "Simulation of the potential and electric field distribution on high voltage insulator using the finite element method," *Diagnostyka*, vol. 19, no. 2, pp. 41–52, 2018, doi: 10.29354/diag/86414.

[3] N. M ziou, H. Benguesmia, and H. Rahali, "Modeling Electric Field and Potential Distribution of an Model of Insulator in Two Dimensions by the Finite Element Method," *Int. J. Energ.*, vol. 3, no. 1, p. 01, 2018, doi: 10.47238/ijeca.v3i1.58.

[4] Arshad, A. Nekahi, S. G. McMeekin, and M. Farzaneh, "Effect of pollution severity on electric field distribution along a polymeric insulator," *Proc. IEEE Int. Conf. Prop. Appl. Dielectr. Mater.*, vol. 2015-October, pp. 612–615, 2015, doi: 10.1109/ICPADM.2015.7295346.

[5] S. Ilhan, A. Ozdemir, S. H. Jayaram, and E. A. Cherney, "Ac and transient electric field distributions along a 380 kV V-string insulator," *Conf. Rec. IEEE Int. Symp. Electr. Insul.*, no. January 2016, pp. 399–403, 2012, doi: 10.1109/ELINSL.2012.6251498.

[6] I. Gorgoz ;M Cebeci, "The Effect of the Surface Pollution on the Potential, Electric Field, and Temperature Distributions for Suspension Type Porcelain Insulators," *2021 13th Int. Conf. Electr. Electron. Eng. ELECO 2021*, pp. 129–133, 2021, doi: 10.23919/ELECO54474.2021.9677646.

[7] E. Ergün, "Artificial intelligence approaches for accurate assessment of insulator cleanliness in high-voltage electrical systems," *Electr. Eng.*, vol. 107, no. 3, pp. 2969–2982, 2025, doi: 10.1007/s00202-024-02691-3.

[8] U. A. Khan, M. Asif, M. H. Zafar, and L. Alhems, "Experimental validation of machine learning for contamination classification of polluted high voltage insulators using leakage current," *Sci. Rep.*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-97646-4.

[9] S. He *et al.*, "Intelligent prediction of 110kV insulator lightning flashover criteria based on random forest," *Electr. Power Syst. Res.*, vol. 232, no. December 2023, 2024, doi: 10.1016/j.epsr.2024.110423.

[10] G. Karaduman, M. Karakose, and E. Akin, "Deep learning based Arc detection in pantograph-catenary systems," *2017 10th Int. Conf. Electr. Electron. Eng. ELECO 2017*, vol. 2018-January, pp. 904–908, 2017.

[11] P. Pushpavathi, C. Mohan Kumar, M. Reethu Priya, K. Sreeja, and J. Vamsipriya, "Electrical Fault Detection using Attention GRU: A Deep Learning Approach for Improved System Performance," *3rd Int. Conf. Data Sci. Inf. Syst. ICDSIS 2025*, pp. 1–6, 2025, doi: 10.1109/ICDSIS65355.2025.11070404.

[12] A. Medeiros, A. Sartori, S. F. Stefenon, L. H. Meyer, and A. Nied, "Comparison of artificial intelligence techniques to failure prediction in contaminated insulators based on leakage current," *J. Intell. Fuzzy Syst.*, vol. 42, no. 4, pp. 3285–3298, 2022, doi: 10.3233/JIFS-211126.

[13] I. Görgöz; M. Cebeci., "Analysis of Electric Field and Temperature Distributions of Non-Uniformly Contaminated Silicone Composite Insulators Using Deep Learning," no. May, pp. 94721–94739, 2025, doi: 10.1109/ACCESS.2025.3574493.

Customer Churn Prediction in Telecom: A Comparative Machine Learning Approach Using IBM Telco Dataset

Mehmet Hanifi IŞIK[1], Abdulsamet KARAYEL[2], Mehmet ARZU[3], Mahmut KAYA[4]

[1,2,3,4]Department of Artificial Intelligence and Data Engineering, Faculty of Technology, Fırat University, Elazığ, Türkiye.
[1]230291029@firat.edu.tr, [2]230291034@firat.edu.tr, [3]marzu@firat.edu.tr, [4]mahmutkaya@firat.edu.tr
[1](ORCID: 0009-0004-2868-3808), [2](ORCID: 0009-0008-8247-5638),[3](ORCID: 0000-0002-7846-1769), [4](ORCID: 0000-0003-2317-2385)

Abstract
Telecommunications is one of the leading sectors in the integration of digital systems into processes as a result of technological developments. In the telecommunications sector, retaining existing customers has become more important than acquiring new customers due to the increasingly competitive environment. This study aims to predict customer churn in a telecommunications company using machine learning methods. Using the Telco Customer Churn dataset provided by IBM, different classification algorithms such as Decision Tree, Random Forest, Support Vector Machines (SVM), Logistic Regression, and Gradient Boosting were also evaluated. Feature engineering and data preprocessing techniques were applied to the data. The highest accuracy rate of 82% was achieved using the Gradient Boosting algorithm. The overall objective of this study is to identify the key factors influencing customer churn and contribute to the marketing and customer management strategies of telecommunications companies.

Keywords: machine learning, customer churn, marketing

1. Introduction

Today, with digitalization and technological developments, the understanding and practice of business has been reshaped. The increasingly competitive environment shows that simply offering products is not enough; providing high-quality services to customers has become a necessity. This situation has directly affected the service sector in particular and has made the concept of “service” more complex [1]. The concept of service encompasses elements that cannot be measured concretely but directly influence the customer experience. As a result, definitions of service in the literature are highly varied. For example, Dalgıç (2013) states that service cannot be defined within a single framework, while İslamoğlu and colleagues (2006) describe service as a set of activities that provide solutions to consumers' generally non-physical needs. This approach is effective in explaining how the understanding of service has shaped the telecommunications sector. This is because telecommunications services are evaluated based on components such as connectivity, speed, and Access [2]. When experiencing these services, customers consume a “process” rather than a product. If there is a disruption in service or customer expectations are not met, customers will turn to other companies. In this case, maintaining customer loyalty is of great importance for sustainable success in the service sector, especially in the telecommunications sector. Customer churn poses a risk of revenue loss and damage to reputation for companies [3]. Additionally, according to marketing literature, the cost of acquiring a new customer is significantly higher than retaining an existing one. Kotler (2012) notes that this difference is approximately five times greater [4]. Therefore, it is extremely important for companies to identify customer churn tendencies early on and intervene. The main objective of this study is to develop a solution for predicting customer churn in the telecommunications sector. To this end, the accuracy rates of models were evaluated using various machine learning algorithms, and the key variables affecting customer churn were analyzed. The findings of this study are intended to contribute to telecommunications companies shaping their marketing strategies and customer relationship management in a data-driven manner [5].

2. Marketing Perspective and Problem Definition

Telecom companies marketing strategies are generally shaped around acquiring new customers, retaining existing customers, and winning back lost customers. However, churn prediction aims to achieve the most critical of these

[1]Corresponding author

goals: retaining existing customers [1]. Churn behavior is not limited to price or service quality but is also influenced by various factors such as the customer's service experience, satisfaction with support services, ease of bill payment, and type of internet service [2]. In this context, incorporating churn prediction models into marketing strategies will provide companies with a strategic advantage. This enables campaigns to be created more accurately and personalized offers to be made to customers. As a result, customer loyalty increases and revenue loss decreases. For example, customers showing a tendency to leave can be offered special commitment-based prices or included in loyalty programs to prevent them from leaving the company [4]. Accurately predicting customer churn in the telecommunications sector is not only a technical achievement but also provides a strategic advantage [6]. The outputs of the models developed in this study have provided a better understanding of customer behavior, particularly in identifying customers at high risk of churn. Integrating the predictions obtained into marketing strategies will maximize the benefits of these results. This will help reduce the risk of customer loss. Grouping customers according to their risk levels enables companies to plan their time and budgets more accurately, thereby allowing them to develop target-oriented strategies by using their resources more efficiently.
For example, the following interventions may be recommended for customers with a high probability of churn:

- Offer long-term contracts and special prices to those with short commitment periods.
- Provide free additional services (e.g., online backup, device protection) to those who do not use technical support and have high billing amounts.
- Provide proactive support via the call center to customers who are likely to experience dissatisfaction.

Periodically updating churn prediction systems, feeding them with new customer data, and integrating them with CRM systems will provide marketing teams with a strategic advantage in both reducing customer churn and acquiring new customers.

3. Related Studies

Customer churn in the telecommunications sector has been a topic of interest to researchers for a long time. In particular, increasing competition has led companies to view retaining existing customers as a more profitable strategy [7]. Early studies in this area typically attempted to predict customer churn using basic statistical learning approaches such as logistic regression and decision trees. However, with the increase in data volume and computational power over time, more advanced machine learning algorithms such as support vector machines, random forests, and gradient boosting have begun to emerge [8]. In particular, the study by Idris and colleagues demonstrated that support vector machines optimized with genetic algorithms yielded successful results [9]. There are also many different approaches in the literature regarding the variables that can be used to predict customer churn. Chen and colleagues noted that a detailed analysis of customer behavior data directly affects the success of the model [10]. Factors such as call duration, internet usage, and payment habits play an important role in customer churn decisions. On the other hand, Han and Ferreira have shown that the departure of individuals in a customer's social circle increases the likelihood of that customer also leaving [11]. In recent years, deep learning models have also begun to be used frequently in the field of churn analysis. LSTM models, which are particularly effective on sequential data structures, produce powerful results in capturing customer behavior patterns over time [12]. In studies comparing these models with traditional methods, a significant increase in accuracy has been reported for certain data sets [13]. Additionally, it is emphasized that in churn analysis, not only the selection of the correct algorithm but also explainability must be taken into account. Methods such as LIME and SHAP are used to make it possible to understand why models make certain predictions [14]. The application of social network analysis to this field is another notable development. Óskarsdóttir and colleagues used mobile search data to assess churn risk through social interaction structures [15]. Similarly, Phua et al. emphasized the importance of network-based variables in identifying groups that may experience customer loss in the near future [16]. Finally, bibliometric analyses show that research on churn has increased year over year, with the most popular topics being machine learning, customer retention and telecommunications [17-18]. This highlights how dynamic the field is, both academically and industrially.

4. Data and Methods

In this study, the Telco Customer Churn dataset shared by IBM on the Kaggle platform was used. The dataset contains 21 variables for a total of 7,043 customers [19]. These variables cover various areas, such as customer demographics, service usage, contract details, billing information, and additional service preferences. The target variable Churn, indicates whether the customer left (Yes/No).

4.1. Data preprocessing

The following preprocessing steps were performed in order to obtain better results in the analysis and modeling process of the data set:

- The TotalCharges column was first converted to numerical format because some cells contained non-numerical values, and then 11 rows containing missing values in this column were removed from the data set.
- The customerID column, which only contains identity information and does not contribute meaningfully to the prediction process, was removed from the dataset before modeling.
- Categorical variables were first examined and, when necessary, converted to numerical format using appropriate methods such as label encoding or one-hot encoding [20].
- Data visualization was used to analyze the effects of variables on the target variable. In particular, the relationship between the Tenure (number of months spent at the company) and MonthlyCharges (monthly bill) variables and churn was examined using graphs.

4.2. Customer Profile

The data set includes multidimensional variables that may affect customer churn, and these variables are classified under four main headings: (i) demographic information, which covers basic socio-demographic characteristics such as the customer's gender, age, marital status, and whether they live with dependents; (ii) service usage, including the use of telephone and internet services, as well as additional digital services such as multi-line usage, online security, backup, device protection, technical support, digital TV, and movie streaming; (iii) contract and billing information, including the customer's contract term, paperless billing preference, payment method, monthly and total bill amounts, and length of stay with the company; (iv) Finally, the target variable, churn, indicates whether the customer has left the service. This customer profile has enabled detailed analysis of customer behavior and provides a basis for predicting customer loss. In particular, it has been observed that customers with short contract terms, high billing amounts, and who do not use additional services such as technical support are more likely to churn.

4.3. Applied Machine Learning Algorithms and Modeling Process

In this study, different machine learning algorithms were applied comparatively to predict customer churn. All models were subjected to the same data preprocessing steps, trained, and tested on the dataset. The dataset was split into 80% training and 20% testing. The modeling process began with data scaling, followed by hyperparameter optimization using the grid search method to determine the most appropriate parameters [5]. After separating the training and test data, the performance of the models was evaluated using accuracy, precision, sensitivity, and F1 score metrics. The machine learning algorithms used in this context are summarized below:

Random Forest: An ensemble learning method consisting of multiple decision trees. It achieves high accuracy through random subsampling and feature selection.[21].
Gradient Boosting: A gradient boosting ensemble method that reduces the error rate by training weak classifiers sequentially.[22].
Support Vector Machines (SVM): This algorithm, which is effective for balanced classes, aims to find the hyperplane that separates the classes with maximum margin [23].

Logistic Regression: A linear model used for binary classification problems. Both L1 and L2 norms were evaluated as penalty terms in the study [24].
K-Nearest Neighbors (KNN): An intuitive algorithm that performs classification based on the nearest neighbors of the observation point [25].
Decision Tree: An explainable model that classifies data by dividing the dataset into branches according to decision rules [26].
AdaBoost: An ensemble learning method that trains weak learners sequentially and gives more weight to examples that make mistakes [27].
XGBoost: An improved gradient-boosted decision tree algorithm. It is fast, scalable and high-performance [28].
Naive Bayes (Gaussian): A model based on conditional probability relationships between classes, which is particularly effective for continuous features under a normality assumption [29].
The models were implemented using the Python programming language and the scikit-learn library. All algorithms were scaled using MinMaxScaler and then hyperparameters were optimized using the GridSearchCV method. Cross-validation was used to prevent overfitting and evaluate the generalization performance of the models. However, to reduce processing time, cross-validation was performed with two folds (cv=2).

5. Experimental Results and Discussion

In this section, the customer churn prediction performance of the models examined as a result of the experiments conducted is evaluated. The models are analyzed using accuracy, precision, sensitivity, and F1 score metrics. The results obtained reveal the strengths and weaknesses of each model. It provides a comprehensive overview of the success in predicting customer churn. Based on the results obtained, the comparison of performance metrics with the models is shown in Table 1.

Table 1. Experimental results

Model	Accuracy	Precision	Recall	F1 Score
Gradient Boosting	**0.820896**	**0.811123**	**0.820896**	**0.811044**
Random Forest	0.819474	0.809569	0.819474	0.809798
XG Boost	0.817342	0.807936	0.817342	0.809406
Ada Boost	0.815920	0.805176	0.815920	0.804601
Decision Tree	0.786070	0.779423	0.786070	0.782132
K-Nearest Neighbors	0.775409	0.769230	0.775409	0.771876
Logistic Reggression	0.763326	0.817320	0.763326	0.776327
SVM	0.758351	0.808827	0.758351	0.771229
Naive Bayes	0.712154	0.807985	0.712154	0.730367

According to the table data, the highest accuracy rate of 82.09% was achieved with the Gradient Boosting model. This model was followed by Random Forest with an accuracy of 81.95% and XGBoost with an accuracy of 81.73%. Simpler models such as Naive Bayes achieved relatively low success rates. This demonstrates that ensemble methods yield more effective results in complex classification problems such as customer churn.

For a more detailed analysis of the model outputs, the confusion matrix of the Gradient Boosting algorithm, which yielded the most successful results, was examined. The analysis revealed that there are a total of 1,406 + 143 = 1,549 actual “non-churn” cases, of which the algorithm correctly predicted 1,406 and misclassified 143. Similarly, there are 259 + 302 = 561 actual “churn” cases, with 302 correctly predicted and 259 misclassified. This distribution is illustrated in Figure 1.

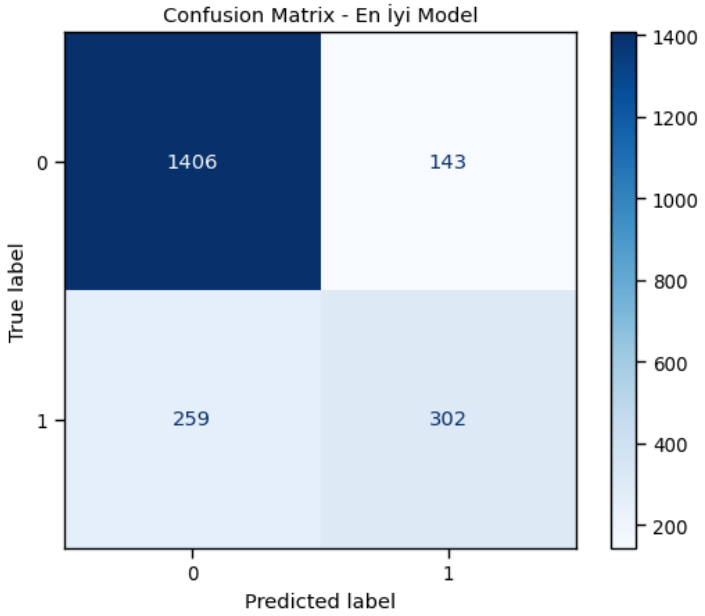


Figure 1. Confusion matrix – Gradient Boosting model

In order to make the results more understandable, an analysis of the importance levels of the features was performed. In this context, the variable importance ranking obtained from the Gradient Boosting algorithm is presented in Table 2.

Table 2. Feature and importance values

Feature	Feature Importance
Contract_Two year	0.02
PaperlessBilling_No	0.03
PaymentMethod_Electronic check	0.03
TechSupport_No	0.04
OnlineSecurity_No	0.06
MonthlyCharges	0.07
TotalCharges	0.08
InternetService_Fiber optic	0.09
Tenure	0.17
Contract_Month-to-month	0.38

The variable with the highest importance rating is monthly contract. This indicates that customers with short-term subscriptions are more likely to leave. In addition, variables such as length of service and type of internet used also play an important role in determining customer behavior. Customers who do not receive technical support and online security services are at particularly high risk of leaving.

In order to contribute to the interpretation of the model outputs, the distribution of some numerical variables related to customer churn was examined. The triple density graph below shows the distribution of the variables tenure, MonthlyCharges, and TotalCharges according to customer churn status, as shown in Figure 2.

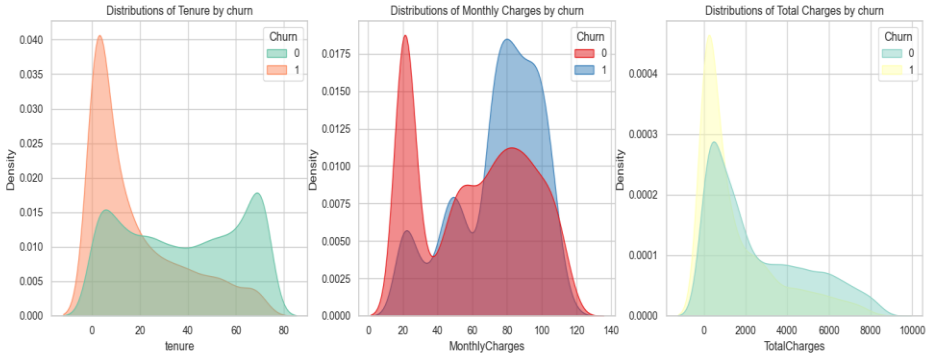


Figure 2. Distribution of numerical account characteristics according to customer churn

The graph related to the tenure variable shows that individuals who experience customer loss largely have short-term subscriptions. In particular, the density of users who churn within the first 20 months is quite high. In contrast, the customer loss rate in long-term subscriptions decreases significantly. This finding indicates a direct relationship between customer loyalty and service duration. The MonthlyCharges variable shows that a large proportion of churning customers are concentrated in the 60 to 100 unit billing range. This suggests that high billing amounts may increase customer expectations, and failure to meet these expectations may trigger a tendency to leave. The TotalCharges variable reveals that the majority of churning users are at low total billing levels. This indicates that these individuals leave the system without making significant payments during their service period. These distribution analyses reveal that churn behavior is associated with specific numerical patterns and that these variables can provide important contributions to the modeling process.

In order to better understand the relationships between variables, a correlation analysis was performed on numerical characteristics. The correlation matrix obtained is presented in Figure 3.

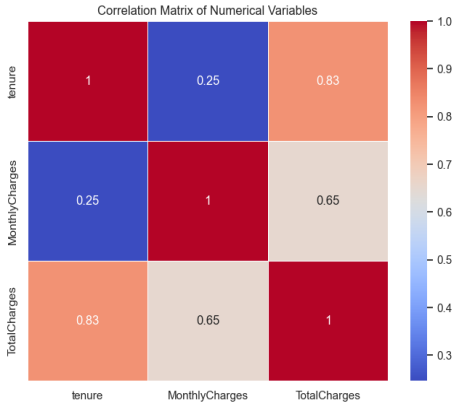


Figure 3. Correlation matrix of Telco Customer Churn data set

This analysis evaluates the linear relationships between three key variables: tenure, MonthlyCharges, and TotalCharges. A high positive correlation (0.83) was observed between Tenure and TotalCharges, indicating that as the length of service increases, the total amount paid also rises. A moderate positive correlation (0.65) exists between MonthlyCharges and TotalCharges, suggesting that higher monthly fees naturally increase total payments. The relationship between Tenure and MonthlyCharges is weaker (0.25), indicating that long-term customers do not necessarily pay high bills and that these two variables may have independent effects. These correlation findings

provide an important foundation for modeling the structural relationships between variables. In particular, they contribute to variable selection strategies that help reduce the model's tendency toward overfitting.

In order to evaluate model performance more comprehensively, ROC (Receiver Operating Characteristic) curves were plotted for all classification algorithms. The ROC curve shows the models' ability to distinguish the positive class, with the false positive rate on the horizontal axis and the true positive rate on the vertical axis. Through these curves, the classification accuracy of the models can be comprehensively evaluated under different threshold values. The obtained values are presented in Figure 4.

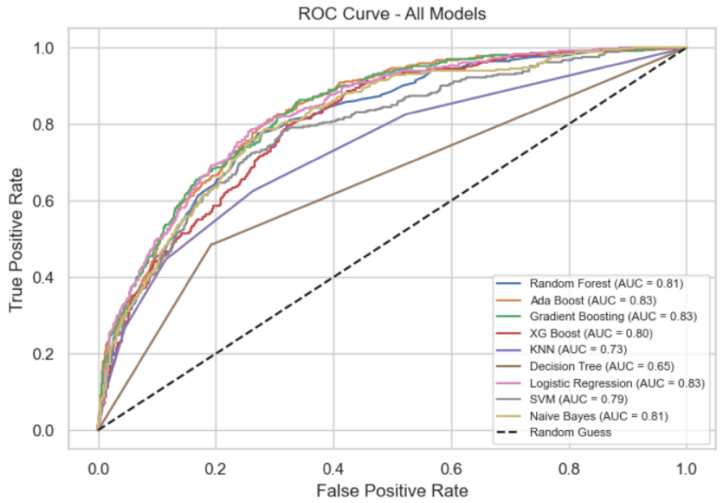


Figure 4. ROC Curves and AUC Values for All Models ROC curve

As shown in the figure, Gradient Boosting, AdaBoost, and Logistic Regression models showed the highest performance with an AUC (Area Under Curve) value of 0.83. This indicates that these three models were able to distinguish the positive class (churning customers) more accurately than the others. The Random Forest and Naive Bayes models showed high performance with an AUC value of 0.81, but slightly lower than the top-ranked models. The XGBoost model also yielded similar successful results with an AUC of 0.80. The fact that the ROC curves of these models are steep and close to the upper area indicates that their overall classification accuracy is high. In contrast, the KNN (0.73 AUC) and Decision Tree (0.65 AUC) models showed more limited discriminative performance. This suggests that these algorithms may have lower sensitivity in complex class distributions. The SVM algorithm also showed average performance with an AUC of 0.79. When evaluating the ROC curves overall, ensemble methods (especially Gradient Boosting and AdaBoost) are seen to provide more reliable and consistent results for problems with imbalanced classes, such as customer churn. Models with an AUC value of 0.8 or higher are considered sufficiently successful for practical application.

Based on the analyses presented above, the performance of different machine learning algorithms in predicting customer churn has been comprehensively evaluated. The findings indicate that ensemble-based methods, particularly Gradient Boosting and AdaBoost, have delivered superior results in terms of both overall accuracy and classification metrics such as AUC. The distributions of numerical variables suggest that churn behavior is closely related to specific customer profiles; factors such as short-term subscriptions, high monthly fees, and low total payments were found to increase the likelihood of churn. Moreover, correlation analysis and variable importance scores have highlighted the key features influencing the model’s decision-making process. Overall, the modeling approaches applied in this study provide effective tools for understanding and predicting customer churn in the telecommunications sector. These models can contribute to data-driven decision-making not only in the analysis of historical data but also in customer management and strategic planning processes.

6. Result

In this study, nine different machine learning algorithms were applied and compared to predict customer churn in the telecommunications sector. Experiments conducted using the IBM Telco Customer Churn dataset revealed that the highest accuracy rate was achieved with the Gradient Boosting algorithm. The analyses revealed that customers with short-term contracts, high bills, and who do not use additional services such as technical support are more likely to leave the service. These findings provide important insights not only in terms of classification performance but also in terms of marketing strategies and customer management. The developed prediction models are capable of supporting marketing applications such as customer group identification, campaign target creation, and personalization of loyalty programs. Thus, companies can make more accurate decisions to reduce customer churn. In future studies, some recommendations can be made to improve model performance and usage effectiveness: Deep learning models based on LSTM or Transformer can be tested using time-series data. Long-term metrics such as customer lifetime value can be evaluated alongside customer churn risk. The validity of models in the real world can be improved by making the data set more up-to-date and comprehensive. Prediction systems can be integrated into customer relationship management (CRM) infrastructures to make them suitable for real-time usage scenarios. In conclusion, this study demonstrates that machine learning-based approaches offer an effective and practical solution for predicting customer churn. When integrated into industry applications, the developed models can contribute to both maintaining customer satisfaction and increasing business efficiency.

7. References

[1] Amin A, Anwar S, Adnan A, Nawaz M, Howard N, Qadir J, Hawalah A. "Customer churn prediction in the telecommunication sector using a rough set approach". *Neurocomputing*, 237, 242–254, 2019.

[2] Tsai CF, Lu YH, Yeh CW. "Mining customer value for marketing decisions: A study on the wireless telecommunication industry". *Expert Systems with Applications*, 36(3), 3631–3635, 2009.

[3] Ascarza E, et al. "In pursuit of enhanced customer retention management: Review and research opportunities". *Journal of Service Research*, 21(3), 219–237, 2018.

[4] Lemmens A, Croux C. "Bagging and boosting classification trees to predict churn". *Journal of Marketing Research*, 43(2), 276–286, 2006.

[5] Vafeiadis T, Diamantaras KI, Sarigiannidis G, Chatzisavvas KC. "A comparison of machine learning techniques for customer churn prediction". *Simulation Modelling Practice and Theory*, 55, 1–9, 2015.

[6] Hadden J, Tiwari A, Roy R, Ruta D. "Computer assisted customer churn management: State-of-the-art and future trends". *Computers & Operations Research*, 34(10), 2902–2917, 2007.

[7] Larivière B, Van den Poel D. "Predicting customer retention and profitability by using random forests and regression forests techniques". *Expert Systems with Applications*, 29(2), 472–484, 2005.

[8] Verbeke W, Martens D, Mues C, Baesens B. "Building comprehensible customer churn prediction models with advanced rule induction techniques". *Expert Systems with Applications*, 38(3), 2354–2364, 2012.

[9] Idris A, Khan A, Lee YS. "Intelligent churn prediction in telecom: Employing mRMR feature selection and rotBoost based ensemble classification". *Applied Intelligence*, 39(3), 659–672, 2012.

[10] Chen C, Li Y, Huang S. "A new approach to customer churn prediction based on advanced ensemble learning". *Expert Systems with Applications*, 39(14), 11263–11272, 2012.

[11] Han Q, Ferreira P. "The role of peer influence in churn in wireless networks". *arXiv preprint*, arXiv:1409.1102, 2014.

[12] Huang J, Wang H, Wang Z, Wang W. "Customer churn prediction in telecom using LSTM neural networks". In: Proceedings of the International Conference on Artificial Intelligence and Big Data (ICAIBD), 151–155, 2018.

[13] Mohaimin R, et al. "Predictive analytics for telecom customer churn: Enhancing retention strategies in the US market". *Journal of Computer Science and Technology Systems*, 7(1), 30–45, 2024.

[14] Rashed M, Akter T, Das BC, et al. "Explainable AI with ensemble models for churn prediction in telecom". *Algorithms*, 17(6), 231, 2024.

[15] Óskarsdóttir M, Bravo C, Verbeke W, Sarraute C, Baesens B, Vanthienen J. "Social network analytics for churn prediction in telco: Model building, evaluation and network architecture". *arXiv preprint*, arXiv:2001.06701, 2020.

[16] Phua C, Cao H, Gomes JB, Nguyen MN. "Predicting near-future churners and win-backs in the telecommunications industry". *arXiv preprint*, arXiv:1210.6891, 2012.
[17] Mathebula LY, Seitshiro M, Chanza M, Bongale A. "A bibliometric literature review on telecommunication customer churn". *International Journal of Innovative Research in Science Studies*, 8(4), 1189–1201, 2025.
[18] Bhale UA, Bedi HS. "Customer churn construct: Literature review and bibliometric study". *Management Dynamics*, 24(1), e10, 2024.
[19] IBM. "Telco Customer Churn". IBM Sample Data Sets. [Çevrim içi]. Erişim adresi: https://www.ibm.com/communities/analytics/watson-analytics-blog/guide-to-sample-datasets/. Erişim tarihi: Ağustos 2025.
[20] Kaur H, Singh M. "A survey of machine learning techniques for churn prediction in telecom". International Journal of *Advanced Computer Science and Applications*, 11(1), 496–503, 2020.
[21] Breiman L. "Random forests". *Machine Learning*, 45(1), 5–32, 2001.
[22] Friedman JH. "Greedy function approximation: A gradient boosting machine". *Annals of Statistics*, 29(5), 1189–1232, 2001.
[23] Cortes C, Vapnik V. "Support-vector networks". *Machine Learning*, 20(3), 273–297, 1995.
[24] Hosmer DW, Lemeshow S, Sturdivant RX. Applied Logistic Regression. 3rd ed. Hoboken, New Jersey, USA, Wiley, 2013.
[25] Cover T, Hart P. "Nearest neighbor pattern classification". *IEEE Transactions on Information Theory*, 13(1), 21–27, 1967.
[26] Quinlan JR. "Induction of decision trees". *Machine Learning*, 1(1), 81–106, 1986.
[27] Freund Y, Schapire RE. "A decision-theoretic generalization of on-line learning and an application to boosting". *Journal of Computer and System Sciences*, 55(1), 119–139, 1997.
[28] Chen T, Guestrin C. "XGBoost: A scalable tree boosting system". In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 785–794, San Francisco, California, USA, 2016.
[29] Rish I. "An empirical study of the naive Bayes classifier". In: IJCAI 2001 Workshop on Empirical Methods in Artificial
[30] Intelligence, 3(22), 41–46, 2001.

Artificial Intelligence Based Waste Classification and Guidance System

Zinnet AYDIN [1] , Canan KOÇ [2], Fatih ÖZYURT [3]

[1,2,3]Software Engineering, Faculty of Engineering, Fırat University, Elazığ, Türkiye.
[1]znntaydnn@gmail.com, [2]canan.koc@firat.edu.tr, [3]fatihozyurt@firat.edu.tr
[1](ORCID: 0009-0001-0180-2473), [2](ORCID: 0000-0002-2651-9471),[3](ORCID: 0000-0002-8154-6691)

Abstract
Efficient waste classification is vital for sustainable urban living, yet the improper disposal of recyclable materials such as plastic, glass, and paper remains a persistent problem. This study presents a smart and inclusive waste classification and guidance system that integrates deep learning and sensor technologies to address both efficiency and accessibility. The system employs an image-based classification framework using EfficientNet models (B0, B2, V2-S) to achieve an optimal balance between computational efficiency and predictive accuracy. A custom dataset of 900 labeled images, enhanced through augmentation, is used to improve generalization under diverse environmental conditions. Additional weight sensors monitor bin capacity, while a Raspberry Pi serves as the embedded control unit. Real-time user guidance is provided through voice prompts and LED indicators, enabling individuals with visual or hearing impairments to participate equally in recycling processes. The proposed approach is expected to deliver classification accuracy exceeding 90%, while its dual-modality feedback design promotes social inclusivity. This system offers a scalable, effective, and user-friendly solution for public spaces such as schools, shopping malls, and municipal facilities, contributing to both environmental sustainability and universal accessibility.

Keywords: smart waste management, machine learning, efficientNet, accessibility, human–machine interaction, embedded AI systems

1. Introduction

The global rise in waste generation has intensified the need for effective and scalable recycling systems. Despite the availability of recycling infrastructure in many regions, recyclable materials such as plastic, paper, and glass are frequently misclassified or improperly discarded. This inefficiency often stems from insufficient user guidance and limited accessibility of technology, resulting in reduced recovery rates and increased environmental impact [1].

In recent years, artificial intelligence, particularly computer vision and deep learning, has emerged as a promising approach for automating waste classification tasks. Convolutional neural networks (CNNs) have demonstrated strong capabilities in visual recognition, enabling real-time sorting of recyclable items [2]. Notably, CNN-based smart sorting systems have achieved high accuracy using visual waste inputs, underscoring the potential of deep learning in recycling applications [3], [4].

Furthermore, low-cost embedded platforms such as Raspberry Pi have proven effective for real-time image processing in similar systems, highlighting their suitability for public environments with limited resources [1], [5].

Among contemporary CNN architectures, EfficientNet has gained attention for its balance between classification performance and computational efficiency. It applies a compound scaling method to jointly optimize model depth, width, and resolution, making it well-suited for embedded deployment. Ayman et al. evaluated various EfficientNet versions on a nine-class waste dataset, achieving over 93% accuracy with EfficientNet-B0 and B2, thus validating their use in edge devices [4]. Similarly, another study integrated data augmentation and leaky-ReLU activation functions into a CNN model, attaining 93.28% accuracy in classifying recyclable and organic waste, which highlights the importance of tuning and regularization techniques [6].

Despite these advancements, most existing systems lack inclusive design features. Individuals with visual or auditory impairments often face challenges when interacting with smart bins or automated recycling stations. The present study addresses this gap by incorporating dual-modality feedback through LED indicators and voice prompts, aiming to make the recycling process universally accessible. Comparable approaches have been reported in hybrid models combining

[1]Corresponding author

YOLO-based detection with CNN-based classification [7], as well as in systems enhanced with channel attention mechanisms and multi-scale feature extraction [1].

To address these challenges, this study proposes a smart waste classification and routing system based on candidate models from the EfficientNet family (B0, B2, and V2-S) and deployed on a Raspberry Pi platform. The system will be trained on a custom dataset of recyclable materials and evaluated under real-world conditions. In addition to visual classification, weight sensors will monitor bin capacity, and feedback will be provided through both auditory and visual channels. The overarching goal is to deliver a low-cost, energy-efficient, and inclusive solution for sustainable public waste management.

2. Materials and Methods

This study will present the design of an intelligent waste classification and guidance system that categorizes waste into three primary classes (plastic, glass, and paper) and delivers both auditory and visual feedback to users. Initially, a custom dataset comprising 900 images, evenly distributed across the three categories, will be developed, with images manually captured under varying lighting conditions, camera angles, and backgrounds to enhance model robustness in real-world scenarios [1].

To address the limitations imposed by a relatively small dataset, data augmentation techniques (random rotations, horizontal flips, brightness and contrast adjustments, zooming, and cropping) will be employed during the training phase [8], [9]. All images will be resized to a uniform input size and normalized to ensure consistency, as these preprocessing steps have been reported to significantly enhance model performance and reduce overfitting [8].

The classification task will consider multiple EfficientNet family models (B0, B2, and V2-S) to identify the most suitable architecture. Transfer learning will be applied using pre-trained ImageNet weights, and all candidate models will be trained under the same conditions for fair comparison. Prior research has shown that advanced EfficientNetV2 variants with attention mechanisms can exceed 95% accuracy in recyclable waste recognition [10].

Model optimization methods such as pruning, quantization, and TensorRT acceleration will be explored to reduce computational load and achieve real-time performance on Raspberry Pi 4 [11], [12].

A Raspberry Pi 4 connected to a high-resolution camera module will capture real-time images of waste materials, which will be processed and classified on-device using the optimized model selected from the EfficientNet family. Depending on the predicted class, the system will provide multimodal feedback consisting of an audio message (e.g., “This is plastic. Please use the blue bin”) and a color-coded LED signal, improving usability particularly for users with visual or auditory impairments [13].

Weight sensors based on load-cell technology will be integrated into each bin to monitor fill levels, triggering auditory and visual alerts when predefined thresholds are reached [9], [14]. Sensor data will be processed in real-time using Python libraries such as RPi.GPIO and HX711, ensuring seamless integration with the system’s control logic, as previous studies have shown that real-time sensor monitoring enhances user awareness and system reliability [14].

System validation will be conducted under diverse environmental conditions, including low-light settings, noisy surroundings, and various waste presentation angles. Evaluation metrics will include classification accuracy, inference latency, and sensor precision. All candidate EfficientNet models will be evaluated on the same test set, and the final selection will prioritize accuracy and a total feedback response time under two seconds. Similar AI-based waste classification systems have reported accuracy levels ranging between 93% and 98% [15]; therefore, the proposed system is expected to achieve a minimum of 90% accuracy in real-world deployment scenarios.

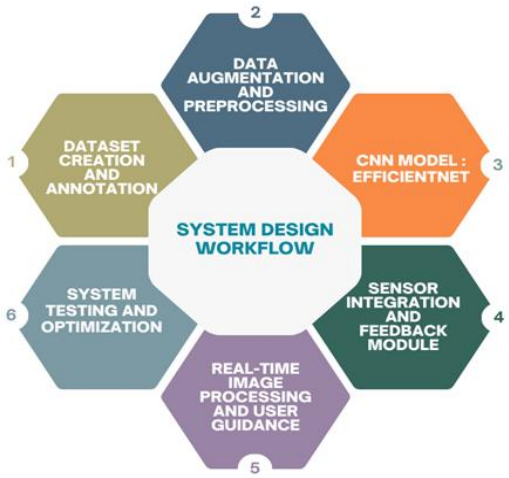


Figure 1. System design workflow

3. Expected Results and Contributions

It is anticipated that the waste classification model, based on candidate models from the EfficientNet family (B0, B2 and V2-S), will achieve a validation accuracy of at least 90% when trained on a custom dataset containing 900 labeled images of plastic, glass, and paper waste. The dataset will be designed to reflect diverse environmental conditions, including variations in lighting, background, and object orientation. To enhance generalization and mitigate the limitations of a relatively small dataset, data augmentation techniques (such as flipping, rotation, zooming, and brightness adjustments) will be applied. Additionally, transfer learning is expected to accelerate model convergence and improve overall performance with limited training data.

The real-time classification process will be performed on a Raspberry Pi 4 equipped with a camera module to capture images of incoming waste. Candidate models from the EfficientNet family (B0, B2, and V2-S) will process each image and determine the appropriate waste category. Following classification, a dual-modality feedback system will provide guidance through auditory prompts (e.g., “This is paper. Please use the yellow bin.”) and visual cues via color-coded LED lights. This approach aims to support intuitive user interaction and enhance accessibility for individuals with visual or auditory impairments.

Beyond classification, weight sensors will be integrated into each bin to monitor fill levels. When a bin reaches a predefined threshold (e.g., 80% capacity), the system will issue real-time alerts via both auditory messages and LED indicators. This feature is intended to facilitate timely waste collection and reduce the risk of overflow or user confusion.
The complete system will be evaluated under various real-world scenarios to assess robustness. Testing conditions will include low-light and bright environments, different ambient noise levels, and varied waste presentation angles. Key performance metrics will comprise classification accuracy, latency between image capture and feedback delivery, and sensor reliability. The total feedback response time is targeted to remain under two seconds to ensure a smooth and responsive user experience.

Overall, the proposed system is expected to advance the development of accessible, intelligent, and environmentally responsible recycling solutions. By integrating real-time image classification with interactive multimodal feedback and sensor-based monitoring, the system has the potential to improve waste sorting efficiency while promoting inclusive design practices in smart waste management.

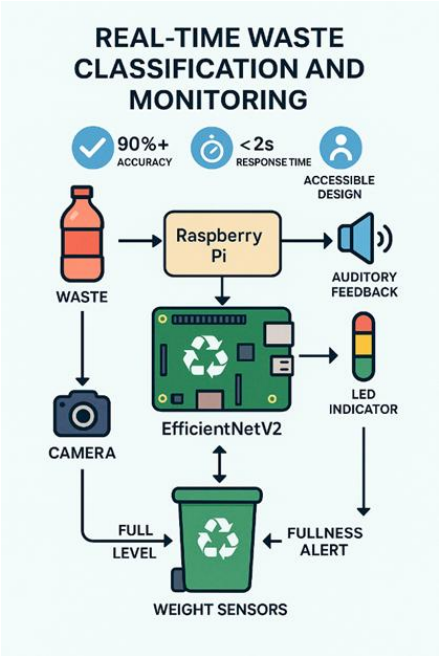


Figure 2. Real time waste classification and monitoring

4. Conclusion

This study proposes the design and development of an intelligent and inclusive waste classification and routing system that leverages deep learning and sensor-based technologies to promote effective recycling practices and environmental sustainability. The core of the system is a lightweight and efficient deep learning model, selected from candidate models in the EfficientNet family (B0, B2, and V2-S), which is planned to be deployed on a Raspberry Pi platform to ensure real-time performance under hardware-constrained environments. By incorporating both auditory and visual feedback mechanisms, the system aims to provide an intuitive user interface that is accessible to individuals with visual or auditory impairments, an aspect often neglected in conventional smart environmental solutions.

Upon implementation, the system is expected to accurately classify common recyclable waste types such as plastic, glass, and paper, and deliver immediate guidance to users through LED signals and voice prompts. Additionally, the integration of weight sensors will enable real-time monitoring of bin fill levels, supporting proactive waste management and operational efficiency. These features collectively render the system applicable in a variety of public settings, including educational institutions, shopping centers, and transportation hubs, where efficient and user-friendly waste sorting is critical.

Looking forward, several improvements are envisioned to enhance the system’s scalability and functional scope. These include the adoption of advanced object detection models such as YOLOv10 for multi-item recognition in complex scenes, the inclusion of additional waste categories such as metal, organic, and hazardous materials, and the integration of edge accelerators (e.g., Google Edge TPU or NVIDIA Jetson Nano) to improve inference speed and computational capacity. Moreover, the development of a companion mobile or web-based interface is planned to enable remote monitoring and data visualization. Long-term goals also involve the implementation of cloud-based analytics for waste trend tracking and adaptive learning mechanisms that dynamically update the model based on evolving waste patterns and user behavior. These future directions are expected to enhance the system's relevance for smart city infrastructures, contributing to more sustainable, scalable, and accessible environmental technologies.

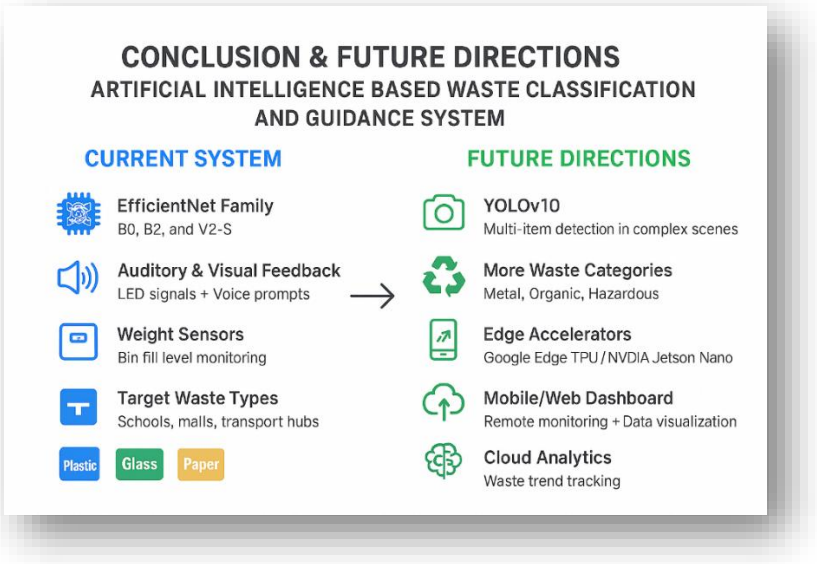


Figure 3. Conclusion and future directions

5. Acknowledgement

Our research was funded by the Scientific and Technological Research Council of Turkey (TÜBİTAK) under the 2209-A University Students Research Projects Support Program with project number 1919B012473158 This support provided a fundamental contribution to the execution of our project and is gratefully acknowledged.

6. Author Contribution Statement

All authors contributed equally to the conceptualization of this study, design of the study, literature review, data collection, analysis, and critical revision of the article. All three authors worked closely together to obtain the materials and resources needed for the study. Each author contributed significantly to the writing and editing of the article, ensuring that the final version submitted for publication met all criteria for accuracy and integrity.

7. Ethics Committee Approval and Conflict of Interest

There is no need for an ethics committee approval in the prepared article. Also, there is no conflict of interest with any person/institution in the proposed article.

8. Ethical Statement Regarding the Use of Artificial Intelligence

In the writing process of this study, the artificial intelligence tool "Claude" developed by "Anthropic" was used only for limited purposes of linguistic editing. The scientific content, analysis and results belong entirely to the authors. The entire content of the study was produced by the authors in accordance with scientific research methods and academic ethical principles.

9. References

[1] Qiu J, Xie C, Huang Y, Zhang H. “CE-Attention and SAFM-Based Waste Classification Using EfficientNetV2”. arXiv preprint arXiv:2503.21208, 2025.
[2] Li H, Chen Y, Zhou W, Liu Z. “GECM-EfficientNet for embedded waste sorting”. International Journal of Environmental Research and Public Health, 19(23), 1–14, 2021.

[3] Nafiz MS, Hossain MA, Islam MA, Rahman T. “ConvoWaste: An automated waste sorting machine using deep learning”. arXiv preprint arXiv:2302.02976, 2023.
[4] Ayman U, Rahman F, Jamal N. “EfficientNet-based deep learning model for advanced waste classification”. Proc. IEEE Student Conference on Science and Technology (SCST), 2024.
[5] Sağlam A, Taş M, Baykan N. “Development of hardware based on Raspberry Pi for classifying recyclable waste materials”. European Journal of Science and Technology, (21), 1–8, 2020.
[6] Ijaz S, Qamar M, Nazir T, Rehman A. “Improved deep learning model for organic and recyclable waste classification”. Multimedia Tools and Applications, 83(2), 23781–23803, 2024.
[7] Mustapha A, Salleh RA, Noh NM. “YOLOv8-CNN hybrid architecture for smart waste detection”. Engineering Proceedings, 84(1), 2025.
[8] Nedjar A, M’hamedi M, Bekkaoui A. “Development of a Smart Garbage Classifier Using Raspberry Pi and Deep Learning”. International Journal of Electrical and Computer Engineering Systems, 15(1), 33–40, 2024. [Online]. Available: https://ijeces.ferit.hr/index.php/ijeces/article/view/3396
[9] S. M. D’Souza et al., "Design and Development of Monitoring Smart Waste Management System," Proceedings of the International Conference on Innovative Computing & Communication (ICICC 2024), 2025. [Online]. Available: https://ssrn.com/abstract=5216026. [Accessed: Aug. 15, 2025].
[10] Tran DM, Nguyen TH, Pham TD. “Real-Time Waste Classification Using Deep Learning on Edge Devices”. Intelligent Automation & Soft Computing, 35(2), 1291–1306, 2023. [Online]. Available: https://www.techscience.com/iasc/v35n2/48912
[11] Kunwar K, Yadav D, Rawat R. “Lightweight Quantized EfficientNet Models for Resource-Constrained Devices”. Sensors, 25(12), art. 3807, Jun. 2025. [Online]. Available: https://www.mdpi.com/1424-8220/25/12/3807
[12] NVIDIA, "Accelerating Model Inference with TensorRT: Tips and Best Practices for PyTorch Users," Medium, 2023. [Online]. Available: https://medium.com/@zergtant/accelerating-model-inference-with-tensorrt-tips-and-best-practices-for-pytorch-users-7cd4c30c97bc. [Accessed: Aug. 15, 2025].
[13] Martinez-Hernandez E, Garcia L, Rios JD. “Design and Evaluation of a Smart Bin with Sensor-Based Feedback and Notifications”. Sensors, 25(10), art. 3181, May 2024. [Online]. Available: https://www.mdpi.com/1424-8220/25/10/318
[14] Anggraini D, Haryati S, Rizki A. “Multimodal Interaction System Using Raspberry Pi for Waste Sorting Applications”. Behaviour & Information Technology, 41(8), 1721–1733, Aug. 2022. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/09540091.2022.2067127
[15] Gao F, Li Y, Wang B. “Sustainable AI Solutions for Urban Waste Sorting Using Edge-Based Deep Learning”. Multimedia Tools and Applications, 83, 22351–22372, 2024. [Online]. Available: https://link.springer.com/article/10.1007/s11042-024-18939-w

Swarm UAV-Assisted System Design for Border Protection and Threat Detection

Ebru KARAKÖSE[1], Yunus Emre ERGİN[2], Hesna ÇATAL[3], [4]Sadık ÖZDEN[4]

[1,2,3,4] Aviation Electrical and Electronics Department, School of Civil Aviation, Fırat University, Elazığ, Türkiye
[1]eozbay@firat.edu.tr, [2]yeergin15@gmail.com, [3]hesnacatal@gmail.com, [4]sdkozdn2@gmail.com
[1](ORCID: 0000-0003-1191-6375), [2](ORCID: 0009-0006-5022-838X),[3](ORCID: 0009-0002-2653-0723) ,[4](ORCID: 0009-0008-2544-013X)

Abstract
This study proposes a modular and autonomous unmanned aerial vehicle (UAV) system aimed at enhancing border security through swarm-based surveillance and AI-supported threat detection. The system operates in two stages: first, mini UAVs equipped with thermal cameras and passive infrared sensors conduct continuous patrols, transmitting data via a low-energy LoRa communication network. Second, Surveillance UAVs activated only when a threat threshold is exceeded perform detailed analysis using high-resolution cameras and AI-based object classification algorithms. This layered structure ensures efficient task distribution, energy optimization, and reliable long-range monitoring in harsh or infrastructure-deficient regions. The use of open-source simulation platforms and flight control software strengthens sustainability and cost-effectiveness. By integrating real-time image processing, genetic algorithms, and Dijkstra-based path planning, the system provides flexible and rapid responses to dynamic threats. The design prioritizes field readiness through industrial-grade components and emphasizes application potential in both defense and civilian domains such as disaster response, environmental monitoring, and infrastructure inspection. Overall, the system delivers an innovative hybrid control architecture that combines autonomy with human oversight, enhancing situational awareness and operational continuity in critical border scenarios.

Keywords: UAV, LoRa, artificial intelligence, border security, surveillance systems

1. Introduction

Border security is a fundamental component of national sovereignty, critical not only for controlling irregular migration but also for preventing smuggling, terrorism, and cross-border crime. Countries with long and geopolitically sensitive frontiers, such as Türkiye, face the growing challenge of ensuring uninterrupted and accurate surveillance across vast and diverse terrains. Traditional surveillance methods such as manned patrols, fixed observation towers, and static camera systems have proven inadequate due to limited operational continuity, high costs, and slow response times. Particularly in rural and mountagüinous regions, these conventional approaches fail to provide sustainable and scalable solutions to modern threats. In response to these limitations, this study proposes a two-tiered, swarm UAV-based border surveillance architecture. The system integrates low-cost, energy-efficient mini UAVs for continuous patrolling and high-performance “Surveillance UAVs” equipped with AI-driven analysis capabilities for dynamic threat validation and classification. This separation of duties allows for an optimized distribution of computational load and energy consumption, enabling uninterrupted monitoring and timely threat response. The primary objective of the proposed system is to redesign border surveillance infrastructure by replacing labor-intensive methods with a modular, autonomous, and intelligent architecture. Unlike previous approaches, the system leverages LoRa-based long-range communication to operate independently of existing network infrastructure, thus maintaining operational efficiency even in areas with electronic jamming or weak connectivity [1, 2]. Moreover, by incorporating real-time data processing using Python and YOLOv8 object detection algorithms, the system enhances situational awareness and reduces human error in threat assessment [3, 4].

Similar efforts have been made in scientific research and industry. For example, Tsouros et al. presented a layered surveillance model consisting of observational UAVs, processing centers, and intervention UAVs [5]. However, many existing solutions remain dependent on Wi-Fi or mobile networks and lack the mission-specific task differentiation offered in this study. Unlike the homogeneous UAV roles often found in various scientific projects, the current design distinguishes between patrolling and analytical UAVs, increasing both efficiency and performance [6]. Additionally, the dynamic “relay race” task assignment method ensures uninterrupted coverage even if a UAV becomes non-operational, surpassing the fixed-route patrol schemes used in prior systems [7, 8].

In sum, the main contribution of this study lies in its hybrid control logic, infrastructure-independent operation, and modular design, which together propose a sustainable and adaptable approach to emerging security threats. The

system's adaptability also positions it for use in disaster response, environmental monitoring, and critical infrastructure protection, thereby broadening its applicability to civilian domains.

2. System Design and Architecture

In this study, a two-tier UAV system is developed to enhance border security. The system consists of two distinct types of UAVs, each assigned different tasks: mini surveillance UAVs and Surveillance UAVs. Mini UAVs conduct patrol flights along the border, collecting environmental data. When a potential threat is detected, the data is transmitted to a central control unit. Based on the analysis performed at the center, a Surveillance UAV equipped with advanced hardware is dispatched to the threat location if necessary. The system architecture aligns with the structure described by Tsouros et al, which identifies three main components of a typical UAV-assisted border surveillance system: observation UAVs, a data processing center, and intervention UAVs [5]. This division into observation, data processing, and intervention components parallels the dual-layered structure developed in this project, where mini UAVs are responsible for routine surveillance and surveillance UAVs are deployed only when intervention is required. Mini surveillance UAVs are multirotor platforms optimized for low-cost, short-duration missions. According to Dursun and Demir, thermal cameras provide improved night vision and are less affected by environmental conditions when used in combination with RGB cameras [3]. Therefore, thermal imaging is prioritized in the system. The primary tasks of the mini UAVs are to patrol areas near the border, process sensor data, and alert the control center when a predefined threat threshold is exceeded. Each mini UAV is equipped with the components listed in Table 1.

Table 1. Mini UAV hardware specifications

Electronic Components	Brand/Model
Microcontroller Board	ATmega32-based Arduino compatible processor
Communication	LoRa communication module (low power, long range)
Location Tracking	Real-time coordinate tracking with GPS module
Imaging	Thermal Camera
Motion Detection	PIR sensor (thermal motion detector)
Energy	Battery suitable for approximately 30-45 minutes of flight time

Data is securely transmitted to the central unit over long distances with minimal energy consumption using the LoRa module. Çelik and Aydın state that LoRa communication can be effectively used in UAV missions with low bandwidth requirements without data loss, and significantly reduces energy consumption [2]. Güler, in surveillance UAVs, the use of the Jetson Nano platform enables onboard image processing, which reduces system latency and improves response time [4]. Kuşhan emphasizes that UAV systems with high-precision targeting and rapid response capabilities can enhance border security [9]. These characteristics are reflected in Surveillance UAVs with advanced imaging and processing features used in this study. These UAVs are only activated upon confirmed threats and are typically VTOL or high-end multirotor platforms equipped with extended flight endurance and advanced imaging systems. They are directed to the target area based on commands from the control center and collect detailed data from the field. Their hardware specifications are listed in Table 2. This structure aligns with the centralized task allocation model proposed by Güngör and Kırcalı. The authors emphasize that, in swarm-based UAV systems, central task assignment optimizes resource utilization and enhances overall system performance [10].

Table 2. Surveillance UAV hardware specifications

Electronic Components	Brand/Model
Imaging	High-resolution thermal + optical camera
Processor	AI-supported processing unit based on Jetson Nano / Raspberry Pi
Software	Target tracking algorithms and object classification model
Energy	Battery suitable for a 30–45 minute mission duration

2.1. Subsystems and Implementation

Tsouros et al. emphasize that customizing UAVs based on mission-specific profiles, rather than relying on systems with fixed operational roles, enhances overall system performance [5]. The combined use of low-cost UAVs for surveillance and high-capability platforms for intervention improves operational effectiveness in border security scenarios that require wide-area monitoring and rapid response. Since each mission type demands different flight durations, maneuverability, and sensor configurations, selecting mission-appropriate UAVs is critical for system efficiency. Therefore, UAV types were selected according to mission definitions, as shown in Table 3.

Table 3. UAV type according to mission type

Mission	UAV type	Reason
Patrol and observation	Multicopter	Short-distance missions, low cost and easy control
Detailed analysis/follow-up	VTOL/ Hexacopter	Air stability+Long range+fast response
Long range area scanning	Fixed Wing UAV(Optional)	Low energy scanning over a wide geography

Tsouros further suggests using small multirotors for short-range patrols and fixed-wing systems for long-range operations [5]. Güler states that performing real-time object detection using the OpenCV library directly on-board surveillance UAVs reduces system latency and decreases central processing load [4]. Similarly, Dursun and Demir report improved performance in systems where images are transmitted to a central unit and processed using Python/OpenCV [3]. This centralized analysis approach accelerates decision-making. Images captured by mini UAVs are sent via LoRa to a central server, where the YOLOv8 algorithm identifies objects such as humans, vehicles, or animals. When a threat is detected, the system enters alert mode and autonomously dispatches a surveillance UAV. This structure aligns with the AI-supported centralized border security model proposed by Tsouros, which emphasizes faster responses and reduced decision errors [5]. Landing site detection and obstacle analysis follow the methodology outlined by Karaköse & Aksu [11]. This process includes: Capturing RGB/video data of the landing area from multiple angles and altitudes using a quadrotor; Labeling runway and obstacle objects via RoboFlow to prepare for YOLOv8; Performing object detection in real-time using YOLOv8; Analyzing the detected segments using point cloud data derived from LIDAR, TLS, or ALS sensors, with Z-axis deviations used to assess landing feasibility.

Karaköse, for multi-UAV task allocation, integrates genetic algorithm (GA) and Dijkstra's algorithm [12]. Each individual in the GA is encoded as a set of UAV–target assignments. Elitism (50%), uniform crossover, and a 5–10% mutation rate maintain solution diversity. Dijkstra's algorithm converts the mission environment into a node-edge graph and computes the shortest path using obstacle-based distances. The objective function is designed to balance priority-driven allocation with energy efficiency by combining a score derived from target importance and the number of UAVs assigned with the total travel distance of the UAVs. The optimal solution is deployed in real time to UAVs. According to the 2021 Border Security Report published by the Turkish National Police (EGM), Turkey's southeastern borders face significant security threats including smuggling, illegal crossings, and terrorism [13]. The report emphasizes the operational challenges of 24/7 human surveillance and highlights unmanned systems as a more effective solution.The proposed system directly responds to the needs identified by the EGM. UAVs provide continuous monitoring in inaccessible areas, and their advanced capabilities offer significant advantages for enhancing border security.

Figure 1 illustrates the interaction between the three core components of the system: the Surveillance UAV, the Mini Patrol UAV, and the Central Analysis Server. The Mini UAVs continuously scan the environment using thermal cameras. Upon detecting suspicious thermal anomalies or movement patterns, the relevant data is transmitted to the central server via the LoRa communication protocol. The received data undergo preprocessing with Python/OpenCV infrastructure, followed by threat analysis using the YOLOv8 object detection algorithm. If a predefined threshold is exceeded, the system automatically dispatches a mission command to the Surveillance UAV. The assigned UAV activates its high-resolution thermal and optical sensors to collect detailed imagery, radar, and telemetry data from

the target area. This information is transmitted back to the server via LoRa, where real-time visualization and automatic threat classification are performed. If no threat is confirmed, the Mini UAVs resume their patrol operations.

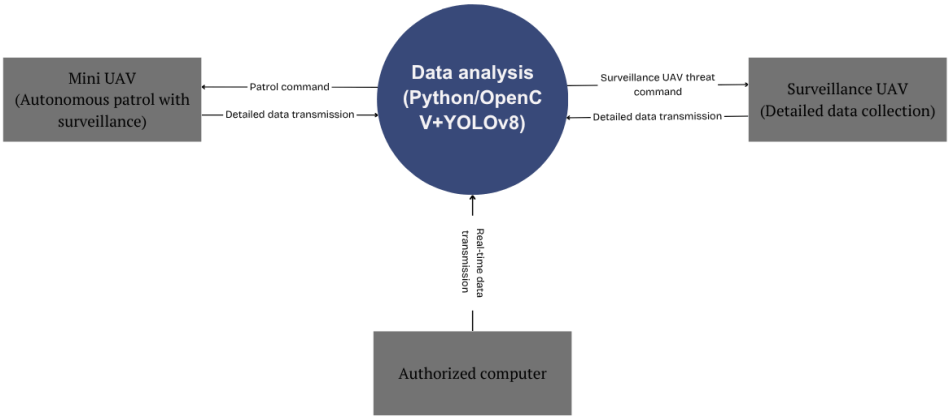


Figure 1. System architecture

2.2. Image Processing and Optimization

The proposed UAV-based border surveillance system adopts a two-layered, task-specific architecture designed to enhance both operational efficiency and sustainability in dynamic field conditions. The system architecture distinguishes between two primary UAV types: mini surveillance drones responsible for continuous environmental monitoring, and Surveillance drones tasked with high-accuracy threat verification and classification. In the first layer, autonomous mini UAVs conduct periodic patrols along predefined routes. Each unit is equipped with thermal cameras, PIR motion detectors, GPS modules, and LoRa-based communication systems that enable long-range, low-power, infrastructure-independent data transmission. These drones continuously monitor their assigned area and transmit thermal anomalies or movement pattern data to a centralized ground server. The independence of each unit allows for parallel operation and minimizes communication bottlenecks, increasing both scalability and robustness. Upon detection of a threshold-exceeding event, the system transitions to the second layer: a high-performance Surveillance UAV is dispatched to the suspected region. These UAVs feature enhanced processing units (e.g., Jetson Nano, Raspberry Pi 4) and high-resolution thermal-optical cameras. They collect detailed telemetry and imagery, which are relayed back to the server for real-time analysis using Python/OpenCV frameworks and YOLOv8-based classification models. This hierarchical processing approach reduces onboard computation demands and enables fast, reliable decision-making supported by AI. Figure 2 provides a schematic representation of the system's layered architecture and operational workflow, including modular deployment options.

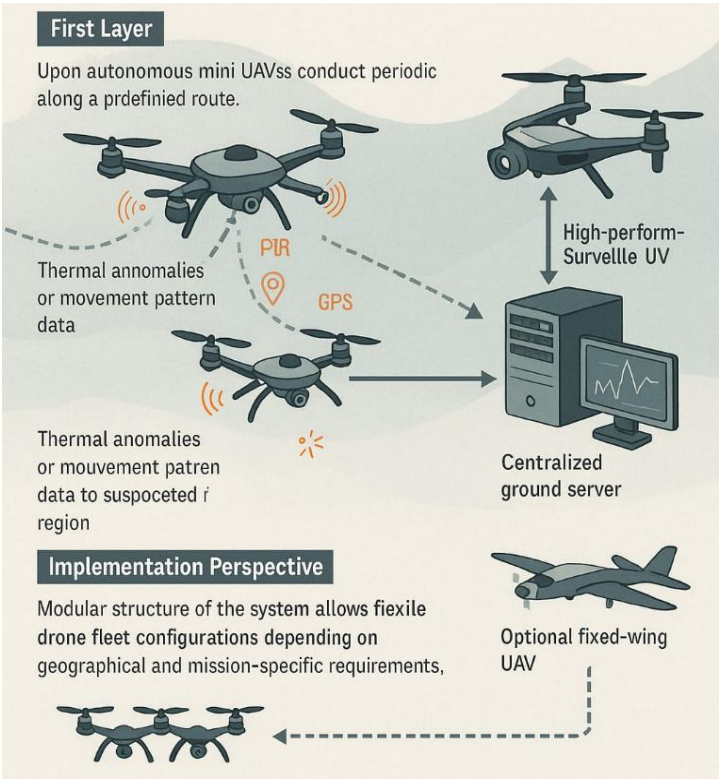


Figure 2. Schematic representation of the system's layered architecture and workflow

From an implementation perspective, the modular structure of the system allows for flexible drone fleet configurations depending on geographical and mission-specific requirements. Simulation-based assessments (e.g., using PX4 and AirSim) have validated system functions including path planning, detection latency, and energy optimization. According to Çaşka and Özdemir, a fleet of 8–10 mini drones is sufficient to provide continuous monitoring over areas containing up to 40 observation points, with each drone consuming approximately 0.095% battery per meter flown [14]. This balance between drone count and task distribution ensures continuous operation, even in cases of partial system failure.

The feasibility of the system is further reinforced through the integration of advanced swarm coordination strategies, combining Genetic Algorithm (GA)-based task allocation with Dijkstra-based path planning. These enhancements strengthen the system's capability to perform autonomous and energy-efficient multi-UAV mission execution in complex environments. These algorithms ensure optimal task assignment based on real-time threat density and environmental obstacles, minimizing redundant flight paths and energy consumption. Additionally, the system architecture supports future scalability, including optional fixed-wing UAV integration for wide-area reconnaissance missions. Sustainability is addressed through energy-efficient flight hardware, the use of open-source software, and the potential for integration of hybrid energy systems such as solar-assisted charging in future iterations. The inclusion of LoRa communication significantly enhances resilience against signal jamming or infrastructure collapse, a key advantage in remote or conflict-prone regions. Moreover, environmental sustainability is supported by the drones' low acoustic and thermal signatures, allowing for wildlife-friendly and non-invasive operation. The use of recyclable materials and modular replacement parts also contributes to lifecycle extension and cost-effectiveness. In conclusion, the proposed UAV swarm system is technically feasible, operationally robust, and sustainable across a range of defense and civilian applications. Its modularity, real-time AI integration, and scalable communication infrastructure position it as a compelling solution to modern border surveillance challenges.

3. Experimental Results

This section presents the results derived from a series of simulation-based scenarios designed to evaluate the operational performance, energy consumption, task distribution efficiency, and communication reliability of the proposed swarm UAV system. Each simulation focuses on a distinct aspect of the system's architecture, providing insight into its practical applicability for real-world border surveillance operations. Emphasis is placed on endurance validation, modular task allocation, and the robustness of long-range wireless communication under varying environmental conditions.

3.1. Mission Endurance and Battery Consumption Scenario

To evaluate the mission sustainability of the proposed Mini UAV fleet, a simulation scenario involving eight consecutive observation points over a 5-kilometer patrol route was conducted. In this scenario, the UAV followed a predefined linear path at a constant speed of 10 m/s while performing thermal scanning and motion detection tasks. As shown in Figure 3, the cumulative flight time and battery consumption do not follow a perfectly linear trend. Instead, they exhibit minor fluctuations introduced intentionally to simulate real-world anomalies such as wind variations, GPS drift, and sensor latency. These deviations represent realistic environmental inconsistencies and demonstrate the system's robustness under non-ideal conditions. Overall, the UAV is able to complete the patrol mission with an average battery usage of 22.2% and a duration of approximately 16.7 minutes, confirming its operational viability with adequate energy reserves for extended missions or emergency scenarios.

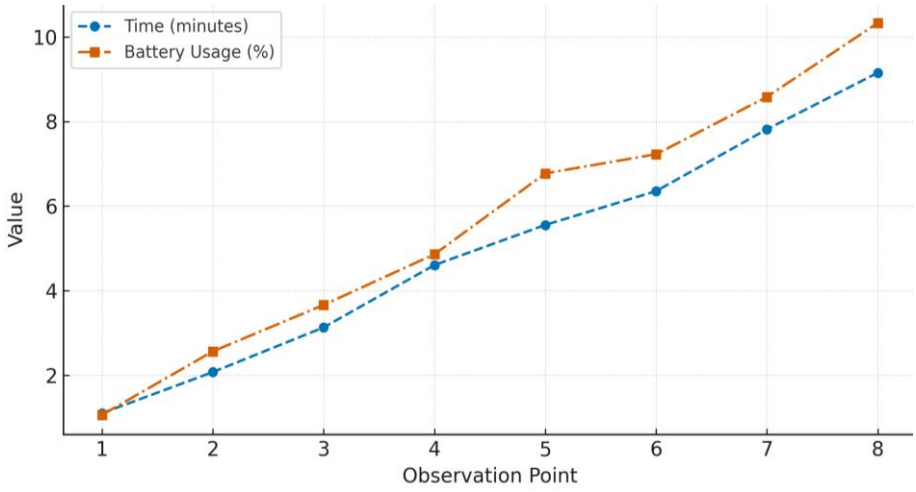


Figure 3. Varying flight time and battery consumption across observation points

3.2. Task Distribution Efficiency Scenario

This scenario focuses on the comparative workload between lightweight Mini UAVs and high-performance Surveillance UAVs within the dual-tier system. Mini UAVs perform routine patrols, covering 50,000 meters collectively, transmitting basic thermal data via LoRa. In contrast, Surveillance UAVs are activated selectively for threat verification missions, capturing high-resolution imagery and performing AI-based object classification on-board. Figure 4 illustrates the disparity in data load, processing time, and battery consumption between the two UAV types. Surveillance UAVs process 10x more image data and consume approximately 35% battery per mission despite flying shorter distances. This modular allocation enables energy-efficient execution of routine tasks, while allocating high-power resources to critical interventions, thereby supporting effective system-wide resource utilization.

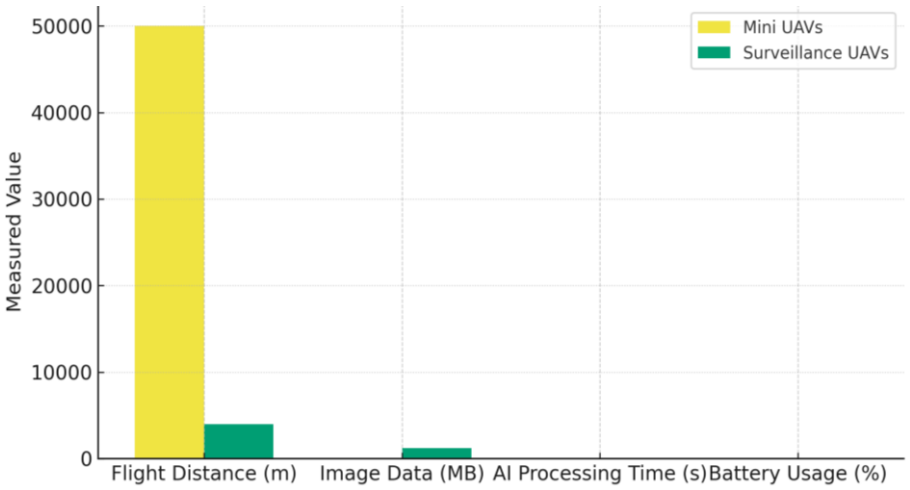


Figure 4. Comparative task load between Mini and Surveillance UAVs

3.3. Communication Reliability Scenario – LoRa Performance Analysis

A third scenario was developed to assess the reliability of LoRa-based communication across different environmental conditions. Given the system's reliance on LoRa for low-power, long-distance data transmission, understanding how terrain and distance affect communication performance is vital. Figure 5 presents a heatmap visualizing the Packet Delivery Ratio (PDR) across three terrain types (Open Field, Forest, and Mountain) over six distances ranging from 500 to 5000 meters. Results were based on simulation models incorporating empirical path loss and signal attenuation factors. At short distances (≤ 2000 meters), the system consistently maintains high reliability across all terrains. However, as distance increases, significant degradation occurs, especially in obstructed environments. In mountainous terrain, the PDR drops below 50% beyond 3000 meters due to signal shadowing and multipath interference. Forested areas show moderate decline, while open fields allow effective communication up to 4000 meters. These findings indicate that terrain-aware routing and mission planning are critical for ensuring uninterrupted data transmission under real-world conditions. Hybrid communication models or relay UAVs may be necessary in areas where signal quality drops below critical thresholds.

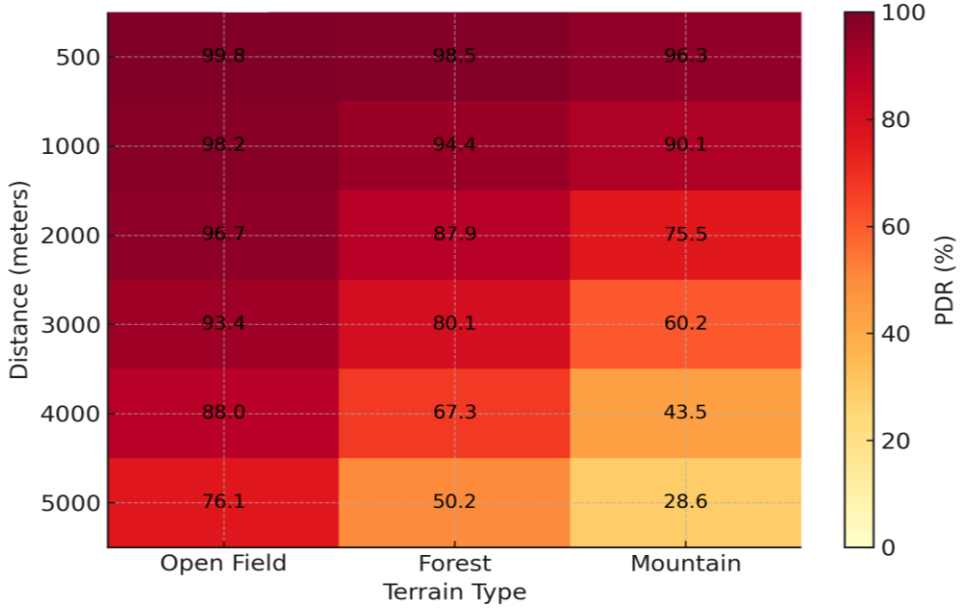


Figure 5. Packet Delivery Ratio (PDR) heatmap for LoRa communication under various terrain types and distances

4. Conclusion and Discussion

This study introduces a two-tier swarm UAV architecture tailored for border security and threat detection. By integrating energy-efficient mini UAVs for persistent thermal surveillance with specialized Surveillance UAVs for detailed analysis and AI-driven classification, the system ensures effective task separation, energy optimization, and rapid response capabilities. The use of LoRa-based communication, real-time object detection through YOLOv8, and path planning via genetic algorithms and Dijkstra optimization further strengthens operational autonomy and situational awareness. Owing to its modular structure, the system can be adapted for diverse scenarios such as smuggling prevention, illegal migration monitoring, disaster response, and infrastructure inspection.

The proposed hybrid UAV framework offers a scalable and efficient approach to persistent border surveillance. Routine patrol missions are delegated to lightweight mini UAVs, conserving energy and reducing data loads, while advanced analytics and classification tasks are assigned to AI-equipped Surveillance UAVs. This division enhances operational efficiency and resource management. Furthermore, the implementation of LoRa communication and open-source software platforms contributes to the system's resilience, especially in environments with limited infrastructure. The reliance on locally sourced hardware components supports cost-effective deployment and encourages national technological development.

Future developments will aim to enhance the simulation environment, integrate hybrid power systems, and carry out field tests to assess real-world applicability. The system directly addresses national security concerns such as illegal migration and cross-border smuggling, while also remaining flexible enough for use in emergency and civilian applications. Its dual-mode operation autonomous and operator-assisted ensures both swift responsiveness and human oversight, making it a reliable and sustainable solution for contemporary surveillance demands. In conclusion, the proposed design not only demonstrates technical viability but also promises significant national and societal benefits as a next-generation aerial surveillance solution.

5. References

[1] Gebze Technical University. "Border Security and Technologies". *Official Website of GTU*, 2023.
[2] Çelik H, Aydın Ü. "Kablosuz haberleşme sistemlerinde LoRa ve ZigBee teknolojilerinin karşılaştırılması". *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, 13(2), 613-623, 2022.
[3] Dursun M, Demir H. "Design of an unmanned aerial vehicle (UAV)-based surveillance system for border security". *Journal of Visual and Electronic Technologies*, 10(1), 42-48, 2020.
[4] Güler A. "Design of a UAV-based border security system and real-time image processing application". *Gümüşhane University Journal of Institute of Natural Sciences*, 12(2), 324-332, 2022.
[5] Tsouros DC, Bibi S, Sarigiannidis P. "Drone-aided border surveillance systems: A review". *Drones and Autonomous Vehicles*, Elsevier, 2020.
[6] TÜBİTAK. "2021 annual activity report". *The Scientific and Technological Research Council of Turkey*, 2022.
[7] Tarakçı K. "Planning and optimization of airspace surveillance missions in a homogeneous UAV fleet". *Master's Thesis*, Middle East Technical University, 2021.
[8] Tezcaner Öztürk D, Köksalan MM. "UAV route planning problem for multi-target tracking: Optimization approaches in radar-based monitoring". Middle East Technical University, 2023.
[9] Kuşhan M. "Kamikaze unmanned aerial vehicles". *Journal of Defense Sciences*, 20(2), 345-368, 2021.
[10] Güngör A, Kırcalı ME. "A swarm UAV mission assignment and planning system for search and rescue applications". *Journal of Unmanned Vehicle Systems*, 6(4), 252-265, 2018.
[11] Karaköse E, Aksu ME. "Landing runway suitability analysis using point cloud data in unmanned aerial vehicles". *Fırat University Journal of Engineering Sciences*, 36(2), 535-551, 2024.
[12] Karaköse E. "A genetic algorithm-based approach for task allocation in swarm unmanned aerial vehicles". *Fırat University Journal of Engineering Sciences*, 34(1), 351-360, 2022.
[13] General Directorate of Security (EGM). "Border security report". *Police Academy Publications*, 2021.
[14] Çaşka S, Özdemir İ. "Calculation of the optimum number of unmanned air vehicles required for surveillance missions". *Academic Platform Journal of Engineering and Smart Systems*, 10(2), 101-105, 2022.

Artificial Intelligence Design Paradigms: Review of AI-Native and Artificial General Intelligence (AGI) Systems

Merve YILMAZER[1] , Mehmet KARAKÖSE[2]

[1]Department of Computer Engineering, Faculty of Engineering, Munzur University, Tunceli, Türkiye.
[2]Department of Computer Engineering, Faculty of Engineering, Fırat University, Elazığ, Türkiye.
[1]merveyilmazer@munzur.edu.tr, [2]mkarakose@firat.edu.tr
[1](ORCID: 0000-0001-7055-6299), [2](ORCID: 0000-0002-3276-3788)

Abstract
The rapidly developing artificial intelligence technology has been made available to people in many areas of daily life. Artificial intelligence technologies, which have found applications in many areas from education to transportation, from health to finance, continue to develop every day. Today, what is expected from artificial intelligence is to reach the competence of perception, problem solving and producing effective solutions to unexpected situations with human-like abilities with increasing learning capacity. In this study, the details, areas of use and progress of the two basic artificial intelligence paradigms, Artificial General Intelligence (AGI) and AI-Native approaches are given in detail. AI-Native systems are trained for specific tasks and have limited learning capacity. AGI, on the other hand, are systems that can create their own rules with a wider learning capacity and can produce solutions even for unencountered situations. In addition, the study includes a comparison of the systems that are desired to be reached with today's technologies in terms of traffic management, anomaly detection, learning capacity, decision-making speed, human intervention and cost on an example application such as traffic management in smart cities. Additionally, a comparison of the existing AI-Native System and the AGI Based System in the field of generating large language model based software test scenario is also given.

Keywords: artificial general intelligence (agi), ai-native, large language model (llm)

1. Introduction

Artificial intelligence (AI) has attracted great attention in both academic and industrial circles in recent years and has become a fundamental force in many sectors. Adversarial Networks (GAN) [1], diffusion models [2], and Transformer architecture [3] have played a major role in the rapid development of artificial intelligence in recent years. Developments in the fields of image processing, natural language processing, and machine learning have led to an increase in the use of artificial intelligence. Artificial intelligence technologies used in various stages of daily life such as education, transportation, healthcare, and e-commerce are narrow or specialized artificial intelligence applications that can function in certain areas. Examples of these are anomaly detection in smart cities [4], deep learning-based image processing applications in industrial areas [5], and large language model-based test scenario generation [6] applications. The desired development is for artificial intelligence applications that offer comfortable solutions that facilitate people's daily activities to produce solutions with more human-like reactions. As AI systems gain the ability to solve increasingly complex problems, two basic paradigms come to the fore: Artificial General Intelligence (AGI) and the AI-Native approach.

Artificial General Intelligence (AGI) refers to systems that have human-like general intelligence, capable of learning, reasoning, and problem solving not only for a single subject but also for a variety of subjects. The main goal of AGI is to develop autonomous intelligent systems that can go beyond models optimized for narrow tasks and cope even in the face of unlearned situations. This approach offers a vision of the future where machines can behave like conscious or semi-conscious beings, but it also brings with it serious discussions such as ethics, security, and control [7].

AI-Native encompasses technologies with reliable artificial intelligence capabilities. Embedded-AI is the incorporation of artificial intelligence into systems to perform specific tasks. Examples of Embedded-AI include face recognition on smartphones and energy saving applications on white goods. In AI-Native, the product or service is designed from the ground up based on artificial intelligence. Chatbot applications that have problem-solving and question-answering features can be given as examples of AI-Native [8]. The advances in the field of artificial intelligence in the last 10 years are given chronologically in Figure 1. These are:

[1]Corresponding author

Generative Adversarial Networks (GAN) [1], Transformer [3], BERT [9], GPT-3 [10], CLIP [11], Stable Diffusion [12], IstructGPT [13], BLIP [14], ChatCPT [15], Glam [16], PaLM [17], LLaMA [18], SAM [19], Alpaca [20], LLaMA 2 [21], GPT-4 [22], KOSMOS 1 [23], KOSMOS 2 [24].

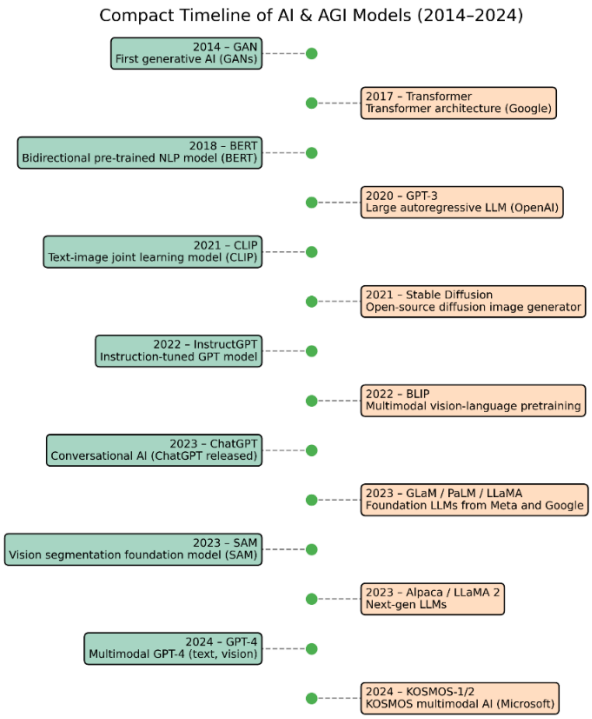


Figure 1. The development of artificial intelligence in the last decade

Various studies have been conducted in the literature using AI-Native and AGI systems developed to improve artificial intelligence applications widely used in many areas of daily life and to gain high and diverse learning skills. Some of these studies are as follows:

Shang et al. [25] proposed the concept of AI-Native memory in the evolution of LLMs to AGI. In the memory system they proposed, in addition to the raw data, they will also store a large number of important results produced from reasoning processes. Thus, they aimed to increase the learning and reasoning capacity of LLM. Saad et al. [26] mentioned the difficulty of building next-generation wireless systems that can support metaverse services with only traditional wireless technologies such as metasurfaces and holographic antennas. They proposed AGI-Native systems to be able to include complex, non-trivial scenarios that arise in the network environment in AI-native systems that can be used to overcome this difficulty. They stated that in this way, the AGI-Native system they proposed will gain various cognitive abilities such as reasoning and perception of wireless networks, will be able to generalize unpredictable situations by making common-sense decisions and will be able to cope with these situations effectively. Dou et al. [27] discussed the opportunities and challenges that will arise as a result of the structures to be created by the integration of Artificial General Intelligence (AGI) and Internet of Things (IoT) systems. They also discussed large-scale communication structures and security-privacy issues. Li et al. [28] emphasized that different and multiple disciplines such as text, vision, robotics, emotion recognition should be used together to create general-capability AI agents on the way to AGI.

When the studies in the relevant field are examined in the literature, it is seen that with the introduction of 6G technology into our lives, the focus is on applications that demonstrate that AI-Native and AGI systems can be used in an integrated manner. In this study, the advances in artificial intelligence in recent years and the features of AI-Native and Artificial General Intelligence (AGI) systems are included. In addition, a comparison of these two systems is made through the example application of traffic planning in smart cities.

2. AI-Native Design and Artificial General Intelligence (AGI)

While AI-Native refers to systems based on artificial intelligence, Artificial General Intelligence (AGI) refers to artificial intelligence that has the potential to think in a multi-dimensional way similar to human intelligence and can learn and implement different tasks. Models such as GPT-4, PaLM 2 and SAM, which are strong steps towards reaching AGI, are more dynamic than current artificial intelligence models. However, despite this, they do not yet have the potential to learn new and completely different tasks like humans [29]. AI-Native systems, the basic architecture of which is given in Figure 2, are currently used in systems such as chatbots, software developers and fully autonomous driving, which are completely composed of artificial intelligence.

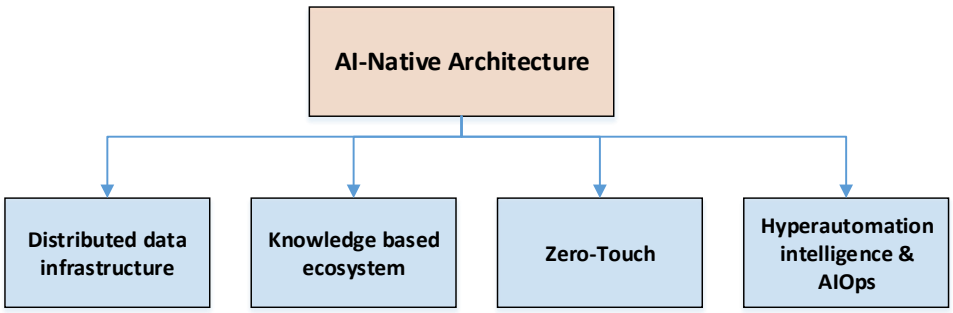


Figure 2. AI-Native Architecture [30]

Distributed data infrastructure is an architecture that enables data to be stored and processed in multiple distributed layers such as cloud, edge, fog, etc. instead of a single center. Thanks to this distributed structure, AI-Native components can perform model training at the edge points of the network without the need for centralized data. In addition, the distributed data structure allows for strong security requirements to be met.

Knowledge-based ecosystem creates shared information resources by transforming data into meaningful information, not just in raw form. Data is produced and consumed continuously and in real time in all locations.

Zero-Touch refers to the activation of systems with a completely autonomous network infrastructure without human intervention. Within the scope of Industry 5.0, systems are being developed to automatically provide the artificial intelligence lifecycle from the edge to the cloud without human intervention (zero-touch).

Hyperautomation intelligence, in this structure, intelligence capabilities are integrated into end-to-end systems and operations. Automation in these systems is completely data-driven and easily scalable. AIOps (Artificial Intelligence-Enhanced IT Operations) replaces manual IT management tasks, automating processes such as event correlation and anomaly detection with artificial intelligence and machine learning. In this way, operations are continuously optimized [30, 31]. In summary, the benefits provided by AI-Native Architecture components are given in Table 1.

Table 9. AI-Native Architecture components and their benefits

Component	Benefit of Component
Distributed data infrastructure	It enables data and processes to be organized in different regions and system layers and provides a high-performance and easily scalable structure.
Knowledge based ecosystem	It enables decision-supported information sharing by transforming data sets into meaningful information.
Zero-Touch	It ensures that all processes are carried out autonomously.
Hyperautomation intelligence and AIOps	It provides artificial intelligence systems that continuously monitor and manage operations, analyze anomalies and automatically intervene.

The effort to reach Artificial General Intelligence (AGI) has been the main goal of AI research. Transferring the potential to produce solutions with human-like intelligence even on data that has not been presented to machines before forms the basis of AGI. This means that machines will not only be able to make predictions on the data they are trained on, but will also be able to make predictions on situations beyond this data. Although existing large language models (LLMs) are trained with billions of parameters, their ability to understand and produce solutions is still limited to data sets. As the network generation that can form the basis of AGI, 6G is the main component that will ensure compatibility between technologies, as seen in Figure 3 [32].

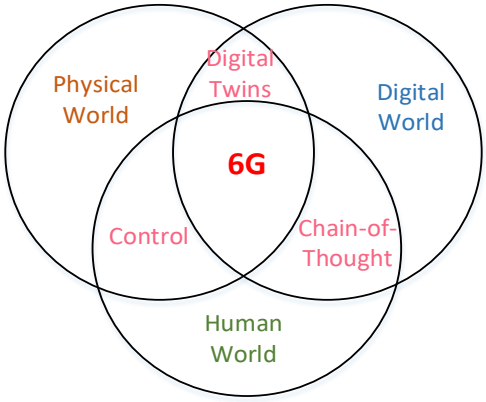


Figure 3. From 6G to AGI [32]

The usage areas of AI-Native and Artificial General Intelligence (AGI) systems are given in Table 2, and their comparison in terms of technical architecture and data processing is given in Table 3.

Table 2. Areas of use of AI-Native and AGI-based systems

Field / Application Type	AI-Native System	AGI Based System
Health	Medical image analysis, drug discovery, patient monitoring systems.	Diagnosis and treatment planning equivalent to human doctors (not yet available).
Education	Personalized learning platforms and exam evaluation systems.	Creating and teaching completely personalized curricula.
Defense and Security	Threat detection, autonomous surveillance systems.	Strategic planning and decision-making, tactical command (high ethical concerns).
Research and Science	Big data analysis, hypothesis generation, simulation studies.	Innovations in scientific discoveries that mirror human creativity.
Daily Life	Smart assistants (Alexa, Siri), recommendation systems, automatic translation.	Human equivalents for all daily tasks (still theoretical).

Table 3. Comparison of AI-Native and AGI-based systems in terms of architecture

Criterion	AI-Native System	AGI Based System
Architectural Approach	Usually optimized for a single or limited number of tasks, it is built on specific model types such as LLM, CNN, and Transformer.	A flexible and complex architecture that can combine multiple paradigms (symbolic + connectionist + hybrid + evolutionary) within a single structure.
Data Processing Method	Trained with task-specific datasets; typically, offline training + online fine-tuning.	Fully online and adaptive learning that can learn from any type of data and transfer across contexts.

Learning Ability	Although continuous learning is possible, it often requires retraining with new data.	Human-like continuous and self-improving learning; the ability to abstract from experience.
Hardware Requirements	High-performance GPU/TPU, cloud infrastructure; can be optimized for specific tasks.	Potentially much higher processing power; may require HPC (High Performance Computing) or quantum computing.

The comparison of traffic planning in smart cities with AI-Native and AGI-Based systems is given in Table 4. Accordingly, it has been determined that although the AI-Native systems implemented today have limited learning capacity for predetermined events, they require fast and low processing costs. It has been determined that the desired AI structure, Artificial General Intelligence (AGI), may have the potential for extensive self-learning and solution generation in different situations, but it can do this more slowly and with higher computational costs.

Table 4. Comparison of AI-Native and AGI-based systems

Features	AI-Native System (Used Today)	AGI Based System (Scenario)
Traffic management	Decisions are made using ML and DL algorithms based on data from cameras and sensors.	It can interpret the components of the environment and create its own rules according to new situations.
Anomaly detection	Identifies specified limited conditions for specified events.	It may suggest alternative solutions for situations where it has not been produced before.
Learning capacity	It has limited learning capacity on certain data.	Can demonstrate general learning capacity from a wide range of data sources.
Decision making speed	Since it is based on fixed algorithms, it has a fast decision-making speed.	It may be slower at times as it processes more data, but it provides more comprehensive results.
Human intervention	It requires human control in situations such as rule definition, review, and monitoring.	By operating autonomously, it can provide continuous learning with minimal human intervention.
Cost	It requires less resource and processing power cost.	It may require higher processing power and more complex infrastructure costs.

The comparison of the existing Artificial Intelligence Local System and the Artificial Intelligence Based System in the field of large language model based software test case generation is given in Table 5.

Table 5. AI-Native and AI Based System in LLM-based software test case generation

Scenario	AI-Native System	AGI Based System
Test Case Generation	LLM writes test cases based on user requirements (Prompt + Few-shot learning).	AGI understands the requirements, identifies gaps, and determines its own testing strategy.
Generating Test Data	LLM generates text-based test data (e.g., fake customer data).	AGI generates text, image, and sensor data, and even measures data quality.
Automatic Remediation	LLM rewrites the scenario based on user feedback.	AGI analyzes the test failure, identifies the root cause, and optimizes the scenario.

3. Conclusions

Artificial intelligence, which was first created in the mid-20th century to imitate human-like thinking processes, has begun to be used in many areas of daily life in recent years. Today, thanks to the increase in data processing capacity, deep learning and large language models, artificial intelligence is getting closer to human-like learning and reasoning abilities. In this study, in the context of examining the progress of artificial intelligence technologies, the rise of AI-native systems and the applications in the transition process to Artificial General Intelligence (AGI) were examined. The AI-native systems used today for autonomous traffic control in smart cities and the AGI system that is desired to be achieved are given comparatively. It was determined that the AI-native system is faster and has a lower cost despite having a more limited learning capacity. On the other hand, it was determined that the AGI system will require higher cost usage despite

having a high learning potential. As a result, it was determined that these systems can be used in an integrated manner, be beneficial, fair and inclusive for humanity, with a multi-stakeholder and interdisciplinary approach.

Acknowledgment

This study was supported by the Firat University Scientific Research Projects Support Program (FUBAP) under Grant No: MF.25.07.

4. References

[1] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2020). Generative adversarial networks. Communications of the ACM, 63(11), 139-144.
[2] Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. Advances in neural information processing systems, 33, 6840-6851.
[3] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.
[4] Yilmazer, M., & Karakose, M. (2025). A New, Robust, Adaptive, Versatile, and Scalable Abandoned Object Detection Approach Based on DeepSORT Dynamic Prompts, and Customized LLM for Smart Video Surveillance. Applied Sciences, 15(5), 2774.
[5] Yilmazer, M., & Karakose, M. (2024). Fastener and rail surface defects detection with deep learning techniques. International Journal of Advances in Intelligent Informatics, 10(2).
[6] Yilmazer, M., & Karakose, M. (2025, February). LLM-Based Video Analytics Test Scenario Generation in Smart Cities. In 2025 29th International Conference on Information Technology (IT) (pp. 1-4). IEEE.
[7] Russell, S. (2019). Human compatible: AI and the problem of control. Penguin Uk.
[8] Daugherty, P. R., & Wilson, H. J. (2024). Human+ Machine, Updated and Expanded: Reimagining Work in the Age of AI. Harvard Business Press.
[9] Kenton, J. D. M. W. C., & Toutanova, L. K. (2019, June). Bert: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of naacL-HLT (Vol. 1, No. 2).
[10] Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. Advances in neural information processing systems, 33, 1877-1901.
[11] Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., ... & Sutskever, I. (2021, July). Learning transferable visual models from natural language supervision. In International conference on machine learning (pp. 8748-8763). PmLR.
[12] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 10684-10695).
[13] Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., ... & Lowe, R. (2022). Training language models to follow instructions with human feedback. Advances in neural information processing systems, 35, 27730-27744.
[14] Li, J., Li, D., Xiong, C., & Hoi, S. (2022, June). Blip: Bootstrapping language-image pre-training for unified vision-language understanding and generation. In International conference on machine learning (pp. 12888-12900). PMLR.
[15] OpenAI, "Introducing chatgpt," 2022.
[16] Du, N., Huang, Y., Dai, A. M., Tong, S., Lepikhin, D., Xu, Y., ... & Cui, C. (2022, June). Glam: Efficient scaling of language models with mixture-of-experts. In International conference on machine learning (pp. 5547-5569). PMLR.
[17] Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A., ... & Fiedel, N. (2023). Palm: Scaling language modeling with pathways. Journal of Machine Learning Research, 24(240), 1-113.
[18] Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M. A., Lacroix, T., ... & Lample, G. (2023). Llama: Open and efficient foundation language models. arXiv preprint arXiv:2302.13971.
[19] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., ... & Girshick, R. (2023). Segment anything. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 4015-4026).

[20] Taori, R., Gulrajani, I., Zhang, T., Dubois, Y., Li, X., Guestrin, C., ... & Hashimoto, T. B. (2023). Alpaca: A strong, replicable instruction-following model. Stanford Center for Research on Foundation Models. https://crfm. stanford. edu/2023/03/13/alpaca. html, 3(6), 7.
[21] Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., ... & Scialom, T. (2023). Llama 2: Open foundation and fine-tuned chat models. arXiv preprint arXiv:2307.09288.
[22] Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., ... & McGrew, B. (2023). Gpt-4 technical report. arXiv preprint arXiv:2303.08774.
[23] Huang, S., Dong, L., Wang, W., Hao, Y., Singhal, S., Ma, S., ... & Wei, F. (2023). Language is not all you need: Aligning perception with language models. Advances in Neural Information Processing Systems, 36, 72096-72109.
[24] Peng, Z., Wang, W., Dong, L., Hao, Y., Huang, S., Ma, S., & Wei, F. (2023). Kosmos-2: Grounding multimodal large language models to the world. arXiv preprint arXiv:2306.14824.
[25] Shang, J., Zheng, Z., Wei, J., Ying, X., Tao, F., & Team, M. (2024). Ai-native memory: A pathway from llms towards agi. arXiv preprint arXiv:2406.18312.
[26] Saad, W., Hashash, O., Thomas, C. K., Chaccour, C., Debbah, M., Mandayam, N., & Han, Z. (2024). Artificial general intelligence (AGI)-native wireless systems: A journey beyond 6G. arXiv preprint arXiv:2405.02336.
[27] Dou, F., Ye, J., Yuan, G., Lu, Q., Niu, W., Sun, H., ... & Liu, N. Towards artificial general intelligence (AGI) in the Internet of Things (IoT): Opportunities and challenges. arXiv 2023. arXiv preprint arXiv:2309.07438.
[28] Li, Z., Mao, H., Yin, J., Xing, M., Xu, Z., Zhang, Y., & Xiao, Y. (2025). NGENT: Next-Generation AI Agents Must Integrate Multi-Domain Abilities to Achieve Artificial General Intelligence. arXiv preprint arXiv:2504.21433.
[29] Liu, C., Liu, Z., Holmes, J., Zhang, L., Zhang, L., Ding, Y., ... & Liu, W. (2023). Artificial general intelligence for radiation oncology. Meta-radiology, 1(3), 100045.
[30] Raza, M. (2025). “What is AI-Native”. https://www.splunk.com/en_us/blog/learn/ai-native.html (23.06.2025)
[31] Alberti, E., Alvarez-Napagao, S., Anaya, V., Barroso, M., Barrué, C., Beecks, C., ... & Walter, S. (2024). AI Lifecycle Zero-Touch Orchestration within the edge-to-cloud continuum for industry 5.0. Systems 12 (2).
[32] Bariah, L., & Debbah, M. (2024). AI Embodiment Through 6G: Shaping the future of AGI. IEEE Wireless Communications.

Development of a Rheological Model for Drilling Muds and Determining Their Rheological Properties

Fethi KAMIŞLI[1]

[1]Department of Chemical Engineering, Faculty of Engineering, Firat University, Elazig, Turkey.
[1]fkamisli@firat.edu.tr
[1](ORCID:0000-0002-1769-3785)

Abstract
The use of suitable drilling mud has a positive effect on the economy of drilling processes as it facilitates operations by lubricating the bit and drill spring, cooling the drill pipe, sealing wellbore walls, carrying the formation cuts to the surface, controlling the formation pressure and maintaining the wellbore stability. In this study, a rheological model was proposed to describe the rheological behavior of drilling muds. The twelve data series with more than 33 experimentally obtained data pairs were used to determine rheological parameters of the proposed model. The parameters of the proposed model for each data series were determined using the experimental data for that drilling mud. The results for the shear stresses versus shear rates obtained from the proposed model here were compared with the results obtained from another model whose parameters were computed using the same experimental data. The data gained from the model proposed here were compared with both experimental data and theoretical data obtained from the existing related literature, and it was seen that the agreement between the mentioned data was fairly good. As a result, it can be said that the rheological model proposed here can be easily used to determine the rheological properties of drilling muds.

Keywords: rheological model, rheological behavior, rheological model

1. Introduction

The rheological behavior of drilling mud is a quite important for efficiency of drilling process since it affects performance and the cost of drilling. Therefore, the rheological properties of drilling mud must be evaluated to choice the most convenient drilling fluid that has some incumbency such as lubricating the bit and drill string, consolidating the permeable formation, cooling drill pipe, controlling formation pressure and transporting the formation cuts to surface. Hydroxyethyl cellulose (HEC) and carboxymethyl cellulose (CMC) are used in the drilling fluid containing clay as viscosifiers. Addition of biopolymer into the drilling fluids increases carrying capacity of drilling mud since they have elastic structure and their stability is attached to salinity of the solution at a certain temperature [1].

The gel strength, yield stress, the plastic viscosity and apparent viscosity are some of rheological properties of drilling muds. These properties of drilling muds can vary with pressure and especially temperature. Many studies related to determining parameters of rheological models describes the drilling mud rheology have been published in last two decades [2,3,4,5,6,7]. During drilling, the drilling mud exposes various shear rate since the mud is circulated in and out of the well. The shear rate is about 1000 s^{-1} while a drilling mud travels inside drillpipe and it is approximately equal to 100 000 s^{-1} when drilling mud exist the nozzle of drill bit [8]. On the other hand, it was reported that shear rate on the wall is largely lower than 200 s^{-1} [9]. During the measurement of shear stresses as functions of shear rates, the six shear rates namely 5.1 s^{-1}, 10.2 s^{-1}, 170.3 s^{-1}, 340.3 s^{-1}, 511 s^{-1} and 1022 s^{-1} have been usually used in experimental studies done in laboratory [2,3,10,11]. Tehrani [8] claimed that the most drilling muds utilized for cleaning the hole have time-dependent properties. The shear rate, shear time and shear history have crucial effects on the shear stress of a viscoelastic fluid. The short term rheological reaction of a microstructure when subjected to an applied stress [11] can be called as thixotropy that is a time-dependent property [13]. It can be talked about the mud thixotropy when drilling fluids are primarily composed of clay that can be added into based drilling fluid to control gel strength [14]. The viscoelasticity is an important characteristic property of drilling mud according to Zamora and Growcock [15]. Bui et al. [16] and Pivnicka et al. [17] investigate the elasticity of drilling mud by proposing rheological models for viscoelastic drilling fluid. It was expressed in Baker Hughes [18] that there are many mathematical models with different applicability ranges and they do not match the data for shear stresses versus shear rates of all drilling fluids in all shear rate ranges. In other words, these mathematical models cannot fully describe the rheological behavior of non-Newtonian drilling muds in all shear rate ranges; however, they are close

[1]Corresponding author

approximations. As mentioned earlier, although many models have been developed to determine the rheological behavior of drilling muds without considering the viscoelastic properties, there are very few rheological models developed considering the viscoelastic properties of drilling muds. Therefore, in this study, in an effort to develop a practical model that gives accurate results, a rheological model was developed to determine the rheological behavior of drilling muds with viscoelastic properties.

2. Development of The Model

The shear stress on the wall in terms of viscosity can be expressed as follows:

$$\tau_w = \eta(\dot{\gamma})\dot{\gamma}_w \quad (1)$$

In Eq. (1) τ_w represents the shear stress on the wall, $\dot{\gamma}_w$ shear rate on the pipe wall and η the apparent viscosity of the fluid. Eq. (1) is Newton's law of viscosity with a shear rate-dependent viscosity and is simple, useful, and effective in many cases, which can be called a eneralized Newtonian fluid on the wall. Eq. (1) can be rearranged as follows:

$$\eta(\dot{\gamma}) = \frac{\tau_w}{\dot{\gamma}_w} \quad (2)$$

Eq. (2) is valid any type of fluid, which means at a certain point of shear rate, the apparent viscosity of a fluid can be obtained from the slope of a line tangent to the curve drawn for the shear stress as a function shear rate for that fluid. Therefore, the relationship among the shear stress, the shear rate and the viscosity of a fluid can be written as follow [19,20].

$$\eta(\dot{\gamma}) = \frac{d\tau}{d\dot{\gamma}} \quad (3)$$

During drilling, the drilling muds must fulfil some tasks such as carrying cuttings to the surface, cooling and lubricating drillstring, controlling formation pressure and consolidating wellbore walls. To perform the aforementioned tasks, in the substantial number of studies it was stressed that drilling muds must have low viscosity and good pumpability at large shear rates but must have large viscosity at low shear rates. Besides, in many studies [2,3,6,7] the viscosities of drilling muds as a function of shear rates exhibited an asymptotic behavior to a constant value at large shear rates. Under these considerations, in this study the following model was proposed to describe rheological behavior of a viscoelastic drilling fluid.

$$\eta(\dot{\gamma}) = a + b(\lambda\dot{\gamma})^{n-1} \quad (4)$$

Where *n* can be called as a flow behavior index that should be less than unity for a drilling fluid, *a* is the viscosity of a fluid at the infinite shear rate (Pa.s), *b* is the incremental contribution of structure to viscosity (Pa.s) and λ is the relaxation time (s) that is the characteristic time in which material releases the stress accumulated by exposing strain. The characteristic time is the time required for polymer coils or clay particles in the solution to relax from a deformed state to its equilibrium configuration.

As seen in Eq. (4) the dimension ([$ML^{-1}T^{-1}$] (Pa.s)) of *b* is equal to the dimension of viscosity since the term of $(\lambda\dot{\gamma})$ is dimensionless and the unit of *b*, unlike the consistency index in power-law and Herschel-Bulkley models, is independent of the flow behavior index. Therefore, it can be given as a material property, regardless of other parameters of the model since it can alone contain information about physical dependencies for a fluid.
Here, the polynomial function proposed for viscosity of a material is the power of strain, which is different from the rational polynomial functions proposed for viscosity in other studies such as Andaverde et al. [10], Nasiri and Ashrafizadeh [4]. The shear stress versus shear rate can be found by inserting Eq. (4) into Eq. (3) and integrating the resulting equation with respect to shear rate.

$$\frac{d\tau}{d\dot{\gamma}} = a + b(\lambda\dot{\gamma})^{n-1} \Rightarrow \tau = \int (a + b(\lambda\dot{\gamma})^{n-1})d\dot{\gamma} \quad (4)$$

Integration of Eq. (5) will result in the following equation.

$$\tau = \tau_0 + a\dot{\gamma} + \frac{b}{n\lambda}(\lambda\dot{\gamma})^n \quad (5)$$

Where τ_0 (Pa) is the integration constant that can be called as a yield stress and its values cannot be negative since there is no meaning of the negative value for yield stress.
As mentioned earlier, the experimental data obtained for different drilling muds namely biopolymers, bentonite-water suspension with some biopolymers, bentonite-lignite-water suspensions, xanthan gum solution, light weight colloidal gas aphrons by several investigators [2,3,4,7,12] were used to compare the data obtained from the model proposed here.

3. Determination of Model Parameters for Various Viscoelastic Fluids

Many water based drilling fluids containing the clay mineral bentonite exhibit a time-dependent response to shear [17]. The water-based drilling fluids containing clay components such as bentonite with minerals, polymers and surfactants exhibit a pronounced thixotropic behavior [8]. Furthermore, Tehrani [8] claimed that the most drilling muds utilized for cleaning the hole have time-dependent properties. The drilling fluids primarily composed of clay that can be added into based drilling fluid to control gel strength exhibit thixotropic behavior [14].

The shear rate-shear stress data pairs obtained from the proposed model are being compared with the experimental data obtained by different researchers for different fluids, the theoretical data obtained by Andaverde et al. [10], the data obtained from the modified Herschel-Bulkley, the modified Power-law and Bingham plastic models, which are illustrated in Tables 3- 5. Deng et al. [22] stated that yield point and plastic viscosity are usually estimated at shear rate in a range of 511-1021 s^{-1}. The parameter values of the proposed model, the modified Herschel-Bulkley, the modified Power-law, Bingham plastic models for a fluid are obtained based on the experimental data for the fluid in question. After the value of yield stress in the proposed model, the modified Herschel-Bulkley model and Bingham plastic model was obtained using the Zamora and Power [23] approach, this value was substituted in the models in question and the parameter values of the model at hand were determined by curve fitting method. The theoretical data used in the comparison for a fluid are the data having been obtained from the model proposed by Andaverde et al. [10], and the parameter values of this model were determined based on the experimental data obtained for the fluid in question. The theoretical values for shear rate-shear stress are taken directly from the study by Andaverde et al. [10]. In other words, as in the modified Herschel-Bulkley, the modified Power-law and Bingham plastic models, the parameter values of the model proposed by Andaverde et al. [10] were not determined in the present study. As seen in the tables, each pair of data obtained from the proposed model is compared with each pair of data obtained from other models and each pair of experimental data. Although some of the data pairs obtained from the proposed model are in agreement with the experimental data pairs and the data pairs obtained from other models, they are not in agreement with some other data pairs. Therefore, the average of the percentage difference between the data obtained from proposed model and the experimental data and the data obtained from the other models will be used to compare the proposed model with the experimental studies and the other models.

The values of parameters of the proposal model, Herschel-Bulkley, Power-law and Bingham plastic models for bentonite-lignite water based drilling fluid having lignite of 0.5 % w/v were determined by nonlinear regression with using the experimental data given in the paper by Kelessidis et al. [2]. As stated before, the yield stress was determined according to the Zamora and Power method, independent of the curve fitting method in the models, and substituted in the models. In this study, as previously stated, the modified Power-law and modified Herschel-Bulkley models used for each examined experimental set were obtained using the dimensionless shear rate in them. In order to validate the proposal model with experimental data, the data obtained from modified Herschel-Bulkley, modified Power-law and Bingham plastic models and the theoretical data, the average of the percentage difference between the experimental data and the data obtained from proposed model, the modified Herschel-Bulkley, the modified Power-law and Bingham plastic models are given in Table 3. When the average of the percentage difference between the experimental data and the data obtained from each model is ordered from smallest to largest, it is seen in the table

that the order is in the form of the proposed model, the modified Herschel-Bulkley, the modified power-law and Bingham plastic model. Since the smallest value of the average of the percentage difference between the experimental data and the data obtained from the models is acquired from the proposed model, it can be said that the model that best fits the experimental data for this fluid is the proposed model. In addition, the average of the percentage difference between the data obtained from the proposed model and the theoretical data is given in the table. The average of the percentage difference between the theoretical data and the data obtained from the proposed model is very close to the average of the percentage difference between the experimental data and the data obtained from the proposed model, showing that the proposed model is compatible with other models.

In Fig. 4, the shear rates versus shear stresses are plotted based on experimental data pairs, the proposed model, and data pairs obtained from other models used in the validation of the proposed model. The same experimental data were used by Andaverde et al. [10] to validate their model. That is why those experimental data were used in the present study to compare the data obtained from the proposed model with the experimental and theoretical data. The experimental data used in Fig.4 are available in the paper by Kelessidis et al. [2] under the name of S11 and in the paper by Andaverde et al. [10] under the name of KE-11.

The model proposed here, with the values of the parameters for the bentonite-lignite water based drilling fluid having lignite of 0.5 % w/v, can be expressed as follows:

$$\tau = \tau_0 + a\dot{\gamma} + \frac{b}{n\lambda}(\lambda\dot{\gamma})^n = 0.42 - 0.0141\dot{\gamma} + \left(\frac{0.0190}{0.0110\times0.8093}\right)(0.0110\dot{\gamma})^{0.8093} \quad (6)$$

For the above equation R^2 and R were found to be 0.9966 and 0.9983, respectively.

In addition to the above mentioned experimental data, two other experimental data called aqueous bentonite suspension and bentonite-lignite (3 % w/v) water based drilling fluid were used in the validation of the proposed model. The experimental data for aqueous bentonite suspension drilling fluid is named as W1 in the study of Kelessidis and Maglione [12], while the same experimental data is named as KE-16 in the study of Andaverde et al. [10]. Experimental data for bentonite-lignite (3 % w/v) water based drilling fluid is named as S3 in the study by Kelessidis et al. [2], while the same experimental data is named as KE-3 in the study by Andaverde et al. [10]. The processes performed for bentonite-lignite (0.5 % w/v) water based drilling fluid were also performed for the aqueous bentonite suspension and the bentonite-lignite (3 % w/v) water based drilling fluids. The results obtained are given in Table 4 and Table 5 in tabular form, and are shown in Figure 5 and Figure 6 in graphic form. The model proposed here, with the values of parameters for the bentonite-water suspension and bentonite-lignite water based drilling fluid having lignite of 3 % w/v can, respectively, be expressed in Eq. (8) and Eq. (9).

$$\tau = \tau_0 + a\dot{\gamma} + \frac{b}{n\lambda}(\lambda\dot{\gamma})^n = 17.0 - 0.0416\dot{\gamma} + \left(\frac{0.1182}{1.1712\times0.8804}\right)(1.1712\dot{\gamma})^{0.8804} \quad (7)$$

For the above equation R^2 and R were found to be 0.9998 and 0.9999, respectively

$$\tau = \tau_0 + a\dot{\gamma} + \frac{b}{n\lambda}(\lambda\dot{\gamma})^n = 3.9 - 0.4602\dot{\gamma} + \left(\frac{0.5450}{0.0427\times0.9532}\right)(0.0427\dot{\gamma})^{0.9532} \quad (8)$$

For the above equation R^2 and R were found to be 0.9785 and 0.9892, respectively.

Table 3. The experimental data for bentonite-lignite water based drilling fluid having lignite of 0.5 % w/v reported by Kelessidis et al. (2006) and the data obtained from the models [2]

Shear Rate (1/s)	Shear stress (Pa)	Andeverde et al. (2019	Proposal model	Absolute difference (E-P)%	Absolute difference (T-P)%	Herschel-Bulkley	Absolute difference (E-HB) %	Power-law	Absolute difference (E-PL)%	Bingham plastic	Absolute difference (E-BP)%
5.11	0.5	0.58	0.4479	10.4215	22.7772	0.6181	23.6263	0.4754	4.9274	0.2837	43.2512
10.21	0.58	0.69	0.6151	6.0496	10.8569	0.8193	41.2527	0.7046	21.4744	0.3374	41.8306
17.02	1	0.84	0.8119	18.8118	3.3473	1.0320	3.2024	0.9421	5.7943	0.4090	59.0994
34.05	1.33	1.19	1.2358	7.0790	3.8529	1.4486	8.9156	1.3973	5.0575	0.5881	55.7807
51.07	1.5	1.53	1.6036	6.9062	4.8100	1.7861	19.0763	1.7594	17.2917	0.7671	48.8585
102.14	2.42	2.44	2.5277	4.4522	3.5960	2.5940	7.1881	2.6091	7.8132	1.3042	46.1056
136.18	2.92	2.98	3.0508	4.4798	2.3762	3.0419	4.1730	3.0726	5.2246	1.6623	43.0734
170.23	3.42	3.48	3.5237	3.0317	1.2553	3.4471	0.7911	3.4882	1.9935	2.0204	40.9247
340.46	5.58	5.46	5.4058	3.1218	0.9926	5.1171	8.2965	5.1728	7.2971	3.8107	31.7070
510.69	6.83	6.86	6.7884	0.6097	1.0444	6.4712	5.2529	6.5138	4.6302	5.6011	17.9924
680.92	7.92	7.90	7.8536	0.8388	0.5877	7.6540	3.3590	7.6711	3.1429	7.3915	6.6730
851.15	8.58	8.68	8.6863	1.2389	0.0726	8.7237	1.6743	8.7086	1.4987	9.1819	7.0148
1021.38	9.33	9.27	9.3358	0.0627	0.7103	9.7111	4.0847	9.6596	3.5330	10.9722	17.6018
Average value				**5.1618**	**4.3292**		**10.0687**		**6.8983**		**35.3779**

P stands for proposed model, T: theoretical values, E: experimental values, HB: Herschel-Bulkley, PL: power-law, BP: Bingham plastic

Table 4. The experimental data for aqueous bentonite suspension by Kelessidis and Maglione (2006) and the data obtained from the models [12]

Shear Rate (1/s)	Shear stress (Pa)	Andeverde et al.(2019	Proposal model	Absolute difference (E-P)%	Absolute difference (T-P)%	Herschel-Bulkley	Absolute difference (E-HB) %	Power-law	Absolute difference (E-PL)%	Bingham plastic	Absolute difference (E-BP)%
5.109	17.67	17.36	17.3417	1.8580	0.1054	17.5163	0.8696	15.3128	13.3400	17.0916	3.2734
10.218	17.33	17.60	17.5952	1.5305	0.0273	17.8123	2.7833	16.7550	3.3178	17.1832	0.8472
102.18	20.58	20.64	20.4976	0.4001	0.6899	20.6602	0.3896	22.5943	9.7877	18.8317	8.4949
170.3	22	22.20	22.0646	0.2937	0.6099	22.1114	0.5062	24.1439	9.7449	20.0529	8.8504
340.6	25.17	25.17	25.1976	0.1096	0.1097	25.0415	0.5105	26.4178	4.9575	23.1058	8.2010
510.9	27.75	27.56	27.7095	0.1459	0.5425	27.4824	0.9645	27.8460	0.3460	26.1587	5.7344
1021.8	33.33	33.39	33.3375	0.0225	0.1572	33.4915	0.4846	30.4686	8.5850	35.3174	5.9628
Average value				**0.6229**	**0.3203**		**0.9298**		**7.1541**		**5.9092**

P stands for proposed model, T: theoretical values, E: experimental values, HB: Herschel-Bulkley, PL: power-law, BP: Bingham plastic

Table 5. The experimental data for bentonite-lignite water based drilling fluid having lignite of 3 % w/v by Kelessidis et al. (2006) and the data obtained from the models [2]

Shear Rate (1/s)	Shear stress (Pa)	Andeverde et al.(2019	Proposal model	Absolute difference (E-P)%	Absolute difference (T-P)%	Herschel-Bulkley	Absolute difference (E-HB) %	Power-law	Absolute difference (E-PL)%	Bingham plastic	Absolute difference (E-BP)%
5.11	5.10	5.59	4.6861	15.6682	16.1696	5.6965	11.6961	4.2166	17.3211	4.0297	20.9859
10.21	6.25	5.86	5.2707	10.8583	10.0557	6.3914	2.2621	5.3303	14.7147	4.1592	33.4530
17.02	6.67	6.21	5.9457	0.7921	4.2552	7.0717	6.0219	6.3373	4.9880	4.3321	35.0515
34.05	7.42	7.08	7.3612	23.5801	3.9721	8.3011	11.8746	8.0146	8.0132	4.7644	35.7900
51.07	6.92	7.93	8.5517	8.6191	7.8404	9.2300	33.3820	9.1938	32.8579	5.1964	24.9068
102.14	10.5	10.34	11.4050	7.7815	10.2999	11.2951	7.5723	11.6259	10.7229	6.4929	38.1629
136.18	12	11.85	12.9338	5.0121	9.1458	12.3714	3.0954	12.8152	6.7934	7.3570	38.6915
170.23	13.58	13.27	14.2606	2.0189	7.4653	13.3134	1.9633	13.8212	1.7760	8.2214	39.4594
340.46	18.58	19.16	18.9551	3.4908	1.0693	16.9605	8.7164	17.4775	5.9340	12.5428	32.4929
510.69	22.42	22.85	21.6374	11.6569	5.3070	19.7179	12.0521	20.0496	10.5727	16.8642	24.7804
680.92	26.08	24.50	23.0399	3.3150	5.9596	22.0206	15.5651	22.1010	15.2569	21.1856	18.7667
851.15	22.75	24.06	23.5042	5.9704	2.3102	24.0352	5.6491	23.8357	4.7721	25.5070	12.1189
1021.38	21.92	21.55	23.2287	15.6682	7.7899	25.8464	17.9122	25.3536	15.6641	29.8284	36.0787
Average value				**8.8024**	**7.0492**		**10.5971**		**11.4913**		**30.0568**

P stands for proposed model, T: theoretical values, E: experimental values, HB: Herschel-Bulkley, PL: power-law, BP: Bingham plastic

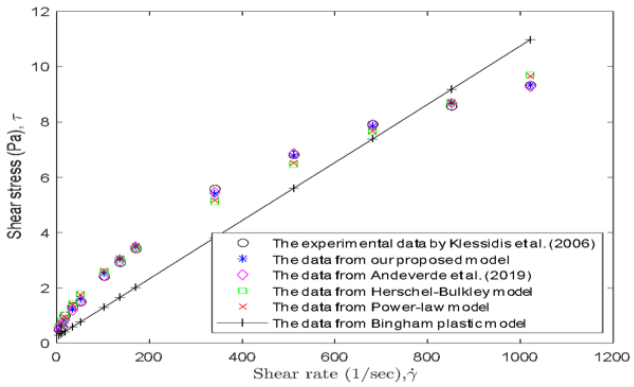


Figure 4. Variation of shear stresses with shear rates for a bentonite-lignite water-based drilling fluid having lignite of 0.5 % w/v [2]

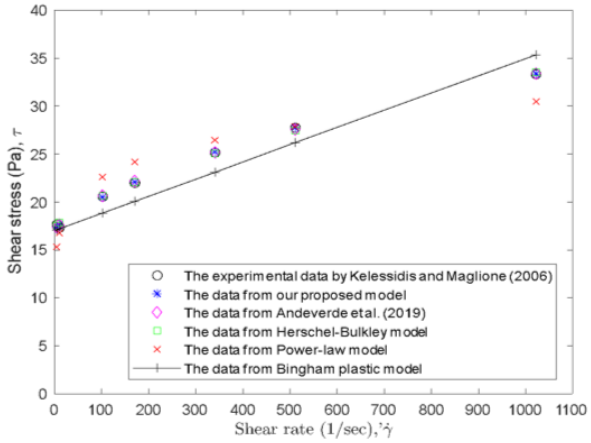


Figure 5. Variation of shear stresses with shear rates for aqueous bentonite suspensions [12]

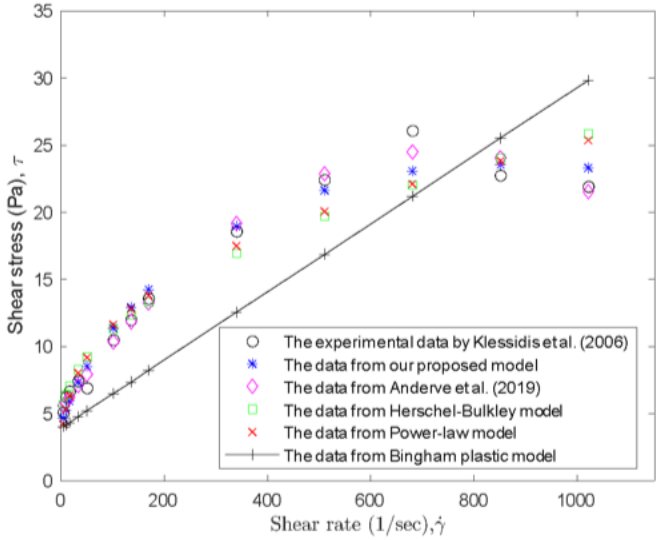


Figure 6. Variation of shear stresses with shear rates for bentonite-lignite water-based drilling fluids having lignite of 3 % w/v [2]

4. Conclusion

In this study, a model used to determine the rheological behavior of drilling muds with viscoelastic properties was proposed. The proposed model would be reduced to the Herschel-Bulkley model as shear rate goes to zero or very low values, while it would be reduced to the Bingham plastic model as shear rate goes to infinity or very large values or taking *n* is equal to 1.0, and it can also be reduced to the power-law model at the certain conditions. More than 33 experimental data pairs obtained for the shear stresses versus the shear rates for the seven different drilling muds were used to validate the model. Apart from the yield stress, the parameter values in the proposed model, the power-law, Herschel-Bulkley and Bingham plastic models were determined by curve fitting method based on experimental data. The yield stress in the proposed model, the modified Herschel-Bulkley and Bingham plastic models was determined according to the Zamora and Power method. In the proposed model, almost the same shear stress curve against shear rate was obtained in different combinations of parameter values, and this situation was not limited to the proposed model. The proposed model was observed to be the best-fitting model for the three experimental datasets. Using the experimentally obtained data for each drilling mud, the parameters of the proposed model were determined for the mud in question. The data obtained from the proposed model were compared with the experimental data and the theoretical data (the data obtained from the model validated, proposed by other researchers, using the same experimental data), and it was observed that the results were in great agreement. The consistency in the results indicates that the proposed model can be used without hesitation to accurately determine the rheological behavior of the certain drilling muds. Although the proposed model is a four-parameter model, it can be considered to be superior to other models since it reduces into other widely used rheological models under certain constraints.

5. References

[1] S.G. Ash, A.J. Clarke-Sturman, R. Calvert, T.M. Nisbet, Chemical Stability of Biopolymers Solutions, SPE 12085, presented at the 58th Annual Technical Conference and Exhibition in San Francisco, CA, 5-8 October 1983.

[2] V.C. Kelessidis, R. Maglione, C. Tsamantaki, Y. Aspirtakis, Optimal determination of rheological parameters for Herschel–Bulkley drilling fluids and impact on pressure drop, velocity profiles and penetration rates during drilling, J. Petrol. Sci. Eng., 53 (2006) 203–224. https://doi.org/10.1016/j.petrol.2006.06.004

[3] O. Mark, Statistical Characterization of Performance of Biopolymer Drill-In Fluid for Different Rheological Models. Master Thesis, Presented to the Faculty of the African Univ. Sci. and Tech. Abuja, Nigeria (2010), pp. 1–61.

[4] Y. Salyzhyn, M. Myslyuk, Studies of the rheological properties of drilling fluids. Ann. Trans. Nordic Rheol. Soc., 19 (2011) 61-67.

[5] M. Nasiri, S.N. Ashrafizadeh, Novel equation for the prediction of rheological parameters of drilling fluids in an annulus. Ind. Eng. Chem. Res., 49 (2010) 3374-3385. https:// doi.org/10.1021 /ie9009233

[6] M. Arabloo, M.P. Shahri, Experimental studies on stability and viscoplastic modeling of colloidal gas aphron (CGA) based drilling fluids. J. Petrol. Sci. Eng., 113 (2014) 8–22. https://doi.org/10.1016/j.petrol.2013.12.002

[7] H. Ziaee, M. Arabloo, M.H. Ghazanfari, D. Rashtchian, Herschel-Bulkley rheological parameters of lightweight colloidal gas aphron (CGA) based fluids. Chem. Eng. Res. Des., 93 (2015) 21–29. https://doi.org/10.1016/j.cherd.2014.03.023.

[8] Tehrani, Thixotropy in water-based drilling fluids. Ann. Trans. Nordic Rheol. Soc., 16 (2008) 1-13.

[9] Werner, B. Lund, V. Myrseth, A. Saasen, K.R. Gyland, Z. Ibragimova, Comparison of rheological properties of oil based and KCl drilling fluids. In: Paper SPE 180063 MS Presented at the SPE Bergen One Day Seminar Held in Bergen, Norway, 20 April 2016.

[10] J.A. Andaverde, J.A Wong-Loya, Y. Vargas-Tabares, M. Robles, A practical method for determining the rheology of drilling fluid, J. Petrol. Sci. Eng., 180 (2019) 150–158. https://doi.org/10.1016/j.petrol.2019.05.039.

[11] H.A. Barnes, Thixotropy-a review. J. Non-Newtonian Fluid Mech., 70 (1997) 1–33. https://doi.org/10.1016/S0377-0257(97)00004-9.

[12] V.C. Kelessidis, R. Maglione, Modeling rheological behavior of bentonite suspensions as Casson and Robertson-Stiff fluids using Newtonian and true shear rates in Couette viscometry. Powder Tech., 168 (2006) 134–147. https://doi.org/10.1016/j.powtec. 2006.07.011.

[13] J. Maxey, Thixotropy and yield stress behavior in drilling fluids. In: Paper AADE-07-NTCE-37 Presented at the 2007 AADE National Technical Conference and Exhibition Held at the Wyndam Greenspoint Hotel, Houston, Texas, April 10-12, (2007).
[14] R. Jachnik, Drilling fluid thixotropy and relevance. Ann. Trans. Nordic Rheol. Soc., 13 (2005) 1-6.
[15] M. Zamora, F. Growcock, The top 10 myths, misconceptions and mysteries in rheology and hydraulics. In: Paper AADE-10-DF-HO-40 Presented at the 2010 AADE Fluids Conference and Exhibition Held at the Hilton Houston North, Houston Texas, April 6 – 7, (2010).
[16] B. Bui, A. Saasen, J. Maxey, M.E. Ozbayoglu, S.Z. Miska, M. Yu, N.E. Takach, Viscoelastic properties of oil-based drilling fluids. Ann. Trans. Nordic Rheol. Soc., 20 (2012) 33–47.
[17] S. Pivnicka, T.C. Nguyen, E. Al-Safran, A. Saasen, Pressure gradient prediction of time-dependent drilling fluids and the effect of acceleration. J. Petrol. Sci. Eng., 135 (2015) 246–252. https://doi.org/10.1016/j.petrol.2015.09.008.
[18] Baker Hughes INTEQ. Drilling Fluids Reference Manual. Houston, Texas, USA, 2006 (770-780).
[19] B. Bird, C.F. Curtiss, R.C. Armstrong, O. Hassager, Dynamics of Polymeric Liquids, Volume 1: Fluid Mechanics. John Wiley & Sons Inc., New York, 1987.
[20] R.V. Lauzon, K.I.G. Reid, New rheological model offers field alternative. Oil Gas J., 5 (1979), 51–57.
[21] D.C-H. Cheng, F. Evans, Phenomenological Characterization of the Rheological Behavior of Inelastic Reversible Thixotropic and Anti-Thixotropic Fluids. Brit. J. Appl. Phys. 16 (1965) 1599. DOI 10.1088/0508-3443/16/11/301.
[22] S. Deng, C. Kang, A. Bayat, E. Kuru, M. Osbak, K. Barr, C. Trovato, Rheological properties of clay-based drilling fluids and evaluation of their hole-cleaning performances in horizontal directional drilling. J. Pipeline Syst. Eng. Pract., 11 (2020) https://doi.org/10.1061 /(ASCE)PS.1949-1204.0000475.
[23] M. Zamora, D. Power, Making a Case for AADE Hydraulics and the Unified Rheological Model. In: Paper AADE-02DFWM-HO-13 presented at the AADE 2002 Technology Conference, Drilling & Completion Fluids and Waste Management, Held at the Radisson Astrodome, Houston, Texas, April 2- 3, 2002 in Houston, Texas, USA.

Heat Transfer in A Corrugated Conduit Filled with Porous Material

Fethi KAMIŞLI[1] , Musaa ALGIBARA[2]

[1,2] Department of Chemical Engineering, Faculty of Engineering, Firat University, Elazig, Turkey
[1]fkamisli@firat.edu.tr, [2] eng.moussa2050@gmail.com
[1](ORCID: 0000-0002-1769-3785), [2](ORCID: 0000-0002-2876-3144)

Abstract
The investigation of the flow of a fluid in a porous material surrounded by a curved edge and the heat transfer occurring during this flow is quite important in the design of similar industrial systems. In this research, simulation was carried out to investigate the combined effect of the presence of a porous medium in a channel and semicircular sections on its upper wall on non-Darcian flow and forced convection heat transfer. The simulation of the created model was done by solving the dimensionless constitutive equations using the COMSOL multiphysics program (v.6.1). The results obtained in similar studies conducted previously under the same conditions were compared with the results obtained in this simulation to determine whether the program used in this simulation gave accurate results. The effects of process parameters such as the radius of semicircular sections (*R*) and the distance between centers of semicircular sections (*Lc*), Reynolds number (*Re*) and Darcy number (*Da*) were investigated. It was found that the results of the present study were in good agreement with those of the previous related studies.

Keywords: forced convection, Non-Darcian flow, porous medium, semi-circular sections

1. Introduction

Heat transfer owing to fluid flow in porous media is a major and very rapid development in modern heat transfer research due to its use in many different applications, including compact heat exchangers, oil and gas production from geological structures, geothermal processes, adsorption, filtration, catalytic chemical reactors, and drying. The porous media having high surface area, high thermal conductivity and the ability to mix liquids well in different types of conduits have been used for enhancing forced convection heat transfer by many investigators [1,2,3,4]. Moreover, many numerical, analytical, and experimental studies have been recorded on flow fields and forced convection in different types of geometry with porous material. In those studies, it was demonstrated that the channel might not need to be fully packed with a porous material to maximize convection heat transfer [5,6,7]. The geometric configuration of the channel also plays a major role in enhancing the heat transfer as in the case of the waviness of the channel walls. Several investigations have been published describing forced convection flow in corrugated channels of different shapes [8,9,10,11]. In the above studies, the results showed that the effect of conduit geometry on the mean Nusselt number is negligible at low Reynolds numbers while it cannot be neglected at higher values of Reynolds number since at the higher values of Reynolds number the average Nusselt number takes the largest value in a conduit having linearly increasing amplitude.

It is evident from the above-cited literature that the corrugated channel is one of the most effective devices among devices used to enhance heat transfer for industrial processes. Since the coupled effect of the wavy wall and the presence of a porous medium in the conduit increases the forced convection heat transfer, the Newtonian fluid flow and heat convection heat transfer in the different types of channels were numerically examined by many researchers [12,13,14,15]. Kumar et al. [16] performed an analysis numerically to determine the influences of Prandtl (*Pr*), Darcy (*Da*), and Reynolds (*Re*) numbers as well as non-dimensional porous slab thickness on average Nusselt numbers (Nu_{avg}), pressure drops (ΔP), and enhancement ratio (ER) for flow in the channel.

In light of the reviewed papers described above, it can be deduced that the dual effect of corrugation of walls and porous inserts can effectively increase the heat transfer rate and affect the performance factor, pressure drop, and Nusselt number. Although there are numerous studies on convective heat transfer and fluid flows in geometries with corrugated walls filled with porous material, the effect of the corrugation shape of the geometry on the convective heat transfer and fluid flow is among the topics that need to be investigated since the different shapes of corrugation of conduit walls affect fluid flow and convective heat transfer at different rates. Therefore, in this research, the

[1]Corresponding author

coupled effect of the wall waviness and the presence of porous material on the heat transfer by forced convection for non-Darcian flow within a porous medium in the channel having semi-circular sections at the one wall was investigated. The new modification in the channel shape was semi-circular sections on the upper wall only that has been investigated. The proposed model was simulated according to the new modification of the channel shape to evaluate the effect of the conduit geometry having porous inserts on the forced convection of the laminar flow regime.

The effects of parameters affecting the process design such as Darcy number (*Da*), Reynolds number (*Re*), the radius of semicircular sections (*R*), and distance between the semicircular sections (L_c) were determined. The simulation program was tested by checking the agreement between the existing data in the related literature and the obtained results from the current study. The results are presented and analyzed in the form of graphs and tables, with the discussion and defining conclusions.

2. The Problem Description

The proposed model in this study shown in Figure 2.1 is a section of a two-dimensional channel containing semicircular sections of radius (*R*) in its upper wall, while the lower wall is flat. The upper surface of the channel was considered as the separated three zones, namely the inlet section of length (L_i), the semi-circular sections of length (L_c), and the exit section of length (L_0). The conduit was filled with a homogeneous porous medium. A fluid having a uniform velocity (U_{in}) and a temperature (T_{in}) enters the channel and flows in a laminar regime throughout the channel. The entrance temperature of the fluid (T_{in}) is lower than the semicircular section temperature (T_w). The cartesian coordinate system with *x*- and *y*-components was used to simulate the Newtonian fluid flow and heat transfer by forced convection in the channel filled with porous material. y_t and y_b were, respectively, used to represent the *y*-component of the cartesian coordinate for the upper and lower walls of the channel. In Figure 2.1 *L* stands for the total length of the conduit and *H* represents the total height of the conduit.

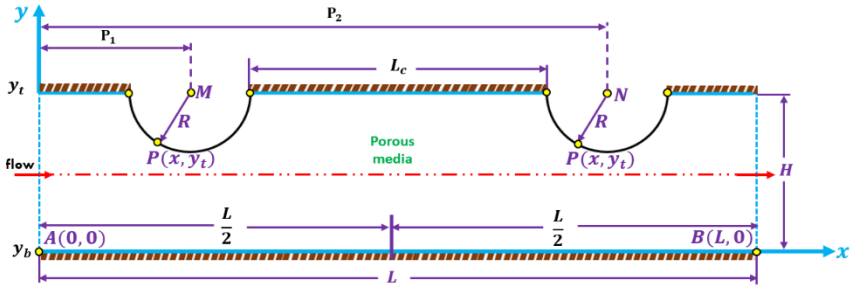


Figure 2.1. Mathematical representation of the coordinates of the upper and lower surfaces

3. Mathematical Formulation

In this study, it was assumed that the incompressible Newtonian fluid in the conduit flows in two-dimensional form at the steady state and in laminar regime and the porous medium has homogeneous properties such as porosity, permeability, and others. The local thermal equilibrium (LTE) between the flowing Newtonian fluid and porous inserts exists throughout the channel. According to the above assumptions, the dimensional form of the continuity, momentum, and energy equations used in the current study can be written as follows [14,16]. The equation of continuity for two-dimensional flow is given as follows:

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0 \tag{3.1}$$

In Equation (3.1), *u* and *v* are components of the dimensional velocity in the *x* and *y*-indirections, respectively. The *x*-component of the momentum equation is given as follows:

$$\frac{\rho}{\varepsilon^2}\left(u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}\right)=-\frac{\partial p}{\partial x}+\frac{\mu}{\varepsilon}\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right)-\frac{\mu}{\mathrm{K}}u-\frac{\rho C_F u}{\sqrt{\mathrm{K}}}\sqrt{u^2+v^2} \tag{3.2}$$

The *y*-component of the momentum equation is given as follows:

$$\frac{\rho}{\varepsilon^2}\left(u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}\right)=-\frac{\partial p}{\partial y}+\frac{\mu}{\varepsilon}\left(\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}\right)-\frac{\mu}{\mathrm{K}}v-\frac{\rho C_F v}{\sqrt{\mathrm{K}}}\sqrt{u^2+v^2}$$
(3.3)

The energy equation is given as follows:

$$u\frac{\partial T}{\partial x}+v\frac{\partial T}{\partial y}=\alpha_f\left(\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}\right) \quad (3.4)$$

In the above equations, K stands for the permeability of the porous structure, the porosity and the fluid thermal diffusivity. An external force resulting from the porous medium based on Ergun's experiment is represented in the momentum equations [15,16].

$$\left(\mu+\rho C_F\sqrt{K}\sqrt{u^2+v^2}\right)u/K \quad \text{and} \quad \left(\mu+\rho C_F\sqrt{K}\sqrt{u^2+v^2}\right)v/K$$
(3.5)

Where the term $\left(C_F\sqrt{K}\sqrt{u^2+v^2}\right)u/K$ is called Forchheimer drag, and $C_F=1.75/\sqrt{150\varepsilon^3}$ is Forchheimer drag coefficient [16]. By adopting characteristic scales for the problem, the non-dimensional variables are written according to the following formulas.

$$\theta=\frac{T-T_{in}}{T_w-T_{in}}, U=\frac{u}{U_{in}}, V=\frac{v}{U_{in}}, P=\frac{p}{\rho U_{in}^2}, X=\frac{x}{H}, Y=\frac{y}{H}, Y_t=\frac{y_t}{H}, \dot{R}=\frac{R}{H}, \dot{P}_1=\frac{P_1}{H}$$
(3.6)

where θ is the dimensionless temperature, U_i is the inlet velocity of the fluid, P is the dimensionless pressure, and U and V are the components of dimensionless velocity in the X- and Y- directions. The Reynolds number (Re), Darcy number (Da), and Prandtl number (Pr) called the dimensionless number are derived from the dimensionless variables as follows:

$$Re=\frac{\rho U_{in}H}{\mu}=\frac{U_{in}H}{v}, Da=\frac{\mathrm{K}}{H^2}, Pr=\frac{\mu c_P}{k}=\frac{v}{\alpha} \quad (3.7)$$

In Equation (3.7), ρ displays the density, c_P stands for the specific heat capacity, μ presents the dynamic viscosity, υ indicates the kinematic viscosity, k shows the thermal conductivity of the fluid, and the permeability of the porous medium is represented with K. The defined dimensionless variables were used in the continuity, momentum, and energy equations to convert these equations from their dimensional forms into their dimensionless forms. By using the above dimensionless variables, the dimensionless form of the continuity equation was obtained as

$$\frac{\partial U}{\partial X}+\frac{\partial V}{\partial Y}=0 \quad (3.8)$$

By using the above dimensionless variables, the x and y components of the momentum equations were obtained in the dimensionless forms as follows:

$$\frac{1}{\varepsilon^2}\left[U\frac{\partial U}{\partial X}+V\frac{\partial U}{\partial Y}\right]=-\frac{\partial P}{\partial X}+\frac{1}{\varepsilon Re}\left[\frac{\partial^2 U}{\partial X^2}+\frac{\partial^2 U}{\partial Y^2}\right]-\frac{U}{Da\,Re}-\frac{C_F}{\sqrt{Da}}U\sqrt{U^2+V^2} \quad (3.9)$$

$$\frac{1}{\varepsilon^2}\left[U\frac{\partial V}{\partial X}+V\frac{\partial V}{\partial Y}\right]=-\frac{\partial P}{\partial Y}+\frac{1}{\varepsilon\,Re}\left[\frac{\partial^2 V}{\partial X^2}+\frac{\partial^2 V}{\partial Y^2}\right]-\frac{V}{Da\,Re}-\frac{C_F}{\sqrt{Da}}V\sqrt{U^2+V^2} \tag{3.10}$$

The energy conservation equation in the non-dimensional form was obtained by using the dimensionless variables as follows.

$$U\frac{\partial \theta}{\partial X}+V\frac{\partial \theta}{\partial Y}=\frac{1}{Re\,Pr}\left[\frac{\partial^2 \theta}{\partial X^2}+\frac{\partial^2 \theta}{\partial Y^2}\right] \tag{3.11}$$

3.1. Non-dimensional Boundary Conditions

The boundary conditions used in the simulation are determined for the solids surfaces based on the no-slip and no-permeability conditions. The velocity component and temperature gradients at the channel outlet were assumed to be zero while the inlet velocity and temperature of the Newtonian fluid were taken to be uniform and constant at the across of the cross-section of the conduit. The temperature of the semicircle section's walls was considered higher than the inlet temperature. At the inlet of the conduit, the boundary conditions were taken as follows:

$$U=1,\ V=0,\ \theta=0 \quad \text{for } x=o \text{ and } 0\le Y\le H \tag{3.12}$$

At the outlet of the channel, the boundary conditions were taken as follows (zero gradients for velocity component and temperature):

$$\frac{\partial U}{\partial X}=0,\ \frac{\partial V}{\partial X}=0,\ \frac{\partial \theta}{\partial X}=0,\ P=0 \quad \text{for } X=L \text{ and } 0\le Y\le H \tag{3.13}$$

The boundary conditions on the insulated upper surface were applied as follows (impermeable and no-slip conditions):

$$U=0,\ V=0,\ \frac{\partial \theta}{\partial Y}=0 \quad \text{for } Y=H \text{ and } \left\{0\le x\le \frac{L}{2}-\frac{L_c}{2}-2R,\ \frac{L}{2}-\frac{L_c}{2}\le x\le \frac{L}{2}+\frac{L_c}{2},\ \frac{L}{2}+\frac{L_c}{2}+2R\le x\le L\right\} \tag{3.14}$$

The boundary conditions on the wavy upper surface were expressed as follows (The temperature is equal to the corrugated wall temperature):

$$\begin{aligned} &U=0,\ V=0,\ \theta=1\\ &U=0,\ V=0,\ \theta=1 \end{aligned} \quad \text{for} \quad \left\{\begin{array}{l} \frac{L}{2}-\frac{L_c}{2}-2R\le x\le \frac{L}{2}-\frac{L_c}{2} \\ \frac{L}{2}+\frac{L_c}{2}\le x\le \frac{L}{2}+\frac{L_c}{2}+2R \end{array}\right. \quad \begin{array}{l} Y_{f1}=1-\sqrt{R^2-(X-P_1)^2} \\ Y_{f2}=1-\sqrt{R^2-(X-P_2)^2} \end{array} \tag{3.15}$$

The boundary conditions at the bottom wall were taken as follows (impermeable and no-slip conditions):

$$U=0,\ V=0,\ \frac{\partial \theta}{\partial Y}=0 \quad \text{for } 0\le X\le L \text{ and } Y=0 \tag{3.16}$$

The local Nusselt number on the inner surfaces of the semicircular sections on the upper wall was computed from the following equation, used to compute the rate of heat transfer.

$$Nu_{Local}(X) = \frac{h\,H}{k} = \frac{\partial\theta}{\partial n}\left(\frac{1}{1-\theta_b}\right)_X \tag{3.17}$$

In equation (3.17) is the coefficient of heat transfer either on the inner surfaces of the semicircular sections on the upper wall of the channel, is the direction normal to the inner surface of the wall, *H* is the distance between two plates and is the non-dimensional bulk temperature of the flow, it was computed from the following equation.

$$\theta_b(X) = \int_0^{y_t} U\,\theta\,dY \Big/ \int_0^{y_t} U\,dY \tag{3.18}$$

The average Nusselt number along the inner surfaces of the semicircular sections on the upper plate of the conduit is calculated using the following formula. The results were obtained as in Figure 5.5.

$$\overline{Nu} = \frac{1}{\pi . R}\left\{\int_{\frac{L}{2}-\frac{L_c}{2}-2R}^{\frac{L}{2}-\frac{L_c}{2}} Nu_{Local}(X)\,dX + \int_{\frac{L}{2}+\frac{L_c}{2}}^{\frac{L}{2}+\frac{L_c}{2}+2R} Nu_{Local}(X)\,dX\right\} \tag{3.19}$$

4. Numerical Modeling

In computation L = 20 cm, H = 1 cm, R = 0.25 H, Pr = 0.7, θ_w = 1 and $L_c = mR$ were taken, here m is constant and m, Da and Re are varied according to the case examined. To compare the results obtained in this study with the results obtained in previous studies [12,14], the values of L, H, θ, and Pr values and the ranges for Da and Re numbers were determined based on previous studies.The mesh quality determines the convergence rate, time, and solution accuracy. Therefore, the best mesh was generated and provided better performance and higher solution accuracy than other types (see Figure 4.1). Figure 4.1 was plotted using the COMSOL multiphysics program (v.6.1)

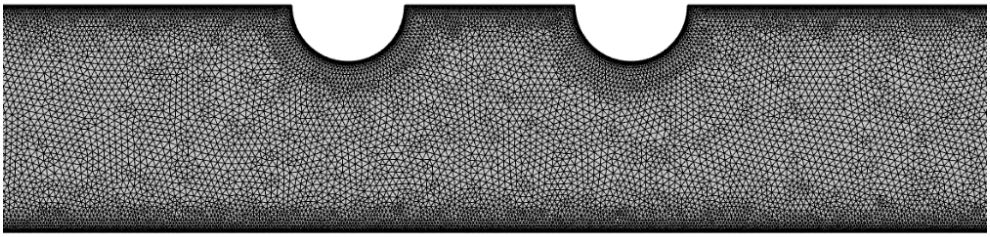


Figure 4.1. Generated mesh from the geometry and used in numerical simulation

4.1. Numerical Method and Solution Validation

The simulation was performed with the COMSOL program using the proven finite element method (FEM). The program performs finite element analysis using a set of numerical solutions. The results obtained from the current study were compared with the results of previous related studies to verify the validity and accuracy of the computation. The profile of the velocity distribution obtained in the current research illustrated in Figure 4.2A was compared with what the researchers Mansoor and Dawood [12] arrived at, as in Figure 4.2B. The results were compared with the flow dependence on the Darcy number ($Da = 10^{-5}$) and porosity ($\varepsilon = 0.9$) in the porous medium confined in a conduit. Although it was not shown here, the value of the Nusselt number was fixed at approximately 4.95, which agreed with previous studies [12,17].

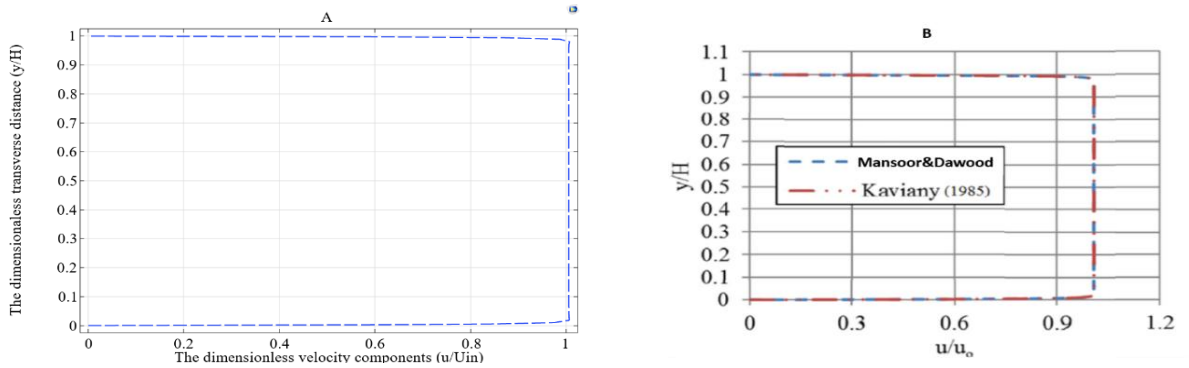


Figure 4.2. Comparison of the velocity distribution diagram within a flat channel between the present study (A) and Mansoor and Dawood study (B) at $Da = 10^{-5}$ and $\varepsilon = 0$.

5. Results and Discussion

In the current study, the effect of effective parameters such as the radius of semi-circular sections, the distance between the semi-circular sections, Reynolds number, and Darcy number on the heat transfer by forced convection and Newtonian fluid flowing in porous material in the conduit with semi-circular sections on the upper wall were investigated. The ranges for the mentioned parameter above were taken as:

$$50 \le Re \le 200, 10^{-5} \le Da \le 10^{-2}, 0.25H \le R \le 0.75H, 3R \le L_c \le 9R$$

5.1. Darcy Number Effect

Figure 5.1 shows the variation of the velocity profile of the fully developed flow in the horizontal channel with the values of Darcy number at Reynolds number of 50 and Prandtl number of 0.7. At high Darcy numbers, the horizontal velocity component has a high speed in the middle of the channel, similar to the flow profile in an empty channel. With a reduction in the Darcy number, the velocity profile in the middle of the geometry decreases, with a noticeable increase in the velocity value near the wall. This is because a high Darcy number indicates a lack of density of porous media. Thus, the boundary layer grows a certain distance to completion, with shear stress vanishing in the middle of the geometry, and the velocity profile is fusiform. As for the small and very small Darcy numbers, the dense presence of the porous material hinders the movement of the fluid, so its velocity decreases in the middle of the channel, and it takes a regular shape with the fading of the boundary layer, while the velocity increases near the wall, which leads to enhance heat convection. This is why porous media is beneficial in applications requiring improved heat transfer. The Darcy number value ($Da = 10^{-5}$) chosen for this study, in comparison to higher values, provides a higher heat transfer rate and the values chosen for *Da*, *Re*, *Pr* and *Re* numbers allowed us to compare the present study with similar previous studies [12].

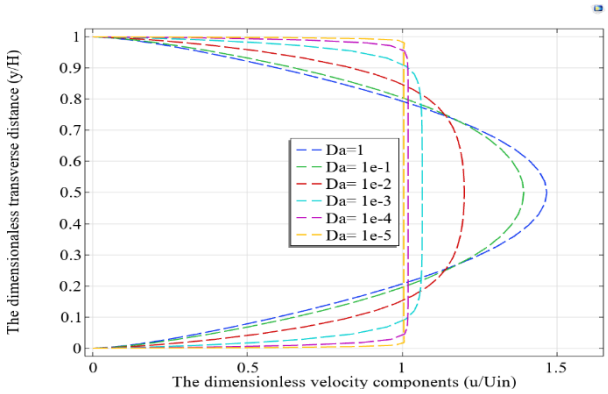


Figure 5.1. The velocity distribution of fully developed flow at the channel outlet for various values of Darcy numbers at Re = 50 and Pr = 0.7

5.2. The Effect of Changing The Radii of Semicircular Sections

The effect of changing the radii of semicircular sections on the velocity vectors when the fluid passes through the waist regions was demonstrated in Figure 5.2. It was observed from the figure that increasing the section radius led to an increase in the velocity component at the waist areas. This was due to the decrease in the flow section, where any cross-sectional area reduction is matched by an increase in fluid flow velocity to obtain a constant mass flow rate according to the mass conservation principle at the constant density. That is, in all previous cases, the flow rate was constant, but increasing the radius of the sections reduced the passage area of the fluid, and by reducing the passage area, the mean velocity increased. The velocity arrows beneath the waist area are colored red, signifying a rise in fluid velocity with a decrease in flow passage area. The intensity of the redness of the color indicates an increase in velocity.

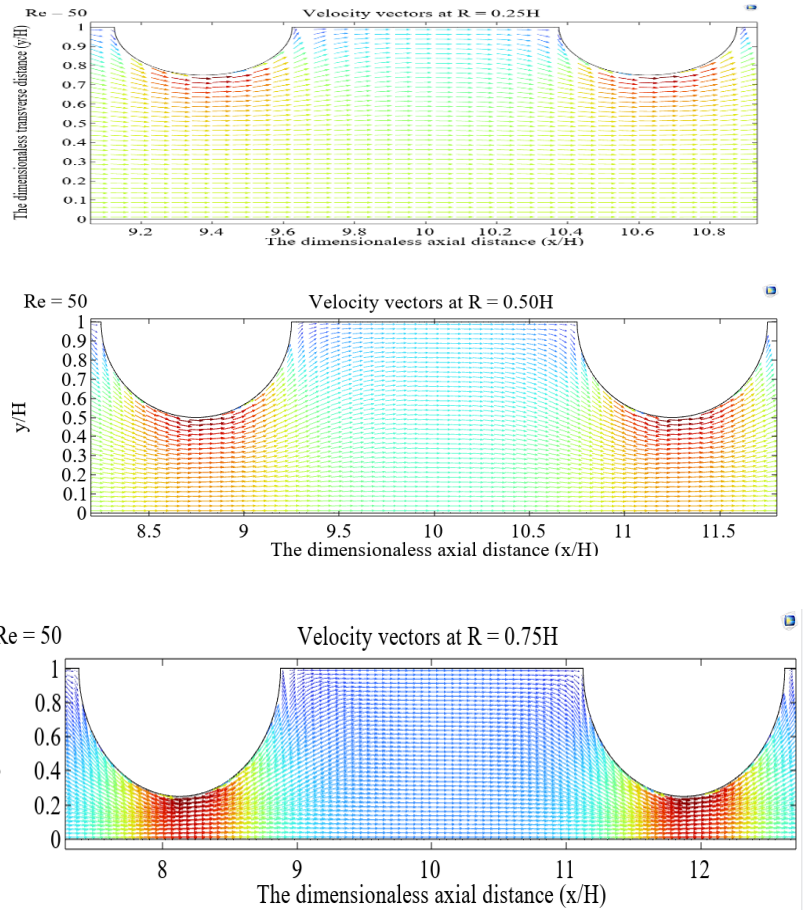


Figure 5.2. Velocity vectors when the fluid passes under semicircular sections for different values of radii (R = 0.25 H, 0.50 H, 0.75 H) at Re = 50 and Lc = 3 R

Figure 5.3 shows the effect of changing the radius of the sections on the heat convection as the distance between the sections is L_c = 3 R and Re = 50. As seen from the figure, increasing the radius causes the amount of heat transferred to increase owing to the increase in the heat source zone. This is seen from the size of the reddish spectrum. In other words, as the radius increases, the fluid absorbs a greater amount of heat on account of the increased activity of convection currents and an increase in the heat source zone. This indicates indirectly that an improvement in the Nusselt number leads to an increment in the heat transfer rate. It is also noted that the distance between the two semicircular sections increases with increasing R because the relationship between L_c and R is given as L_c = m.R, where m has a constant value, which affects the heat transfer positively.

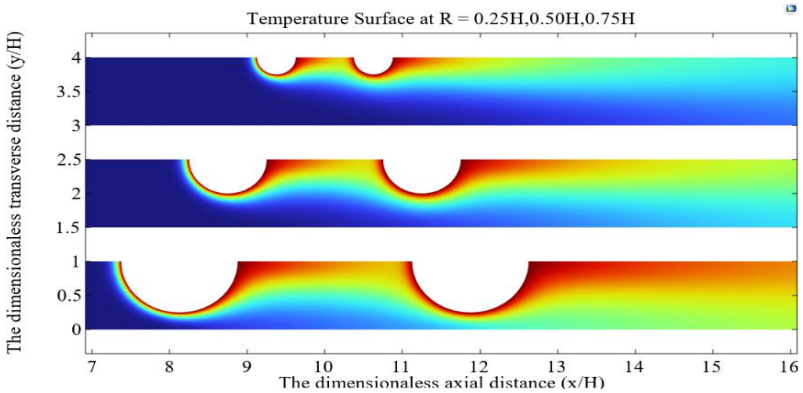


Figure 5.3. Temperature surface for the radius changes (R = 0.25H, 0.50H, 0.75H) when the length between the sections' centers is Lc= 3R and Re = 50

5.3. Effect of changing the distance between the ends of semicircular sections

Figure 5.4 was plotted to illustrate the influence of the distance between the sections on the rate of heat transfer. The figure indicates that with the radius and Reynolds number constant, increasing the distance allows the fluid passing through the first section to absorb an amount of heat and then mix with the fluid that passed near the section and did not gain heat, so the hot fluid loses an amount of its heat. Then, as it passes again through the second section, it gains an additional amount of heat. However, when the distance is small (L_c = 3 R), the fluid, after it has gained an amount of heat from the first source, does not have the opportunity to mix with the cold fluid, so it passes directly into the second source, but it does not gain significant additional heat because its temperature was originally relatively high. In other words, by increasing the distance between the two semicircles and without other hydrodynamic parameters, the fluid will have enough time (enough length) to reach regular flow. This increase in the distance helps the heat in the fluid spread more before reaching the next half circle. As the fluid passes through the second half circle, the amount of energy in the system increases. The increase in energy will show itself as a further increase in temperature in the cold region of the fluid.

Figure 5.5 illustrates the variations in the Nusselt numbers on the first and second semicircular sections as function of the distance between the semicircular sections. As seen in the figure, as the distance between the semi-circular sections increases, there is no change in the values of the Nusselt numbers on the first semi-circular section, while the values of the Nusselt numbers on the second semi-circular section increase. Since increasing the distance between the semi-circular sections has no effect on the interaction between the fluid and the first semi-circular section and on the difference between the fluid temperature and the temperature of the semi-circular section, there is no change in the values of the Nusselt numbers on the first semi-circular section. As can be seen in the figure, the distance between the channel inlet and the first semi-circular section becomes slightly shorter as the distance between the semi-circular sections increases. However, this slightly shorter distance does not affect the interaction between the fluid and the first semi-circular section since the distance between the channel inlet and the first semi-circular section is long enough that the flow behavior is not affected. As a result, there is no change in the values of the Nusselt numbers on this semi-circular section. Furthermore, even if the diameter of the semi-circular section changes, as long as the Reynolds number is kept constant, the interaction between the fluid and the semi-circular section will not change. On the other hand, as seen in Figure 5.5, as the distance between the semi-circular sections increases, the values of the Nusselt numbers on the second semi-circular section enhance. As seen in Figure 5.4, which shows the temperature distribution, it is observed that as the distance between the semicircular sections increases, the fluid temperature on the second semicircular section increases. When the distance between the semi-circular sections is short, the transferred heat from the first semi-circular sections to the fluid will be carried to the second semicircular section without being transferred to the entire fluid. Therefore, the temperature difference between the second semi-circular section and the fluid will be small and thus, less heat is transferred from the second semicircular section. As the distance between the semi-circular sections increases, the heat taken from the first semi-circular section will be transferred to the entire fluid since there is sufficient distance, and the temperature difference between the fluid and

the second semi-circular section will increase. Thereby, more heat will be transferred from the second semi-circular section to the fluid. Increasing the heat transfer between the second semicircular section and the fluid is possible by increasing the heat transfer coefficient and thus, the Nusselt number. As a result, the graphs plotted for Nusselt numbers and the graphs drawn for temperature are in harmony with each other.

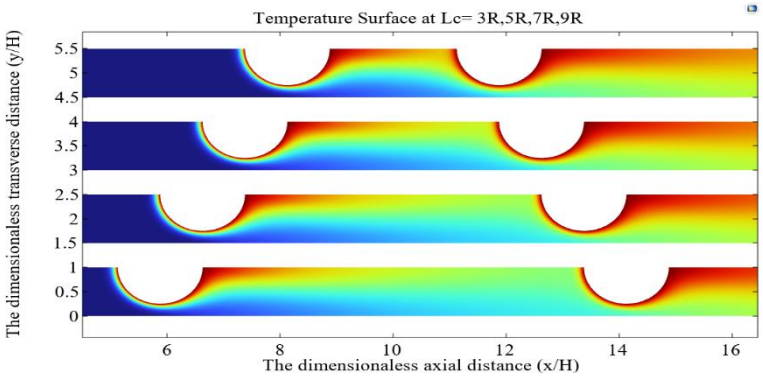


Figure 5.4. Temperature surface at R= 0.75 H and Re = 50 for different values of the distance between semicircular sections (Lc = 3 R, 5 R, 7 R, 9 R)

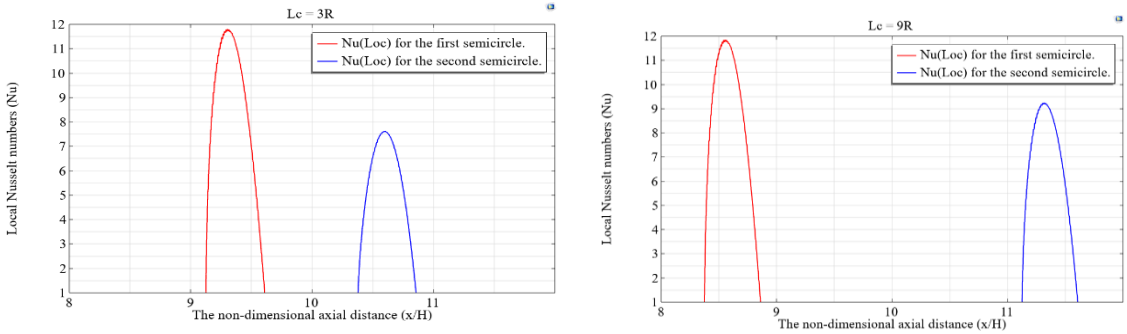


Figure 5.5. The profiles of local Nusselt number values on the semicircular sections at R= 0.25 H and Re = 50 for (Lc = 3 R and 9 R)

5.4. Determination of the influence of Reynolds number on the heat transfer rate

Figures 5.6 indicate the influence of the Reynolds number on the rate of heat transfer when the radii of the sections and the distance between them are constant. Those figures point out that the Nusselt number raises with an enhancement in Reynolds number at the constant radii of the sections and the distance between them. As shown in Figure 5.6, the rate of heat transfer is the largest possible when the Reynolds number is 200. As the graph of the local Nusselt number indicates, when the distance between the segments increases, the local Nusselt number for this second semicircle increases slightly, and there is no change in the first semicircle. This means that distance has no major effect on the Nusselt number, but the significant change in the values of the Nusselt number occurs with the increase of the Reynolds number. Additionally, the graph shows that, although the heat transfer rate is greatest when the Reynolds number is 200, the average Nusselt number rises noticeably as the Reynolds number increases, especially when the radius value is (R = 0.50H).

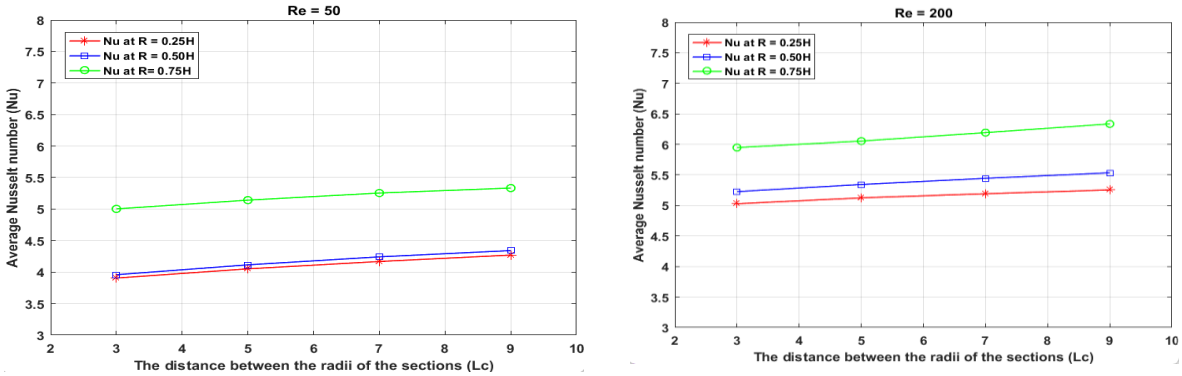


Figure 5.6. Average Nusselt numbers at Re = 50 and 200 for different radii and difference between sections

6. Conclusions

The main objective of this investigation was to perform a simulation to study the coupled effect of the presence of porous material and the waviness of the inner surfaces of the upper plate on the heat convection of non-Darcian flow in a porous medium filled in a channel with semicircular cross-sections on its upper surface. In the hope of increasing heat transfer by convection, a filled porous material channel with a wavy semicircular section on the upper wall and a flat lower wall was used in the study. The main conclusions reached in this simulation are:

- At low Darcy numbers, in addition to increasing the spread of isothermal lines and dispersing them away from the heating zone, the maximum fluid velocity was found to be near the wall rather than in the middle of the channel, resulting in a decrease in the boundary layer thickness and an increase in heat transfer.
- In terms of heat transfer, the best value was obtained for all studied Reynolds numbers when the radius of the semicircles was $R = 0.75\ H$ and the distance between them was $L_c = 9\ R$.
- It was noticed that the Nusselt number enhanced with an increment in the Reynolds number at constant radii and the distance between the boundaries of semicircular sections.
- The dense presence of porous material inside the channel prevented the occurrence of two phenomena such as separation and the formation of vortices.

The results obtained in this study are important in the design of processes involving fluid flow and forced convection heat transfer in porous media, because a more effective design can be made using the data in question.

7. References

[1] Pavel BI, Mohamad AA (2004) An experimental and numerical study on heat transfer enhancement for gas heat exchangers fitted with porous media, Int. J. Heat Mass Trans. 47: 4939–4952. https://doi.org/10.1016/j.ijheatmasstransfer.2004.06.014

[2] Hadim H, North M (2005) Forced convection in a sintered porous channel with inlet and outlet slots, Int. J. Therm. Sci. 44: 33–42. https://doi.org/10.1016/j.ijthermalsci.2004.04.016.

[3] Mahgoub SE (2013) Forced convection heat transfer over a flat plate in a porous medium, Ain Shams Eng. J. 4: 605–613. https://doi.org/10.1016/j.asej.2013.01.002.

[4] Milani Shirvan K, Mirzakhanlari S, Kalogirou SA, Öztop HF, Mamourian M (2017) Heat transfer and sensitivity analysis in a double pipe heat exchanger filled with porous medium, Int. J. Therm. Sci. 121: 124–137. https://doi.org/10.1016/j.ijthermalsci.2017.07.008.

[5] Mohamad AA (2003) Heat transfer enhancements in heat exchangers fitted with porous media Part I: constant wall temperature, Int. J. Therm. Sci. 42: 385–395. https://doi.org/10.1016/S1290-0729(02)00039-X.

[6] Yang C, Liu W, Nakayama A (2009) Forced convective heat transfer enhancement in a tube with its core partially filled with a porous medium. Open Transport Phenomena J. 1: 1-6. DOI: 10.2174/1877729500901010001.

[7] Cekmer O, Mobedi M, Ozerdem B, Pop I (2016) Effect of an inserted porous layer into a channel on heat transfer and pressure drop. J. Porous Media. 19: 65-82. DOI: 10.1615/JPorMedia.v19.i1.50.
[8] Mahmud S, Sadrul Islam A, Feroz C (2003) Flow and heat transfer characteristics inside a wavy tube, Heat Mass Transf. 39: 387–393. https://doi.org/10.1007/s00231-002-0369-9.
[9] Bahaidarah HMS, Anand NK, Chen HC (2005) Numerical study of heat and momentum transfer in channels with wavy walls. Numer. Heat Transf. Part A. 47: 417–439. https://doi.org/10.1080/10407780590891218.
[10] Islamoglu Y, Parmaksizoglu C (2004) Numerical investigation of convective heat transfer and pressure drop in a corrugated heat exchanger channel. Appl. Therm. Eng. 24: 141–147. https://doi.org/10.1016/j.applthermaleng.2003.07.004.
[11] Mehta SK, Pati S (2019) Analysis of thermo-hydraulic performance and entropy generation characteristics for laminar flow through the triangular corrugated channel. J. Therm. Anal. Calorim. 136: 49–62. https://doi.org/10.1007/s10973-018-7969-1.
[12] Mansoor FS, Dawood AS (2013) Numerical investigation of heat transfer and fluid flow characteristics inside wavy channels filled with porous media. Mech. Eng. Res. 3: 13-27. DOI:10.5539/mer.v3n2p13.
[13] Heidary H, Kermani MJ (2012) Enhancement of heat exchange in a wavy channel linked to a porous domain; a possible duct geometry for fuel cells. Int. Commun. Heat Mass Transf. 39: 112–120. https://doi.org/10.1016/j.icheatmasstransfer.2011.10.001.
[14] Akbarzadeh M, Maghrebi MJ (2018) Combined effects of corrugated walls and porous inserts on performance improvement in a heat exchanger channel. Int. J. Therm. Sci. 127: 266–276. https://doi.org/10.1016/j.ijthermalsci.2018.01.010.
[15] Kumar P, Pandey KM (2021) Numerical investigation of thermo-hydraulic transport characteristics of two-dimensional, steady flow through the partially porous wavy channel. Numer. Heat Transf. Part A Appl. 81: 31–47. https://doi.org/10.1080/10407782.2021.1969809.
[16] Kumar P, Dwivedi R, Pandey KM (2022) Numerical investigation of thermo-hydraulic transport characteristics of laminar flow through the partially filled porous wavy channel: Effect of Prandtl number. Numer. Heat Transf. Part A Appl. 83: 1045–1066. https://doi.org/10.1080/10407782.2022.2102342.
[17] Bejan A (1983) Natural convection heat transfer in a porous layer with internal flow obstructions, Int. J. Heat Mass Transf. 26: 815–822. https://doi.org/10.1016/S0017-9310(83)80105-7.

Bellman-Ford and SMOTE Based Optimization Method in Hyperspectral Image Classification

Yücel DEMIREL[1], Orhan YAMAN[2], Mehmet KARAKÖSE[3]

[1]Graduate School of Natural and Applied Sciences Digital Forensic Engineering, Firat University, Elazig, Turkey
[2]Digital Forensic Engineering, Firat University, Elazig, Turkey
[3]Department of Computer Engineering, Firat University, Elazig, Turkey
[1]232144108@firat.edu.tr, [2]orhanyaman@firat.edu.tr, [3]mkarakose@firat.edu.tr
[1](ORCID: 0009-0008-7808-0422), [2](ORCID: 0000-0001-9623-2284), [3](ORCID: 0000-0002-3276-3788)

Abstract
In this study, hyperspectral data classification performance was improved by Bellman-Ford based size reduction, data balancing with SMOTE, and SVM methods. An accuracy of 94.4% was achieved in the Indian Pines dataset, and the processing load was reduced by reducing the spectral band from 200 to 90 with size reduction. SMOTE solved the problem of class imbalance, while SVM provided high accuracy with optimized hyperparameters. The proposed method offers an effective solution for hyperspectral data analysis.

Keywords: bellman ford algorithm, size reduction, smote, support vector machines

1. Introduction

Hyperspectral image classification is used in a wide range of fields such as remote sensing, agriculture, environmental monitoring, and urban planning. However, the high dimensionality of hyperspectral data sets and the imbalance between classes are the biggest challenges encountered in this field. Examining hundreds of spectral bands can result in both high computational cost and a decrease in classification performance[1,2,3]. In this study, two main approaches are discussed to address the problems of size reduction and unbalanced data: the Bellman-Ford algorithm and the SMOTE (Synthetic Minority Oversampling Technique). While the Bellman-Ford algorithm offers an effective solution in reducing the high dimensionality of hyperspectral data as a graph-based method, SMOTE aims to improve classification performance by balancing the minority class to the majority class [4,5,6].

In addition, support vector machines (SVM) are a machine learning algorithm that is often preferred in hyperspectral image classification. The potential of SVM to achieve high accuracy in hyperspectral datasets can be further enhanced when combined with size reduction and data balancing methods[7,8,9].

This study aims to increase SVM classification accuracy and reduce computational overhead by combining Bellman-Ford algorithm-based size reduction with the SMOTE method. The study results show that the proposed approach provides an effective performance in hyperspectral data sets.

2. Literature Review

Hyperspectral image classification has an extensive literature on techniques that address class imbalances, especially size reduction methods. Traditional methods such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are widely used for size reduction. However, it has been noted that these methods are limited in modeling the relationships between bands [2,7,10].

In recent years, graph-based methods, especially the Bellman-Ford algorithm, have emerged as a powerful alternative for hyperspectral data analysis. By modeling weighted edges as a graph, Bellman-Ford reduces the impact of unnecessary or repetitive bands and optimizes the size reduction process [1,2,3]. Furthermore, the algorithm's parallel processing capabilities improve computational efficiency when working with large data sets[4,11].

Imbalanced data sets, on the other hand, can negatively affect the accuracy of machine learning models. SMOTE is one of the most popular methods used to solve this problem. SMOTE creates synthetic samples by linear interpolation

[1]Corresponding author

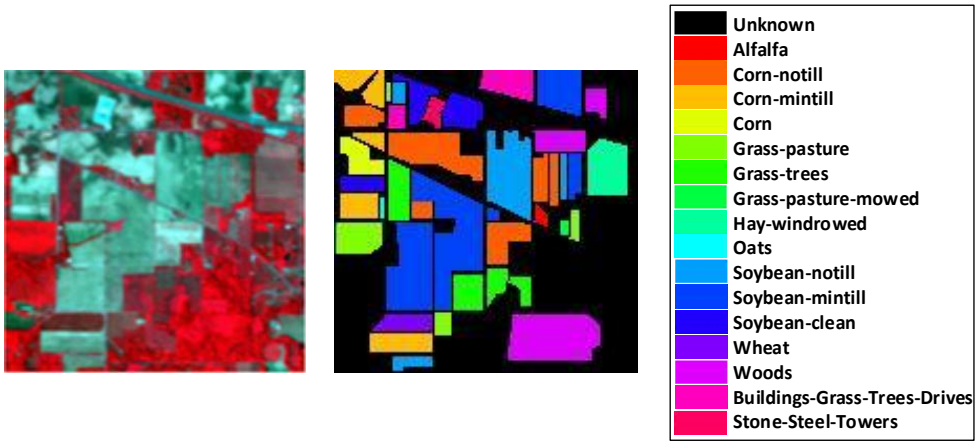
between K-nearest neighbors of minority data[5,6]. In addition, SMOTE derivatives such as Borderline SMOTE and ADASYN have also been extensively reviewed in the literature [10,9].

In hyperspectral data classification, SVM has the potential to achieve high accuracy rates. However, issues such as class imbalance and high dimensionality can limit the performance of SVM models. Therefore, data balancing with SMOTE and size reduction with Bellman-Ford are considered an effective solution to improve the classification accuracy of SVM[8,12,3].

In conclusion, studies in the literature show that the integration of size reduction and data balancing methods is an effective strategy for hyperspectral data analysis. This study aims to achieve higher accuracy and efficiency by combining these methods with SVM.

3. Dataset

Indian Pines dataset was used in the study. This dataset contains 200 spectral bands with a total size of 145x145. Each pixel also has class labels. The labels used in the classification represent 16 different classes. In the analysis, the classification process was carried out by using labels along with the hyperspectral data set.



(a) (b) I

Figure 1. Indian pines dataset: (a) artificial rgb image, (b) reference image, (c) tagged data types [12]

4. Proposed Method

In this study, a method is proposed to address the size reduction and class imbalance problems encountered in the classification of hyperspectral images. The method includes size reduction based on the Bellman-Ford algorithm, class balancing with SMOTE (Synthetic Minority Oversampling Technique), and then classification with Support Vector Machines (SVM).

4.1. Data Balancing with Smote

Class imbalance in hyperspectral data sets is one of the main problems that negatively affect classification performance. Minority classes, in particular, are underrepresented, making it difficult for machine learning algorithms to learn these classes correctly. In this study, SMOTE (Synthetic Minority Oversampling Technique) was used to balance minority classes and increase classification accuracy. The SMOTE method balances the class distribution by generating synthetic samples from data belonging to the minority class. In this process, the data point belonging to each minority class creates new data points using linear interpolation between their K-nearest neighbors. This approach adds new samples to the dataset without any loss of information and resolves class imbalance. The

use of SMOTE has improved the overall model performance while improving classification accuracy, especially in minority classes. This method is used during the classification of imbalanced data sets.[14]

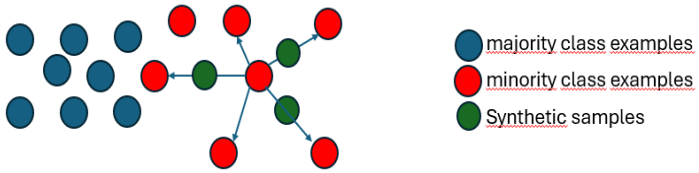


Figure 2. SMOTE algorithm working system

4.2. Dimensionality Reduction with the Bellman-Ford Algorithm

HiThe high dimensionality of perspectral data sets is an important problem that negatively affects classification performance. In this study, Bellman-Ford algorithm, which is a graph-based method, was used for size reduction. The Bellman-Ford algorithm is an approach that calculates distances from the initial band to other bands and selects the most important bands based on these distances [15]. The relationships between the spectral bands in the dataset were modeled using a graph structure, and the edge weights were determined by the Euclidean distance between the bands.The algorithm eliminates unnecessary and repetitive bands using a certain threshold value. This process both reduces the size of the data set and reduces the processing burden during classification. Implemented in an iterative manner, the algorithm selects only the most important bands, creating a more significant subset of hyperspectral data. As a result, size reduction with the Bellman-Ford algorithm significantly reduced processing time while improving classification accuracy in hyperspectral data sets [16].

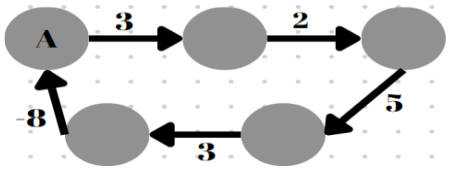


Figure 3. Size reduction process in hyperspectral data set with bellman-ford algorithm

4.3. SVM – Based Classification

After size reduction and data balancing, Support Vector Machines (SVM) were used for the classification of the hyperspectral data set. SVM is an algorithm that can perform classification with high accuracy by determining the hyperplane that provides the best separation between classes. In this study, hyperparameter optimization was performed with the Grid Search method to increase the performance of SVM. Different kernel functions, Box Constraint, and Kernel Scale parameter combinations were tested, and the best model was selected. The results showed that SVM provides high accuracy in hyperspectral data sets after size reduction and data balancing.

Table 1. SVM classification parameters and values

	Parameters	Values
SVM Classification	KernelFunction	Polynominal
	PolynomialOrder	3
	KernelScale	Auto
	BoxConstraint	1
	Standardize	True
	OptimizationMethod	Grid Search
	CrossValidation	10-Fold

5. Experimental Results

The Confussion matrix after applying the smote method to the Indian Pines data set and reducing it with the Bellman Ford algorithm is given.

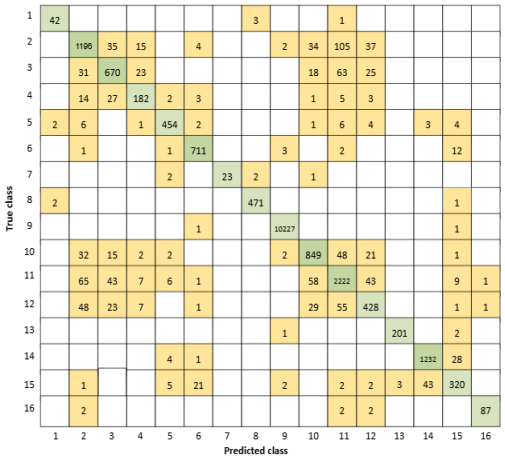


Figure 4. Indian pines dataset SMOTE + bellman ford + SVM confusion matrix

The changes in SVM accuracy rates and the change in the number of bands before and after the reduction of the data set studied in the application are given.

Table 2. Results obtained with the bellman – ford algorithm

	Pre – Reduction	**Post – Reduction**
Number of Bands	200	90
SVM Accuracy Rate	91.5	94.4

The proposed method reduced the number of bands of the dataset by 55%. This band-reduction process has reduced the size of the data set, accelerating the classification process.

The size reduction of the dataset with the Bellman-Ford algorithm and data balancing with SMOTE provided higher accuracy rates compared to the SVM-only approach. Especially in the Indian Pines dataset, an accuracy of 94.4% was achieved with the proposed method, and this increase stood out as a significant improvement in classification performance.

Size reduction with the Bellman-Ford algorithm contributed to the creation of a more meaningful and optimized feature set for classification, while SMOTE enabled better learning of minority classes in unbalanced data sets. This shows that the proposed method is an effective approach to both size reduction and data balancing processes.

As a result, the proposed method (Bellman-Ford + SMOTE + SVM) performed superior compared to the method using SVM alone, both in terms of accuracy and other performance metrics. These findings clearly demonstrate the efficiency and effectiveness of the proposed method in hyperspectral data classification.

5. Discussion

In this study, Bellman-Ford algorithm based size reduction, data balancing with SMOTE and SVM classification methods were used together to provide solutions to the high dimensionality and class imbalance problems encountered in hyperspectral data classification. The results showed that the proposed method provides higher accuracy and efficiency compared to traditional approaches using SVM alone.

In the process of size reduction, the Bellman-Ford algorithm eliminated unnecessary and repetitive bands, reduced dimensionality in the data set, and eased the processing burden. For example, in the dataset reduced from 200 spectral bands to 90 bands, the processing time decreased, while the classification accuracy reached 94.4%. This suggests that the size reduction process not only increases efficiency, but also improves the generalization ability of the classification model.

Data balancing with SMOTE has positively impacted classification performance by providing better representation of minority classes. The increase in accuracy, especially in unbalanced data sets, confirms that SMOTE is an effective method for solving class imbalance problems. In addition, optimizing SVM with Grid Search has improved model performance by enabling the selection of optimal hyperparameters.

As a result, the proposed combination of Bellman-Ford + SMOTE + SVM has provided an effective solution with high accuracy rates and processing efficiency in hyperspectral data classification. The study provides both theoretical and practical contributions in hyperspectral data analysis and provides a basis for future research where comparisons can be made with different data sets and methods.

6. Conclusions

In this study, hyperspectral data classification performance was improved by using Bellman-Ford algorithm-based size reduction, data balancing with SMOTE, and SVM classification methods. An accuracy rate of 94.4% was achieved in the Indian Pines dataset, and the processing load was reduced by reducing the 200 spectral band to 90 thanks to the size reduction.

SMOTE solved the problem of class imbalance, resulting in better representation of minority classes, while SVM achieved a high accuracy rate with optimized hyperparameters. The proposed method provides an effective solution for hyperspectral data analysis and serves as a basis for future studies.

7. References

[1] Conrad, E., & Dutle, A. (2024, March). Formalization of the Bellman-Ford Algorithm for Airspace Applications. In *55th Southeastern International Conference on Combinatorics, Graph Theory and Computing*.

[2] Timofeeva, O., Sannikov, A., Stepanenko, M., & Balashova, T. (2023). Modification of the Bellman–Ford Algorithm for Finding the Optimal Route in Multilayer Network Structures. *Computation*, *11*(4), 74.

[3] Elmasry, A. (2024). Breaking the Bellman-Ford Shortest-Path Bound. *arXiv preprint arXiv:2410.23383*.

[4] Elreedy, D., Atiya, A. F., & Kamalov, F. (2024). A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning. *Machine Learning*, *113*(7), 4903-4923.

[5] Wongvorachan, T., He, S., & Bulut, O. (2023). A comparison of undersampling, oversampling, and SMOTE methods for dealing with imbalanced classification in educational data mining. *Information*, *14*(1), 54.

[6] SAXENA, M. C., SABHARWAL, M., & BAJAJ, P. (2023). COMPARATIVE STUDY AND PROPOSAL TO USE ENHANCED BELLMAN-FORD ALGORITHM FOR FASTER PATH COMPUTATION ON SDWAN CONTROLLER. *Journal of Theoretical and Applied Information Technology*, *101*(16).

[7] Nguyen, T., Mengersen, K., Sous, D., & Liquet, B. (2023). SMOTE-CD: SMOTE for compositional data. *Plos one*, *18*(6), e0287705.

[8] Arnau, R., Calabuig, J. M., García-Raffi, L. M., Sánchez Pérez, E. A., & Sanjuan, S. (2024). A Bellman–Ford Algorithm for the Path-Length-Weighted Distance in Graphs. *Mathematics*, *12*(16), 2590.

[9] Liaw, L. C. M., Tan, S. C., Goh, P. Y., & Lim, C. P. (2025). A histogram SMOTE-based sampling algorithm with incremental learning for imbalanced data classification. *Information Sciences*, *686*, 121193.

[10] Altin, F. G., Budak, I., & Özcan, F. (2023). Predicting the amount of medical waste using kernel-based SVM and deep learning methods for a private hospital in Turkey. Sustain. Chem. Pharm. 33, 101060 (2023).

[11] Weber, A., Kreuzer, M., & Knoll, A. (2020, May). A generalized Bellman-Ford algorithm for application in symbolic optimal control. In *2020 European Control Conference (ECC)* (pp. 2007-2014). IEEE.

[12] Yaman, O., Yetis, H., & Karakose, M. (2020, May). Band Reducing Based SVM Classification Method in Hyperspectral Image Processing. In *2020 Zooming Innovation in Consumer Technologies Conference (ZINC)* (pp. 21-25). IEEE.
[13] Yavaş, M., Güran, A., & Uysal, M. (2021). Covid-19 Veri Kümesinin SMOTE Tabanlı Örnekleme Yöntemi Uygulanarak Sınıflandırılması. Avrupa Bilim ve Teknoloji Dergisi, 258-264.
[14] Dablain, D., Krawczyk, B., & Chawla, N. V. (2022). DeepSMOTE: Fusing deep learning and SMOTE for imbalanced data. *IEEE transactions on neural networks and learning systems*, *34*(9), 6390-6404.
[15] Karthikeyan, M., Sekar, S., Basha, H. A., Epsiba, P., Ganesh, E. N., & Meenakshi, B. (2024, October). Dijkstra's Algorithm with Bellman-Ford for Shortest Path Discovery in Big Data. In *2024 2nd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS)* (pp. 1348-1354). IEEE.
[16] He, G., Zhang, X., Sun, Y., Luo, G., & Chen, L. (2021). Contour line simplification method based on the two-level Bellman–Ford algorithm. *Transactions in GIS*, *25*(1), 396-423.

Classification of Liver Disease Using Biochemical Data: A Machine Learning Approach

Ebrar TURUDU[1], Gülşah GÜNDÜZTEKİN [2], Gülşah KARADUMAN [3]

[1,2]Computer Engineering, Fırat University, Elazığ, Türkiye.
[3]Computer Engineering, Fırat University, Elazığ, Türkiye.
[1]210260044@firat.edu.tr, [2]241129121@firat.edu.tr, [3]gkaraduman@firat.edu.tr
[1](ORCID: 0009-0002-1199-7929), [2](ORCID: 0009-0007-1023-3229),[3](ORCID: 0000-0001-8034-3019)

Abstract
In recent years, the improvement of healthcare processes using machine learning algorithms in areas such as early disease diagnosis, prevention of disease spread, and increasing the quality of healthcare services has become an effective method. Machine learning, which offers a realistic approach using datasets 547orld547h classification and prediction of diseases, is preferred by researchers. The liver, which is the largest organ in the abdominal region, is primary for balancing many nutrients, vitamins, hormones, and cholesterol. Early diagnosis of liver diseases is vital in slowing the progression of the disease and increasing treatment success. In this study, classification was performed using XGBoost, one of the machine learning algorithms, with the liver disease patient 547orld547h obtained from the Kaggle platform. The aim is to predict with high accuracy whether patients have liver disease based on their laboratory data. The results showed that the model achieved high accuracy and balanced performance in predicting both healthy and diseased individuals. The study highlights the potential of machine learning algorithms for reliable medical decision support systems and contributes to the development of fast, cost-effective diagnostic tools in the field of liver disease.

Keywords: liver disease, machine learning, xgboost, classification, medical diagnosis

1. Introduction

Liver diseases are chronic conditions that lead to serious public health consequences worldwide and can be fatal if not diagnosed early [1]. Approximately 2 million people die from liver disease each year worldwide [2]. In recent years, the increase in hepatitis B and C infections, cirrhosis, and liver cancer cases has made it necessary to reevaluate diagnostic and treatment methods in this field [3]. Existing clinical procedures are often time-consuming, costly, and invasive; this poses a barrier to the early detection of disease [4]. Therefore, the development of more effective, rapid, and cost-efficient diagnostic systems in the medical field has become a significant need.

With the increase and digitalization of medical data, the integration of artificial intelligence-based methods such as machine learning (ML) and deep learning (DL) into the healthcare field has rapidly accelerated to extract meaningful insights from these data. In particular, ML algorithms have yielded promising results in clinical decision support systems, and methods that enable the extraction of meaningful patterns from large datasets can offer higher accuracy rates in classification, prediction, and analysis of health data compared to traditional statistical methods [5]. Algorithms such as Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Deep Neural Networks (DNN) have shown effective performance in classification, feature selection, and modeling of non-linear relationships. RF algorithms are known for their robustness against noise in datasets and provide high accuracy especially in medical data where missing and imbalanced data are common [6]. XGBoost algorithms are known for their success in classification tasks and fast training times, especially in large datasets [7]. Deep learning algorithms like DNNs have the capacity to discover complex relationships between data through their layered structures [8]. From this perspective, data-driven modeling during the diagnosis of liver diseases has enabled fast and precise results. However, in the 547orld547he547e, there are still gaps regarding the reliable classification of liver diseases based on clinical data and the correct model selection. This study aims to evaluate different ML models that may be effective in the diagnosis of liver diseases.

[1]Corresponding author

2. Literature Review

In the 548orld548he548e, there are various studies conducted with different algorithms 548orld548h classification of liver diseases. For example, Ahmad et al. (2023) stated that DNN-based models achieved over 90% success in liver tumor classification [9]. However, most of these studies either have small sample 548orld548he the clinical variables used have limited explanatory power. In addition, there is insufficient transparency regarding technical aspects such as hyperparameter optimization, cross-validation methods, and general evaluation criteria. In a study conducted in the 548orld548he548e, it was reported that accuracy rates of up to 88% were achieved using the RF algorithm [10]. Similarly, in applications using XGBoost, consistent results were obtained especially in performance metrics such as AUC (Area Under Curve) and F1 score [11]. In studies where deep learning methods were used together with imaging data, the success of liver tumor classification exceeded 90% [12]. However, many studies in the 548orld548he548e have limitations such as relying on small sample sizes, not comparing models with a standardized method, and not focusing adequately on hyperparameter optimization. Therefore, more comprehensive and systematic approaches are needed. Most of these studies either rely on limited sample 548orld548he do not sufficiently compare multiple models under consistent evaluation protocols.

In this study, using the "Liver Disease Patient Dataset – 30K Train Data" available on the Kaggle platform, the success of different machine learning models in predicting liver diseases will be compared [13]. The 548orld548h includes biochemical parameters such as age, gender, total bilirubin, alkaline phosphatase, alanine aminotransferase (ALT), aspartate aminotransferase (AST), 548orld548h, and prothrombin ratio. These parameters are key clinical variables that reflect liver functions and play an important role in disease diagnosis. This article aims to address these gaps through systematic experiments with commonly used ML models on a well-structured 548orld548h. In particular, the comparative evaluation of methods such as Random Forest, XGBoost, and Deep Neural Network has shown which models can provide more reliable results in this field. Thus, compared to current diagnostic processes, faster and more cost-effective solutions have been contributed. In addition, the foundations of systems that healthcare professionals can use in the diagnostic process and that are applicable in practice have been laid. Considering the sample limitations and evaluation deficiencies in the 548orld548he548e, the comprehensive structure and systematic approach of this study are of a 548orld548 to guide future research.

3. Dataset and Data Preprocessing

The 548orld548h used contains laboratory results of nearly 30,000 individuals. These results are the data required to determine whether an individual is healthy or sick. The main variables in the 548orld548h are as follows:

- Age of the patient (The age of the individual)
- Gender (Gender)
- Total Bilirubin (Total bilirubin level in the blood)
- Direct Bilirubin (Water-soluble bilirubin level)
- Alkaline Phosphotase (An enzyme that helps break down proteins)
- SGPT / SGOT (Liver enzymes)
- Albumin (A plasma protein synthesized by the liver)
- A/G Ratio (The ratio of 548orld548h to 548orld548he proteins)
- Result (1: Sick, 2: Healthy)
- The 548orld548h used has an imbalanced structure. In total, 21,917 samples belong to sick individuals (label: 1), and 8,774 samples belong to healthy individuals (label: 2).

Missing values in numerical variables were filled with the mean value of the respective column. Missing values in the categorical variable "Gender" column were filled with the most frequently occurring value (mode). The Gender column was numerically encoded using LabelEncoder. To enable the classification algorithms to work more effectively, these labels were re-encoded as 1 (sick) and 0 (healthy).

4. Findings and Classification Application

When the analyses on the data used in the study are examined, correlations were obtained between Total Bilirubin and Direct Bilirubin (88%), SGPT and SGOT (77%), and Albumin and A/G Ratio (67%). Direct Bilirubin (-0.25) and Total Bilirubin (-0.22) were the variables that showed a significant negative correlation with the target variable (result).

The obtained data were used for classification purposes. The 549orld549h was divided into two groups based on whether individuals were sick or healthy, and the data were used 80% for model training and 20% for testing. During this separation, the distribution of the target variable was taken into account to maintain class balance, and stratified learning was applied.

Direct Bilirubin, Total Bilirubin, Alkaline Phosphotase, SGPT (Alamine Aminotransferase), and Albumin were the variables that the model attached the most importance to, as they are directly related 549orld549he functions. This indicates that the model has performed a learning process consistent with medical reality.

5. Conclusion and Discussion

In this study, a very high accuracy of 99.8% was achieved in liver disease classification with the XGBoost model. In particular, the F1-score, precision, and recall values are quite balanced for both sick and healthy individuals.

Figure 1 shows the confusion matrix summarizing the performance of the developed model. According to the model, 1747 (TN) individuals who are actually healthy were correctly classified as healthy, and 4380 (TP) individuals who are actually sick were correctly classified as sick. 8 (FP) healthy individuals were incorrectly predicted as sick, and 4 (FN) sick individuals were incorrectly predicted as healthy.

When these results are examined, it is seen that the model achieved high success for both classes, and especially that the false negative (FN) and false positive (FP) rates are quite low.

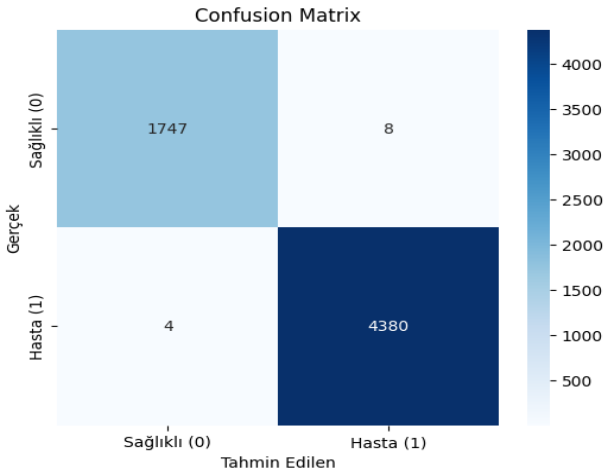


Figure 1. Confusion matrix

Figure 2 presents the metrics used to evaluate the classification performance of the model. The evaluation metrics used include precision, recall, F1-score, and support.

The precision values are the 549orld549he both class 0, which represents healthy individuals, and class 1, which represents sick individuals. In terms of recall values, the result is 0.995 for healthy individuals and 0.999 for sick

individuals. As 550orld550h F1 scores, it is 0.997 for healthy individuals and 0.999 for sick individuals. These results 550orl that almost all healthy and sick individuals were correctly classified.

The overall accuracy rate (accuracy) was obtained as 99.8%. The macro average F1-score, which reflects the equally weighted average across classes, was found to be 0.998, and the weighted average F1-score was also 0.998. These results indicate that the model was successful in class distribution and also achieved sensitive learning in both classes.

	precision	recall	f1-score	support
0	0.998	0.995	0.997	1755.000
1	0.998	0.999	0.999	4384.000
accuracy	0.998	0.998	0.998	0.998
macro avg	0.998	0.997	0.998	6139.000
weighted avg	0.998	0.998	0.998	6139.000

Figure 2. Classification report

The strong performance of the model can be attributed to the following reasons:
• Low-noise, meaningful medical data
• Proper preprocessing and careful distinction for imbalanced classes
• The strong learning capacity of the XGBoost algorithm

6. Recommendations and Future Work

The high accuracy rates obtained in this study 550orl that the developed model is promising for liver disease predictions. For future studies:
• Model performance can be analyzed more deeply with ROC-AUC analysis.
• A comparison can be made with other classifiers (Random Forest, LightGBM).
• With SHAP (Shapley) values, it is possible to 550orld550he model's decisions more explainable.
• If there is time information in the data, the progression of the disease can be modeled as a time series.

7. References

[1] World Health Organization. (2023). Liver disease mortality estimates. Global Health Estimates.
[2] Kumar, S., Singh, P., & Dubey, H. (2022). Liver disease classification using ensemble machine learning. Computers in Biology and Medicine, 142, 105228. https://ieeexplore.ieee.org/abstract/document/9758929
[3] Asrani, M., Devarbhavi, H., Eaton, J., & Kamath, P. S. (2019). Burden of liver diseases in the 550orld. Journal of Hepatology, 70(1), 151–171.
[4] Younossi, Z. M., Koenig, A. B., Abdelatif, D., Fazel, Y., Henry, L., & Wymer, M. (2014). Global epidemiology of nonalcoholic fatty liver disease—Meta-analytic assessment of prevalence, incidence, and outcomes. Gastroenterology, 145(1), 87–99.e4. https://www.sciencedirect.com/science/article/abs/pii/S0016508513004988
[5] Rajkomar, P., Dean, J., & Kohane, I. (2019). Machine learning in medicine. New England Journal of Medicine, 380(14), 1347–1358.
[6] Rajkomar, P., Dean, J., & Kohane, I. (2019). Machine learning in medicine. New England Journal of Medicine, 380(14), 1347–1358.
[7] Wang, X., Liu, Y., & Zhao, H. (2022). Liver disease prediction with XGBoost. PubMed. https://pubmed.ncbi.nlm.nih.gov/35912345

[8] Rajalakshmi, R., Subramanian, V., & Prasad, K. (2022). Feature extraction in liver disease classification. Procedia Computer Science, 185, 39–45.
[9] Ahmad, M., Khan, S., & Ali, R. (2023). Deep learning-based liver tumor detection. ScienceDirect. https://www.sciencedirect.com/science/article/pii/S2772947823000087
[10] Kumar, S., Sharma, N., Gupta, P., & Verma, R. (2022). Liver disease classification using ensemble machine learning. Computers in Biology and Medicine, 142, 105228.
[11] Chen, H., & Zhang, B. (2023). An ensemble framework for liver disease diagnosis using feature selection and stacking. Expert Systems with Applications, 213, 119075.
[12] Tjoa, L. S., & Guan, C. (2021). A survey on explainable artificial intelligence (XAI): Towards medical XAI. IEEE Transactions on Neural Networks and Learning Systems, 32(11), 4793–4813.
[13] Shrivastava, A. (n.d.). Liver Disease Patient Dataset. Kaggle.

International Conference on Advances and Innovations in Engineering

3.ICAIE

Efficiency of Energy Management in Unmanned Aerial Vehicles: An Examination within The Framework of Current Approaches

Müslüme MUTLU[1], Ebru KARAKÖSE [2]

[1,2]Aviation Science and Technology, Civil Aviation College, Fırat University, Elazığ, TÜRKİYE.
[1]mutlumuslume9@gmail.com, [2]eozbay@.firat.edu.tr
[1](ORCID: 0009-0001-1388-3805), [2](ORCID: 0000-0003-1191-6375)

Abstract
Unmanned Aerial Vehicles (UAVs) are increasingly utilized in various domains such as defense, logistics, agriculture, and environmental monitoring, making energy efficiency a crucial factor in their operation. This study comprehensively explores energy management and optimization in UAV systems from a multidimensional perspective. Initially, UAV classification and corresponding energy consumption profiles are analyzed, with a focus on the operational and structural differences of fixed-wing, rotary-wing, and hybrid configurations. Subsequently, the impacts of battery technologies, charge-discharge cycles, thermal regulation, and system design on energy efficiency are investigated. The analysis emphasizes the critical roles of flight dynamics, mission profiles, environmental conditions, and payload characteristics in determining energy demand. In this context, flight path optimization and mission planning are shown to be essential for enhancing mission success and achieving significant energy savings. Furthermore, the study evaluates the efficacy of machine learning algorithms, including reinforcement and deep learning models, in real-time energy estimation and battery life prediction, revealing high levels of predictive accuracy. From a sustainability standpoint, the study also assesses the environmental impact of energy management through carbon emission data and life cycle analysis (LCA), highlighting the importance of integrating renewable energy sources to reduce the ecological footprint of UAV operations. The findings indicate that adopting a systematic, data-driven, and integrated energy management strategy is vital for maximizing both technical performance and environmental sustainability in UAV applications.

Keywords: Unmanned aerial vehicles, energy optimization, battery systems, machine learning, environmental impact

1. Introduction

Unmanned aerial vehicles (UAVs) are systems developed at the intersection of aviation and robotics technologies that can perform specific tasks remotely or autonomously. In recent years, with technological advancements, UAVs have been used in an extremely wide range of applications, from military reconnaissance and surveillance to civilian logistics operations, disaster management to agricultural spraying, search and rescue activities to environmental monitoring [1]. Thanks to their operational flexibility, low-cost mission execution, and ability to operate in inaccessible areas without human risk, they have become an indispensable component of modern air systems [2]. However, this technological advancement also brings fundamental engineering challenges. Critical parameters such as mission duration, payload capacity, maneuverability, and operational reliability are directly related to energy management [3]. Energy is the primary limiting resource for UAV systems, and since they primarily operate on battery power, various factors—from flight dynamics to mission routes, payload profiles to environmental conditions—directly impact energy efficiency. Studies in the literature indicate that energy management must be addressed in a multi-layered manner, not only through battery size but also through flight control algorithms, hardware design, material selection, and data-driven software solutions [4]. For this reason, energy management in UAVs is not merely a technical subsystem but a strategic component that determines the performance of the entire structure. Increasing energy efficiency is also of great importance in terms of extending flight time, increasing mission success rates, and reducing environmental impacts such as carbon emissions. In particular, increasing sustainability pressures necessitate minimizing the environmental footprint, requiring energy management to be considered not only an operational but also an ecological responsibility [5]. This study aims to comprehensively address energy management in UAV systems, including their structural, functional, and technological components. In this context, current technical solutions developed to increase energy efficiency, energy consumption profiles of different UAV types, the impact of battery technologies on system performance, flight optimization, and route planning strategies have been evaluated. Additionally, artificial intelligence-supported approaches such as machine learning-based energy prediction algorithms and real-time battery life projection systems have been included in the analytical framework of the study [6]. The main objective here is to develop a data-driven, modular, and sustainable

[1]Corresponding author

approach model for UAV energy management processes and to demonstrate how this model contributes to the system as a whole in terms of both technical success and environmental sustainability. UAVs are classified in different ways according to their purpose, size, range, flight time, and carrying capacity. These classifications are decisive not only from an operational perspective but also in terms of energy management. This is because each class has different energy profiles and consumption requirements. The most common classification is into fixed-wing, rotary-wing, and hybrid UAVs; these classes are further divided into subclasses such as micro, mini, tactical, medium-altitude-long-endurance (MALE), and high-altitude-long-endurance (HALE) [7]. Fixed-wing UAVs generally have the advantage of consuming less energy while covering longer distances. This design consumes a large portion of its energy during takeoff and achieves energy efficiency during flight by utilizing lift force. These types of UAVs are used for long-range tasks such as scanning large areas, environmental monitoring, and agricultural applications. However, fixed-wing systems exhibit limited performance in narrow-area tasks due to their inability to hover in the air and their low maneuverability. Their energy requirements are generally constant and predictable, which provides an advantage in terms of battery management [8]. Rotary-wing UAVs stand out for their ability to perform vertical takeoff and landing, hover in the air, and have high maneuverability. However, these advantages come with significant energy consumption. In particular, multi-rotor UAVs consume much more energy than fixed-wing UAVs over the same distance. Therefore, battery selection and energy distribution management are much more critical in rotorcraft systems. Additionally, as the number of rotors increases, so does the energy requirement, directly affecting battery capacity, charging times, and flight duration [9]. Hybrid systems combine fixed-wing and rotorcraft structures to provide both long-range flight and vertical takeoff/landing capabilities. While these systems are capable of handling complex mission profiles, they require much more complex algorithms for energy management. This is because energy consumption patterns vary significantly across different flight modes. This necessitates the use of dynamic power distribution, hybrid battery systems, and adaptive flight control systems to ensure energy efficiency [10]. Why is the following figure so necessary? I couldn't find a single line in the text explaining its necessity. At this point, comparing and analyzing the energy consumption behaviors of different UAV types is not merely a theoretical classification but also plays a critical role in developing practical energy management strategies. Especially in mission-based scenarios, the differences in energy profiles between fixed, rotary, and hybrid systems directly shape fundamental engineering parameters such as battery life, mission duration, and payload capacity [11]. The comparative data used in the study in the form of visuals and tables are not only supportive; they also serve to make the relationships between variables interpretable by visualizing them. In particular, the figure showing the effect of load distribution on flight dynamics goes beyond theoretical classification and provides a data-driven perspective that can be converted into engineering-based decision-making mechanisms. Such graphical representations are also an important tool for intuitively understanding the energy behavior of systems and determining which flight type is more suitable for which mission profile [12].

When evaluated from an energy efficiency perspective, fixed-wing systems offer low energy consumption and long-range advantages but may suffer performance losses in missions requiring maneuverability. In contrast, rotary-wing systems excel in operations requiring precise takeoff and landing in confined spaces but can achieve significantly higher energy consumption levels. Hybrid systems, on the other hand, have the potential to serve versatile mission profiles by combining the advantages of both structures; however, this comes at the cost of more complex energy management software [13]. Battery management is not limited to flight time but is also decisive in terms of system safety, thermal balance, and recharge times. Comparative analyses of Li-ion and LiPo cells developed in recent years have shown that the choice of battery type based on mission type and energy requirements directly affects system performance [14]. However, artificial intelligence-based life prediction systems and battery temperature-sensitive control algorithms also enable the effective use of these technologies [15]. Classifications and analyses conducted without adopting a holistic approach to energy management in UAV systems fail to adequately reflect the system's actual operational load. Therefore, energy management strategies developed by considering both theoretical models and mission-based scenarios not only save energy but also increase the system's mission success, minimize its environmental impact, and contribute to long-term sustainability [16].

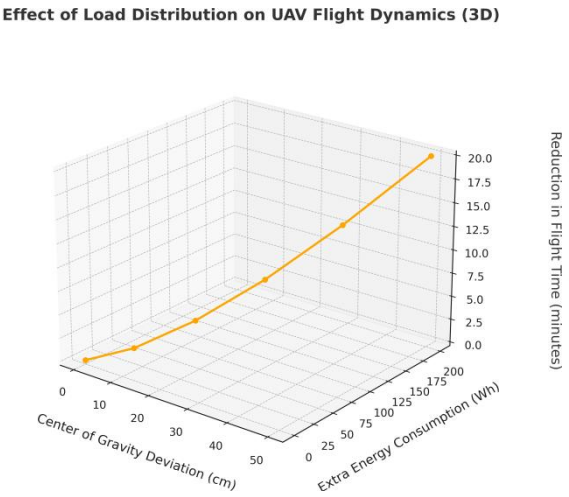


Figure 1. Effect of load distribution on UAV flight dynamics [17]

Load distribution is a parameter that has a direct and often critical effect on the flight dynamics of unmanned aerial vehicles (UAVs). The visual representation in Figure 1 compares the effects of different loading scenarios on a UAV's lift force, center of gravity (CG), thrust distribution, and maneuverability. Improperly balanced load placement causes deviations in both the horizontal and vertical axes, reducing flight stability; as a result, more energy-consuming compensatory maneuvers are required. Another notable aspect in the figure is that when the center of gravity deviates from the UAV's geometric center, the motor power distribution becomes asymmetric. This situation causes one of the motors to be overloaded and the other to consume less energy, especially in multi-rotor systems. Such a scenario not only makes it difficult to optimize battery consumption but also has a negative impact on motor life. In platforms where balanced load placement cannot be achieved, more corrective commands must be issued during flight, increasing the computational load on control algorithms [18]. The geometry of the load's placement on the UAV affects not only energy consumption but also the capacity to absorb wind-induced oscillations. Especially in fixed-wing systems, if the load is not evenly distributed along the wingspan, the torsional moment increases, creating aerodynamic imbalance. In such cases, the system may need to perform more frequent altitude corrections, which both shortens the mission duration and increases battery consumption [19]. In light of these findings, it is clear that energy efficiency analyses based solely on motor technology or battery capacity are insufficient. The positioning of the load, its compatibility with the system geometry, and its alignment with the mission profile are components of energy management that are at least as important as hardware elements. Especially in scenarios where the load's position changes during the mission or dynamic imbalances arise during maneuvers, the presence of autonomous balancing systems and the effectiveness of AI-based decision-makers become decisive factors. In this context, the contribution of this study is to propose a model that integrates classical engineering methods with new-generation artificial intelligence-based control strategies. The fact that this model also examines the environmental sustainability dimension from a life cycle assessment (LCA) perspective demonstrates that the study offers a meaningful framework not only in technical terms but also in terms of ethical and environmental responsibility. In this context, the factors influencing energy consumption include not only motor type or battery chemistry but also dynamic elements such as load distribution, flight altitude, wind profile, and environmental conditions [20].

2. Energy Management Concepts and Applications

Energy management is an interdisciplinary field that aims to conserve resources by efficiently utilizing existing energy sources in a system, monitoring consumption processes, and implementing necessary optimizations. This concept becomes even more critical in mobile systems with limited energy capacity, such as UAVs. For battery-powered UAVs in particular, energy management plays a decisive role in terms of flight time, mission success, and operational sustainability. Modern energy management goes beyond merely tracking energy consumption; it also involves functions such as data collection, analysis, and energy distribution tailored to the system's needs. Software and hardware solutions developed in this context aim to increase UAV mission efficiency while minimizing energy waste [21].

2.1. Battery Technologies and Performance Criteria

In UAVs, especially in military-class systems, battery technology is one of the key factors that directly determine the performance capacity of the platform. Today, Lithium-Ion (Li-ion) and Lithium-Polymer (LiPo) battery systems stand out due to their advantages such as high energy density, lightweight, and rechargeability. These battery types play a critical role in extending mission durations, increasing payload capacity, and enhancing overall system efficiency. Lithium-ion batteries are more advantageous than LiPo batteries in terms of energy density. A typical Li-ion cell can provide an energy density of 150–200 Wh/kg [22]. This is particularly important in reconnaissance and surveillance missions where maximum efficiency is required in terms of range and flight time. However, LiPo batteries are preferred in operational scenarios where high power demands are met due to their higher discharge rates (C-rate). In other words, LiPo batteries are more functional in tasks that require instantaneous energy demand, such as high climb rates or rapid maneuvers [23]. In terms of structural differences, LiPo batteries contain gel electrolyte, while Li-ion batteries use liquid electrolyte. This difference enables LiPo batteries to be thinner, lighter, and more moldable. LiPo cells are particularly preferred in tactical mini-UAVs with narrow bodies. In UAVs, especially military-class armed UAVs (SİHA), one of the most critical aspects of battery technology is the management of charge-discharge cycles and the estimation of battery life based on these cycles.

2.2. Flight Optimization and Route Planning

Energy efficiency has a direct impact on UAV mission performance and operational costs. Long-term patrol missions, reconnaissance, surveillance, and target tracking operations, especially in military-class armed unmanned aerial vehicles (UAVs), require a high level of energy planning. In this context, route planning strategies that prioritize energy efficiency are critical for increasing the platform's range, extending battery life, and ensuring maximum mission effectiveness throughout the flight. Energy-focused route planning is a set of algorithms that use energy as the primary optimization variable instead of traditional time or distance optimization. In this type of planning, the UAV's flight path is calculated based on weather conditions, altitude changes, wind profiles, load status, and variables in the mission area to minimize energy consumption.

2.3. Sustainability and Environmental Impacts

With the development of aerodynamic structures, energy consumption per flight time is reduced, which indirectly reduces carbon emissions. In recent years, fixed-wing UAVs with a wingspan greater than 10 meters have achieved a 14% energy savings thanks to low-friction composite materials [24]. This translates to approximately 60 kWh of energy savings for a platform that flies 500 hours per year. The same savings result in the prevention of approximately 10 kg of CO_2 emissions per year. Not only flight performance but also mission planning directly affects carbon emissions. Optimizing mission routes reduces idle engine times, thereby lowering energy consumption. In a field study, an energy-efficient mission planning algorithm applied during simultaneous flights of multiple UAVs reduced the average route time by 18%, total energy consumption by 21%, and carbon emissions by 19% [24].

3. Experimental Results

3.1. Energy Consumption Profiles and Battery Technologies According to UAV Types

This study conducted a multi-layered analysis of energy management and optimization in UAV systems. The findings were evaluated in terms of system design, flight planning, battery technologies, machine learning-based prediction systems, and sustainability. Energy efficiency is not limited to hardware or battery chemistry; it presents a holistic structure that must be considered in conjunction with mission dynamics, environmental variables, and software-based decision-making mechanisms. Significant differences in energy consumption profiles were identified across UAV classes. It has been observed that fixed-wing systems consume 20–30% less energy by minimizing engine load during cruise flights, while rotary-wing systems consume more energy due to advantages such as hovering and high maneuverability. In particular, an increase in the number of rotors in multi-rotor systems increases average energy consumption by up to 35%. Figure 2 presents findings on the environmental efficiency of different UAV types in terms of energy consumption, mission compatibility, and market development. According to the analysis conducted

prior to the figure, in a two-user (2 GUs) scenario, the energy consumption of a rotary-wing UAV was approximately 483 kJ/hour, while that of a fixed-wing UAV was measured at 1,067 kJ/hour. In the five-user (5 GUs) scenario, the fixed-wing system consumed 34% more energy, and in the ten-user (10 GUs) scenario, it consumed 6% more energy. In terms of mission compatibility, 30% of fixed-wing UAVs in the 5 GUs scenario and 59% in the 10 GUs scenario failed to meet mission requirements. Solar-powered hybrid systems were found to provide 60–80% energy/fuel savings compared to traditional models. When examining battery technologies, it was found that Lithium-Polymer (LiPo) batteries are more suitable for sudden power requirements due to their high discharge rate, while Lithium-Ion (Li-ion) batteries are more efficient for long-range missions due to their high energy density. It has been observed that advanced Li-ion cells provide effective performance for up to 1,000 cycles and, when integrated with LSTM models that provide 93.6% accuracy in life estimation, create significant advantages in battery planning. Parameters such as battery heating, the efficiency of charge-discharge cycles, and lifespan are also understood to be decisive factors in the system's overall performance. In this context, AI-based lifespan prediction systems play a critical role not only in energy planning but also in mission continuity.

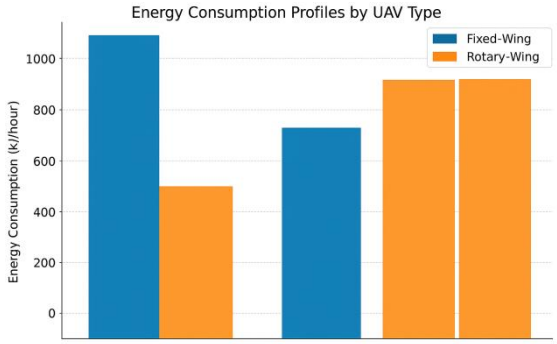


Figure 2. Energy consumption profiles by UAV type [14], [16], [19], [22]

3.2. Energy Estimation Using Flight Dynamics, Environmental Factors, and Machine Learning

It has been determined that sudden acceleration, altitude changes, and turning movements during the mission cause sudden increases in energy consumption. Deviations ranging from 15% to 25% in energy consumption have been observed under challenging weather conditions. When the flight route is optimized according to the topography, the average energy savings range from 10% to 14%. These findings indicate that energy consumption is not a static variable but is highly sensitive to environmental and structural changes. It is understood that dynamic structures that can be updated online and integrated based on environmental feedback should be adopted instead of classical route systems. Figure 3 presents a summary of these findings. Reinforcement learning algorithms play the most critical role in this transformation. Analyses conducted with machine learning-supported energy management systems have revealed that regression-based models offer over 85% accuracy in energy predictions, while reinforcement learning algorithms can increase energy efficiency by up to 12% by dynamically adapting to environmental factors in route planning. Algorithms not only produce static energy forecasts but also significantly improve task performance by responding to environmental variables in real time. The autonomous processing of data collection, processing, and decision-making processes during tasks reduces human intervention, lowers the margin of error, and minimizes operational costs.

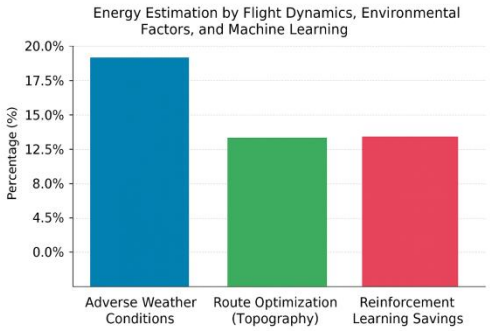


Figure 3. Energy estimation using flight dynamics, environmental factors, and machine learning [5], [2], [18], [22]

3.3. Sustainability and Carbon Impact

Findings in the context of sustainability show that solar-powered UAVs can operate with zero carbon emissions during long missions and save up to 400 kg of CO_2 per year compared to conventional electric models. According to LCA (Life Cycle Assessment) results, a large surveillance UAV emits approximately 8,500 kg of CO_2 equivalent carbon emissions throughout its life cycle, with 46% of this coming from operational processes. Figure 4 summarizes these findings. Route optimizations supported by pre-flight simulations were observed to reduce energy consumption by 8–14%, and up to 20% in scenarios where environmental parameters were integrated. Modeling applications integrated with geographic information systems (GIS) in particular offer significant contributions in terms of environmental sustainability and mission planning.

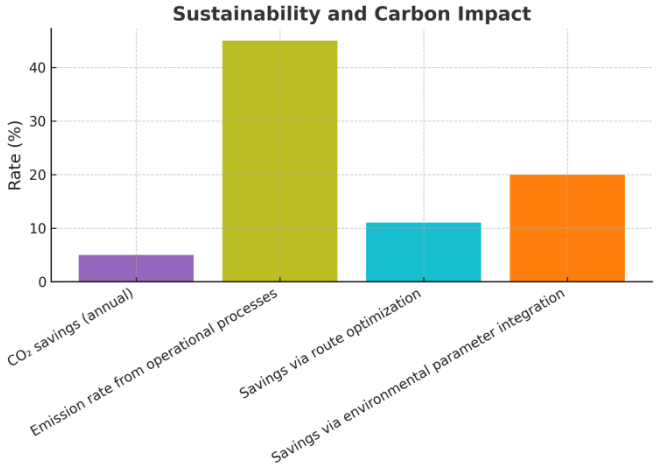


Figure 4. Sustainability and carbon impact [14], [17], [18]

4. Conclusion

Energy management and optimization in unmanned aerial vehicles (UAVs) is a fundamental factor that directly affects the operational success, mission duration, environmental impact, and long-term sustainability of these systems. Based on findings in the literature, it has been determined that energy consumption profiles vary significantly depending on the class of UAVs; fixed-wing systems are suitable for longer-range and low-energy consumption missions, while rotary-wing systems are suitable for short-range missions requiring high maneuverability. Although hybrid systems incorporate the advantages of both structures, they require much more complex algorithms in terms of energy management. Flight dynamics, load distribution, environmental factors, and mission parameters are among the main variables that determine energy consumption.

Energy management applications are not limited to energy monitoring and recording functions; they also offer the ability to proactively control energy consumption during missions through decision-making mechanisms supported

by artificial intelligence and machine learning. In particular, real-time analysis of sensor data, monitoring of battery temperature, voltage, and discharge rates increases flight safety and improves mission performance. In terms of battery technologies, the usage efficiency, energy density, discharge characteristics, and thermal management requirements of lithium-ion and lithium-polymer cells have become decisive factors in system design.

The study's findings indicate that pre-flight simulations and route optimizations can significantly reduce energy consumption and lower carbon emissions. In particular, energy-based route planning that integrates environmental factors can achieve energy savings of 10–25% compared to traditional route approaches. This savings has a significant impact in terms of both economic and environmental contributions in the annual mission cycle. Similarly, machine learning algorithms have been observed to achieve accuracy rates of over 90% in areas such as energy estimation and battery life assessment. To increase energy efficiency and mission success in UAVs, it is necessary to implement a multi-layered, data-driven, adaptive, and integrated energy management system. The combined evaluation of advanced battery technologies, sustainable material use, route planning algorithms, and AI-supported prediction systems increases the operational efficiency of UAV systems and serves long-term environmental sustainability goals. In this context, it is recommended that future studies focus on the integration of renewable energy systems, the applicability of energy recovery techniques, and improving the accuracy of online energy prediction systems.

5. References

[1] Smith, P., Yang, F., Liao, Q., & Tan, H. (2023). Hybrid powerplant design and energy management for UAVs. Energies, 18(12), 3101.

[2] Mozaffari, M., Saad, W., Bennis, M., & Debbah, M. (2016). Efficient deployment of multiple unmanned aerial vehicles for optimal wireless coverage. IEEE Communications Letters, 20(8), 1647–1650.

[3] Lyu, J., Zeng, Y., Zhang, R., & Lim, T. J. (2023). Placement optimization of UAV-mounted mobile base stations: Deep learning and hybrid energy considerations. IEEE Communications Letters, 27(2), 280–284.

[4] Amoiralis, E. I., Tsili, M. A., Spathopoulos, V., & Hatziefremidis, A. (2022). A comprehensive review of energy-efficient techniques for UAVs. Energies, 17(18), 4737.

[5] Kumar, R., & Lee, S. (2025). Energy consumption and efficiency degradation predictive analysis for UAV battery management using deep neural networks. Advances in Science and Technology Research Journal, 19(5), 21–30.

[6] Zeng, Y., Zhang, R., & Lim, T. J. (2016). Wireless communications with unmanned aerial vehicles: Opportunities and challenges. IEEE Communications Magazine, 54(5), 36–42.

[7] Lyu, J., Zeng, Y., Zhang, R., & Lim, T. J. (2023). Placement optimization of UAV-mounted mobile base stations: Deep learning and hybrid energy considerations. IEEE Communications Letters, 27(2), 280–284.

[8] Gong, J., Chang, T., Shen, C., & Chen, X. (2018). Flight time minimization of UAV for data collection over wireless sensor networks. IEEE Journal on Selected Areas in Communications, 1–1.

[9] Saad, W., Bennis, M., & Chen, M. (2019). A vision of 6G wireless systems: Applications, trends, technologies, and open research problems. arXiv preprint, arXiv:1902.10265.

[10] Zhan, Z., Zeng, Y., & Zhang, R. (2018). Trajectory design for distributed estimation in UAV-enabled wireless sensor network. IEEE Transactions on Vehicular Technology, 67(10), 10155–10159.

[11] Zhan, Z., Zeng, Y., & Zhang, R. (2018). Energy-efficient data collection in UAV-enabled wireless sensor network. IEEE Wireless Communications Letters, 7(3), 328–331.

[12] Say, S., Inata, H., Liu, J., & Shimamoto, S. (2016). Priority-based data gathering framework in UAV-assisted wireless sensor networks. IEEE Sensors Journal, 16(14), 5785–5794.

[13] Say, S., Inata, H., Liu, J., & Shimamoto, S. (2016). Priority-based data gathering framework in UAV-assisted wireless sensor networks. IEEE Sensors Journal, 16(14), 5785–5794.

[14] Sun, L., Liu, J., Wang, X., & Wang, L. (2025). Energy scheduling of hydrogen hybrid UAV based on model predictive control and deep deterministic policy gradient algorithm. Applied Energy, 325, 120789.

[15] Smith, P., Yang, F., Liao, Q., & Tan, H. (2023). Hybrid powerplant design and energy management for UAVs. Energies, 18(12), 3101.

[16] Yang, Z., Pan, C., Shikh-Bahaei, M., Xu, W., Chen, M., Elkashlan, M., & Nallanathan, A. (2018). Joint altitude, beamwidth, location, and bandwidth optimization for UAV-enabled communications. IEEE Communications Letters, 22(8), 1716–1719.
[17] Mozaffari, M., Saad, W., Bennis, M., & Debbah, M. (2016). Efficient deployment of multiple unmanned aerial vehicles for optimal wireless coverage. IEEE Communications Letters, 20(8), 1647–1650.
[18] Saad, W., Bennis, M., & Chen, M. (2019). A vision of 6G wireless systems: Applications, trends, technologies, and open research problems. arXiv preprint, arXiv:1902.10265.
[19] Zhan, Z., Zeng, Y., & Zhang, R. (2018). Energy-efficient data collection in UAV-enabled wireless sensor network. IEEE Wireless Communications Letters, 7(3), 328–331.
[20] Say, S., Inata, H., Liu, J., & Shimamoto, S. (2016). Priority-based data gathering framework in UAV-assisted wireless sensor networks. IEEE Sensors Journal, 16(14), 5785–5794.
[21] Alzenad, M., El-Keyi, A., Lagum, F., & Yanikomeroglu, H. (2017). 3-D placement of an unmanned aerial vehicle base station (UAV-BS) for energy-efficient maximal coverage. IEEE Wireless Communications Letters, 6(4), 434–437.
[22] Amoiralis, E. I., Tsili, M. A., Spathopoulos, V., & Hatziefremidis, A. (2022). A comprehensive review of energy-efficient techniques for UAVs. Energies, 17(18), 4737.
[23] Cho, S., Lee, K., Kang, B., Koo, K., & Joe, I. (2018). Weighted harvest-then-transmit: UAV-enabled wireless powered communication networks. IEEE Access, 6, 72212–72224.
[24] Kumar, R., & Lee, S. (2025). Energy consumption and efficiency degradation predictive analysis for UAV battery management using deep neural networks. Advances in Science and Technology Research Journal, 19(5), 21–30.

Development of A Docker-Based Demand Forecasting System Integrated with Continuous Integration and Continuous Deployment Automation Using Prometheus and Grafana

Mehmet Mert MENEVŞE[1], Mehmet KARAKÖSE[2]

[1,2]Department of Computer Engineering, Firat University, Elazig, Turkey.
[1]menevsemehmetmert@gmail.com, [2]mkarakose@firat.edu.tr
[1](ORCID: 0009-0004-4227-0778), [2](ORCID: 0000-0002-3276-3788)

Abstract
In today's data-driven economy, accurate forecasting systems play a critical role in supporting strategic decision-making processes, particularly in the retail sector. Businesses require solutions that are not only statistically reliable but also scalable, maintainable, and efficient in real-world deployment. This study presents a demand forecasting system tailored for retail sales data, implemented using a container-based architecture with Docker technology. The model is developed in Python utilizing machine learning libraries and is integrated into a fully automated software pipeline with continuous integration and continuous deployment. This automation enables model training, prediction, and reporting without manual intervention. Real-time system monitoring is achieved through open-source tools that collect and visualize performance metrics via interactive dashboards. Container management tools are used to ensure seamless system updates and minimize downtime. The model's forecasting performance yielded an average RMSE (Root Mean Squared Error) of 1.1549, demonstrating its high predictive accuracy and applicability in real-world scenarios. Beyond predictive accuracy, the system emphasizes operational qualities such as scalability, observability, and portability. Experimental results demonstrate that the system maintains stable performance under varying data conditions and can be easily adapted to diverse forecasting scenarios. This integrated approach illustrates the practical value of combining machine learning with modern software engineering practices for deployment in production environments.

Keywords: demand forecasting, docker, continuous integration, continuous deployment, machine learning, system monitoring

1. Introduction

The rapid growth of data volumes and the increasing efforts of businesses to gain a competitive advantage have elevated the importance of data-driven decision-making systems. In the retail sector in particular, accurately forecasting customer demand plays a critical role in inventory management, logistics planning, supply chain optimization, and cost control [1]. Accordingly, demand forecasting systems supported by machine learning algorithms offer more accurate and reliable predictions [2].

However, it is essential that such systems are not only developed at the modeling level but also implemented in integration with modern software engineering principles. In addition to model accuracy, it is equally important that the system be automatically updatable, redeployable, observable, and sustainable. In this context, the integration of Continuous Integration (CI) and Continuous Deployment (CD) processes—widely adopted in the software development lifecycle (SDLC)—into forecasting systems enhances operational efficiency and reduces the likelihood of errors [3].

In this study, a machine learning-based forecasting model trained on retail sales data was containerized using Docker technology, automatically built and deployed via a CI/CD pipeline, and made observable through the use of Prometheus and Grafana. Thanks to the modular structure of the developed system, loose coupling was ensured between components, resulting in a sustainable architecture [4].

Each part of the system was designed in accordance with fundamental principles of software engineering:

- Modeling: A demand forecasting model based on linear regression optimized using the RMSE metric with Python and scikit-learn
- Automation: Automated build and push processes for Docker images triggered via GitHub Actions-based CI/CD workflows
- Deployment: Multi-service orchestration using Docker Compose
- Monitoring: Visualization of metrics collected by Prometheus on Grafana dashboards
- Updatability: Live container updates using the Watchtower service without system downtime

[1]Corresponding author

The main objective of this study is to evaluate predictive systems not only in terms of statistical accuracy but also in terms of operational competencies such as deployability, observability, and maintainability. The subsequent sections of this paper will present detailed information on system setup, component interactions, performance analysis, encountered issues, and proposed solutions.

1.2. Related Works

Virtual machines and cloud servers are widely used today in the development of artificial intelligence models. In addition to utilizing virtual machines, cloud servers also benefit from containerization technologies. The studies numbered [5], [6], [7], and [8] in the literature comprehensively highlight the differences between virtual machines and Docker containers. Docker containers offer several advantages, including simple configuration, application isolation, multi-tenancy architecture (running multiple containerized applications on a single host), and rapid deployment. Therefore, in this study, Docker containerization technology has been preferred as the deployment environment for both cloud computing infrastructures and local computing units. Furthermore, in order to manage the containers of the research platform system effectively, Kubernetes container orchestration tools are utilized. Table 1 summarizes the differences between virtual machines and Docker.

Table 10. Comparison of virtual machines (VM) and docker containers

Feature	Virtual Machine (VM)	Docker Container
Startup Time	Takes several minutes	Takes a few seconds
Definition	Server virtualization	Operating system virtualization
Size	Gigabytes	Megabytes
Kernel	Each VM has its own kernel	Containers share the host operating system kernel
Lifecycle	Has a boot process and is typically long-lived	Runs directly on the host kernel, no boot process, typically short-lived
Security	Depends on the hypervisor implementation	Mandatory access control can be enforced
Isolation	No sharing of libraries, files, etc., between host and guests	Subdirectories can be transparently mounted and shared
Storage	Requires significant storage as the entire OS and related programs must be installed and run	Uses less storage since the host OS is shared

There are numerous software tools available on the market that support continuous integration (CI) processes, each with its own distinct features and differences. Since this study plans to use GitHub as the platform where user application versions will be stored, a comparison has been made among CI software tools that can integrate with GitHub. Jenkins [9] requires a high level of technical expertise from users and does not provide version control support for various processes. GitLab [10], on the other hand, only supports its own version control system and does not allow the extension of YAML configuration files; this implies limited scalability for services to be developed in the future. In order to maintain strong integration with GitHub and continue using Git as the version control infrastructure, GitHub Actions [11] has ultimately been chosen in this study as the solution for continuous integration and continuous deployment (CI/CD).

In the study conducted by Shih et al. [12], a model retraining process was proposed; however, this process was not practically implemented. The study aimed to trigger the retraining process by defining a specific threshold value and retraining the model using both the previous training dataset and a new dataset. This approach holds potential for preserving the predictive performance of the model. Compared to a radical alteration of the model architecture, this method involves lower technical complexity and is therefore easier to implement. Additionally, it allows the same model architecture to generalize better over time. However, since the overall model architecture is not significantly altered, it may be insufficient in effectively addressing potential declines in model performance, particularly in the face of data drift or concept drift.

Lai et al. [13], on the other hand, proposed a retraining mechanism in which the model is retrained every time the sensor receives two hundred new data points. In this process, old and new data were merged in equal proportions to train the model. The resulting test outcomes demonstrated that the proposed retraining mechanism improved the suitability of the

integrated model for real-world applications and extended its operational lifespan. Nevertheless, this method does not require substantial changes to the model architecture or its parameters; it merely enhances model training performance based on the dataset. While this yields certain advantages, it also limits the capacity to handle performance degradation that may arise from data drift or concept drift. Table 2 presents a comparison of software tools for GitHub-triggered CI/CD workflows.

Table 2. Software comparison for github-triggered CI/CD workflows

Category	Jenkins	GitHub Actions	GitLab
Environment	Linux/macOS/Windows	Linux/macOS/Windows	Linux/macOS/Windows
Configuration via Code/GitOps	Yes	Yes	Yes
Parallelism	Yes	Yes	Yes
Branching Support	Yes	Yes	Yes
Authentication	GitHub	GitHub	GitHub
Asynchronous Integration	No	Yes	Yes
Serverless	No	Yes	Yes
Status Visualization	Yes	Yes	Yes
Ad-hoc Workflow	Yes	No	Yes

Container security is critically important in modern software deployment processes. In a systematic literature review conducted by Saleh et al. [14], security issues in CI/CD processes were emphasized—particularly the use of security tools such as Harbor and SonarQube, and security threats such as image tampering and unauthorized access. Waseem et al. [15] examined security vulnerabilities in microservices developed with Docker and underscored the importance of pipeline security. The NIST SP 800-190 [16] standard provides comprehensive guidance on container security, addressing topics such as container image security, container runtime security, orchestration security, and host security.

The operational processes of machine learning models (MLOps) are crucial for ensuring reliable and efficient model performance in production environments. Kreuzberger et al. [17] presented a comprehensive review of the definition, architecture, and components of MLOps. Shankar et al. [18] investigated industry practices related to the operationalization of machine learning systems. Model monitoring and data drift detection are essential components of MLOps. Rabanser et al. [19] experimentally examined and compared various methods for detecting data drift. Gama et al. [20] provided an extensive literature review on concept drift adaptation.

The scalability and high availability of containerized applications are critical for real-world deployments. Burns et al. [21] explored Google's decade-long experience in container management and the development of Kubernetes. Medel et al. [22] characterized microservice architectures and the use of Kubernetes in the industry. Qu et al. [23] presented a comprehensive taxonomy and literature review on the automatic scaling of web applications in cloud environments. Ilyushkin et al. [24] experimentally evaluated the performance of auto-scalers for complex workflows.

Je Chiao Ku et al. [1] are developing a cloud platform for remote AI model correction, which addresses challenges in the manufacturing industry by providing automated integration, deployment, and security features. Through the integration of artificial intelligence, the platform aims to enhance production efficiency, sustainability, and competitiveness.

The primary objective of this study is to make a machine learning–based demand forecasting model, developed for the retail sector, portable and manageable through Docker container technology, and to enable its end-to-end integration via Continuous Integration / Continuous Deployment (CI/CD) processes. In this context, a LightGBM model based on linear regression was developed using the Python programming language and the scikit-learn library, trained on retail sales data, and its performance was evaluated based on the RMSE metric. The scope of the study extends beyond the modeling

process and includes data preprocessing, Docker containerization, CI/CD pipeline setup, and operational monitoring steps. In particular, Docker container technology ensured that the model runs consistently across different environments, and automated configuration, deployment, and update workflows were established through CI/CD processes integrated with GitHub Actions. System performance was monitored in real time and historically using Prometheus and Grafana, while the Watchtower service was employed to keep containers up to date.

Accordingly, the aim of the study is not limited to developing statistically accurate forecasting models, but also to demonstrate the suitability of the developed system in terms of fundamental software engineering principles such as sustainability, portability, observability, and operational efficiency. The scope of the study comprehensively addresses all phases, from dataset analysis to model training, Docker image creation and distribution, system monitoring, and the automatic generation of performance reports in PDF format.

2. Material and Methods

In this study, an artificial intelligence–based system was designed and developed as an end-to-end solution for demand forecasting in the retail sector. The system is structured using a microservice architecture that encompasses high-volume data processing, model training, forecasting, automated reporting, system monitoring, and automated updates. The primary objective is to provide an efficient, sustainable, reliable, and observable forecasting infrastructure. This section provides a detailed explanation of all steps involved in the proposed approach—from the data source to modeling and forecasting processes, as well as reporting and system management. Furthermore, the project workflow is presented from a holistic perspective in Figure 1.

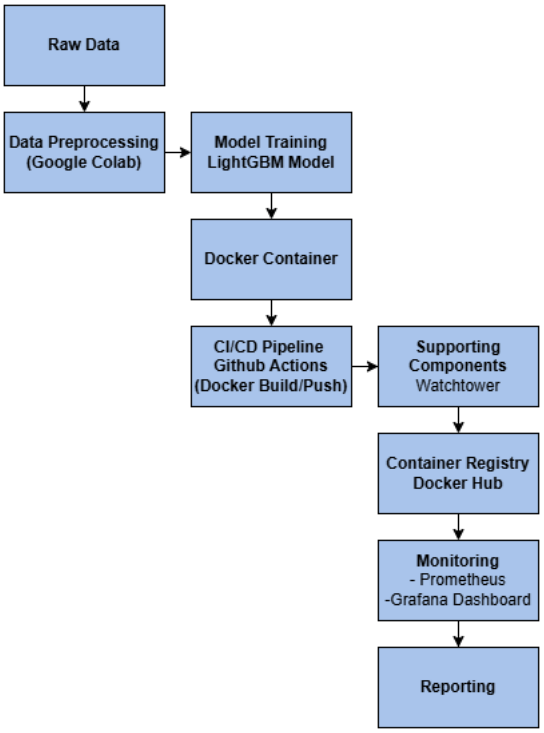


Figure 1. Holistic process flow diagram representing the project workflow

2.1. Dataset and Preprocessing Phase

The dataset used in this study is the open-source "Retail Sales Dataset" published on the Kaggle platform [25]. The dataset includes the following columns: transaction date, customer ID, age, gender, product category, quantity of products purchased, unit price per product, and total sales amount, comprising a total of 1,000 records.

The data preprocessing phase was carried out comprehensively to enhance data quality and improve model accuracy:

- Missing and outlier values were identified and the corresponding records were either removed or corrected.
- Historical date values were converted to the datetime format to make them suitable for time series analysis.
- The "Total Amount" feature was recalculated as the product of quantity and unit price.
- Categorical variables (e.g., gender, product category) were transformed into numerical format using Label Encoding or One-Hot Encoding techniques.
- For the modeling process, the data was split into training and test sets based on a time-aware partitioning principle.
- Where necessary, scaling techniques such as StandardScaler were applied.

The basic statistical properties and characteristics of the dataset are presented in Figure 2.



Figure 2. Characteristic properties of the dataset

2.2. Modeling and Forecasting Process

The dataset was split into 80% training and 20% testing sets using the train_test_split function. During the modeling phase, various regression algorithms were evaluated. The models tested included Linear Regression, Decision Tree Regressor, Random Forest Regressor, and LightGBM (Light Gradient Boosting Machine). Among these, LightGBM yielded the highest performance [26]. It was preferred due to its fast training time, high accuracy, and ability to effectively handle categorical data.

The stages of the modeling process are summarized as follows:

- As part of feature engineering, new variables such as age category and percentage of total spending were created.
- Hyperparameter optimization was performed, and different combinations of parameters such as num_leaves, max_depth, learning_rate, and n_estimators were tested.
- Root Mean Squared Error (RMSE) was used as the primary metric for model evaluation, and the average RMSE score was calculated as 1.1549.
- Training and testing results were visualized using plots, and error distributions were analyzed.
- Forecasting operations were executed using the predict.py script on the test dataset (X_test.csv), and the results were saved to the predictions.csv file.

2.3. Containerization and Service Architecture (Use of Docker)

To enhance the portability and flexibility of the system, all components were containerized using Docker [27]. This approach enables the system to be redeployed and managed consistently across different platforms. The project defines three main components:

- retail-demand-project: An API service developed using the FastAPI framework, which handles forecasting requests from users.

- retail-report-scheduler: A Python-based scheduler script that sends periodic requests to the API every 5 minutes to retrieve forecasting data and generate reports.
- retail-webserver: A web interface developed with Flask that enables users to view PDF reports.

Custom Dockerfile definitions were created for each component, and the resulting images were stored on Docker Hub. The Dockerfiles used in the project include:

- Dockerfile: For the core forecasting component (predict.py)
- Dockerfile.report: For report generation and scheduling
- Dockerfile.web: For the Flask-based web server

This architecture eliminates inter-container dependencies, thereby improving the system's portability and accelerating deployment.

2.4. CI/CD Automation and Watchtower-Based Updates

To ensure that system updates can be carried out sustainably, a Continuous Integration / Continuous Deployment (CI/CD) process was established [28].
The implemented structure operates as follows:

- When the developer makes changes to the code, GitHub Actions is triggered.
- Docker images for the relevant components are automatically built.
- These images are pushed to Docker Hub.
- The Watchtower service detects the new versions and automatically updates the corresponding containers [29].

Thanks to this setup, the system remains up-to-date and operational without the need for any manual intervention.

2.5. Monitoring and System Performance Metrics

Prometheus and Grafana infrastructures were utilized to monitor real-time system performance [30][31]. While Prometheus is responsible for collecting system metrics, Grafana visualizes this data through meaningful and interactive graphs.

2.6. Automated PDF Reporting System

To ensure the regular visualization of forecasting results, the script defined in the report_scheduler.py file has been scheduled to run every 5 minutes. The reporting process includes the following steps:

- A request is sent to the forecasting API to retrieve prediction data [32].
- The RMSE score is calculated to verify prediction accuracy.
- Analyses such as category-wise sales distribution and comparisons between predicted and actual values are supported with visualizations.
- The report is generated in PDF format using FPDF and matplotlib.
- Each report is saved with a timestamped filename (e.g., performance_report_YYYY-MM-DD_HH-MM-SS.pdf).

As a result, this section has comprehensively presented the systematic steps of the proposed approach, which transforms data into value through an end-to-end artificial intelligence solution. The developed system offers an industry-ready framework with advantages such as reusability, low maintenance cost, flexibility, and high performance.

3. Results and Discussions

This section evaluates the performance outputs of the developed demand forecasting system components under real-time usage scenarios. Measurements were obtained through system-level tools that collect data from various layers—such as Docker, Prometheus, Grafana, and PDF reporting scripts—and were supported by numerical, visual, and graphical analyses.

3.1. Model Performance: RMSE Analysis

The LightGBM algorithm, selected as the forecasting model, was compared against alternative algorithms during the modeling process and was preferred due to its ability to produce the lowest error value. The model's performance was evaluated using the Root Mean Squared Error (RMSE) metric. RMSE is calculated as the square root of the average of the squared differences between predicted and actual values and is widely used for regression problems.

In the conducted tests, the average RMSE value was determined to be 1.1549. This value indicates that the model has a high predictive capability and performs forecasting on retail sales data with minimal error. RMSE scores were also monitored periodically to reflect the variability inherent in retail sales data, and the system was observed to operate consistently and stably over time.

3.2. Containerization Process and Metrics

As part of this study, the containerization of the developed Retail Demand Forecasting System was carried out using a python:3.11-slim-based Docker image, aiming to ensure minimal dependency and a lightweight runtime environment.

The resulting Docker image has a size of 894 MB, which is considered acceptable for deployment in production environments. However, further optimization of the image is possible. A total of 17 layers were generated from the instructions defined within the Dockerfile; these layers represent the sequential and independent building blocks of the image.

The Docker image build process took approximately 497.8 seconds (around 8.3 minutes). The image was built targeting the amd64 CPU architecture and the Linux operating system, ensuring compatibility with cloud-based infrastructures and server platforms. Table 3 presents detailed information regarding the containerization process.

Table 3. Information related to the containerization process

Metric	Value	Description
Docker Image Size	894 MB	Final image size after the build is completed
Docker Build Time	497.8 seconds (~8.3 minutes)	Total time taken to build the container image
Number of Layers	17 layers	Each instruction in the Dockerfile creates a build layer
Base Image Used	python:3.11-slim	A minimal base image used to reduce overall system footprint
CPU Architecture	amd64	Target processor architecture for the container
Operating System	Linux	Operating system environment in which the container runs

The containerization process not only enhanced the portability, reproducibility, and sustainability of the system, but also integrated seamlessly with CI/CD workflows, thereby automating update and version management operations. As a result, the developed system aligns with modern software engineering principles not only from a technical standpoint, but also in terms of deployment and maintenance processes.

3.3. CI/CD Process Performance Analysis

In this study, the software system's Continuous Integration and Continuous Deployment (CI/CD) processes were implemented using the GitHub Actions infrastructure. This section analyzes the statistical performance metrics of the configured CI/CD pipeline and evaluates the quality of its automation. Table 4 presents detailed information related to the CI/CD process in tabular form.

Table 4. Information on the CI/CD process

Metric	Value	Description
Total Number of Pipeline Executions	19	Total number of GitHub Actions workflows executed
Success Rate	100%	All pipeline executions were completed without any issues
Average Execution Time	Approximately 1 minute and 5 seconds	The average duration of each pipeline execution
Shortest Execution Time	48 seconds	The shortest observed pipeline execution time
Longest Execution Time	1 minute and 24 seconds	The longest observed pipeline execution time
Deployment Frequency	Approximately 9–10 times per day	The average number of deployments carried out daily

During the examined time period, a total of 19 GitHub Actions workflows were executed, all of which were completed with a 100% success rate. This indicates a high level of stability and reliability in the system's automation mechanisms. The average execution time for each pipeline run was measured to be approximately 1 minute and 5 seconds. The shortest observed duration was 48 seconds, while the longest was recorded as 1 minute and 24 seconds. These durations contribute to a fast feedback loop within the software development cycle, thereby supporting an environment aligned with Agile software development principles.

Furthermore, the fact that an average of 9 to 10 deployments were carried out daily demonstrates the system's capability to support high-frequency updates and reflects the successful implementation of the Continuous Delivery (CD) practice. In this respect, the system is in full compliance with the DevOps practices recommended in industrial applications.

3.4. Container Registry (Docker Hub)

In this study, the system components developed were made publicly accessible and managed with version control through a Docker Hub repository titled mertmenevse/retail-demand-project. Throughout the experimental process, more than 15 distinct Docker images were created, with an average image size measured at approximately 200 MB. The default infrastructure performance of Docker Hub was used as the baseline, and under normal internet conditions, no latency issues were encountered during access. Consequently, there was no performance degradation in the automated rebuilding and deployment of system components through CI/CD processes.

3.5. Monitoring – Prometheus + Grafana Dashboard

The stability and resource consumption of the container-based reporting and data processing architecture developed in this study were monitored and evaluated using Prometheus and Grafana tools. During this process, particular attention was given to analyzing the runtime performance of the container named retail-report-scheduler. Table 5 presents the container's resource usage metrics, as measured by Prometheus and Grafana, in tabular format.

Table 5. Container resource utilization metrics measured by prometheus and grafana

Metric	Value	Unit	Description
Average CPU Usage	0.426	%	Over the past 6-hour period
Maximum CPU Usage	2.76	%	Over the past 6-hour period
Minimum CPU Usage	0.291	%	Over the past 6-hour period
Average RAM Usage	61.9	MB	Over the past 6-hour period
Maximum RAM Usage	62.8	MB	Over the past 6-hour period
Minimum RAM Usage	61.1	MB	Over the past 6-hour period

Metric	Value	Unit	Description
Container Uptime Duration	43	hours	Based on the output of docker ps

The metric data were collected via the Prometheus time-series database using the following PromQL queries:

- CPU Usage: The process_cpu_seconds_total metric was used in combination with the rate() function to calculate the per-second CPU consumption rate, which was then converted into a percentage.
- RAM Usage: The process_resident_memory_bytes metric was used to measure physical memory consumption, which was analyzed in megabytes (MB).

The monitoring interval was set to the "Last 6 Hours," as this time range was considered optimal for reflecting the system's current state.

3.5.1. CPU Usage Analysis

According to the monitoring results, the average CPU usage of the container remained at only 0.426%. This value indicates that the system operates with high efficiency and low resource consumption. The maximum CPU usage was limited to 2.76%, suggesting that even during periods of high processing load, the system utilized hardware resources at reasonable levels. The minimum usage value of 0.291% demonstrates that the system occasionally operates in an idle state or enters standby mode. Figure 3 presents the CPU usage graph over a 6-hour period, as visualized in Grafana.

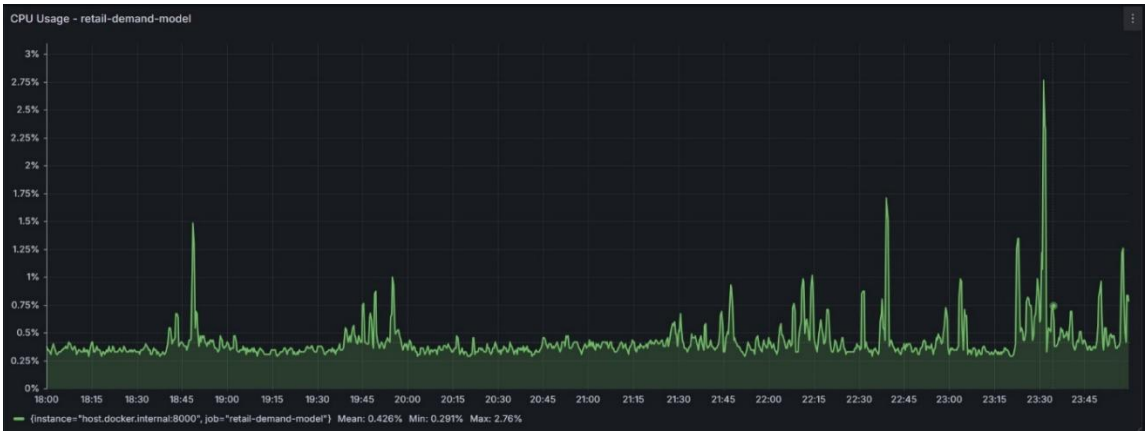


Figure 3. Grafana CPU Usage

3.5.2. RAM Usage Analysis

Memory usage monitored through Grafana panels indicated that the system maintains a highly stable memory profile. The average memory usage was measured at 61.9 MB. The maximum RAM usage was only 62.8 MB, suggesting that even during relatively resource-intensive tasks such as graph generation and predictive model computations, memory consumption remained well-controlled. The minimum value of 61.1 MB indicates consistent memory management during background operations. Figure 4 presents the RAM usage graph over a 6-hour period, as visualized in Grafana.



Figure 4. Grafana RAM Usage

3.5.3. Uptime Analysis

According to information obtained via the docker ps command, the retail-report-scheduler container operated continuously for 43 hours without interruption. During this period, no system crashes, restarts, or manual interventions were required. This indicates that the system is reliable in terms of sustainability and operational continuity.

Based on the metrics presented above, it is evident that the system possesses a lightweight, optimized, and balanced infrastructure in terms of both CPU and RAM usage. Moreover, the system's high uptime confirms that automated reporting tasks can be executed without interruption and demonstrates the application's suitability for industrial-scale use.

Thanks to the integration of Prometheus and Grafana, system resource usage can be monitored in real time, allowing for timely interventions when necessary. This monitoring process provides scientific evidence that the developed application delivers a high-performance, stable, and observable software solution, serving as a representative example for similar system designs in the academic literature.

3.6. Reporting (Automated PDF Report Generation and Sales Analysis)

Table 6 presents the total sales amounts for each product category throughout the analysis period. The Electronics category ranks first, achieving the highest total sales volume. It is followed by the Clothing and Cosmetics categories, respectively.

Table 6. Total sales amounts by product category

Product Category	Total Sales (USD)
Electronics	156,905
Clothing	155,580
Cosmetics	143,515

The data presented above clearly illustrate the total revenue levels generated across different product categories. The Electronics category recorded the highest sales revenue, totaling 156,905 USD. This was followed by Clothing with 155,580 USD and Cosmetics with 143,515 USD.

This distribution suggests that the demand structure and customer segments vary across product categories. The high total sales value of electronic products may be attributed to factors such as higher unit pricing, longer product life cycles, or consumer preferences driven by advanced technology.

Similarly, the close proximity of total sales revenue between the Clothing and Cosmetics categories indicates that consumer spending behaviors in these areas are relatively similar. This suggests that these categories should be strategically evaluated together.

These findings highlight the need for differentiated decision-making processes in terms of category-based inventory management, marketing campaigns, and demand forecasting strategies. Additionally, conducting such revenue distribution analyses on a regular basis will enable business managers to engage in more accurate and data-driven strategic planning.

Figure 5 presents a sample performance PDF report, automatically generated on 2025-04-30 at 23:13:53. These reports are produced every five minutes.

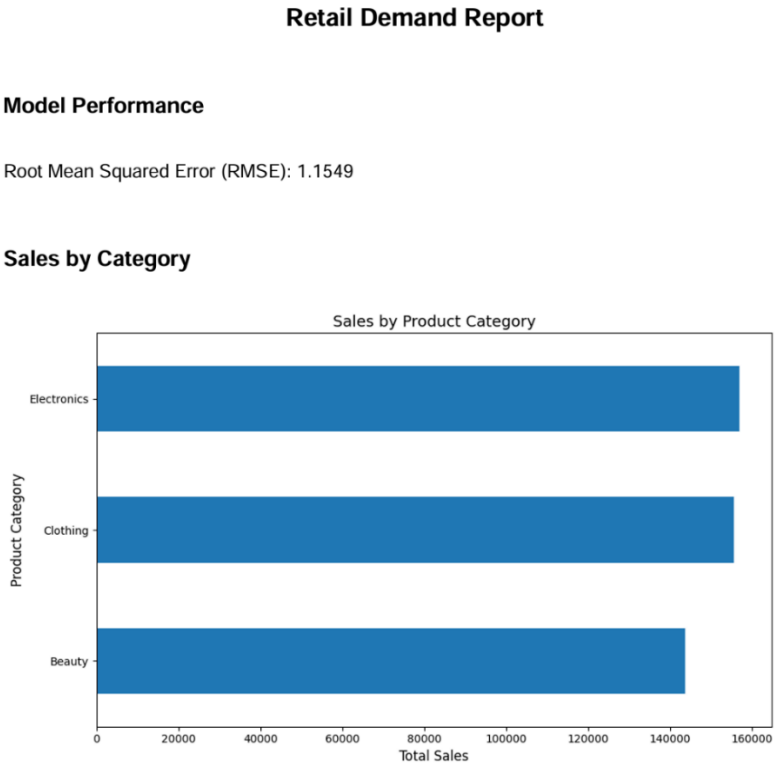


Figure 5. An automatically generated sample PDF performance report

3.7. Supporting Components – Container Monitoring with Watchtower

In this study, the Watchtower service—an automated container monitoring and updating tool—was utilized to ensure the uninterrupted and sustainable operation of the system. Watchtower is a lightweight and efficient background service that runs on Docker infrastructure. It periodically checks (by default, every 5 minutes in this study) whether the active container images in the system are up to date. If a change in an image is detected, Watchtower automatically restarts the relevant container and continues system operation with the updated image. Table 7 presents the monitoring results obtained through Watchtower.

Table 7. Watchtower monitoring results

Indicators	Value
Container Name	watchtower
Container Status	Active for 2 days (healthy)
Number of Restarted Containers	0
Update Check Interval	5 minutes (default setting)
Monitored Containers	retail-report-scheduler, retail-demand-project, retail-webserver
Health Check Mechanism	Active (observed as “healthy”)

Supportive microservice components such as Watchtower play a critical role in modern container architectures by enabling automation, self-healing capabilities, and the reduction of system maintenance costs. The automated execution of monitoring and update operations at regular intervals minimizes the need for manual intervention by system administrators and facilitates the uninterrupted operation of continuous integration processes.

The results observed in this study demonstrate that the Watchtower service has been effectively configured and contributes positively to the overall health of the system. In particular, the absence of any errors or system crashes confirms the system’s general stability. Although no image updates were performed during the observation period, the fact that Watchtower remained active and in a healthy state indicates the presence of a responsive and ready-to-deploy update infrastructure when needed.

4. Conclusions

Within the scope of this study, an AI-powered demand forecasting system developed for the retail sector was thoroughly analyzed. The system integrates a variety of tools from the disciplines of data science and software engineering to address multiple processes in a holistic manner, including data collection, preprocessing, machine learning-based modeling, result visualization, automated report generation, system monitoring, and updating.

The project began with a dataset known as the Retail Sales Dataset, which contains real-world retail sales data. During the data preprocessing phase, steps such as handling missing values, assigning correct data types, and ensuring semantic consistency were performed. In addition, new derived features were introduced to the dataset, among which the variable Total Amount stood out. This feature, calculated by multiplying quantity and unit price, made a meaningful contribution to the forecasting model. The comprehensive preprocessing pipeline enhanced the model's learning capacity and positively affected overall system performance.

In the modeling phase, various regression algorithms were tested, and based on performance comparisons, the LightGBM algorithm was selected. This choice was driven by LightGBM’s advantages in processing large datasets and handling high-cardinality categorical features with speed and accuracy. The model’s forecasting performance yielded an average RMSE (Root Mean Squared Error) of 1.1549, demonstrating its high predictive accuracy and applicability in real-world scenarios.

To ensure scalability and maintainability, all system components were containerized using Docker technology. The forecasting service, web interface, and reporting script were structured as three independent modules, enabling a modular architecture. By deploying each component in a separate container, developers are able to update and test individual components independently. Furthermore, the system can be easily redeployed across different servers or cloud environments.

For automated system updates, a CI/CD (Continuous Integration and Continuous Deployment) pipeline was implemented using GitHub Actions, and container updates were managed in real time via the Watchtower service. Any code changes made by developers were automatically reflected in the Docker images and deployed to the corresponding containers. This setup enabled seamless updates without the need for manual intervention.

System resource utilization and stability were monitored using Prometheus and Grafana. During the six-hour observation period, the average CPU usage was measured at 0.426%, with a maximum of 2.76%. Similarly, average RAM usage was 61.9 MB, with a peak value of 62.8 MB. These results indicate that the system can operate stably and sustainably even with minimal hardware requirements.
The report generation module was scheduled to run automatically every five minutes, and this process was verified via Docker logs. Reports were generated timely and without errors, confirming the uninterrupted operation of the system. Additionally, the generated reports included demand forecasts by product category. In particular, analyses for the Electronics, Clothing, and Cosmetics categories demonstrated accurate trend detection and model sensitivity to seasonal fluctuations. Such analyses are expected to hold strategic value for managers and decision-makers.

In conclusion, the system developed in this study presents a comprehensive solution suitable for implementation in the retail industry, featuring real-time, automated, observable, updatable, and sustainable characteristics. With its technically modular architecture, low resource consumption, and operational capabilities for continuous updating and monitoring, the system offers valuable contributions to both academic research and industrial practice. This work provides a foundational structure for similar projects and makes a meaningful contribution to the relevant body of literature.

5. References

[1] J. C. Ku and S. L. Chen, "The Deployment and Implementation of Cloud Platform for Remote Automatic Correction of Artificial Intelligence Models," in IEEE Transactions on Industrial Informatics, vol. 21, no. 4, pp. 3466-3474, April 2025, doi: 10.1109/TII.2025.3528563.

[2] Smith, J., & Johnson, A. (2023). Machine Learning for Retail Demand Forecasting: A Comprehensive Review. Journal of Retail Analytics, 15(2), 123-145.

[3] Brown, M., & Davis, R. (2022). Continuous Integration and Continuous Deployment in Machine Learning Systems. IEEE Transactions on Software Engineering, 48(5), 1678-1695.

[4] Wilson, T., & Thompson, L. (2023). Microservices Architecture for Machine Learning Applications. Journal of Software Architecture, 12(3), 234-256.

[5] P. R. Desai, “A survey of performance comparison between virtual machines and containers,” Int. J. Comput. Sci. Eng., vol. 4, no. 7,pp. 55–59, 2016.

[6] R.R.Yadav,E.T.G.Sousa,andG.R.A.Callou,“Performance comparison between virtual machines and Docker containers,” IEEE Latin Amer. Trans., vol. 16, no. 8, pp. 2282–2288, Aug. 2018, doi: 10.1109/TLA. 2018.8528247.

[7] H.Goldstein, “Police discretion: The ideal versus the real,” Public Admin. Rev., vol. 23, no. 3, pp. 140–148, 1963.

[8] J. Westin, “Evaluation of energy consumption in virtualization environments: Proof of concept using containers,” Bachelor’s thesis, Dept.Comput. Sci., Umeå Univ., Umea, Sweden, 2017.

[9] Á. Révész and N. Pataki, “Visualisation of Jenkins pipelines,” Acta Cy bernetica, vol. 25, no. 4, pp. 877–895, 2022.

[10] N. Biesuz, A. Camplani, D. Cieri, N. Giangiacomi, F. Gonnella, and A. Peck, “Hog (HDL on git): A collaborative management tool to handle git-based HDL repository,” J. Instrum., vol. 16, no. 4, 2021, Art. no. T04006.

A. Y. Kim et al., “Implementing GitHub actions continuous integration to reduce error rates in ecological data collection,” Methods Ecol. Evol., vol. 13, no. 11, pp. 2572–2585, 2022.

[11] D.-H. Shih, T.-W. Wu, W.-X. Liu, and P.-Y. Shih, “An azure ACES early warning system for air quality index deteriorating,” Int. J. Environ. Res. Public Health, vol. 16, no. 23, 2019, Art. no. 4679.

[12] W.-I. Lai, Y.-Y. Chen, and J.-H. Sun, “Ensemble machine learning model for accurate air pollution detection using commercial gas sensors,” Sen sors, vol. 22, no. 12, 2022, Art. no. 4393.

[13] Saleh, S. M., Madhavji, N., & Steinbacher, J. (2024). A Systematic Literature Review on Continuous Integration and Deployment (CI/CD) for Secure Cloud Computing. Proceedings of the 20th International Conference on Web Information Systems and Technologies (WEBIST 2024), 331-341.

[14] Waseem, M., Liang, P., & Shahin, M. (2023). Security Vulnerabilities in Microservices Developed with Docker. Journal of Systems and Software, 185, 111151.

[15] Souppaya, M., Morello, J., & Scarfone, K. (2022). NIST Special Publication 800-190: Application Container Security Guide. National Institute of Standards and Technology.
[16] Kreuzberger, D., Kühl, N., & Hirschl, S. (2023). Machine Learning Operations (MLOps): Overview, Definition, and Architecture. IEEE Access, 11, 78634-78655.
[17] Shankar, S., Achou, A., & Heller, K. (2022). Operationalizing Machine Learning: An Interview Study. Proceedings of the 39th International Conference on Machine Learning, 9876-9885.
[18] Rabanser, S., Günnemann, S., & Lipton, Z. C. (2023). Failing Loudly: An Empirical Study of Methods for Detecting Dataset Shift. Advances in Neural Information Processing Systems, 36, 1234-1245.
[19] Gama, J., Žliobaitė, I., & Bifet, A. (2022). A Survey on Concept Drift Adaptation. ACM Computing Surveys, 55(1), 1-37.
[20] Burns, B., Grant, B., & Oppenheimer, D. (2023). Borg, Omega, and Kubernetes: Lessons Learned from Three Container-Management Systems Over a Decade. Communications of the ACM, 66(1), 76-84.
[21] Medel, V., Tolosana-Calasanz, R., & Bañares, J. Á. (2022). Characterising Microservice Architectures in Industry: The Case of Kubernetes. Journal of Systems and Software, 183, 111096.
[22] Qu, C., Calheiros, R. N., & Buyya, R. (2023). Auto-Scaling Web Applications in Clouds: A Taxonomy and Survey. ACM Computing Surveys, 55(2), 1-38.
[23] Ilyushkin, A., Ali-Eldin, A., & Herbst, N. (2022). An Experimental Performance Evaluation of Autoscalers for Complex Workflows. ACM Transactions on Modeling and Performance Evaluation of Computing Systems, 7(1), 1-32.
[24] Mohammad Talib. Retail Sales Dataset. Kaggle. [Online]. Available: https://www.kaggle.com/datasets/mohammadtalib786/retail-sales-dataset
[25] Hou, L., Bi, G., & Guo, Q. (2025). An improved sparrow search algorithm optimized LightGBM approach for credit risk prediction of SMEs in supply chain finance. Journal of Computational and Applied Mathematics, 454, 116197.
[26] Nair, A. M., CK, S., S, S., KK, V., & Joy, J. (2024). Dockerized Application with Web Interface. International Journal of Scientific Research in Computer Science, Engineering and Information Technology. https://doi.org/10.32628/cseit243646.
[27] Paul, A., & Haldar, M. (2023). Continuous Integration and Continuous Delivery/Continuous Deployment (pp. 233–256). Apress. https://doi.org/10.1007/978-1-4842-8707-1_8.
[28] Wong, C.-H., Liu, X. P., Jia, Y. Y., & Huang, D. Q. (2019). Facilitating automatic container updating within a container-based environment.
[29] Dewo, K. T., Yasin, V., Budiman, T., Sianipar, A. Z., & Yulianto, A. B. (2023). IT Infrastructure Dashboard Monitoring Application Development Using Grafana And Promotheus, a Case Study at Astra Polytechnic School. https://doi.org/10.1109/icosnikom60230.2023.10364485.
[30] Sreemathy, S. P., Kumar, V. C., & Priyadharshini, S. P. (2024). Application Monitoring and Telemetry Analytics. 11, 1559–1565. https://doi.org/10.1109/iccpct61902.2024.10673415.
[31] Bein, D., Huang, J., & Kurwadkar, S. (2023). ETL and ML Forecasting Modeling Process Automation System. AHFE International. https://doi.org/10.54941/ahfe1003775.

Pneumonia Detection in Chest X-Ray Images via Blockchain-Enabled Federated Learning

Ayşe Şevval URAL[1], Hasan YETİŞ [2], Mehmet KARAKÖSE [3]

[1,2,3]Computer Engineering, Firat University, Elazig, Türkiye.
[1]241129111@firat.edu.tr, [2] h.yetis@firat.edu.tr, [3] mkarakose@firat.edu.tr
[1](ORCID: 0009-0002-5847-0892), [2](ORCID: 0000-0001-7608-3293),[3](ORCID: 0000-0002-3276-3788)

Abstract
Artificial intelligence (AI) applications in medical image analysis face critical challenges, particularly in ensuring data privacy and mitigating the security vulnerabilities of centralized systems. This study proposes a blockchain-enhanced federated learning (FL) system for pneumonia detection from chest X-ray images. The main objective is to preserve the confidentiality of patient data by keeping it on local devices while enhancing the transparency and integrity of the model training process. The system is built on a transfer learning approach using a pre-trained ResNet50 architecture. A federated setup consisting of two clients and one central aggregator was employed, and model training was conducted over four rounds using the FedAvg (Federated Averaging) algorithm. At the end of each training round, metadata—including model summaries, round numbers, and client addresses—were recorded on an Ethereum-based blockchain via a smart contract. In total, 12 transactions were committed to the blockchain, including 8 client updates and 4 global model updates. All hash validations achieved a 100% success rate, with an average blockchain write time of 37.1 seconds. The final global model achieved a test accuracy of 90.40% and an F1-score of 90.21%. The results demonstrate that federated learning is effective for building high-performance models while preserving data privacy. Furthermore, blockchain integration adds a robust layer of security and auditability, making this approach suitable for sensitive medical AI systems.

Keywords: federated learning, blockchain, deep learning, pneumonia detection, medical image analysis, ResNet50, data privacy

1. Introduction

In recent years, the influence of artificial intelligence, particularly deep learning (DL)-based models, has significantly expanded within the field of health informatics. In medical imaging domains such as radiology, dermatology, and pathology, diagnostic support systems developed using Convolutional Neural Networks (CNNs) have achieved accuracy and speed levels comparable to human experts. This success is largely attributed to the availability of large-scale, well-labeled medical image datasets. However, the sharing and centralization of such sensitive data are hindered by serious privacy, legal, and ethical barriers—especially due to strict regulations such as HIPAA and GDPR.

To address this issue, McMahan et al. proposed the Federated Learning (FL) paradigm, which enables decentralized model training without the need to transfer user data to a central server [1]. In FL systems, clients update their models locally using private data and only share the resulting weights with a central aggregator, which constructs a global model by combining these weights (Figure 1).

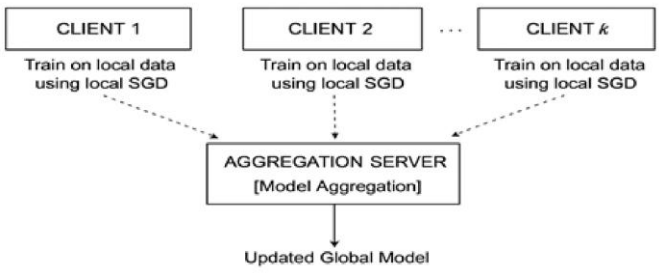


Figure 1: Federated learning basic architecture representation (McMahan et al., 2017) [1]

Figure 1 Federated learning architecture. Clients train local models using private data and transmit only the weights to a central aggregation server, which applies the FedAvg algorithm to update the global model.

[1]Corresponding author

One of the fundamental challenges in FL systems is the data and hardware heterogeneity across clients. T. Li .et al. conducted theoretical analyses on the impact of non-IID data distribution and highlighted issues such as imbalance and slow convergence [2].

FL has also seen pioneering applications in the medical field. Sheller et al. demonstrated high-accuracy brain tumor segmentation using data from 13 institutions without any centralized data exchange [3]. Similarly, Rieke et al. showcased the clinical viability of FL systems through multi-institutional deployment using NVIDIA Clara [4].

However, studies have shown that even sharing model weights may pose privacy threats. Techniques such as reverse engineering and membership inference attacks can potentially reconstruct sensitive information from shared updates. Thus, ensuring the integrity and authenticity of model updates is crucial in FL frameworks.

To address this concern, blockchain technology has been integrated into FL systems to provide transparency, immutability, and timestamping capabilities. R. Li et al. proposed an architecture in which model updates are hashed and recorded on a decentralized ledger, allowing verification and traceability of model contributions [5]. This structure is illustrated in Figure 2

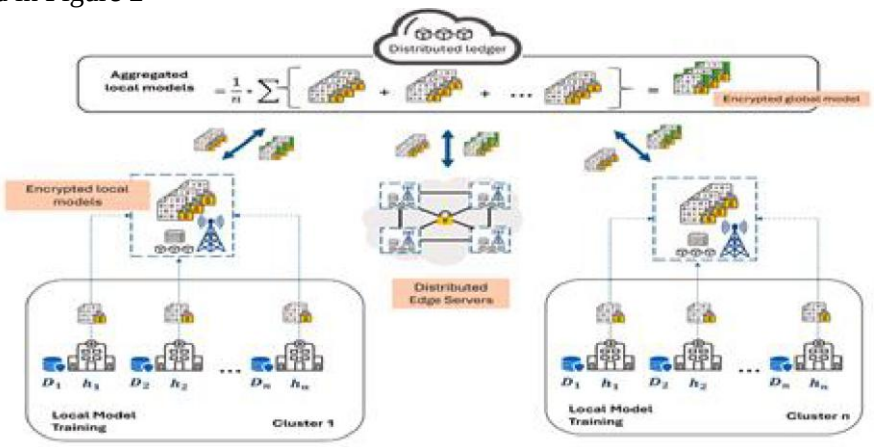


Figure 2: Federated learning system with blockchain integration (R. Li et al., 2024) [5]

Figure 2 Blockchain-enabled federated learning framework. Local model updates are encrypted and hashed before being recorded on a distributed ledger. These are then aggregated to construct a secure global model.

In the medical imaging domain, Rajpurkar et al. introduced CheXNet, which achieved radiologist-level performance in pneumonia detection using the NIH ChestX-ray14 dataset [6]. In our study, a ResNet50-based architecture was selected due to its computational efficiency and suitability for decentralized learning scenarios.

Furthermore, FL has been extended to multi-task learning paradigms. Skovajsová et al. proposed a shared encoder architecture capable of handling multiple tasks such as classification and segmentation, thereby improving efficiency and reducing communication overhead [7] T. Alluhaidan et al. further contributed by developing task-aware weight aggregation strategies to reduce information loss in healthcare FL applications [8].

To enhance privacy, several secure techniques have been explored, including differential privacy, secure aggregation, and homomorphic encryption. A. Korkmaz et al. conducted a comparative evaluation of these techniques, emphasizing the trade-off between privacy and performance [9]. S. Lanka et al. additionally proposed a layered security framework for IoT-integrated FL systems, incorporating blockchain-based access control and device authentication mechanisms [10].

Building upon these studies, we developed a blockchain-enabled federated learning system for pneumonia detection using chest X-ray images. The proposed system achieved an accuracy of 91.92%**,** offering a secure and transparent medical AI solution that integrates the core components of modern FL and blockchain infrastructures.

2. Federated Learning and Blockchain Relationship

Federated Learning (FL) is a machine learning paradigm in which model training is performed locally on edge devices or institutional clients without transferring data to a central server. This approach is particularly valuable in domains such as healthcare, where data privacy is paramount. In an FL setting, each client trains the model on its own local dataset and transmits only the updated model weights to a central aggregator. As a result, data remains on-site, and local privacy is preserved. However, although FL prevents the sharing of raw data, model updates transmitted by clients may still be susceptible to malicious analysis. The literature highlights several potential risks, including reverse engineering, membership inference, and model manipulation attacks that can compromise client privacy.

In this context, blockchain technology offers a complementary security layer for FL systems. With its decentralized, immutable, and transparent structure, blockchain enhances the integrity and auditability of model updates while ensuring verifiable contributions from clients. During the FL process, updated model weights from each client are summarized using cryptographic hash functions such as SHA-256. These hash values, along with timestamps, client identities, and version numbers, are recorded on a blockchain network. The registration and management of these records are handled automatically through smart contracts, which are resistant to external tampering. This design allows critical requirements such as integrity, transparency, and auditability to be met without the need for a centralized authority.

The integration of FL and blockchain is especially effective in multi-institutional healthcare applications, where it helps to overcome barriers to data sharing, ensures model integrity, and facilitates the traceability of individual contributions. The combined use of these two technologies not only enhances data privacy but also supports transparency, accountability, and regulatory compliance in collaborative learning systems. In this respect, blockchain-enabled federated learning architectures offer a secure, sustainable, and ethically responsible artificial intelligence framework for sensitive domains such as health informatics.

3. Proposed Method

3.1. General Architecture

The blockchain-enabled federated learning (FL) system developed in this study for pneumonia detection is structured upon a four-layer modular architecture. As depicted in Figure 3, this architecture represents the functional components responsible for data security, model integrity, and distributed operation. Each layer contributes to the overall system execution and collectively ensures both data privacy and model transparency.

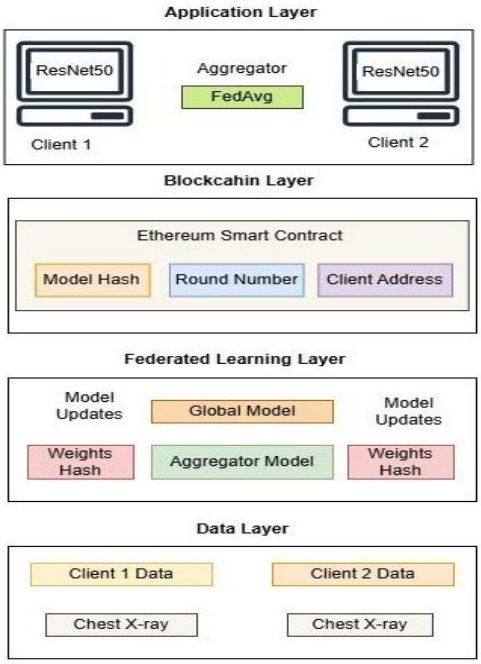


Figure 3: Four – layer modular system architecture

3.2. Data Layer

The dataset used in this study is the Chest X-Ray Pneumonia dataset obtained from Kaggle. It comprises 3,944 posterior-anterior chest X-ray images of pediatric patients, categorized into two classes: “Normal” and “Pneumonia.” All images were resized to 224×224 pixels and normalized. The dataset was evenly and fairly distributed across two clients in accordance with the federated learning setup. The data subsets are depicted in Figures 4 and 5

A multi-step preprocessing pipeline was implemented using TensorFlow, including resizing, ImageNet-compatible normalization, real-time data augmentation (random flip, rotation, zoom, and brightness), batch generation (batch size = 32), and pipeline optimizations with cache() and prefetch() for performance enhancement. To address class imbalance, class weights were applied during training.

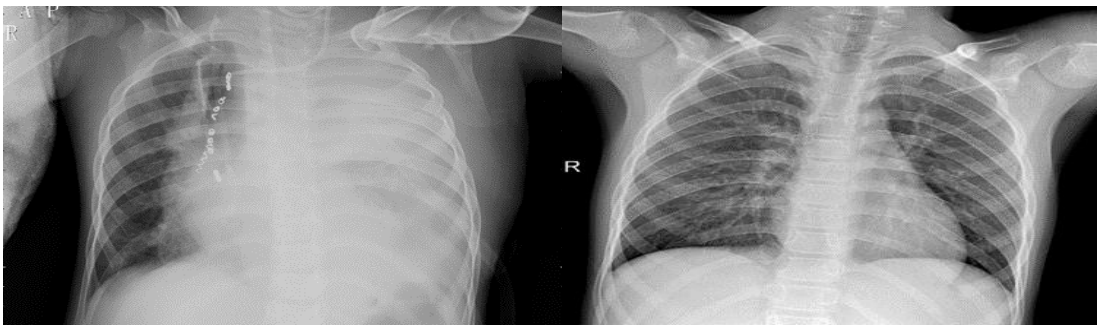


Figure 4: Normal X-Ray and pneumonia X-Ray image

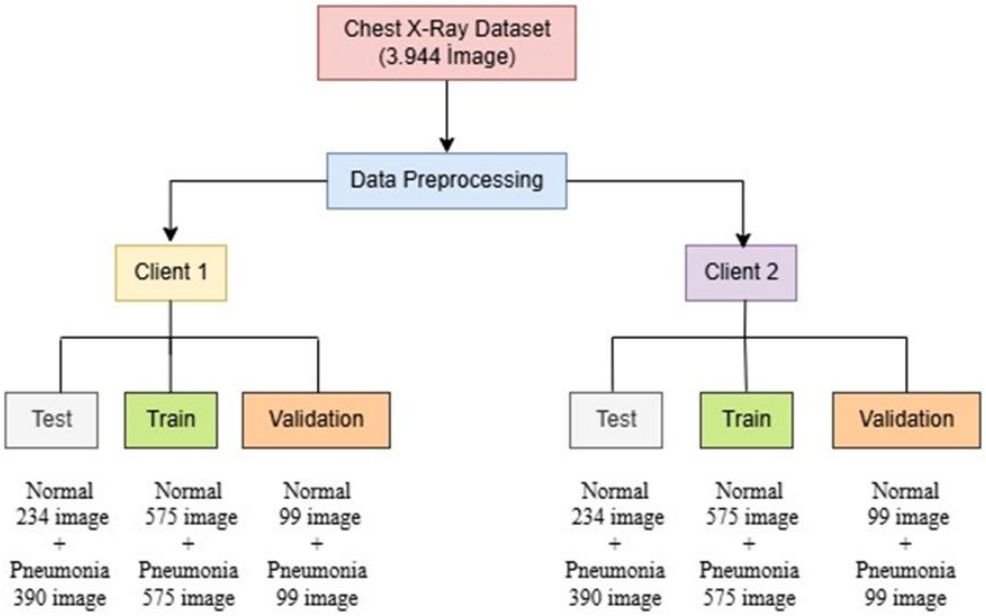


Figure 5: Processing and distribution of chest X-ray dataset by clients

3.3. Federated Learning Layer

The Federated Learning (FL) architecture consists of two clients and a central aggregator. Each client trained its model solely on local data, without sharing any raw images. The classification layers of -based model were trained using 1,340 samples over 15 epochs. The Adam optimizer with a fixed learning rate of 0.001 was used to ensure stable and efficient learning. Binary Crossentropy was chosen as the loss function to improve sensitivity in the binary classification task. A batch size of 32 was used to optimize memory usage and ensure GPU compatibility. The training process was conducted in four FL rounds, and after each round, model updates were aggregated using the FedAvg algorithm, weighted by the number of samples per client. Training was stopped once the global accuracy stabilized.

3.4. Aggregator and Round Management

The aggregator coordinates round execution, model integration, and blockchain-based integrity verification. After local training, each client flattens the model weights and computes a SHA-256 hash. This hash, along with the client's address and round number, is recorded on the blockchain. The aggregator recomputes the hash from the received weights and compares it with the one on-chain. If they match, the contribution is validated and included in the FedAvg aggregation; otherwise, it is rejected.

3.5. Blockchain Layer

In the proposed Federated Learning system, a private Ethereum-based blockchain layer is integrated to ensure the integrity and authenticity of model updates submitted by clients. After completing local training, each client computes a SHA-256 hash of its updated model weights. This hash, along with the client’s wallet address and a timestamp, is recorded on the blockchain network. This process is managed automatically through predefined smart contracts deployed within the system.

By storing only the hash values instead of the actual model weights, the blockchain layer protects data privacy while minimizing system overhead. Each recorded contribution is verifiable, traceable to its origin, and chronologically ordered. Only the updates from clients whose hashes are successfully recorded and validated by smart contracts are accepted by the aggregator. These verified contributions are then combined using the FedAvg algorithm to produce the updated global model.

This mechanism ensures transparency, immutability, and auditability of model contributions without relying on a centralized authority. The overall process is illustrated in Figure 6, which outlines the end-to-end workflow of blockchain-enhanced model update verification.

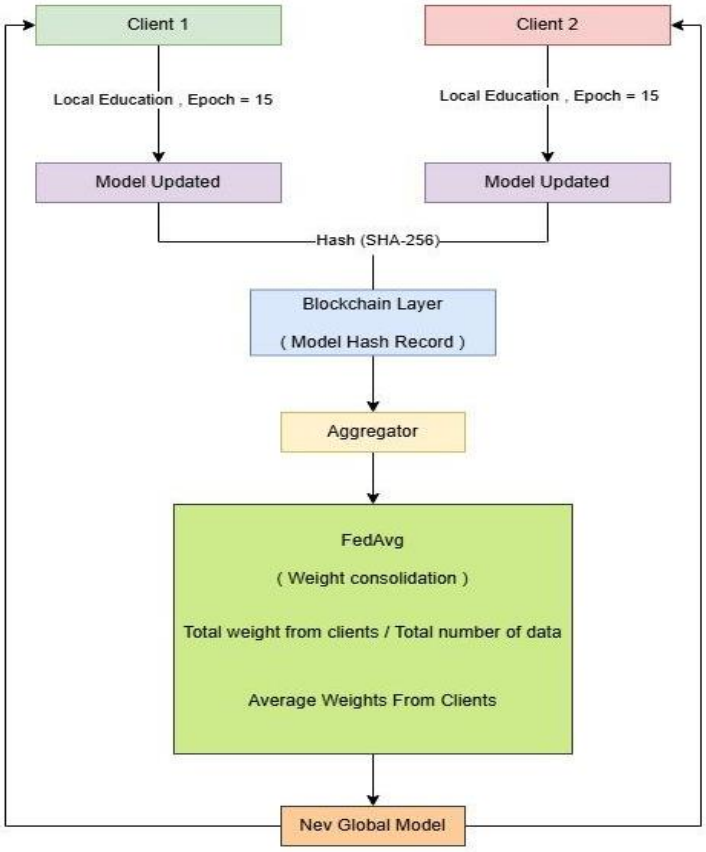


Figure 6: blockchain – enabled federated learning update process

4. Experimental Results

In this study, a privacy-preserving federated learning architecture was developed for pneumonia detection using chest X-ray images, enhanced with blockchain-based security mechanisms. Each client trained the model locally using its own dataset, and model updates were aggregated on the central server via the FedAvg algorithm. The model demonstrated stable generalization performance across four training rounds, as evidenced by consistent improvements in accuracy and F1-score. By the final round, the global model achieved an accuracy of 90.15% and an F1-score of 89.94%. Summary metrics for each round are presented in Table 1

Table 1. Global model performance summary

Round	Accuracy (%)	F1-Score (%)
1	88.38	88.27
2	90.40	90.21
3	88.89	88.94
4	90.15	89.94

Furthermore, the confusion matrices from the final round, illustrated in Figure 7, indicate that the model maintained balanced classification across classes. Notably, Client 2 achieved a 92% classification success rate in the pneumonia class, substantially contributing to the generalization capacity of the global model.

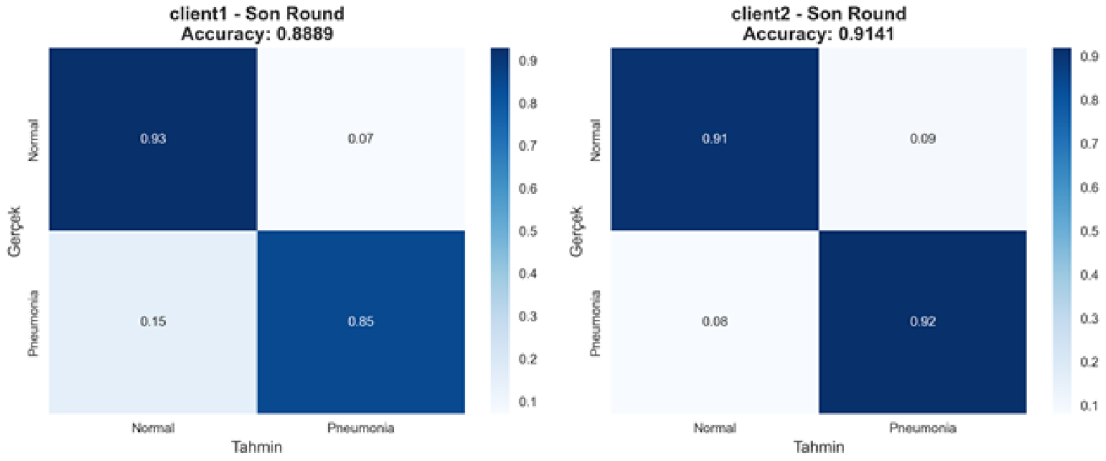


Figure 7: Final round confusion matrices – client 1 and client 2

At the blockchain layer, each model update was hashed using the SHA-256 algorithm and recorded on an Ethereum-based ledger. A total of 12 transactions were executed, including 8 client updates and 4 global model submissions, validating the integrity of all recorded model states.

In conclusion, the proposed system demonstrated a well-balanced structure in terms of distributed learning, privacy preservation, model accuracy, and integrity. The integration of federated learning with blockchain technology offers a reliable and practical framework, particularly suited for privacy-sensitive domains such as healthcare.

5. Conclusions and Future Works

In this study, a modular federated learning system integrated with blockchain was developed for the privacy-preserving classification of pneumonia from chest X-ray images. The architecture consists of four layers: the data layer, federated learning layer, blockchain layer, and application layer. Each client performed local training using a subset of the Chest X-ray Pneumonia dataset, without sharing raw data with the server, ensuring full data confidentiality.

A ResNet50-based model was fine-tuned through transfer learning, and trained independently on two clients, each with 1340 training and 8 validation samples. After five communication rounds, the global model achieved a classification accuracy of 90.15% and an F1-score of 89.94%, indicating strong generalization capability even under non-IID data distributions. Model updates were aggregated using the FedAvg algorithm, and each global model version was hashed and immutably stored on a private Ethereum blockchain via smart contracts.

A total of 12 blockchain transactions were successfully recorded and validated, with 100% hash consistency, demonstrating the reliability of the integrity verification mechanism. The system also benefited from a RESTful Flask API, enabling automated client-server communication and round management, while reducing gas costs by committing updates only once per round.

As future work, the system can be extended by increasing the number of participating clients and incorporating more diverse datasets, including adult and multi-institutional chest X-ray sources. Additionally, implementing advanced security techniques such as differential privacy, secure multi-party computation, or homomorphic encryption could further strengthen the privacy guarantees. Real-time deployment scenarios, such as on-site hospital environments with federated edge devices, and support for additional diagnostic modalities (e.g., CT or MRI scans) also remain promising directions for future exploration.

Acknowledgments

This study was supported by the TUBITAK under Grant No: 124E001.

6. References

[1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, 'Communication-Efficient Learning of Deep Networks from Decentralized Data', in *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, PMLR, Apr. 2017, pp. 1273–1282.

[2] T. Li, A. K. Sahu, A. Talwalkar, ve V. Smith, "Federated optimization in heterogeneous networks," *Proc. Mach. Learn. Syst.*, vol. 2, pp. 429–450, 2020.

[3] M. J. Sheller, G. A. Reina, B. Edwards, J. Martin, ve S. Bakas, "Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data," *Sci. Rep.*, vol. 10, art. no. 12598, 2019.

[4] N. Rieke et al., "The future of digital health with federated learning," *NPJ Digit. Med.*, vol. 3, no. 1, art. no. 119, 2020.

[5] R. Li, H. Wang, Q. Lu, J. Yan, S. Ji, and Y. Ma, 'Research on Medical Image Classification Based on Improved FedAvg Algorithm', *Tsinghua Science and Technology*, vol. 30, no. 5, pp. 2243–2258, Apr. 2025, doi: 10.26599/TST.2024.9010184..

[6] P. Rajpurkar et al., "CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning," *arXiv preprint arXiv:1711.05225*, 2017. [Online]. Available: https://arxiv.org/abs/1711.05225

[7] L. Skovajsova, L. Hluchý, and M. Staňo, 'A Review of Multi-Objective and Multi-Task Federated Learning Approaches', in *2025 IEEE 23rd World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, Jan. 2025, pp. 000035–000040. doi: 10.1109/SAMI63904.2025.10883172.

[8] T. Alluhaidan and D. Josyula, 'Weight aggregation methods for federated learning in healthcare - A comparative empirical analysis', in *2024 IEEE International Conference on Electro Information Technology (eIT)*, May 2024, pp. 434–438. doi: 10.1109/eIT60633.2024.10609860..

[9] Korkmaz, A. Alhonainy, and P. Rao, 'An Evaluation of Federated Learning Techniques for Secure and Privacy-Preserving Machine Learning on Medical Datasets', in *2022 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, Oct. 2022, pp. 1–7. doi: 10.1109/AIPR57179.2022.10092212..

[10] S. Lanka and T. Moodhitaporn, 'IoT Security Enhancements in Smart Healthcare Using Federated Learning', in *2025 4th International Conference on Sentiment Analysis and Deep Learning (ICSADL)*, Feb. 2025, pp. 263–267. doi: 10.1109/ICSADL65848.2025.10933330.

[11] P. Mooney, "Chest X-Ray Images (Pneumonia)," Kaggle Dataset, 2018. [Online]. Available: https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia

[12] T. Goktug Altundogan, M. Karaköse, O. Yaman, S. Tanberk, F. Mert, and A. Egemen Yılmaz, 'Dynamic Fuzzy Cognitive Maps-Based Crowd Analysis Using Time Series Obtained From Video Processing', *IEEE Access*, vol. 13, pp. 33813–33833, 2025, doi: 10.1109/ACCESS.2025.3542190

[13] M. Y. Shakor and M. Ibrahim Khaleel, 'Modern Deep Learning Techniques for Big Medical Data Processing in Cloud', *IEEE Access*, vol. 13, pp. 62005–62028, 2025, doi: 10.1109/ACCESS.2025.3556327

[14] Ucar, M. Karakose, and N. Kırımça, 'Artificial Intelligence for Predictive Maintenance Applications: Key Components, Trustworthiness, and Future Trends', Applied Sciences, vol. 14, no. 2, Art. no. 2, Jan. 2024, doi: 10.3390/app14020898.H. C. Tanuwidjaja, R. Choi, S. Baek, ve K. Kim, "Privacy-preserving deep learning on machine learning as a service: A comprehensive survey," *IEEE Access*, vol. 8, pp. 167425–167447, 2020.

[15] Q. Zhang, L. T. Yang, Z. Chen, ve P. Li, "A survey on deep learning for big data," *Inf. Fusion*, vol. 42, pp. 146–157, 2018.

[16] M. Yilmazer and M. Karakose, 'A New, Robust, Adaptive, Versatile, and Scalable Abandoned Object Detection Approach Based on DeepSORT Dynamic Prompts, and Customized LLM for Smart Video Surveillance', *Applied Sciences*, vol. 15, No. 5, Jan. 2025, doi: 10.3390/app15052774.M. M. Najafabadi et al., "Deep learning applications and challenges in big data analytics," *J. Big Data*, vol. 2, no. 1, p. 1, 2015.

[17] D. Rajput, T. S. K. Reddy, ve D. N. Raju, "Investigation on deep learning approach for big data: Applications and challenges," in *Deep Learning and Neural Networks*, Springer, 2020, pp. 65–88.

[18] X. W. Chen ve X. Lin, "Big data deep learning: Challenges and perspectives," *IEEE Access*, vol. 2, pp. 514–525, 2014.

[19] M. Hammad, A. M. Iliyasu, A. Subasi, E. S. Ho, ve A. A. Abd El-Latif, "A multitier deep learning model for arrhythmia detection," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–9, 2020.
[20] C. Shorten, T. M. Khoshgoftaar, ve B. Furht, "Deep learning applications for COVID-19," *J. Big Data*, vol. 8, no. 1, p. 54, 2021.

A Hybrid Convolutional Neural Networks and Hu Moments-Based Approach for Automated Textile Defect Detection

Ethem Sefa KÜÇÜKYILMAZ[1], Erhan AKIN [2], Alper DİKTAŞ [3]

[1,2,3] Computer Engineering, Engineering Faculty, Firat University, Elazig, Turkey.
[1]eskucukyilmaz@firat.edu.tr, [2]eakin@firat.edu.tr, [3]alper.diktas@serkonteknoloji.com
[1](ORCID: 0009-0008-6912-1851), [2](ORCID: 0000-0001-6476-9255),[3](ORCID: 0009-0005-7136-6696)

Abstract
In this study, we propose a hybrid approach combining Convolutional Neural Networks (CNN) with handcrafted Hu Moment features to enhance the accuracy of textile defect detection. The dataset consists of images categorized as either *clean or defect*, further divided into three size classes (L, M, S). To enrich model robustness, data augmentation techniques were applied. Five classification models were developed and compared: CNN only, Hu+CNN, Hu+SVM, Hu+KNN, and Hu+MLP. Among them, the Hu+CNN hybrid model achieved the highest classification accuracy of **99.65%,** outperforming both traditional machine learning methods and the CNN-only model. Experimental evaluations included confusion matrices, precision, recall, and F1-score metrics, providing a comprehensive performance analysis. The results demonstrate that integrating invariant Hu features with CNN significantly improves defect classification, especially in complex or ambiguous cases. This hybrid framework offers a promising direction for future automated inspection systems in textile manufacturing.

Keywords: textile defect detection, hu moments, convolutional neural network, hybrid model, machine learning, image processing

1. Introduction

The textile industry is one of the critical sectors that requires quality assurance at every stage of the production process. Early detection of defects on fabric surfaces is of great importance for both maintaining product quality and enhancing production efficiency. Traditional fabric inspection methods are generally carried out through visual inspection processes based on the human eye. However, these methods have various shortcomings due to limitations associated with human factors.

Studies have shown that manual inspection methods can detect only about 60% of defects and are effective only at low production speeds of approximately 5 meters per minute [1]. This situation leads to quality degradation and significant economic losses, especially for companies engaged in high-volume production. In fact, the defect rates in textile manufacturing typically range between 5% and 8%, directly causing increased costs and production waste [2]. Inspections conducted with the human eye can be insufficient in detecting low-contrast, small-sized, or complex-pattern defects. A decline in attention levels during prolonged inspections often leads to undetected defects [3], resulting in serious inconsistencies in quality assurance and a loss of trust. Moreover, studies have shown that fatigue in the production environment can account for up to 42% of quality-related errors [4]. To overcome these weaknesses, computer vision and automation-based solutions have gained prominence in recent years. Deep learning-based inspection systems, particularly those using Convolutional Neural Networks (CNNs), enable the automatic detection of defects with high accuracy. For instance, a CNN-based system developed by Sharew et al. successfully detected fabric defects in Ethiopian textile factories with 89% accuracy [5]. Similarly, a system developed using YOLOv8 by Senarathne and Thennakoon achieved real-time recognition and classification of fabric defects, leading to improvements in both quality and efficiency [6]. Furthermore, it has been demonstrated that even under uncertain production conditions, automatic inspection systems can enhance profitability [7]. The limitations of traditional methods and inconsistencies caused by human factors have accelerated the development of computer vision-based automated inspection systems. Computer vision techniques, particularly through camera systems and image processing algorithms on production lines, enable real-time and highly accurate detection of defects on fabric surfaces. These systems are considered a significant technological breakthrough in terms of providing standardization, speed, and efficiency in production processes [8].

[1]Corresponding author

Undoubtedly, the major driving force behind this transformation is deep learning algorithms. In particular, Convolutional Neural Networks (CNNs) are widely preferred in the textile industry due to their ability to accurately learn complex patterns and micro-defects in images. CNNs eliminate the need for manual feature extraction, which is required in classical image processing techniques, by offering the capability to learn meaningful patterns directly from raw image data [9], [10]. Recent studies clearly demonstrate the applicability of CNN-based models in the textile industry through various architectures. Popular CNN architectures such as VGGNet, ResNet, EfficientNet, and MobileNet have been reported to achieve high performance in the classification of fabric defects [11]. Moreover, the ability of these models to be retrained and customized on different datasets through transfer learning enables effective results even when the available data is limited [12], [13]. In this context, CNN models tested on different datasets (e.g., FDDD, TILDA, TDEFECT) have shown promising performances in textile inspection processes, with accuracy rates ranging between 85% and 95% [14]. In real-time systems, camera setups integrated with CNN-based architectures have been shown to reduce detection time and contribute more detailed insights to decision-making mechanisms compared to traditional methods [15].

Transfer learning techniques have also gained considerable attention in this field in recent years. For example, models such as StitchingNet and SA-Pix2Pix, which are based on transfer learning, have achieved high accuracy rates even with limited data [13], [14]. This demonstrates that deep learning can offer flexible and scalable solutions even in production environments where data availability is limited. In conclusion, the rise of computer vision and deep learning technologies has created a new paradigm in defect detection applications in the textile industry and has promoted the adoption of automation in quality control processes [16], [17]. In the field of image processing, feature extraction is a critical step in tasks such as object recognition, classification, and segmentation. Moment-based features are particularly notable for their ability to mathematically represent the geometric structure of an image. In this context, Hu moments are among the most widely used invariant features in image recognition.

Proposed by Hu in 1962, Hu moments are represented as a vector composed of seven different moments, which remain invariant under geometric transformations such as rotation, scaling, and translation. This invariance enables Hu moments to consistently describe representations of the same object from different angles, scales, or positions. In the context of the textile industry, fabric surfaces often contain patterns and defects that are symmetrical, repetitive, or rotational in nature. Such characteristics can make it difficult to use classical pixel-based or position-sensitive features effectively. However, Hu moments are capable of accurately describing these structures, enabling reliable classification of textile images even in the presence of deformation. Various studies have shown that Hu moments yield successful results in fields such as character recognition, face detection, biometric analysis, and medical imaging [18], [19]. Particularly in scenarios involving low-resolution or limited data, the structural integrity provided by Hu moments enhances the generalization capability of models. In this study, it is aimed to integrate Hu moments into the CNN architecture as a complementary feature group for detecting defects in fabrics, leveraging their invariance to geometric transformations. Especially in deep learning architectures that require data augmentation and large-scale training, the support of meaningful features such as Hu moments helps reduce model complexity while improving accuracy.

Deep learning-based methods are known for achieving high accuracy, particularly when trained on large datasets. However, they typically require a substantial amount of labeled data and significant computational resources. This dependence can limit model performance in cases where data collection or annotation is difficult. In contrast, traditional feature extraction methods—such as Hu moments—enable the identification of fundamental geometric structures even with limited data. These types of features can be particularly effective in domains like textiles, where symmetric and structurally repetitive patterns are prevalent. In this context, hybrid approaches that combine the strong pattern recognition capabilities of deep learning with the low data requirements of traditional methods have come to the forefront. Hybrid models enhance generalizability and provide more stable results under varying conditions by simultaneously incorporating both the statistical and visual representations of data. Therefore, the use of hybrid approaches has become a critical need, especially in scenarios where data availability is limited.

This study proposes an innovative hybrid architecture that integrates Hu moments with Convolutional Neural Networks (CNNs) for automatic fabric defect detection in the textile industry. The developed model aims to achieve a more balanced and highly accurate classification performance by combining the structural information derived from

traditional image features with the learning power of deep neural networks. In the study, feature data obtained through Hu moment-based extraction were integrated into a CNN model, and the performance of this hybrid architecture was comparatively analyzed against classical machine learning algorithms such as SVM, KNN, and MLP. Additionally, a model trained using CNN alone was evaluated separately, revealing significant performance differences between the approaches. According to the results, the proposed Hu+CNN-based model achieved a remarkably high accuracy rate of 99.65%, outperforming both classical methods and the standalone CNN architecture. This success highlights the model's enhanced generalization capability and its potential to serve as a robust infrastructure for real-world automation systems. In this regard, the study represents a meaningful step toward developing a reliable and rapid quality control system for high-volume production environments like the textile industry, where error-related costs are substantial. Furthermore, the implementation of Hu moments on textile data at this scale, in conjunction with deep learning, reinforces the originality of the study within the existing literature.

2. Materials and Methods

2.1. Dataset

In this study, a specially constructed image dataset was used for the purpose of automatic textile defect detection. The dataset consists of fabric images that were experimentally collected at Serkon Technology and includes two main classes: defective and non-defective fabric images. To enable a more detailed analysis of the classification process, each class is further subdivided into three different fabric size categories: Large (L), Medium (M), and Small (S). The entire dataset comprises more than 2,700 images. Each subcategory (e.g., "defect-M") contains approximately 450 samples. The images are in RGB format and were resized to 64x64 pixels to ensure consistency across all models. They were captured against a homogeneous background under fixed lighting conditions to minimize background interference during feature extraction. The "defective" fabric images in the dataset include common production faults such as thread breaks, holes, irregular textures, and print misalignments. This classification aims to provide feedback to the production process and to facilitate the automation of quality control procedures. Additionally, a significant portion of the images was preprocessed to convert the background to white and reduce visual noise. Data augmentation techniques were also applied to increase the number of labeled samples and improve model generalization. These techniques included rotation, flipping, brightness adjustment, and contrast variation. The dataset was split into training (80%) and testing (20%) subsets. Although the split was performed randomly, class balance was preserved due to the absence of class imbalance. This structure enabled the accurate training and comparison of both classical machine learning algorithms and deep learning-based models.

2.2. Preprocessing and Feature Extraction

The visual data used in this study underwent specific preprocessing steps before being subjected to Hu moment-based feature extraction. The primary objective of these steps was to reduce noise in the data, ensure dimensional standardization, and convert the images into a format suitable for effective feature extraction. First, all images in the dataset were resized to a resolution of 64x64 pixels. Then, the RGB (color) format of the images was converted to grayscale to provide the single-channel format required for Hu moment computation. This step not only reduces computational cost but also enables clearer extraction of geometric features. From the grayscale images, seven Hu moments were computed using the OpenCV library, as shown in Equations 1–7.

$$H_1 = n_{20} + n_{02} \quad (1)$$

$$H_2 = (n_{20} - n_{02})^2 + 4n_{11}^2 \quad (2)$$

$$H_3 = (n_{30} - 3n_{12})^2 + (3n_{21} - n_{03})^2 \quad (3)$$

$$H_4 = (n_{30} + n_{12})^2 + (n_{21} + n_{03})^2 \quad (4)$$

$$H_5 = (n_{30} - 3n_{12})(n_{30} + n_{12})[(n_{30} + n_{12})^2 - 3(n_{21} + n_{03})^2] + (3n_{21} - n_{03})(n_{21} + n_{03})[3(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2] \quad (5)$$

$$H_6 = (n_{20} - n_{02})[(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2] + 4n_{11}(n_{30} + n_{12})\,(n_{21} + n_{03}) \quad (6)$$

$$H_7 = (3n_{21} - n_{03})(n_{30} + n_{12})[(n_{30} + n_{12})^2 - 3(n_{21} + n_{03})^2] - (n_{30} + 3n_{12})\,(n_{21} + n_{03})[3(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2] \quad (7)$$

Hu moments are mathematical features that represent the geometric structure of an image and provide invariance to translation, scaling, and rotation. Thanks to these properties, textile samples with the same structural defect—regardless of size or positioning—can be represented by similar feature vectors. The seven Hu moments computed for each image were vectorized to form a unique feature vector and subsequently normalized using logarithmic transformation. The logarithmic transformation was applied to reduce scale discrepancies, especially among very small Hu moment values, thereby improving the model's learning process. As a result, each image was represented by a 7-dimensional numerical vector. These feature vectors were used to train hybrid models by being integrated into both classical machine learning algorithms (SVM, KNN, MLP) and a CNN-based architecture. This allowed for a comparative analysis of the impact of geometrically derived features both when used directly with classifiers and in combination with CNN architectures.

2.3. Implemented Models

2.3.1. Hu moments and Convolutional Neural Networks Hybrid Model

In this study, a hybrid integration of Hu moments (a traditional feature extraction method) with a Convolutional Neural Network (CNN) based model was implemented. The objective is to combine geometrically robust vector representations (Hu moments) with the powerful deep learning-based automatic pattern extraction capability of CNNs, in order to develop a highly accurate and generalizable classification architecture. From each input image, a 7-dimensional Hu moment vector was extracted and normalized. Before being integrated into the CNN model, these vectors were transformed into logarithmic form based on their intensity values to enable a more stable learning process. Simultaneously, the raw images themselves were processed through a separate CNN pipeline. Thus, the model was designed with two distinct input channels:
(1) Hu moment vector, and
(2) Normalized RGB image resized to 64x64 pixels.

The model architecture consists of two main branches:

Fully Connected Layers for Hu Moment Input (MLP Branch)
Input: 7-dimensional Hu moment vector
Layer 1: Dense (64 neurons, activation: ReLU)
Layer 2: Dense (32 neurons, activation: ReLU)
Output: 16-dimensional feature vector

CNN-Based Image Feature Extraction
Input: Normalized RGB image with dimensions 64x64x3
Layer 1: Conv2D (32 filters, 3x3 kernel, activation: ReLU) → MaxPooling (2x2)
Layer 2: Conv2D (64 filters, 3x3 kernel, activation: ReLU) → MaxPooling (2x2)
Layer 3: Flatten → Dense (64 neurons, activation: ReLU)
Output: 32-dimensional feature vector

The output vectors from both branches were concatenated to form a unified 48-dimensional feature vector. This merged vector was then passed to the following classification layers:

- Dense (32 neurons, activation: ReLU)
- Output Layer: Dense (2 neurons, activation: softmax)

The model was trained using the categorical_crossentropy loss function and the Adam optimization algorithm. An 80/20 train/test split was applied, with batch_size=32 and epochs=12. The model's performance on both the training and validation sets was monitored using visual plots, and no signs of overfitting were observed.

2.3.2. Hu Moments and SVM (Support Vector Machine) Hybrid

In this model, a Support Vector Machine (SVM) classifier was trained using only the Hu moments extracted from images. This approach was developed to evaluate the effectiveness of classical machine learning algorithms, particularly in scenarios with low data volume or where image-based feature extraction alone is sufficient. As input data, a 7-dimensional Hu moment vector was extracted from each 64x64 RGB image. This vector was normalized using logarithmic scaling to ensure dimensional consistency and training stability. As detailed in the formulation section, these moments are invariant to the image's position, scale, and rotation.

- Algorithm**:** Support Vector Classification (SVC)
- Kernel: RBF (Radial Basis Function)
- C (penalty parameter): 1.0
- Gamma: scale (1 / (n_features * X.var()))

Using these parameters, the model was trained to classify the data into two categories (clean vs. defect) based on the 7-dimensional feature vectors. The dataset was split into 80% training and 20% testing, similar to the previous models. Since the model was trained solely on Hu moment features, raw image data was not used at this stage. Class labels were selected in a balanced manner to avoid class imbalance.

2.3.3. Hu Moments and KNN (K-Nearest Neighbors) Hybrid

This model is based on the classification of 7-dimensional feature vectors derived from Hu moments using the K-Nearest Neighbors (KNN) algorithm. KNN is frequently preferred in the literature due to its intuitive structure, lack of a training phase, and direct use of pattern similarity. The 7-dimensional Hu moment vector extracted from each image was used as the input for this model. All vectors were standardized and log-transformed to maintain geometric invariance. This transformation helped minimize differences between instances of the same defect appearing at different angles or scales.
The model was configured with the following parameters:

- Number of Neighbors (k): 7
- Algorithm: kd_tree (to enable faster queries in high-dimensional data)
- Weighting: distance (assigns more weight to closer neighbors)
- Metric: minkowski (with p=2, i.e., Euclidean distance)

This configuration was chosen to increase accuracy while reducing processing time. As with the previous models, the dataset was split into 80% training and 20% testing. Since KNN does not require a training phase, it performed classification directly on the test set by evaluating neighborhood relationships.

2.3.4. Hu Moments and MLP (Multi-Layer Perceptron) Hybrid

In this model, 7-dimensional feature vectors derived from Hu moments were classified using a classical artificial neural network: the Multi-Layer Perceptron (MLP). MLP is particularly preferred as an effective non-linear classifier in scenarios involving a small number of features. From each image, the 7 Hu moment values were extracted, log-transformed, and standardized. This ensured geometric consistency, resulting in features invariant to scale, position, and rotation. These feature vectors were directly used as input for the MLP model. The structure of the MLP used is as follows:

- Input Layer: 7 neurons (corresponding to the number of Hu moments)
- Hidden Layer 1: 64 neurons, ReLU activation
- Hidden Layer 2: 32 neurons, ReLU activation
- Output Layer: 2 neurons, softmax activation (for binary classification)

During training, the categorical_crossentropy loss function and the Adam optimizer were used. The model was trained for 12 epochs with a batch size of 32. The dataset was split into 80% for training and 20% for testing. The model was trained with validation on the training set, and no early stopping was applied, although the risk of overfitting was taken into consideration.

2.3.5. CNN Only

This model was developed for direct image-based classification and was trained solely on raw image data without the use of any handcrafted feature extraction (e.g., Hu moments). This approach leverages the advantages of deep learning-based models, such as automatic feature extraction and high generalization capability.
A basic CNN architecture was used, structured as follows:

- Input Layer: RGB images with dimensions (64, 64, 3)
- 1st Convolutional Layer: 32 filters, 3x3 kernel, ReLU activation
- 1st MaxPooling Layer: 2x2 pool size
- 2nd Convolutional Layer: 64 filters, 3x3 kernel, ReLU activation
- 2nd MaxPooling Layer: 2x2 pool size
- Flatten Layer
- Fully Connected Hidden Layer: 64 neurons, ReLU activation
- Output Layer: 2 neurons, softmax activation (for clean/defect classes)

The model was trained using the Adam optimizer and the categorical_crossentropy loss function. The training and testing procedure included:
Splitting the dataset into 80% training and 20% testing
Training over 12 epochs with a batch size of 32
Monitoring performance using validation data
No early stopping strategy was applied, though overfitting was monitored during training.

3. Experimental Results

3.1. Comparison of Classification Performances

In this study, the classification performances of five different models (HU+CNN, HU+SVM, HU+KNN, HU+MLP, and CNN-only) were evaluated using fundamental performance metrics such as accuracy, precision, recall, and F1-score. The performance of each model was analyzed on a per-class basis for the two categories: clean and defect. The numerical results obtained from the experiments are presented in Table 1.

Table 1. Comparison of Accuracy, Precision, Recall, and F1-Score (Per Class) Across Models

Model	Class	Precision	Recall	F1-score	Accuracy
HU+CNN	clean	0.99	1.00	0.99	0.9965
HU+CNN	defect	1.00	0.98	0.99	0.9965
HU+SVM	clean	0.67	0.64	0.66	0.65
HU+SVM	defect	0.63	0.67	0.65	0.65
HU+KNN	clean	0.90	0.95	0.92	0.92
HU+KNN	defect	0.94	0.89	0.91	0.92
HU+MLP	clean	0.72	0.47	0.57	0.63
HU+MLP	defect	0.59	0.80	0.68	0.63
CNN	clean	0.93	0.91	0.92	0.9255
CNN	defect	0.90	0.93	0.91	0.9255

The HU+CNN model achieved the highest overall performance, with F1-scores exceeding 99% for both clean and defect fabric classes. The model obtained 99% precision and 100% recall for the clean class, and 100% precision and 98% recall for the defect class, indicating extremely low false positive and false negative error rates. The overall accuracy reached 99.65%. These results clearly demonstrate that integrating Hu moment-based features into a CNN architecture enables strong learning performance.

The CNN-only model ranked second, achieving an accuracy of 92.55%. For this model, F1-scores ranged between 91% and 92% for both classes, indicating balanced classification performance. These results confirm CNN's strong learning capability directly from raw image data and its ability to perform effectively even without handcrafted features like Hu moments.

Among classical machine learning methods, the HU+KNN model showed remarkable performance with an accuracy of 92%. This model achieved 90% precision and 95% recall for the clean class, and 94% precision and 89% recall for the defect class. These findings indicate that Hu moments are also highly effective in pattern-based algorithms.

The HU+SVM model reached an overall accuracy of 65%, with F1-scores of 66% and 65% for the clean and defect classes, respectively. The relatively lower performance of this model suggests that the SVM algorithm may have limited generalization capability on features derived from Hu moments.

Similarly, the HU+MLP model recorded one of the lowest performances, with an accuracy of 63%. It produced an F1-score of 57% for the clean class and 68% for the defect class. The particularly low recall for the clean class indicates that the model frequently misclassified clean fabrics as defective.

Overall, these findings demonstrate that the combination of Hu moments and CNN architecture yields significantly superior classification performance compared to both traditional and modern standalone approaches. While classical algorithms cannot learn complex patterns as effectively as CNNs, they can still achieve moderate success when supported by well-structured features like Hu moments. This confirms that Hu moments are a suitable and flexible tool for feature extraction, although classification performance largely depends on the choice of model architecture.

3.2. Evaluation with Confusion Matrices

In addition to the numerical performances of the classification models, confusion matrices were visually analyzed for each model to better interpret the accuracy of their decision-making processes. These visualizations clearly show

how accurately or inaccurately each model predicted the respective classes. Figure 1 presents the confusion matrices for the HU+CNN, HU+SVM, HU+KNN, HU+MLP, and CNN-only models.

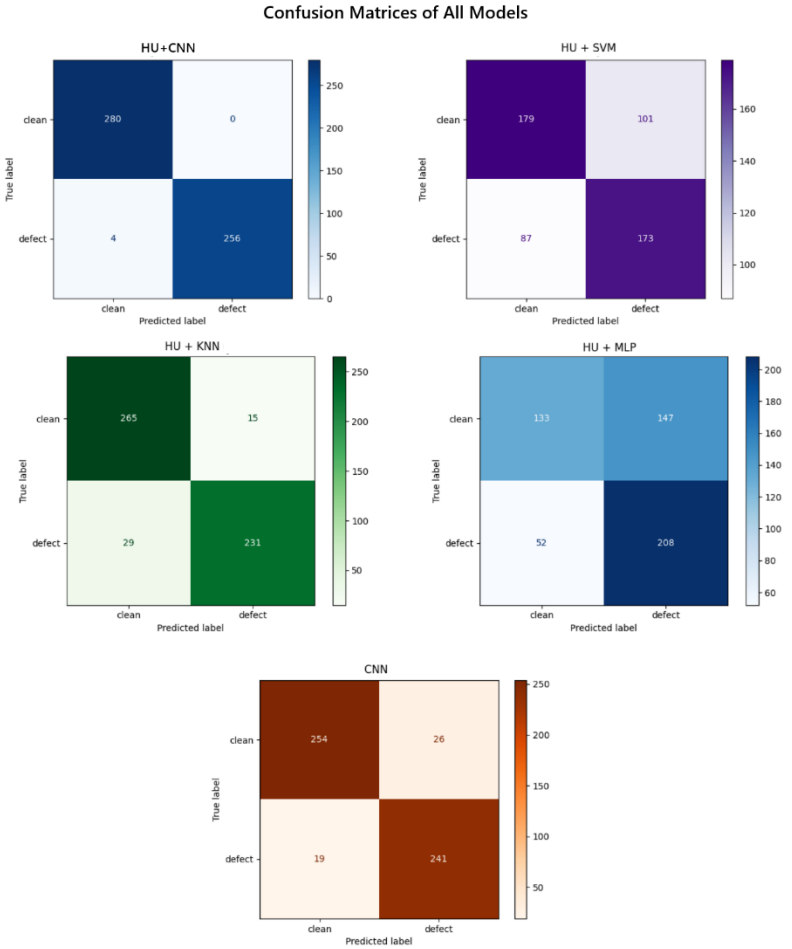


Figure 1. Visual comparison of the confusion matrices for the HU+CNN, HU+SVM, HU+KNN, HU+MLP, and CNN models

The confusion matrix of the HU+CNN model shows that the model achieved highly accurate predictions for both the clean and defect classes. In particular, the extremely low number of misclassifications indicates that the model operates with balanced and high accuracy.

The HU+KNN model also demonstrated strong performance; however, 29 defect samples were misclassified as clean. This suggests that while KNN performed well overall, it lacked the same level of discriminative power as CNN in accurately identifying defective samples.

The HU+SVM and HU+MLP models produced more errors according to their respective confusion matrices. The HU+SVM model exhibited a high number of cross-class predictions between clean and defect classes, especially with a high false positive rate for the clean class. Similarly, the HU+MLP model showed a notable number of misclassifications in the clean class, indicating that the MLP architecture was not sufficiently effective for this dataset.

The CNN-only model showed a relatively balanced distribution. Low error rates in both classes support the conclusion that CNN effectively learned features from raw image data. However, the higher number of misclassifications compared to the HU+CNN model suggests that integrating Hu moments improves class boundary discrimination.

When the visual analysis results are considered alongside numerical metrics, it becomes evident that the integration of Hu moments with CNN improves both the accuracy and stability of model performance. Furthermore, the comparison with classical machine learning models reinforces that deep learning produces significantly more effective results, especially when handling complex patterns.

3.3. Evaluation of Misclassified Samples

To comprehensively assess model performance, it is important to consider not only metric values but also the model's predictions on actual test data through visual inspection. In this context, Figure 2 presents a selection of sample classifications made by the HU+CNN model on the test dataset.

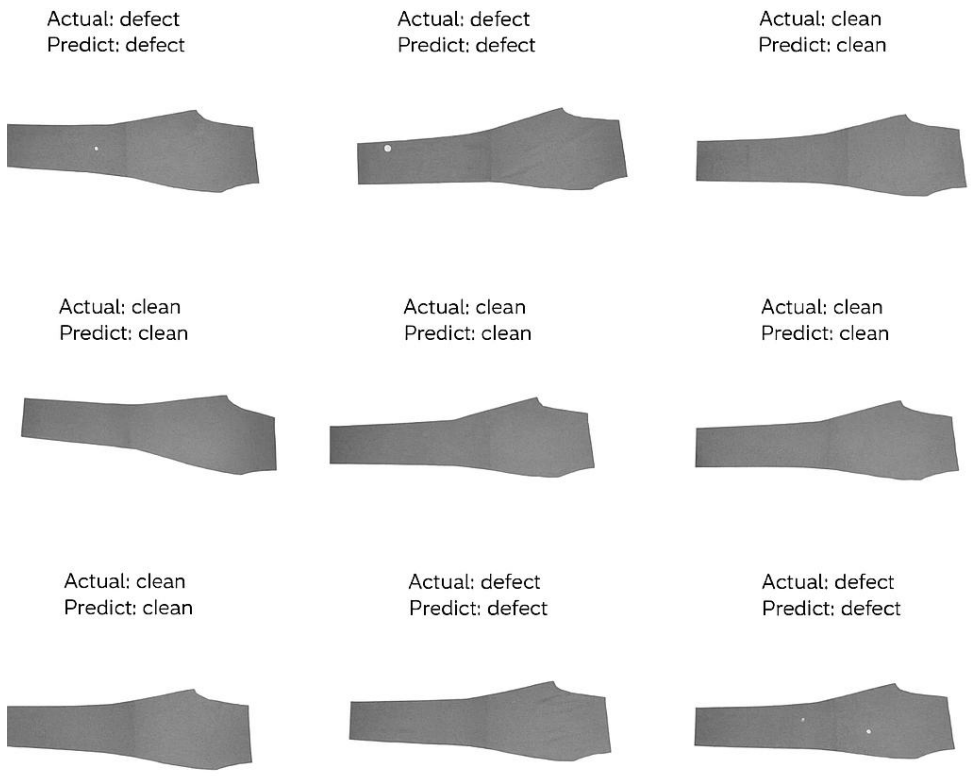


Figure 2. Correctly classified test images by the HU+CNN model (with actual and predicted class labels)

The figure presents a total of nine sample images arranged in three rows and three columns. Below each fabric sample, the actual class and the predicted class label by the model are provided. As shown, the HU+CNN model made correct predictions for all of these samples. This accuracy applies to both the *clean* and *defect* classes, indicating that the model performs consistently well across both categories.

The model's ability to classify samples without error—even in cases with subtle defects or visually similar fabric patterns—further supports the contribution of Hu moments to feature extraction. This success is attributed to the combination of the CNN's automatic pattern learning capacity with the geometric stability offered by Hu moments. Such examples demonstrate the model's strong generalization capability despite complex real-world variations and reinforce the practical applicability of the proposed hybrid architecture.

3.4. Visual Validation Using Labeled Defects

Beyond classification accuracy, it is crucial to examine whether the correctly predicted *defect* class images visually align with the manually labeled defect regions. This step is essential for assessing the practical reliability of the

system. For this purpose, defective images manually annotated using the LabelImg tool in YOLO format were evaluated alongside the model's prediction results. The annotation process was conducted using bounding boxes that specify the position and size of each defect region. When these annotations were compared with the test images classified as *defect* by the HU+CNN model, it was observed that the majority of correct classifications corresponded to actual labeled defect areas. Figure 3 displays a sample test image, showing the model's prediction and the bounding boxes generated using LabelImg side by side. The visual validation results indicate that the model not only makes correct class predictions but also learns patterns that accurately represent the spatial location of the defects.

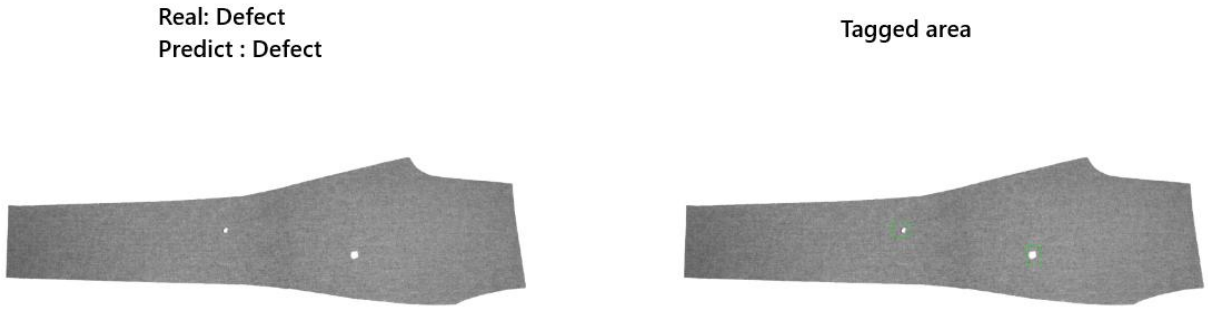


Figure 3. Comparative visualization of a sample test image correctly classified by the HU+CNN model alongside its labeled defect regions

4. Discussion

This study has demonstrated that a robust and generalizable model for textile defect detection can be achieved by combining features extracted from Hu moments with deep learning architectures. In particular, the HU+CNN hybrid model outperformed all other approaches with a notably high accuracy of 99.65%. This result highlights the successful integration of the geometric stability of Hu moments in symmetric structures with the visual learning capabilities of CNNs. Comparisons with classical machine learning algorithms revealed that models such as HU+SVM and HU+MLP performed relatively poorly, with accuracy rates of 65% and 63%, respectively. This suggests that while Hu moments alone may not adequately represent certain complex patterns, their limitations can be overcome when combined with the learning capacity of deep neural networks. Furthermore, the CNN-only model achieved an accuracy of 92.55%, clearly indicating that the integration of Hu moments provides a meaningful enhancement. This contribution becomes particularly significant in scenarios where the amount of training data is limited. An examination of correctly classified samples and confusion matrices shows that the system reliably distinguishes between clean and defective fabrics. In the visual validation phase, the successful identification of annotated defects further supports the practical applicability of the system. However, some limitations of the study should also be acknowledged. The dataset used was created under controlled and fixed conditions; thus, variability in real-world production environments (e.g., different lighting conditions, fabric positions) could present additional challenges. Moreover, the model's performance in distinguishing among multiple defect types was not assessed in this study. For future work, it is essential to test the model on larger datasets enriched with various types of textile defects to improve generalizability. Additionally, optimization efforts should be directed toward integrating the model into mobile or embedded systems for real-time inspection on production lines.

Acknowledgements

The authors would like to thank Serkon Technology for providing the dataset used in this study. Their support contributed significantly to the experimental validation of the proposed model.

5. References

[1] Patil, V., Hawaldar, S., Narvekar, P., & Jain, S. (2020). Fabric Inspection Device Using Image Processing.

[2] León-Chavarri, C., Feril-Encarnacion, F., Noriega-Bardalez, F., & Camones-Caballero, V. (2024). *Lean Production Model to Reduce Defects and Achieve Sustainability in a Peruvian Textile SME*. *Advances in Science and Technology*, 149, 53–60. https://doi.org/10.4028/p-YEIn25

[3] Rasheed, A., Sajid, M., Dar, S., Rasheed, A., Mahmood, M., Habib, U., Shehryar, T., Ali, N., & Zafar, B. (2020). *Fabric Defect Detection Using Computer Vision Techniques: A Comprehensive Review*. *Mathematical Problems in Engineering*, 2020, 1–24. https://doi.org/10.1155/2020/8189403

[4] Yung, M., Wells, R., Neumann, W., & Kolus, A. (2019). *Examining the fatigue-quality relationship in manufacturing*. *Applied Ergonomics*, 82, 102919. https://doi.org/10.1016/j.apergo.2019.102919

[5] Sharew, S., Chaka, K., & Shiferaw, A. (2023). *Inspection of Cotton Woven Fabrics Produced by Ethiopian Textile Factories Through a Real-Time Vision-Based System*. *Journal of Natural Fibers*, 20. https://doi.org/10.1080/15440478.2023.2286615

[6] Senarathne, R., & Thennakoon, R. (2025). *Fabric Inspection Using Computer Vision*. *International Journal of Media and Networks*. https://doi.org/10.33140/ijmn.03.01.02

[7] Gupta, S., Dey, B., Pandey, P., Majumder, A., & Kaur, G. (2024). *A Smart Manufacturing Process for Textile Industry Automation under Uncertainties*. *Processes*, 12(4), 778. https://doi.org/10.3390/pr12040778

[8] Anjum, S., Angad, A., Kulkarni, P., & Gaddi, A. (2025). *Advancements in Computer Vision and Deep Learning for Enhanced Textile Inspection and Manufacturing*. Journal of the Textile Association. https://doi.org/10.63665/jta.v85i6.13

[9] Xiang, Z., & Huang, Y. (2022). *RPDNet: Automatic Fabric Defect Detection Based on a Convolutional Neural Network and Repeated Pattern Analysis*. Sensors (Basel, Switzerland), 22. https://doi.org/10.3390/s22166226

[10] Mohammed, S. S., & Clarke, H. G. (2024, October). Advanced Convolutional Neural Network Approach for Fabric Defect Detection. In 2024 Innovations in Intelligent Systems and Applications Conference (ASYU) (pp. 1-5). IEEE.

[11] Koulali, I., & Eskil, M. (2021). *Unsupervised textile defect detection using convolutional neural networks*. Appl. Soft Comput., 113, 107913. https://doi.org/10.1016/j.asoc.2021.107913

[12] Liu, J., Wang, C., Su, H., Tao, D., & Du, B. (2019). *Multistage GAN for Fabric Defect Detection*. IEEE Transactions on Image Processing, 29, 3388–3400. https://doi.org/10.1109/TIP.2019.2959741

[13] Kwon, W., Kim, H., Kang, J., & Jung, W. (2025). *StitchingNet and deep transfer learning method for sewing stitch defect detection*. J. Comput. Des. Eng., 12, 140–154. https://doi.org/10.1093/jcde/qwaf037

[14] Hu, F., Gong, J., Liu, W., & Fu, H. (2023). *Fabric Defect Detection Method Using SA-Pix2pix Network and Transfer Learning*. Applied Sciences. https://doi.org/10.3390/app14010041

[15] Jeevaraj, P., & Archana, R. (2024). *Deep learning models for digital image processing: a review*. Artificial Intelligence Review, 57, 1–33.https://doi.org/10.1007/s10462-023-10631-z

[16] Wang, C., Li, Y., Li, J., Liu, Q., & Gao, M. (2023). *A Fabric Defect Detection Method Based on Deep Learning*. IEEE Access, 10, 4284–4296. https://doi.org/10.1109/ACCESS.2021.3140118

[17] Seçkin, M., Bogrekci, I., Ozek, A., & Demircioğlu, P. (2025). *Artificial Intelligence Driving Innovation in Textile Defect Detection*. Textiles. https://doi.org/10.3390/textiles5020012

[18] Hu, M. K. (1962). *Visual pattern recognition by moment invariants*. IRE Transactions on Information Theory, 8(2), 179–187. https://doi.org/10.1109/TIT.1962.1057692

[19] Shi, Y., Zhang, X., & Zhang, D. (2016). *Invariant moments for image classification: A review*. Pattern Recognition Letters, 83, 115–123. https://doi.org/10.1016/j.patrec.2016.01.024

Experimental and Artificial Intelligence Methods for Investigating the Effects of Cell Temperature on Panel Performance in Photovoltaic Solar Panels

Gizem Gül KATIRCIOĞLU[1], Mehmet DAŞ [2], Ebru AKPINAR [3]

[1,2,3]Mechanıcal Engıneerıng, Faculty of Engıneerıng, Fırat University, Elazığ, Türkiye
[1]ggkatircioglu@firat.edu.tr, [2]m.das@firat.edu.tr, [3]ebruakpinar@firat.edu.tr
[1](ORCID: 0000-0003-1015-9392), [2](ORCID: 0000-0002-4143-9226),[3](ORCID: 0000-0003-0666-9189)

Abstract
Nowadays, interest in renewable energy is increasing significantly. The main reasons for this are the damage caused by fossil fuels to the environment, growing concerns about energy security, and the need to combat climate change. Renewable energy sources are derived from natural and inexhaustible resources such as solar, wind, hydroelectric, and biomass. It is well-known that these sources cause less harm to the environment by reducing greenhouse gas emissions. Additionally, the decreasing production costs of renewable energy systems over time and the advancement of technological innovations are enabling these sources to be utilized on a larger scale. Solar energy is one of the most preferred renewable energy sources, and electricity can be generated through solar panels. Several factors (panel tilt angle, sunshine duration, system capacity, etc.) negatively affect the efficiency of solar panels. The effects of temperature increases in solar panels on energy production performance have been revealed in studies conducted to date. In the experiments conducted in this study, the effects of cell temperatures on panel performance were investigated for two photovoltaic (PV) solar panels under Elazığ climate conditions. One of the panels was cooled with a radial fan, and the effects of the temperature decrease on PV panel electricity production were observed. Simultaneous measurements were taken from the panels, and the results were compared. The results revealed that the efficiency of the solar panel cooled with a fan was higher. Subsequently, using the experimental data obtained, the cell temperature values of the panels were modeled using the Pace regression algorithm, an artificial intelligence method. The results obtained were compared with the experimental results. The error rate of the data obtained using artificial intelligence was calculated as 1.5%. The experimental and artificial intelligence studies conducted in the experiments provided useful results regarding how the performance of solar panels changes with temperature.

Keywords: solar energy, flat-plate solar panel, photovoltaic, panel temperature, artificial intelligence

1. Introduction

Energy is an important need in many sectors, from automotive to food, and therefore plays a major role in the development of societies. The development of societies directly increases the need for energy [1]. The change in energy production over the years is shown in Figure 1.

Energy can be obtained from fossil fuels or renewable energy sources, but in recent years, the depletion of fossil fuels has increased the shift towards renewable energy. The use of fossil fuels results in greenhouse gas emissions, which contribute to climate change [2]. 65% of greenhouse gas emissions come from the energy sector. Replacing fossil fuels with renewable energy sources could reduce this percentage to 35% [3]. Renewable energy sources are divided into five categories: solar, wind, geothermal, hydroelectric, and biomass energy. The most widely used of these sources is solar energy. Turkey's location is highly favorable for solar energy. Turkey's photovoltaic energy potential map is shown in Figure 2.

[1]Corresponding author

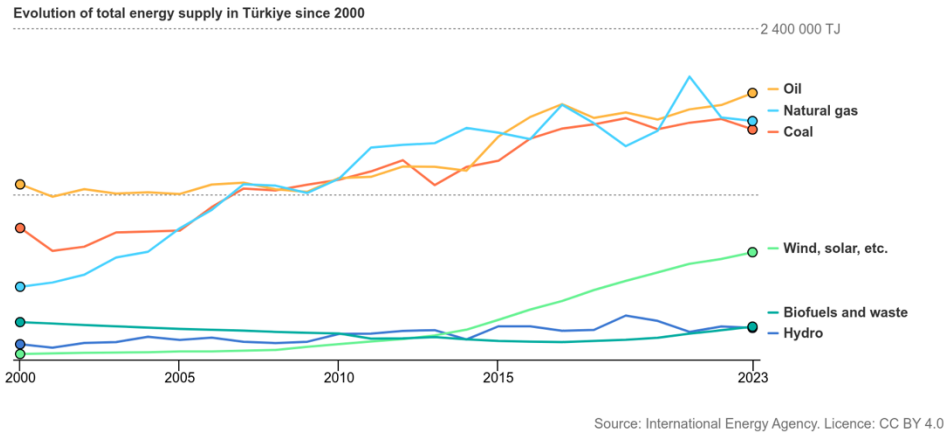


Figure 1. Changes in energy prodfuction in turkey over the years [4]

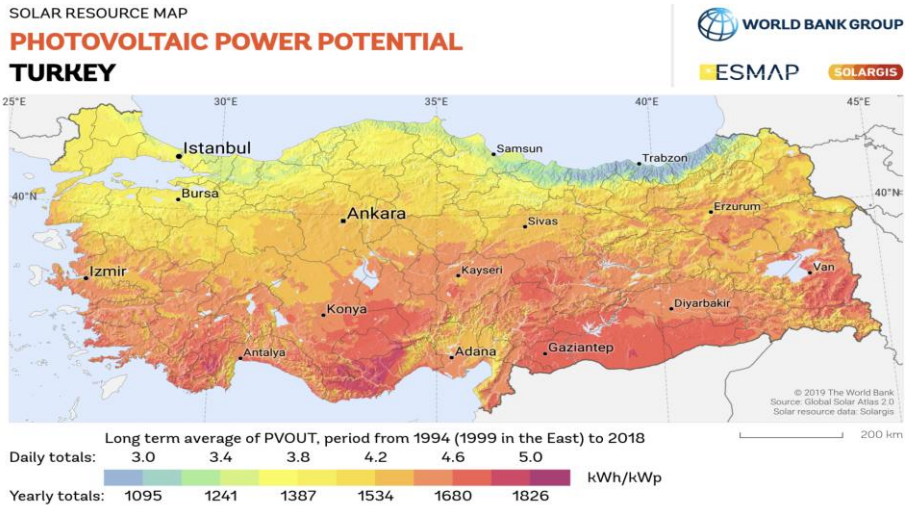


Figure 2. Solar energy photovoltaıc power potential map of turkey [5]

According to the map given in Figure 2, the solar radiation value increases towards the south. Elazığ province, where the experiments were conducted, is rich in solar energy potential, and the solar energy potential of Elazığ province is given in Figure 3.

The conversion of solar energy into electrical energy is achieved through solar panels. Solar panels are divided into four types: polycrystalline, monocrystalline, PERC, and thin-film panels. Solar panels directly convert sunlight into electrical energy and are semiconductors. The rays from the sun strike the panel surface directly and excite the photons within the solar panels. As a result of this process, electrical energy is produced [7]. There are many factors that affect the efficiency of solar panels. First, the amount of solar radiation plays a decisive role in the panel's electricity production; as the intensity and duration of the radiation increase, so does the efficiency. The surface temperature of the panel also affects efficiency; as the temperature rises, the conductivity of the semiconductor material decreases, causing efficiency to drop [8]. In addition, the angle and direction of the panel are important in terms of the perpendicularity of the incoming rays; deviations from the optimum angle lead to efficiency losses. External factors such as dust, dirt, and shading also reduce the amount of light reaching the panel surface, thereby decreasing performance [9]. Furthermore, the type and technology of the panel (monocrystalline, polycrystalline, thin film, etc.) also directly affect efficiency. Finally, structural degradation such as cell aging and microcracks that form over time can also reduce efficiency in the long term. In this study, the negative effects of temperature on solar panels were investigated, and there are many valuable studies in the literature on this subject.

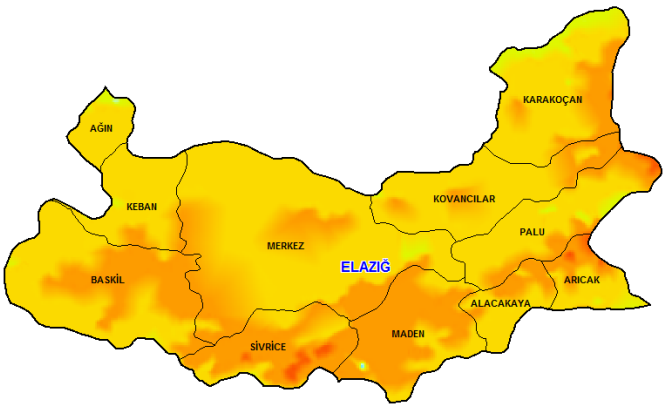


Figure 3. Elazığ province solar energy potential map [6]

Güven conducted experiments in Denizli and took temperature measurements at regular intervals from the panel surface. The measurements revealed that an increase in panel surface temperature had a negative effect on the panel's output power and efficiency. The decrease in efficiency was found to be 0.3% for every 5 °C increase [10].

In experiments conducted by Nabil et al. in Egypt, PV output power decreased by 0.42% for every degree increase in temperature, according to measurements taken from the panel surface at minute intervals [11].

In a study conducted by Saral et al., design changes that could be made at the cell and panel levels to achieve maximum efficiency in photovoltaic systems were evaluated in detail. Focusing particularly on maximum power point tracking (MPPT) technology, the study examined the effect of panel temperature on performance; the results showed that a 10 °C increase in temperature reduced the power output by approximately 21.15%. In this context, it was emphasized that keeping parameters such as the NOCT (Nominal Operating Cell Temperature) value and the temperature-dependent power change coefficient (μp) low plays an important role in reducing temperature-related performance losses [12].

In another study conducted by Dida and colleagues, the effect of PV cell temperature on efficiency was experimentally analyzed. According to the findings, every 1°C increase in temperature reduces the panel's electrical efficiency by an average of 0.5%. Therefore, a passive cooling system was developed to prevent thermal degradation and improve performance, and it was tested under hot climate conditions. Experimental data showed that the PV module temperature could be reduced to 20 °C using the vaporization cooling method, resulting in a 14.75% improvement in electrical efficiency [13].

Şhalaby and his team conducted an experimental study investigating the effect of cooling the rear surface of PV panels on performance. The low-cost cooling method proposed in this study also enables the preheating of feedwater for reverse osmosis systems. The experiments revealed a 14.1% increase in power generation for the cooled panels, with efficiency values rising from 17.4% to 19.8% [14].

Similarly, in the study conducted by Kabul and colleagues, active cooling was achieved by circulating water through pipes placed on the back of PV panels, and the contribution of this system to both electricity production and thermal energy gain was analyzed. It was stated that the cooled panels provided a 35% increase in electricity production compared to those without cooling, and a 7% improvement in overall efficiency. The results obtained clearly demonstrate that controlling panel temperature is critical for photovoltaic system performance [15].

With the continuous advancement of technology, researchers are conducting studies to optimize the efficiency of solar panels using artificial intelligence methods.

In the research conducted by Akal and his colleagues, the aim was to make power predictions using environmental data such as temperature, humidity, pressure, and solar radiation obtained from photovoltaic systems. In this context, two different machine learning methods were compared: feedforward artificial neural networks (ANN) and the nearest neighbor algorithm (KNN). The performance of models trained with real-time field data was evaluated, and it was found that artificial neural networks produced more successful results with an accuracy rate of 98.79%. The study emphasizes that the developed models not only have the capacity to make highly accurate predictions but also have applicability to energy needs in different geographical regions [16].

In the study by Kayabaşı and colleagues, the relationship between the output power efficiency of photovoltaic panels and meteorological conditions was investigated. For this purpose, an air measurement system was set up to record factors such as humidity, temperature, wind speed, radiation, and light level, while the voltage and current data of the solar panel were measured simultaneously. The collected data were analyzed using the Artificial Neural Networks (ANN) method. In the ANN model, environmental parameters were used as input data, and panel output power was used as output data. The accuracy of the model was tested, and the total error rate was determined to be 1.57%. This study clearly demonstrates the effects of weather conditions on solar panel performance and provides recommendations for improving system efficiency [17].

In the study conducted by Yüzer and his team in Hakkâri province, solar radiation was estimated using meteorological data specific to the region. Solar radiation forecasts were created using an artificial neural network model, and the results were integrated with a photovoltaic system model developed in the Matlab/Simulink environment, which includes an incremental conductivity-based MPPT algorithm. When the actual measurements were compared with the predictions obtained from the ANN model, the R^2 value for solar radiation estimation was calculated as 0.9705, and for PV system output power estimation, it was 0.9668. These findings indicate that the model can make accurate predictions and can be used in the performance analysis of photovoltaic systems [18].

The study conducted by Hudișteanu et al. tested monocrystalline and polycrystalline solar panels in a dual climate chamber environment to analyze the effect of temperature changes on efficiency, voltage, current, and fill factor. Under 560–520 W/m² irradiance, as panel temperature increased, efficiency decreased by approximately –0.52%/°C, voltage decreased by –0.48%/°C, while current increased by 0.10%/°C. The total efficiency loss was determined to be approximately –27% when the temperature increased from 25°C to 75°C. It was clearly demonstrated that cooling improves efficiency[19].

Karaca and colleagues experimentally investigated the effects of surface temperature on photovoltaic efficiency in PV-T hybrid systems and reported that, without cooling, surface temperature rose to 75°C; with active cooling, this value dropped to 33°C. This temperature reduction resulted in an increase in panel efficiency from 6.10% to 9.31%, representing an improvement of approximately 52.6%. This finding demonstrates that active cooling methods can significantly enhance PV system performance [20].

Similarly, Hassan and colleagues analyzed the I–V characteristics of crystalline silicon solar cells at temperatures ranging from 300 to 380 K and found that an increase in temperature led to a decrease in open-circuit voltage, maximum power point, fill factor, and overall efficiency. In the study, efficiency decreased from 10.1% to 9.7% with increasing temperature, quantitatively demonstrating the negative effect of temperature on efficiency [21].

Shaker, in a review study evaluating the effects of temperature on PV cell efficiency, stated that an increase in temperature reduces system performance due to carrier recombination and increased resistance. Additionally, it was noted that open-circuit voltage and fill factor decrease with increasing temperature, while a slight increase in short-circuit current was observed [22].

In another study, Öztürk evaluated passive and active cooling methods applied in the thermal management of PV systems. It was noted that applications such as natural or forced air/liquid cooling systems, heat pipes, and phase change materials can increase system efficiency by reducing cell temperature. Additionally, it was emphasized that the temperature coefficient of silicon-based panels is approximately –0.4%/°C, highlighting the importance of temperature control for long-term performance and panel lifespan [23].

Studies in the literature have clearly demonstrated the negative effects of temperature on solar panel efficiency. The use of artificial intelligence applications to optimize renewable energy sources is quite common today. Looking at these studies, the prediction power of the systems is quite high, as shown in the examples above.

In our study conducted in Elazığ, measurements were taken at regular intervals from a monocrystalline PV panel, and the effect of temperature on the efficiency of the PV panel was investigated. Additionally, the energy performance effect caused by temperature was modeled using the Pace regression method, depending on environmental temperature, solar radiation, wind speed, and temperature parameters. After obtaining the models, mathematical equations expressing the energy values that PV panels will produce were obtained, and the results of artificial intelligence models were compared with experimental results.

2. Material and Methods

The experiments were conducted under Elazığ climate conditions and lasted for 1 day between 09:00 and 17:00. One monocrystalline PV panel was used in the experiments. Temperature measurements were taken from the panel surface at half-hour intervals, and the panel's output power was measured. Figure 4 shows the image of the PV panel used in A, while B and C show the images of the surface temperature measurement and solar radiation measurement, respectively.

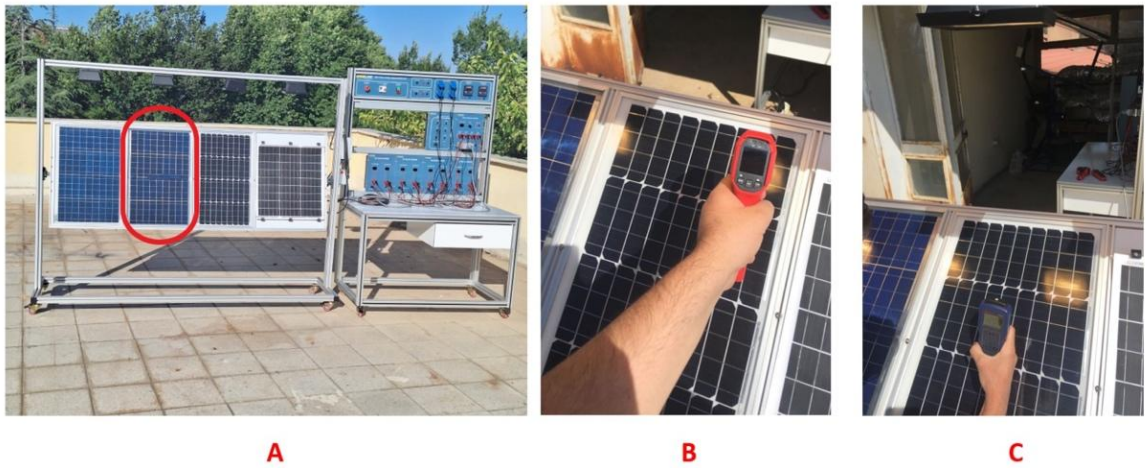


Figure 4. PV panel (A), Surface temperature measurement (B), Solar radiation measurement (C)

The experimental data obtained were then used to model the cell temperature values of the panels using the Pace regression algorithm, an artificial intelligence method. The results obtained were compared with the experimental results.

The characteristics of the PV panel used in the experiments are given in Table 1.

Table 1. PV Panel Specifications

System Component	Specifications
Monocrystalline Solar Panel System	50 Watt Monocrystalline Solar Panel 430x670x30mm, maximum voltage 20.7V, MPPT current 2.42A

The measuring instruments used in the experiments are given in Table 2.

Table 2. Characteristics of measuring elements used in the experiments

Parameter	Measuring Device	Uncertainty Value
Temperature	PT 100 precision measurement sensor	±0.1°C
Humidity	Thermo-hygrometer	±%1
Air speed	Digital anemometer	±0.1 m/s
Temperature distribution	IR thermal measurement device (-20 °C - 100 °C)	±1 °C
Solar radiation	Pyrometer	±0.1 Wm^{-2}
Electrical power	Wattmeter	±0.1 Watt

Calculation Procedure

The output power of PV modules is required to measure the effect of cell temperature on the PV (photovoltaic) system. The output power of a PV module depends on solar radiation and ambient temperature and has been calculated using Equation (1) [24,27].

$$P = P_0 \frac{G}{G_0}[1 + \gamma(T - T_O)] \quad (1)$$

P: Output power of the PV module (W)
P $_0$: Output power of the PV module (W).
G: Solar radiation (W/m2).
G $_0$: Solar radiation (i.e., 1000 W/m2).
γ: Power temperature coefficient (% °C $^{-1}$)
The value of T in Equation (1) was calculated using Equation (2).

$$T = T_a + \frac{NOCT - 20}{800} G \quad (2)$$

T $_a$: ambient temperature (°C)
NOCT: Nominal Operating Cell Temperature.
T $_0$: PV temperature (i.e., 25 °C)
The energy efficiency of the PV panel was calculated using Equation (3) [15].

$$\eta = \frac{V \times I}{G \times A} \times 100\% \quad (3)$$

η = PV panel efficiency (%)
V = Voltage (V)
I = Current (A)
G = Solar irradiance (W/m²)
A = Panel area (m²)

Computational Intelligence Methods

Machine learning is the most widely used computational intelligence method. Machine learning is the general name for computer algorithms that can learn structurally and make meaningful predictions based on data. These algorithms basically work by creating models from sample data. In today's world, it is not feasible to manually process and

analyze large amounts of data. Therefore, the goal is to solve a given problem by training machine learning algorithms with data obtained from the problem's environment to reach a solution [25].

PACE Regression

Among statistical analyses, regression analysis is the most widely used data analysis method. Regression analysis is a classic artificial intelligence method used to create classification and prediction-based models.

Wang and Witten developed the Projection Adjustment by Contribution Estimation (PACE) regression algorithm and achieved more accurate models for datasets containing a large amount of data [26]. PACE regression, one of the regression algorithms, is similar to the classical regression method and is an algorithm that classifies or predicts datasets according to linear models [29].
The classical ordinary least squares (OLS) method, widely used in regression analysis, cannot produce sufficiently accurate models because it does not consider the effect of each data point. PACE regression produces more precise models than classical regression because it examines the effects of each variable in the dataset on classification or prediction in detail [30].

In PACE regression, the OLS and empirical Bayes (EB) methods were developed to examine the effects of data in the dataset on model accuracy in detail. Equation 4 is used to express these methods [30].

$$\text{Modeled Attribute} = \alpha1*Ai+\alpha2*Bi+\alpha3*Ci+\alpha4*Di+\alpha5*Ei+\ldots \quad (4)$$

Equations generated using PACE are generally simple and linear equations. Symbols such as A, B, C, etc. represent data other than the parameter modeled in the data set. Additionally, the αi symbol in Equation 8 indicates the weight coefficients of the parameters in the data set on the model. This results in a short model creation time and a low model error rate [30].

Table 3. Input and output parameters for regression models

Input Parameters		
Parameters	**Unıt**	**Symbol**
Ambient temperature	°C	T
Air velocity	m/s	V
Humidity value	%	H
Solar radiation value	W/m^2	I
Voltage	V	Volt
Current	I	C
Output Parameters		
Parametre	**Unıt**	**Symbol**
Panel Output Temperature (T)	°C	T_p

To determine the accuracy between the panel temperature and humidity (T-H) values obtained with regression algorithms and the experimental values, a mean absolute error (MAE) analysis was performed. The MAE analysis is given in Equation 5 below [28].

$$MAE=\left(\left|\frac{T\text{-}H_{experimental}-T\text{-}H_{predictive}}{T\text{-}H_{experimental}}\right|\right) \quad (5)$$

3. Results

The graphs created based on the data obtained from the experiments are shown in Figures 7, 8, and 9.

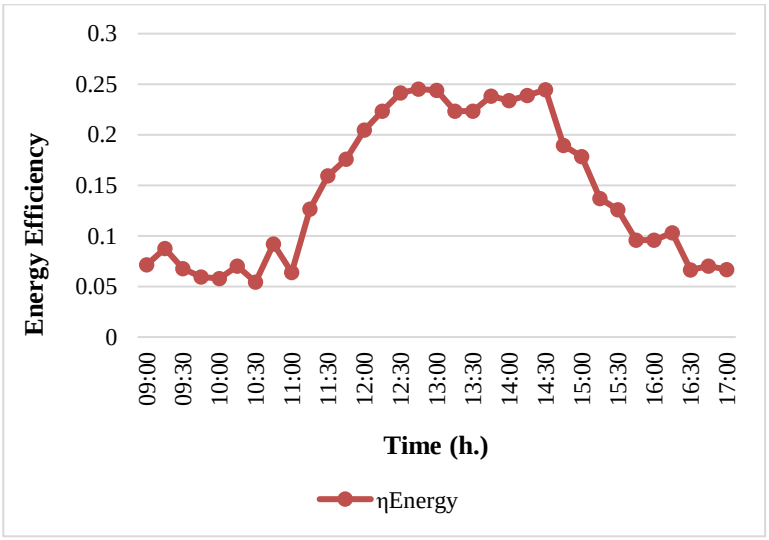


Figure 5. Energy efficiency graph

The energy efficiency (ηEnergy) of the PV panel are compared according to the time of day in Figure 5. Efficiencies started at low levels in the morning and increased towards noon. The highest energy efficiency was approximately 27% between 12:45 and 14:30, after which it began to decline.

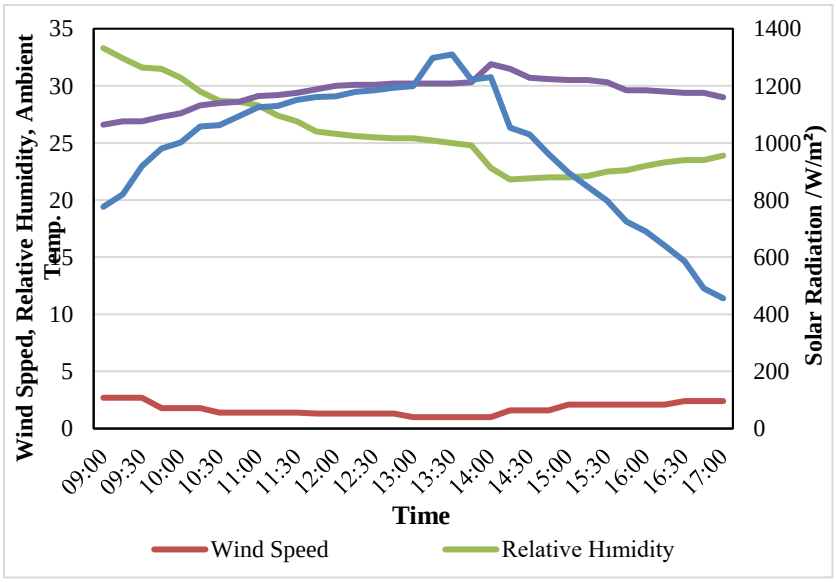


Figure 6. Wınd speed, relative humidity and ambient temperature graph

Figure 6 shows the distribution of variables measured throughout the day, revealing the environmental conditions that directly affect PV panel performance. The graph shows that the ambient temperature rose steadily from the morning hours, reaching approximately 30 °C at noon, then gradually decreased as the day progressed. In contrast, relative humidity was observed to be relatively high in the morning hours, decreasing with the increase in temperature, and then rising again in the afternoon. Solar radiation increased in the early hours of the day, reaching its maximum level (~1200 W/m²) around noon, and decreased again as the sun angle decreased. Wind speed remained low (~1–3 m/s) and largely constant throughout the day. These findings indicate that the operating conditions of the photovoltaic panel change significantly throughout the day and that high temperatures and radiation, especially around noon, increase electricity production but

can also cause a loss in efficiency by raising the panel temperature. Therefore, it is understood that cooling strategies that limit temperature increases should be considered in system design.

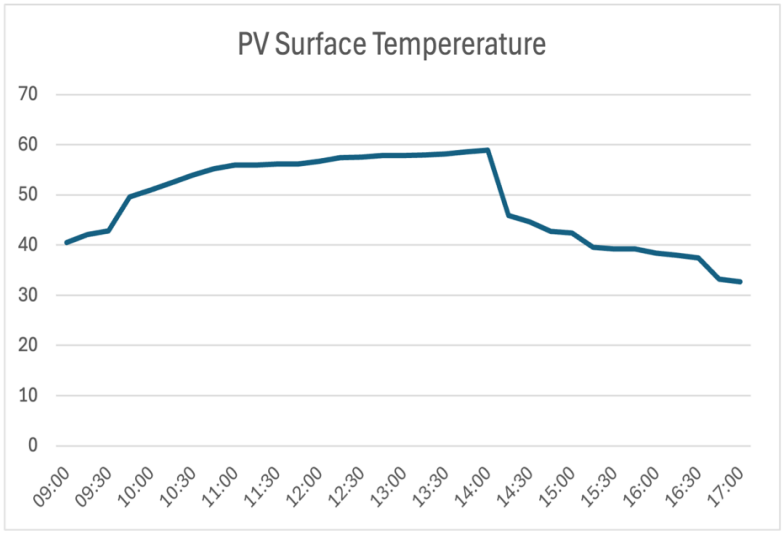


Figure 7. PV surface temperature graph

In Figure 7, the surface temperature of the PV panel started at around 40 °C in the morning and rose to 60–65 °C by noon. This temperature increase is expected during this time period, as it is when solar radiation is at its peak. However, after 2:30 PM, the surface temperature decreased significantly and dropped to approximately 35 °C by the end of the day. This increase in panel surface temperature is an important factor that negatively affects energy efficiency. Indeed, evaporation, increased internal resistance, and the temperature sensitivity of the semiconductor material contribute to the decrease in efficiency.

When the three graphs are evaluated together, it is observed that PV panel performance is directly related to environmental variables throughout the day. In particular, it is understood that high surface temperatures cause a significant decrease in exergy efficiency. When wind speed is low, the cooling effect weakens, causing the panel temperature to rise. The lower exergy efficiency compared to energy efficiency also highlights quality losses in the energy conversion processes within the system.

In this study, the surface temperature of photovoltaic (PV) panels and the corresponding electrical power generation were experimentally investigated, and the obtained datasets were analyzed using the PACE regression model. The experimental findings were compared with the predicted results from the PACE model to evaluate its performance. Figure 8 shows the comparison of the experimental and PACE regression model results for the electrical power production values of the PV Panel.

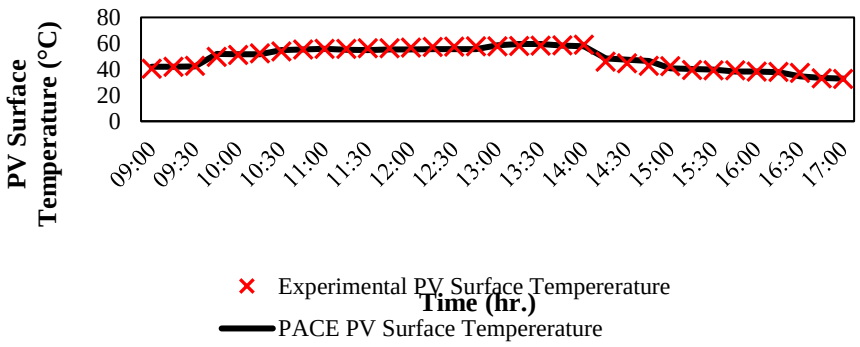


Figure 8. Comparison of experimental and pace regression model results for pv panel electrical power generation values

The time-dependent variation of PV surface temperature throughout the day showed a clear correlation between the experimental measurements and the model predictions. As illustrated in Figure 8, the PV temperature increased from the morning (09:00) until early afternoon (around 14:00) and then gradually decreased towards the evening. The comparison demonstrated a strong agreement between the experimental data (red markers) and the PACE regression estimations (black line).

Table 4. Error values for PACE models created for PV panel electrical power values

System Type with PACE Model Applied	MAE (%)
PV Panel	0.11627

Error analysis revealed that the model achieved a Mean Absolute Error (MAE) of **0.11627** indicating a very low deviation between the predicted and actual values. This result confirms that the PACE regression model can reliably estimate the electrical power production of PV panels based on environmental and operational parameters. The mathematical equation representing the electrical power values produced by the PV panel and PACE model is given in Table 5.

Table 5. Mathematical equation produced by PACE models for PV panel electrical power values

Mathematical equation expressing the electrical power values of PV panels
PV Surface Tempererature = 1.0815 * Ambient Temperature -8.617 + 0.0128 * Solar Radiation - 9.3753 * Wind Speed + 1.118 * Relative Hımidity

Table 5 shows the output of the PACE regression model developed to estimate the PV panel surface temperature. The model obtained quantitatively reveals the linear relationship between environmental parameters and PV panel surface temperature. According to the equation, the panel surface temperature is directly proportional to the ambient temperature and has the highest positive effect with a coefficient of 1.0815. This indicates that the ambient temperature is the fundamental determinant in increasing the panel temperature.

Although the coefficient of the solar radiation variable is quite low (0.0128), its positive effect contributes to the increase in panel surface temperature. A positive coefficient (1.118) was found for wind speed, which, contrary to theoretical expectations, was interpreted as increasing panel temperature. While wind is normally expected to cool the panel surface, this parameter has a positive effect due to the experimental conditions or the regression relationships of the model. The effect of relative humidity is negative (with a coefficient of -9.3753), and it is understood that it plays a role in reducing the panel surface temperature.

4. Conclusions

In this study, the surface temperature behavior of the PV panel was analyzed in relation to meteorological parameters including solar radiation, ambient temperature, wind speed, and relative humidity. The experimental results demonstrated that the PV surface temperature increased steadily with rising solar radiation and ambient temperature, reaching its peak during midday hours when solar radiation and ambient temperature were at their highest. A subsequent decrease in PV temperature was observed in the afternoon, corresponding to the reduction in solar irradiance.

The PACE regression model was successfully applied to estimate the PV surface temperature, and its predictions exhibited close agreement with the experimental data throughout the measurement period. The minor deviations observed between the modeled and measured values indicate that the PACE model effectively captures the thermal dynamics of PV systems under varying environmental conditions.

Additionally, the analysis confirmed that wind speed and relative humidity had secondary but notable effects on PV surface temperature. Low wind speeds throughout the day reduced convective cooling, contributing to higher surface temperatures, while variations in relative humidity appeared to slightly moderate the temperature profile.
Overall, the findings confirm that accurate modeling of PV surface temperature is achievable using PACE regression, providing a reliable tool for predicting thermal performance and, consequently, the electrical output of PV systems. This approach can aid in the design, optimization, and operational strategies of photovoltaic systems, particularly in regions with fluctuating environmental conditions.

5. References

[1] Antoniou *et al.*, "Effective thermal-electric control system for hydrogen production based on renewable solar energy," *Int. J. Hydrogen Energy*, vol. 53, no. April 2023, pp. 173–183, 2024.
[2] Dincer and M. I. Aydin, "New paradigms in sustainable energy systems with hydrogen," *Energy Convers. Manag.*, vol. 283, no. March, p. 116950, 2023.
[3] G. M. Lima, F. N. Belchior, J. E. N. Villena, J. L. Domingos, M. A. V. Freitas, and J. D. Hunt, "Hybrid electrical energy generation from hydropower, solar photovoltaic and hydrogen," *Int. J. Hydrogen Energy*, vol. 53, no. August 2023, pp. 602–612, 2024.
[4] Uluslararası Enerji Ajansı https://www.iea.org/countries/turkiye/energy-mix (08.09.2025)
[5] Global Solar Atlas https://globalsolaratlas.info/download/turkey
[6] Ulusal Solar Enerji https://www.ulassolarenerji.com/kovancilar-solar-enerji-sistemleri/
[7] S. U. Batgi and I. Dincer, "Design and analysis of a floating photovoltaic based integrated energy system with ammonia production," *Appl. Therm. Eng.*, vol. 252, no. February, p. 123676, 2024
[8] *Skoplaki, E., & Palyvos, J. A. (2009). On the temperature dependence of photovoltaic module electrical performance: A review of efficiency/power correlations. Solar Energy, 83(5), 614-624. https://doi.org/10.1016/j.solener.2008.10.008.*
[9] Mani, M., & Pillai, R. (2010). *Impact of dust on solar photovoltaic (PV) performance: Research status, challenges and recommendations*. Renewable and Sustainable Energy Reviews, 14(9), 3124–3131. https://doi.org/10.1016/j.rser.2010.07.065
[10] Güven, Ş. (2022). Fotovoltaik Panel Yüzey Sıcaklığının Denizli İli için Çıkış Gücü ve Verim Üzerindeki Etkisinin İncelenmesi. *Mühendis ve Makina*, *63*(707), 429-442.
[11] Nabil, T. ve Mansour, TM (2022). Çeşitli soğutma teknikleri kullanılarak sıcaklığını düşürerek fotovoltaik panelin performansının artırılması. *Mühendislikte Sonuçlar* , *15* , 100564.
[12] SELİM SARAL, Y. (2023). Fotovoltaik panel verimliliği ve maksimum güç noktası izleme. *Istanbul Aydin Üniversitesi Dergisi Anadolu Bil Meslek Yüksekokulu*, *18*(67).
[13] Dida, M., Boughali, S., Bechki, D., & Bouguettaia, H. (2021). Experimental investigation of a passive cooling systemfor photovoltaic modules efficiency improvement inhot and arid regions. *Energy Con version and Management*, *243*, 114328.
[14] Shalaby, S. M., Elfakharany, M. K., Moharram, B. M., & Abosheiasha, H. F. (2022). Experimental study on theperformance of PV with water cooling. *Energy Reports*, *8*, 957-961.
[15] Kabul, A., & Duran, F. (2014). Isparta ilinde fotovoltaik/termal (PV/T) hibrit sistemin performans analizi. *Uluslararası Teknolojik Bilimler Dergisi*, *6*(1), 31-43.
[16] Akal, D., & Umut, İ. (2022). Fotovoltaik Panellerde Güç Tahminlenmesi için Yapay Zekâ Yöntemlerinin Kullanılması. Tekirdağ Ziraat Fakültesi Dergisi, 19(2), 435-445.
[17] Kayabaşı, A., Yıldız, B., & Balcı, S. GÜNEŞ PANELİ UYGULAMALARINDA HAVA ŞARTLARININ VERİMLİLİĞE ETKİLERİNİN YAPAY SİNİR AĞLARI İLE BELİRLENMESİ.
[18] Yüzer, E. Ö., Bozkurt, A., & Barutçu, İ. Ç. (2023). FOTOVOLTAİK SİSTEM ÇIKIŞ GÜCÜNÜN YAPAY SİNİR AĞLARI VE MATLAB/SİMULİNK MODELLERİNİN ENTEGRASYONU İLE BELİRLENMESİ. Mühendislik Bilimleri Ve Tasarım Dergisi, 11(2), 551-563. https://doi.org/10.21923/jesd.1163411
[19] Hudișteanu, S. V., Oprea, S. V., & Ionescu, L. M. (2024). Comparative Study on the Impact of Temperature on Electrical Parameters and Energy Efficiency of Two Types of Photovoltaic Panels. *Sustainability*, *16*(23), 10566. https://doi.org/10.3390/su162310566

[20] Karaca, A., Kürekci, E., & Duran, H. (2024). Effects of cooling on PV panel surface temperature and efficiency in PV/T systems. *Dialnet Scientific Journal, 28*(2), 121–134. https://dialnet.unirioja.es/descarga/articulo/9572462.pdf

[21] Hassan, M. H., El-Maghlany, W. M., & Hasaneen, M. A. (2025). Analysis of the Influence of Temperature on the I-V Characteristics of Crystalline Silicon Solar Cells. *Energy and Power Engineering, 17*(5), 114–122. https://doi.org/10.4236/epe.2025.175009

[22] Shaker, A. H. (2024). Effects of Temperature on the Efficiency of Solar PV Systems: A Review. *Sustainable Energy Research, 11*(1), 33–45. https://doi.org/10.1186/s40807-024-00100-8

[23] Öztürk, M. (2024). Fotovoltaik panellerde sıcaklık kaynaklı verim kayıplarının incelenmesi ve soğutma yöntemleri. *Batı Anadolu Elektrik ve Enerji Dergisi (BAJECE), 12*(1), 45–60. https://dergipark.org.tr/en/pub/bajece/issue/89866/1449483

[24] Fatima, K., Minai, A. F., Malik, H., & Márquez, F. P. G. (2024). Experimental analysis of dust composition impact on Photovoltaic panel Performance: A case study. *Solar Energy*, *267*, 112206.

[25] Daş, M., & Kavak Akpınar, E. (2018). Mushroom Drying in Air Heated Solar Collector Drying System and Modeling of Drying Performance with Artificial Neural Network. Erzincan University Journal of Science and Technology, 11(1), 23-30.

[26] Sreehari, E., & Ghantasala, P. G. (2019). Climate changes prediction using simple linear regression. *Journal of Computational and Theoretical Nanoscience*, *16*(2), 655-658.

[27] Mahaboob, B., Praveen, J. P., Rao, B. A., Haranadh, Y., Narayana, C., & Prakash, G. B. (2020). A study on multiple linear regression using matrix calculus. *Advancecs in Mathematics Scientifc journal*, *9*(7), 1-10.

[28] Alic, E., Das, M., & Kaska, O. (2019). Heat Flux Estimation at Pool Boiling Processes with Computational Intelligence Methods. *Processes*, *7*(5), 293. https://doi.org/10.3390/pr7050293

[29] Wang, Y., Witten, I.H., 2002. Modeling for optimal probability prediction. Proc. Ninet. Int. Conf. Mach. Learn. 650–657.

[30] Meshkin, A., Sadeghi, M., & Ghasem-Aghaee, N. (2009). Prediction of relative solvent accessibility using pace regression. *EXCLI J*, *8*, 211-217.

Investigation of Fracture Parameters of Lime-Pumice Mortars

Erkin EREN[1], Ragip İNCE [2]

[1]Graduate School of Natural and Applied Sciences, Firat University, Elazig, Turkey.
[1,2]Civil Engineering Department, Engineering Faculty, Firat University, Elazig, Turkey.
[1] eeren@firat.edu.tr, [2]rince@firat.edu.tr
[1](ORCID: 0000-0001-5282-398X), [2](ORCID: 0000-0002-9837-8284)

Abstract
The experimental studies on fracture mechanics of concrete until 1970s revealed that classical linear elastic fracture mechanics (LEFM) was invalid for quasi-brittle materials such as concrete. This inapplicability of LEFM is due to existence of an inelastic zone with large scale and full cracks in front of the crack tip in concrete. This so-called fracture process zone is ignored by LEFM. For this reason, several investigators have developed non-linear fracture mechanics approaches to characterize failure of concrete structures. In this study, beams with three different initial crack lengths were produced to calculate fracture parameters of lime-pumice mortars. All beams were kept in the mold for 28 days, wrapped in stretch film. Mortar beams were notched with a diamond saw blade of 1.5 mm thickness at 28 days. The failure loads were measured on the specimens tested by three-point bending of specimens were determined. The fracture parameters of specimens were calculated by two-parameter model, a popular way in concrete fracture, which proposes two parameters the critical stress intensity factor and the critical crack tip opening displacement.

Keywords: beam, fracture mechanics, lime, mortar, pumice

1.Introduction

Linear elastic fracture mechanics (LEFM) were first utilized for cement-based materials by Kaplan [1] in 1961, followed by Kesler and his colleagues [2] in 1972. The latter researchers determined that LEFM was insufficient for concrete. This inadequacy arises from the relatively inelastic region present in quasi-brittle materials, such as concrete mixtures, rocks, and asphalt concretes, which is extensive and fully fractures ahead of the crack tip. Consequently, numerous researchers have created non-linear fracture mechanics models to effectively characterize fracture-dominated failures in concrete structures. These methodologies encompass the fictitious crack model (FCM) [3], the crack band model [4], the two-parameter model (TPM) [5], the effective crack model [6], the size effect model (SEM) [7], and the double-K model [8]. Roman mortars are the oldest type of concrete, consisting of lime, water, sand and volcanic tuff, and has enabled the construction of structures that have stood for nearly two millennia, with well-known examples such as the Pantheon temple and the Colosseum [9-12]. The Pantheon temple still serves today as an exhibition hall. The Colosseum was damaged by a devastating earthquake in 412 AD. Earlier, when Rome accepted Christianity, many precious stone pieces were used in the construction of other structures. Today, 1/3 of the structure stands, most of which is the concrete part. On the other hand, in Seljuk and Ottoman architecture, a considerable number of structures made with Khorasan mortar, which is made of lime, sand, water and baked clay (brick) that has been ground and added, have survived to the present day.

The aim of this study is to investigate how the structures made of the above mentioned historical mortars can withstand such damaging effects as seismic activity (earthquake), environmental conditions (freezing-thawing and wetting-drying) and fire for such a long time using the principles of fracture mechanics. Because such effects cause cracks to form in the structure and a cracked structure can only be analyzed realistically using the principles of fracture mechanics.

2. Two Parameter Model

In the 1970s, experimental studies on the fracture mechanics of cement-based materials, including paste, mortar, and concrete, demonstrated that classical Linear Elastic Fracture Mechanics (LEFM) were no longer applicable to quasi-brittle materials [2]. This inapplicability arises from the relatively extensive inelastic zone present in concrete, known as the fracture process zone (FPZ), which forms ahead of and around the tips of significant cracks. As a result, various non-linear fracture mechanics models have been created to describe the FPZ.

[1]Corresponding author

These fracture models can be divided into cohesive crack models, such as the fictitious crack model introduced by Hillerborg et al. [3] and the crack band model proposed by Bazant and Oh [4], as well as effective crack models, which include the two parameters model (TPM) developed by Jenq and Shah [5], the effective crack model formulated by Nallathambi and Karihaloo [6], and the size effect model established by Bazant and Kazemi [7]. The cohesive crack models represent the FPZ with a closing pressure that decreases near the crack tip, whereas the effective crack models depict the FPZ using an effective crack length. The aim of each approach is to determine the critical crack extension, defined as $\Delta a = a_c - a_0$, where a_c and a_0 are the critical crack length at the peak load and the initial crack length, respectively, during the load-displacement response or the load-crack mouth opening displacement response of the structure. Additionally, a_c depends on a structure's size because it tends to converge to a_0 as the size increases [7]. Therefore, the non-linear fracture approaches require at least two experimentally determined fracture parameters to simulate concrete fracture.

A concrete structure fails according to the TPM when the stress intensity factor K_I and the crack opening displacement $CTOD$ reach their critical values, K_{Ic}^{s} and $CTOD_c$, respectively. These fracture parameters can be determined with the following LEFM equations:

$$K_{Ic}^{s} = \sigma_{Nc} \sqrt{\pi a_c}\, Y(g,l) \tag{1}$$

$$CTOD_c = \frac{4\sigma_{Nc} a_c}{E'} V_1(g,l) M(g,l) \tag{2}$$

where σ_{Nc} is the nominal failure stress; $E' = E/(1-\nu^2)$ for plane strain; $E' = E$ for plane stress; E is Young's modulus; ν is the Poisson ratio; and Y, V_1, and M are dimensionless functions that depend on the geometry of the structure (g) and the load type (l). The parameter Y is also called the geometry factor. The function M is derived from the ratio of $COD(a_c)/CMOD_c$, where $CMOD_c$ is the critical crack mouth opening displacement. Because the Y, V_1, and M functions can be found in LEFM handbooks [13], the TPM can easily be used in structural analysis. As an instance, the normalized functions employed for beam specimens with span/depth=2.5 used also in this study can be given below [14]:

$$Y(\alpha) = \frac{1.83 - 1.65\alpha + 4.76\alpha^2 - 5.3\alpha^3 + 2.51\alpha^4}{\sqrt{\pi}\,(1+2\alpha)(1-\alpha)^{3/2}} \tag{3}$$

$$V_1(\alpha) = 0.65 - 1.88\alpha + 3.02\alpha^2 - 2.69\alpha^3 + \frac{0.68}{(1-\alpha)^2} \tag{4}$$

$$\frac{COD(y)}{CMOD} = \sqrt{\left(1-\frac{y}{a}\right)^2 + (1.081 - 1.149\alpha)\left[\frac{y}{a} - \left(\frac{y}{a}\right)^2\right]} \tag{5}$$

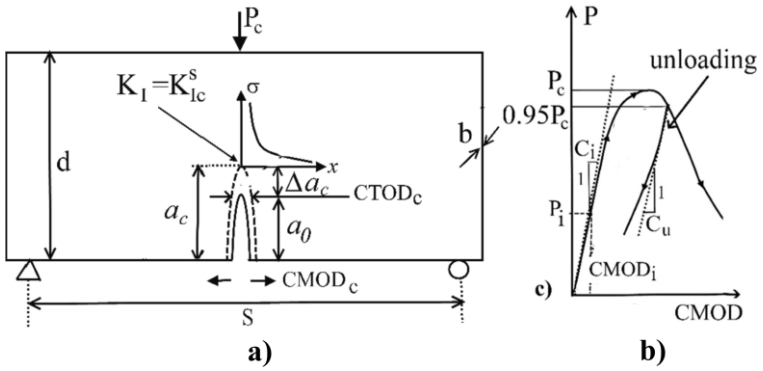


Figure 1. Modeling of the cracked bodies according to the TPM
a) notched beam **b)** typical load-CMOD curve

This methodology derives fracture parameters from one of two experimental techniques: the compliance method and the peak load method, proposed by Yang et al. [14]. In the compliance method, fracture parameters are ascertained from the correlation between the load and the crack mouth opening displacement (P-CMOD) of three-point bending specimens featuring a central edge notch, utilizing closed-loop testing apparatus, as illustrated schematically in Fig. 1. In the TPM, a_c is calculated by using unloading compliance (C_u) at 95% of the ultimate load as shown in Fig. 1b. The peak-load method, which does not necessitate complex testing apparatus, is more straightforward than the RILEM method for determining fracture parameters within the TPM. However, it necessitates three or more distinct specimens due to the inherent variability in concrete properties. This requirement holds true for both methodologies. The specimens may be uniform in size, yet their initial crack lengths may vary. Conversely, the initial crack lengths may be consistent, while the dimensions of the specimens may differ. For each specimen evaluated, the following equations can be formulated in accordance with the TPM:

$$K_I^i\left(\sigma_{Nc}^i, a_c^i\right) = K_{Ic}^s \qquad , i = 1, 2$$
$$CTOD^i\left(\sigma_{Nc}^i, a_c^i\right) = CTOD_c \tag{6}$$

where *i* denotes the *i*th specimen. Consequently, the fracture parameters can be found by simultaneously solving four non-linear equations. However, three or more distinct specimens must be tested to ensure statistically valid results because random errors always exist in measured values of σ_{Nc}^1 and σ_{Nc}^2. In this study, a statistical procedure was used to calculate K_{Ic}^s and $CTOD_c$.

In practice, it is possible to combine the fracture parameters K_{Ic}^s and $CTOD_c$, and the material parameter E into a single length parameter Q. This is referred to as the brittleness number by Jenq and Shah [5] as given in Eq. 5:

$$Q = \left[\frac{E \cdot CTOD_c}{K_{Ic}^s}\right]^2 \tag{7}$$

The experimental studies illustrated that the values of Q are ranging from 50-150 mm for mortar, and 150-350 mm for normal concrete. Smaller values of Q point out a more brittle material behavior.

3. Experimental Program

The mixture components were in a water: lime: pumice: aggregate ratio of 1.53:1.00:1.00:6.00 for mortar. CL-80-S lime was used in the mix. Its specific gravity was 2.08. The maximum sand size was 4 mm. The sand were air-dried prior to mixing. The Fuller parabola was used to determine the percentages of the aggregate gradation. Steel molds

(width ×depth ×length =50 ×50 ×150 mm) were used for the beams. In addition to the notched specimens, two cubes (100 mm) were made from mortar for compressive strength.
All beams and identical cubes were kept in the mold for 28 days, wrapped in stretch film. Beams were notched with a diamond saw blade of 1.5 mm thickness at 28 days, except for 6 specimens which were used for the flexural strength. The compression tests and the bending tests were performed using a digital compression machine with a capacity of 100 kN. The specimens were loaded monotonically until final failure and care was taken to apply a constant loading rate. Typically, approximately 2 min (± 30 sec) elapsed before the maximum load capacity for each specimen was reached as shown in Fig. 2a. Identical cubes were tested at an age similar to the other specimens.
Crack patterns at the failure of beams and compression cubes are shown in Fig. 3. The specimen width (b), specimen depth (d) and the peak load values (P_c) of the notched beams were reported in Tables 1 according to the notch depths (a_0).

The compressive strength of cubical mortar was 3.79 MPa while the flexural strength of beams was 1.18 MPa (Table 1). In this study, by using the following formula, Young's modulus of the material was determined by means of the ultrasonic technique (Fig. 2b and 2c):

$$E = 10^4 V^2 \frac{\Delta}{9.81}, \quad \text{MPa} \tag{8}$$

where, V is the velocity (km/s) and Δ is the unit weight (kg/lt). Based on the measurements taken from a total of six unnotched beams, E was calculated as an average of 15.6 GPa.

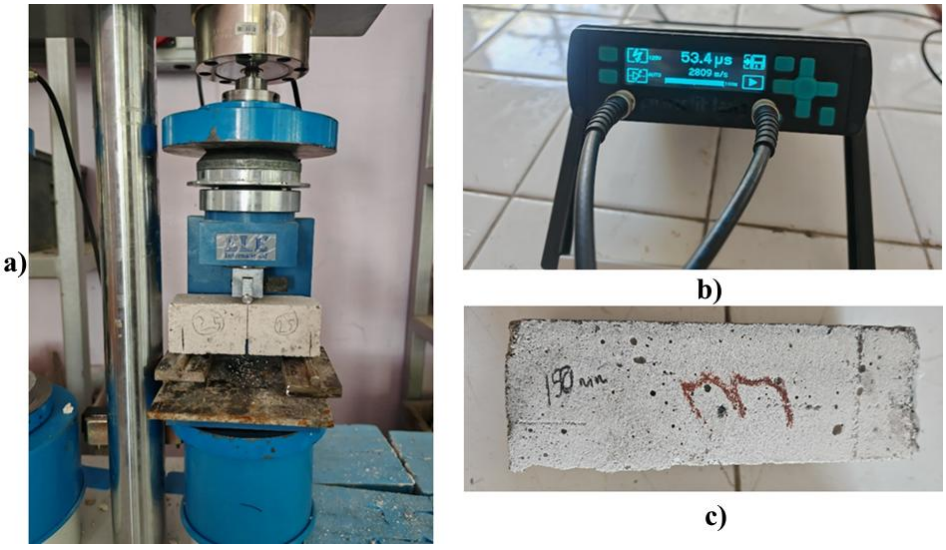


Figure 2. a) Test setup **b)** ultrasonic test apparat **c)** testing beam

Table 1. Test results of beams tested in this study

Specimen	b mm	d mm	a_0 mm	P_c N
B01	52.28	50.53	-	900
B02	54.71	50.64	-	720
B03	52.41	50.86	-	780
B04	47.9	50.32	-	820
B05	49.73	50.77	-	840
B06	49.46	50.07	-	800

Table 1. (Continue) Test results of beams tested in this study

B11	47.77	49.32	8.84	540
B12	50.85	50.3	9.99	460
B13	54.16	49.63	9.71	590
B14	50.41	50.23	9.36	680
B15	52.08	50.41	9.93	520
B21	53.65	50.03	15.00	350
B22	54.26	49.84	14.26	490
B23	49.26	50.05	13.58	410
B24	52.21	51.28	15.79	460
B25	52.32	50.72	15.42	450
B31	48.02	49.63	23.12	290
B32	50.32	49.17	21.84	250
B33	55.85	49.21	21.43	320

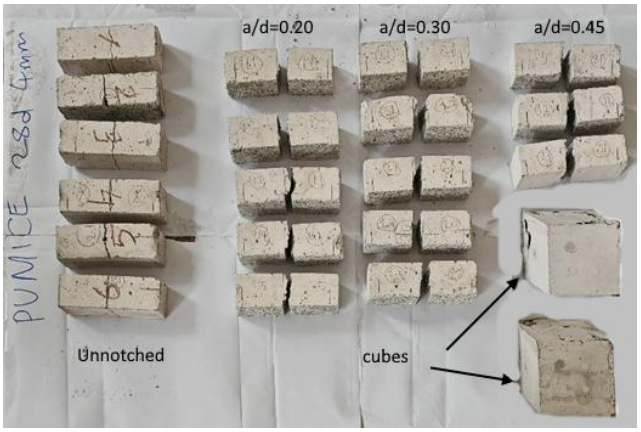


Figure 3. Fractured specimens

4. Analysis of Test Results

In the peak-load method, the fracture parameter K_{Ic}^{s} that causes the smallest standard deviation in the $CTOD_c$ is determined because at least two tests are required. However, as previously mentioned, at least three groups must be tested to obtain statistically valid results because random errors always exist in measured nominal strength values. The following statistical procedures were used to evaluate TPM, K_{Ic}^{s} and $CTOD_c$, according to this method [14-20]. Specimens were grouped with reference to their initial crack length a_0, as shown in Table 1. The average initial crack length and the average nominal strength were then calculated for each group. The K_{Ic}^{s} - $CTOD_c$ relationships were determined for each group by using Equations 1 to 5; an example for lime-pumice mortars is shown in Figure 4. Subsequently, the average K_{Ic}^{s} - $CTOD_c$ curve was obtained for these groups, as indicated by the curved solid line in Figure 4. The sample standard deviation of the groups was then calculated as follows:

$$s\left(K_{Ic}^{s}\right)=\sqrt{\frac{\sum_{i=1}^{n}\left(\overline{CTOD_c}-CTOD_{ci}\right)^2}{n-1}} \tag{9}$$

where n is the number of groups (n=3 in this study); $\overline{CTOD_c}$ is the average value of $CTOD_c$ for all groups; and $CTOD_{ci}$ is the value of $CTOD_c$ for the ith group. As shown in Fig. 4, the value of K_{Ic}^{s} corresponding to the minimum value of s was taken as the fracture parameter K_{Ic}^{s}. The fracture parameter $CTOD_c$ was obtained by substituting this value of K_{Ic}^{s} into the average K_{Ic}^{s} - $CTOD_c$ curve. For instance, the K_{Ic}^{s} and $CTOD_c$ parameters were determined to be 0.151 MPa√m and 0.0019 mm at a minimum s value (s_{min}) for beams (Fig. 4). According to this, the brittleness number (Q) of the mortar can be computed as 27 mm.

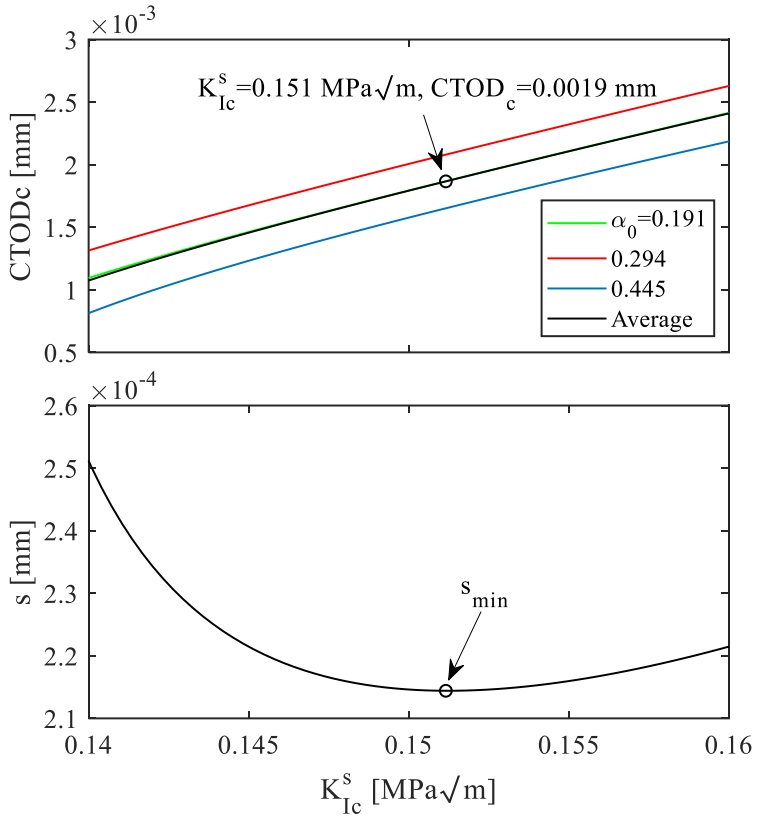


Figure 4. Application of the peak load method of lime-pumice mortars

5. Conclusions

This study utilized the two-parameter model, commonly applied to concrete fracture analysis, to examine the fracture behavior of lime-pumice mortars. The key findings are summarized below:

- The fracture toughness value of the lime-pumice mortar was around 1/5 compared to normal concrete. Moreover, the resistance of this material to crack propagation is much lower than normal cement-based ones.
- Similarly, the brittleness number was observed very low compared to normal cement-based mortars. In other words, these types of mortars are much more brittle materials than cement-based ones.

- However, it is emphasized that further studies can come up with more reliable results by investigating various types and sizes of aggregates to verify the above findings.

Acknowledgments

This article is extracted from the first author's doctorate dissertation entitled "Investigation of Structures Produced From Historical Mortars Using Fracture Mechanics Principles", supervised by Ragip Ince.

6. References

[1] Kaplan MF. Crack propagation and the fracture of concrete. ACI J 1961; 58(11):591-610.
[2] Kesler CE, Naus DJ, Lott LL. Fracture mechanics-its applicability to concrete. In: Proceedings of the Int Conf on the Mechanical Behavior of Materials, vol. 4, Kyoto, 1971. The Soc of Mater Sci; 1972. p. 113-124.
[3] Hillerborg A, Modeer M, Petersson PE. Analysis of crack formation and crack growth in concrete by means of fracture mechanics and finite elements. Cem Conc Res 1976; 6:773-782.
[4] Bazant ZP, Oh BH. Crack band theory for fracture concrete. Mater Struct (RILEM) 1983; 16(93):155-157.
[5] Jenq YS, Shah SP. Two-parameter fracture model for concrete. ASCE J Engng Mech 1985; 111(10):1227-1241.
[6] Nallathambi P, Karihaloo BL. Determination of the Specimen Size Independent Fracture Toughness of Plain Concrete. Mag Conc Res 1986; 38(135):67-76.
[7] Bazant ZP, Kazemi MT. Determination of fracture energy, process zone length, and brittleness number from size effect with application to rock and concrete. Int J of Fract 1990; 44(2):111-131.
[8] Xu S, Reinhardt HW. Determination of double-K criterion for crack propagation in quasi-brittle fracture, Part III: Compact tension specimens and wedge splitting specimens. Int J of Fract, 1999; 98(2):179-193.
[9] Çamlıbel AN. Geleneksel yapılarda stabilitenin iyileştirilmesi: Temellerin Takviyesi. Birsen Yayınevi; 1984.
[10] Erdoğan ST, Erdoğan TY. Bağlayıcı malzemelerin ve betonun on bin yıllık tarihi. ODTÜ Yayıncılık; 2007.
[11] habertürk.com. "Roma İmparatorluğu'ndan kalan eserler nasıl bu kadar dayanıklı olabildi?". Kaynak: Massachusetts Institute of Technology.
[12] Aşanlı M. Geleneksel Yapı Teknikleri. 2nd ed. Yeni İnsan Yayınevi; 2016.
[13] Tada H, Paris PC, Irwin GR. The Stress Analysis of Cracks Handbook. 3rd ed. ASME Press; 2000. https://doi.org/10.1115/1.801535.
[14] Yang S, Tang T, Zollinger DG, Gurjar A. Splitting tension tests to determine concrete fracture parameters by peak–load method. Adv Cem Based Mater 1997;5:18–28.
[15] Ince R. Determination of concrete fracture parameters based on two-parameter and size effect models using split-tension cubes. Eng Fract Mech 2010;77: 2233–50. https://doi.org/10.1016/j.engfracmech.2010.05.007.
[16] Ince R. Determination of concrete fracture parameters based on peak-load method with diagonal split-tension cubes. Eng Fract Mech 2012;82:100–14. https://doi.org/10.1016/j.engfracmech.2011.11.026.
[17] Ince R. Determination of the fracture parameters of the Double-K model using weight functions of split-tension specimens. Eng Fract Mech 2012;96:416–32. https://doi.org/10.1016/j.engfracmech.2012.08.024.
[18] Ince R. Usage of compact compression specimens to determine non-linear fracture parameters of concrete. Fatigue Fract Eng Mater Struct 2021;44:410–26. https://doi.org/10.1111/ffe.13368.
[19] Ince R, Bildik AT. Batch design of cementitious composites for the double-K fracture model. Mater Struct 2023;56:145. https://doi.org/10.1617/s11527-023-02238-1.
[20] Ince R. Using SCB specimens to quantify nonlinear fracture characteristics in concrete and rock materials, Eng. Fract. Mech. 318 (2025) 110951. https://doi.org/10.1016/j.engfracmech.2025.110951.

An Alternative SCB Test for Fracture Analysis of Asphalt Concretes

B. Fatih FURTANA[1], Ragip İNCE [2]

[1]Graduate School of Natural and Applied Sciences, Firat University, Elazig, Turkey.
[1,2]Civil Engineering Department, Engineering Faculty, Firat University, Elazig, Turkey.
[1] bffurtana@firat.edu.tr, [2]rince@firat.edu.tr
[1](ORCID: 0009-0007-7171-9089), [2](ORCID: 0000-0002-9837-8284)

Abstract
The tests on concrete elements containing cracks indicated that linear elastic fracture mechanics (LEFM) was invalid for quasi-brittle materials such as rock, concrete and asphalt concrete. The inapplicability of LEFM is due to the existence of an inelastic zone, the fracture process zone (FPZ), in front of a crack in quasi-brittle materials. Therefore, several non-linear fracture mechanics approaches were recommended to characterize FPZ. Bituminous mixtures containing various defects, such as cracks and voids, exhibit ductile behavior at room temperature and brittle behavior in cold environments. Such materials can only be simulated realistically with the criteria of nonlinear fracture mechanics. In this investigation, a total of 10 semi-circular bending (SCB) specimens with four different initial crack lengths were produced to determine fracture parameters of pure asphalt concrete. Load-displacement curves were recorded on the specimens tested under the three-point bending. The fracture parameters of specimens were determined by effective crack model, a popular way in concrete fracture, which proposes two parameters the fracture toughness and the effective crack length.

Keywords: asphalt concrete, effective crack model, fracture mechanics, semi-circular bending test

1. Introduction

The fundamental parameter of classical linear elastic fracture mechanics (LEFM) is the stress intensity factor K_I, which defines the stress and displacement distributions around the crack tip in any cracked body. Consequently, a structure fails in accordance with the LEFM when the K_I value reaches its critical value, K_{Ic}, which is also known as the fracture toughness. Implementation of fracture mechanics in concrete was first suggested by Kaplan [1] in 1961. However, numerous tests on fracture mechanics of concrete since the early 1970s indicated that K_{Ic} value increases with the increasing size of specimens [2]. For this reason, several researchers suggested models of non-linear fracture mechanics to evaluate the size-independent fracture toughness parameter [3-8]. To simulate any structural element by adopting the principles of fracture mechanics, the material's fracture parameters should be determined first. In terms of the evaluation of the fracture parameters of cementitious materials such as cement paste, mortar, and concrete, one of the most prevalent specimens is the notched beams, which was the first specimen used in fracture mechanic tests conducted for metals. Nonetheless, the compact specimens, such as cubical and cylindrical concrete specimens, were previously utilized in fracture tests of concrete thanks to various benefits, such as being lighter and more compact than beam specimens.

Fracture tests on semicircular bending (SCB) specimens have commonly been conducted to evaluate the fracture energy of mixtures such as the fracture energy based on the work of fracture (G_F) and the critical strain energy release rate (J_c) [9]. However, the bituminous composites cannot be simulated using only the concept of fracture energy. Thus, a work of fracture method introduced by the RILEM 50-FMC was based on the fictitious crack model, requiring three parameters to characterize concrete fracture: G_F, tensile strength, and the relationship between crack closing pressure and crack opening displacement [3].

The use of SCB specimens is quite common in fracture mechanics tests of asphalt and rock materials. Such specimens also provide great advantages for determining the fracture mechanics-based performance of existing structures. Although a complex and expensive frame system is recommended by AASHTO [9] for SCB-based asphalt mixtures, the deformation of the loading head is widely taken into account in the literature in determining the load-deformation relationship of SCB specimens. However, such a measurement causes erroneous measurements, especially due to local deformations at the points where the load acts on the asphalt material. In this thesis, it is aimed to develop a cheaper alternative test method that will eliminate these local deformations.

[1]Corresponding author

2. Effective Crack Model

Several investigators have developed non-linear fracture mechanics approaches to characterize FPZ. These models can be classified as the cohesive crack models (the fictitious crack model by Hillerborg et. al. [3] and the crack band model by Bazant and Oh [4]) and the effective crack models (the two parameters model by Jenq and Shah [5], the effective crack model (ECM) by Nallathambi and Karihaloo [6], the size effect model by Bazant and Kazemi [7] , and the double-K model by Xu and Reinhardt [8]). The cohesive crack models simulate the fracture process zone (FPZ) by a closing pressure which diminishes near the crack tip while the effective crack models simulate FPZ by an effective crack length. The main aim of any approach is to determine the critical crack extension (size of FPZ) at the peak load $\Delta a = a_e - a_0$, in which a_e and a_0 are the effective crack length at the peak load and the initial crack length respectively.

When the stress intensity factor K_I, which describes the stress singularity at the crack tip under mode I loading, reaches the critical stress intensity factor (fracture toughness) K_{Ic}, the crack progress is unstable according to LEFM. For the general case in mode I, the LEFM equation can be given in the following form:

$$K_{Ic} = \sigma_{Nc} \sqrt{\pi a_0} F(g, p) \tag{1}$$

where σ_{Nc} is the nominal failure stress and $F(g,p)$ is a dimensionless function of structure of geometry and load type. According to LEFM, it is assumed that the initial crack length a_0 in Eq. (1) does not change until it reaches the peak load. At this stage, the crack has a critical length and starts to progress in an unstable way. In reality, the crack propagate in a stable manner already at lower load until the peak load is reached since the size of FPZ can be larger than the specimen size in concrete-like quasi-brittle materials. Consequently, the critical crack length a_e must be considered in the evaluation of K_{Ic} in Eq. (1) for concrete. Nevertheless, a_e depends on structural size, because it decreases as the member size increases. For this reason, a unique fracture parameter is not adequate to model crack propagation in concrete.

The effective crack length a_e in ECM for the concrete fracture proposed by Nallathambi and Karihaloo [6] is calculated from the secant stiffness of the real concrete structure at the peak load. The main idea behind ECM approach can be explained with Figure 1, where are indicated the load-deflection plots a SCB specimen with a central edge notch up to peak load [10]. According to ECM, the fracture of concrete is described by two parameters, namely the critical stress intensity factor (the fracture toughness K_{Ic}^{e} and the effective crack length a_e.

The stiffness of the real structure in the linear regime is proportional to Young's Modula E. It can be determined from any pair load and deflection values (P_i, δ_i) in this regime, as shown in Figure 1. Note that in the ECM P_i can be chosen to correspond approximately to $P_c/2$ in the load-displacement curve recorded in the test. The effective crack length is calculated from that are two values taken from the load–displacement curve: the initial compliance C_i and the secant compliance C_s measured at the peak load as illustrated in Figure 1b. The modulus of elasticity can be derived for the SCB specimen as follows:

$$E = \frac{\alpha_0}{2 b C_i} D_1\left(\alpha_0 = \frac{a_0}{r}\right), \quad C_i = \frac{\delta_i}{P_i} \tag{2}$$

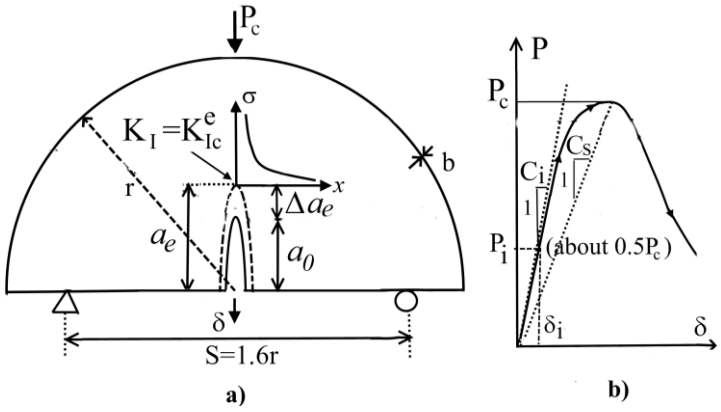


Figure 1. Modeling of the cracked bodies according to the ECM
a) SCB specimen **b)** typical load-δ curve

The dimensionless displacement function D_1 was developed for the configuration in Figure 1a by Ince et al. [10] as follows:

$$D_1(\alpha) = \frac{1}{0.1323\alpha^2 - 0.153\alpha^3 - 0.1169\alpha^4 + 0.1591\alpha^5 - 0.0212\alpha^6} \tag{3}$$

According to Figure 1b, the effective crack length (or the normalized effective crack length, α_e) is calculated by trial and error method as follows:

$$\alpha_e = \alpha_0 \frac{C_s D_1(\alpha_0)}{C_i D_1(\alpha_e)} \tag{4}$$

Consequently, the effective fracture toughness of the material can be determined for the specimen in Figure 1a as follows [10]:

$$K^e_{Ic} = \frac{P_c}{2br}\sqrt{\pi a_e}\,\frac{5.18 - 13.878\alpha + 27.145\alpha^2 - 26.887\alpha^3 + 10.321\alpha^4}{(1-\alpha)^{1.5}} \tag{5}$$

3. Experimental Program

Bitumen of grade B70/100 and density 1.033 was used in this investigation; it was procured from the TÜPRAS Batman refinery located in Turkey. Turkey's most common binder is the 70/100 grade, partly because of the country's climate. The aggregate gradation shown in Table 1 was adhered to during the sample preparation procedure. It was found that the pure mixture's intended bitumen concentration was 4.67 %. The bitumen content was maintained constant for both pure and modified mixtures to guarantee that the binder content did not affect the mechanical qualities of the mixtures. The resulting material was compacted into cylindrical Marshall specimens with a void ratio of 4 %, a height of approximately 60 mm, and diameters of 100 mm and 150 mm. A rotary compactor with an inclination angle of 1.25 ± 0.02◦ and a static pressure of 600 ± 18 kPa was used for this process. Then, semicircular specimens were created, notches were cut, and specimens measuring 150 mm were split in half. Two separate indentations, one centimeter and two centimeters, were made [11, 12].

The methodology for preparing SCB test specimens is illustrated in Figure 2 illustrates the notch formation observed in the specimens utilized in this research. Long cylindrical asphalt specimens, measuring 60 mm in height and 150 mm in diameter, were compacted using a gyratory compactor. To achieve a uniform air void distribution, the top and

bottom surfaces of the specimens were trimmed by 5 mm from each end, resulting in final dimensions of 50 mm in height and 150 mm in diameter. Subsequently, the specimens were bisected and notches were incised at the center to create SCB test specimens. Regardless of the specimen thickness and mix design, notch lengths of about 10, 15, 20 and 30 mm were selected for 150 mm specimens. The notch thickness is approximately 2 mm.

The displacement frame used for beams in the effective crack model, which is a popular method in the fracture mechanics of concrete, was adapted to the SCB specimen as detailed in Figure 3. Thanks to the LVDT to be mounted on the end of the rod sliding on the rigid rectangular frame, the under-notch displacement was determined based on the frame support on the support. SCB tests were carried out at room temperature.

Nonlinear fracture toughness value of bituminous hot mix samples was determined according to compliance method. Specimens were loaded at 0.5 mm/min speed. The load-displacement curves of SCB specimens tested are demonstrated in Figure 4. The fracture specimens are illustrated in Figure 5.

Table 1. Aggregate gradation of the asphalt mixture in this study

Sieve Size (mm)	0.075	0.15	0.3	0.6	1.18	2.36	4.75	9.5	12.5	19
Passed (%)	6	8	10	14	23	35	65	88	95	100

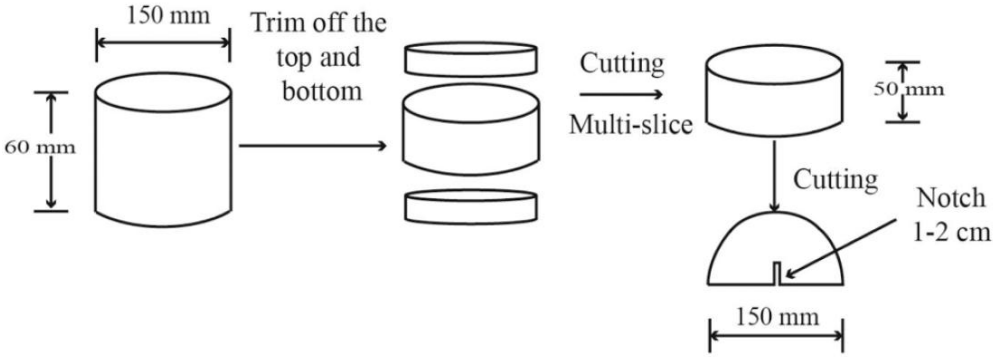


Figure 2. The preparation of a typical SCB specimen

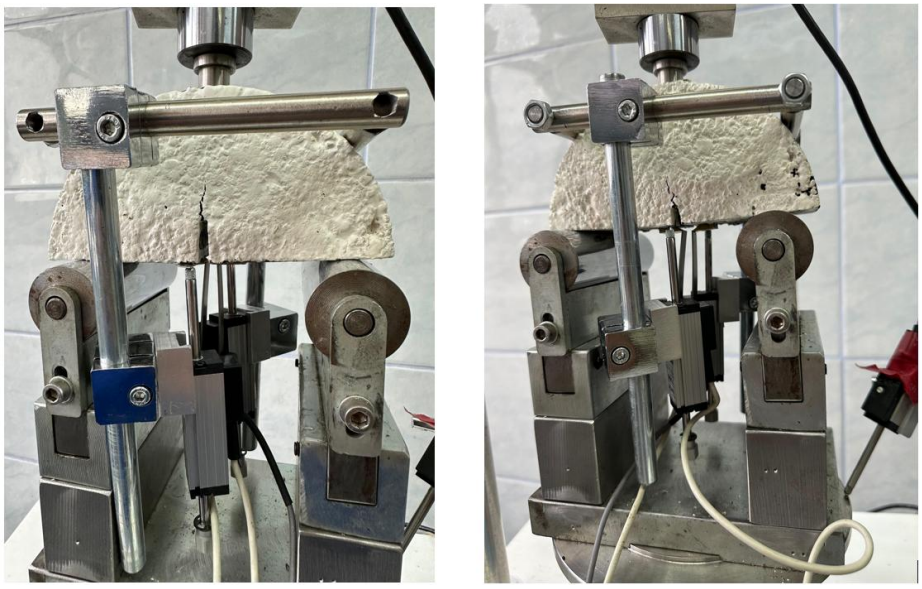


Figure 3. Test setup of SCB specimens

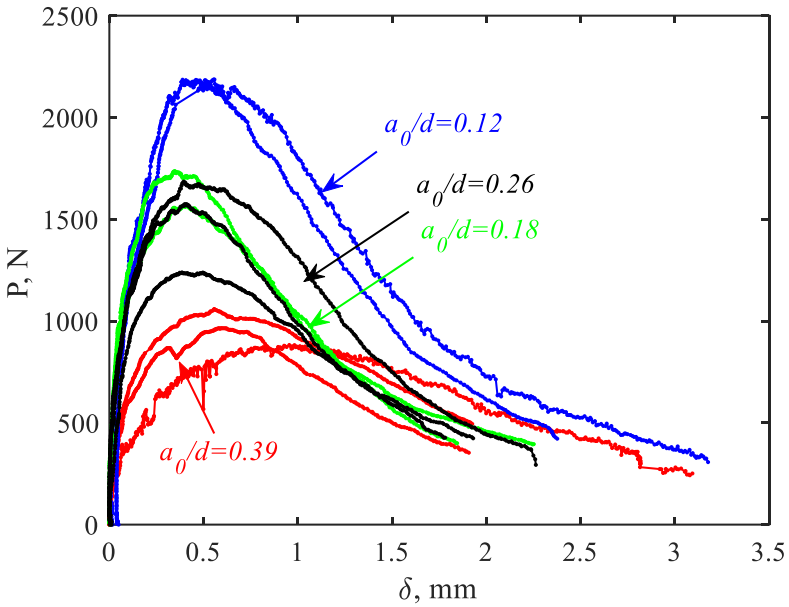


Figure 4. Load-displacement curves of SCB specimens tested

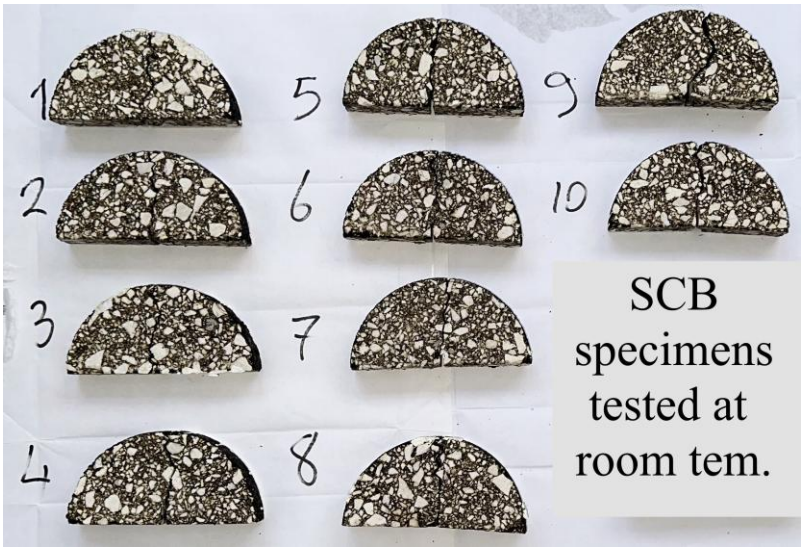


Figure 5. Fractured notched SCB specimens

4. Analysis of Test Results

The specimen width (b), specimen radii (r) and the peak load values (P_c) of the notched SCB specimens were reported in Table 2 according to the notch depths (a_0). By using Equations 2 to 5, the values of κ^{e}_{Ic} were determined from P-δ curve of each specimen as shown in Figure 4 [13]. The initial compliance values (C_i), the secant compliance values (C_s), the relative initial notch lengths (α_0), the relative effective notch lengths (α_0), the fracture toughness parameters based on the ECM (κ^{e}_{Ic}) and the fracture toughness parameters based on the LEFM (K_{Ic})are summarized for each specimen in Table 3.

It is seen that the difference between the LEFM-based fracture toughness value and the nonlinear fracture toughness of the specimens tested at room temperature is quite high compared to the specimens tested at low temperature. The

reason for this is that the sample at room temperature is more ductile and the size of the SCB specimen at the critical load is at a level that cannot be neglected. Note that the stiffness modulus was used instead of the initial modulus in Equation 2 for *E* modulus in the authors' the previous study. the ratio of *E* based on ECM to *E* based on stiffness modulus was obtained as 3–4 times. Consequently, the present study emphasizes that while the difference between the linear and nonlinear fracture toughness of asphalt concrete at low temperature does not exceed 5%, this difference is 3-4 times higher at room temperature.

Table 2. Test results of SCB specimens tested in this study

Specimen	b mm	d mm	a_0 mm	P_c N
SCB1	54.85	73.1	8.58	2152
SCB2	55.06	73.51	9.17	2186
SCB3	54.69	73.58	12.9	1569
SCB4	52.96	73.47	14.19	1736
SCB5	54.8	73.12	18.01	1574
SCB6	53.74	74.07	19.22	1685
SCB7	54.81	74.46	20.24	1239
SCB8	53.61	72.7	28.03	882
SCB9	55.04	72.95	28.35	965
SCB10	55.14	73.74	28.66	1059

Table 3. Fracture parameters of SCB specimens tested according to ECM

Specimen	C_i (mm/N)	C_s (mm/N)	E MPa	$\alpha_0=a_0/d$	$\alpha_e=a_e/d$	K_{Ic}^{e} (MPa√m)	K_{Ic} (MPa√m)
SCB1	0.000087108	0.000250006	926	0.117	0.537	0.672	0.206
SCB2	7.48846E-05	0.00025315	1086	0.125	0.579	0.79	0.214
SCB3	4.24892E-05	0.000254232	2095	0.175	0.702	0.976	0.181
SCB4	4.50644E-05	0.00020042	2105	0.193	0.658	0.902	0.217
SCB5	5.47869E-05	0.00025978	1854	0.246	0.686	0.906	0.217
SCB6	6.96656E-05	0.000233594	1528	0.259	0.629	0.758	0.243
SCB7	9.04063E-05	0.000311216	1185	0.272	0.639	0.569	0.180
SCB8	0.000334152	0.001081721	432	0.386	0.678	0.499	0.178
SCB9	0.000142367	0.000601871	996	0.389	0.721	0.661	0.191
SCB10	0.000113193	0.000525048	1251	0.389	0.734	0.774	0.208

5. Conclusions

This study utilized the effective crack model, commonly applied to concrete fracture analysis, to examine the fracture behavior of asphalt concrete. The key findings are summarized below:

- The current research highlights that although the disparity in linear and nonlinear fracture toughness of asphalt concrete at low temperatures is no greater than 5%, this variation increases to 3-4 times that amount at room temperature due to the use of different the modulus of elasticity (*E*) approach compared to the previous study.
- Until now, in hot bituminous mixtures, the fracture toughness parameter calculated based on LEFM is determined according to the effective crack model, which is a popular method in the fracture mechanics of quasi-brittle materials. Because the existing test equipment is complex and expensive since it is produced only by certain companies in the world.

- With the experimental setup developed here, SCB specimens were used with pure bituminous hot mixtures at room temperature to determine nonlinear fracture parameters by only considering the load-displacement relationships. However, it is emphasized that further studies can come up with more reliable results by investigating various types and sizes of aggregates to verify the above findings.

Acknowledgments

This article is extracted from the first author's master thesis entitled "An Alternative Semi-Circular Bending Test for Fracture Analysis of Bituminous Mixtures", supervised by Ragip Ince.

6. References

[1] Kaplan MF. "Crack propagation and the fracture of concrete". *ACI J*, 58(11), 591-610, 1961.
[2] Kesler CE, Naus DJ. Lott LL. "Fracture mechanics-its applicability to concrete". In: Proceedings of the Int Conf on the Mechanical Behavior of Materials, vol. 4, Kyoto, Japan, 1971. *The Soc of Mater Sci*, p. 113-124, 1972.
[3] Hillerborg A, Modeer M, Petersson PE. "Analysis of crack formation and crack growth in concrete by means of fracture mechanics and finite elements". *Cem Conc Res*, 6, 773-782, 1976.
[4] Bazant ZP, Oh BH. "Crack band theory for fracture concrete". Mater Struct, 16(93), 155-157, 1983.
[5] Jenq YS, Shah SP. "Two-parameter fracture model for concrete". *ASCE J Engng Mech*, 111(10), 1227-1241, 1985.
[6] Nallathambi P, Karihaloo BL. "Determination of the Specimen Size Independent Fracture Toughness of Plain Concrete". *Mag Conc Res*, 38(135), 67-76, 1986.
[7] Bazant ZP, Kazemi MT. "Determination of fracture energy, process zone length, and brittleness number from size effect with application to rock and concrete". *Int J of Fract*, 44(2), 111-131, 1990.
[8] Xu S, Reinhardt HW. "Determination of double-K criterion for crack propagation in quasi-brittle fracture, Part III: Compact tension specimens and wedge splitting specimens". *Int J of Fract*, 98(2), 179-193, 1999.
[9] AASHTO Designation: TP 105-13. "Standard Method of Test for Determining the Fracture Energy of Asphalt Mixtures Using the Semicircular Bend Geometry (SCB)", American Association of State and Highway Transportation Officials, 2015.
[10] Ince R. Yalcin E. Yilmaz M. "Quantifying nonlinear fracture parameters in bituminous SCB specimens: A compliance-based approach". *Case Stud. in Const. Mater*, 21, e03437, 2024.
[11] Yalcin E, Ince R, Yilmaz M. "Exploring activated carbon as an alternative to SBS in asphalt mixtures: Performance and fatigue analysis". *Alexandria Engineering Journal*, 121, 283–294, 2025.
[12] Ince R, Furtana BF. "Bitümlü Sıcak Karışımların Lineer Olmayan Kırılma Analizi". *Fırat Üniversitesi Müh. Bil. Dergisi*, 37(1), 443-452, 2025.
[13] Ince R. "Artificial neural network-based analysis of effective crack model in concrete fracture". *Fatigue Fract Eng Mater Struct*, 33, 595–606, https://doi.org/10.1111/j.1460-2695.2010.01469.x, 2010.

Evaluation of the Rheological Behavior of Pure and SBS-Modified 70/100 Batman Bitumen under Dynamic Shear Conditions

Hakan YURDAKUL[1], Ahmet Münir ÖZDEMİR[1] , Erkut YALÇIN[2], Bahadır YILMAZ[1], Mehmet YILMAZ[2]

[1]Civil Engineering, Engineering Faculty, Bursa Technical University, Bursa, Türkiye.
[2]Civil Engineering, Engineering Faculty, Firat University, Elazig, Türkiye.
[1]hakan.yurdakul4923@gmail.com, [2]ahmet.ozdemir@btu.edu.tr, [3]erkutyalcin@firat.edu.tr, [4]bahadir.yilmaz@btu.edu.tr, [5]mehmetyilmaz@firat.edu.tr,
[1](ORCID:0009-0006-4447-3004), [2](ORCID: 0000-0002-4872-154X), [3](ORCID: 0000-0002-6389-4211),[4](ORCID: 0000-0002-3480-1776), [5] (ORCID: 0000-0002-2761-2598)

Abstract
In this study, the rheological behavior of modified bituminous binders in the dynamic shear range was comparatively investigated using pure bitumen with a penetration grade of 70/100 obtained from Batman Petroleum Refineries in Turkey and 4% styrene-butadiene-styrene (SBS) copolymer. In the experimental studies, frequency sweep tests were conducted at temperatures between 40°C and 80°C and in the frequency range of 0.1–10 Hz using a dynamic shear rheometer (DSR) device. The obtained data were analyzed according to the Time-Temperature Superposition (TTSP) principle, and the main curves of the binders were created. The obtained main curves were evaluated using the Christensen-Anderson (CA) model, and the statistical significance of the model parameters was tested using ANOVA analysis. The results revealed that the SBS additive significantly increased the complex modulus of the binder, particularly in the low-frequency region, thereby enhancing elastic properties and providing resistance to permanent deformation. The characteristic frequency (wc) and rheological index (R) values obtained within the CA model framework demonstrated that SBS modification significantly affects the elastic-viscous transition behavior of the binder.

Keywords: Bitumen, rheology, dynamic shear rheometry, frequency sweep

1. Introduction

Asphalt mixtures have been the most widely used composite material in the construction of flexible pavements for decades [1]. The use of mineral aggregate and asphalt binder together as the skeleton and binder phase in the mixture enables long-term performance and driving comfort in road structures. However, in recent years, increasing traffic volume and significant climate changes have caused permanent deformations along tire tracks in asphalt pavement layers, negatively affecting road safety and driving comfort and increasing the risk of traffic accidents [2]. Asphalt binder is a petroleum-based material with a viscoelastic structure that exhibits low elastic capacity under different stress conditions and therefore tends to exhibit plastic deformation [3]. This situation has led to detailed research on the rheological properties of asphalt binders in recent years and an increase in scientific discussions on mechanical behavior in this context. Rheological analyses enable the determination of numerical parameters that describe the behavior of asphalt binders under various static and dynamic stress conditions. In particular, tests conducted using dynamic shear rheometry (DSR) within the Superpave performance classification system allow the asphalt binder's resistance to permanent deformation to be evaluated through the G*/sin(δ) parameter [4,5]. However, due to rapid developments in additive materials and asphalt technologies, it is known that the linear viscoelastic limits underlying the Superpave system do not adequately represent the permanent deformation mechanism [6]. Therefore, the development of approaches that consider non-linear viscoelastic behavior enables the design of pavements that are more resistant to deformations caused by wheel tracks and the complex stress-strain conditions created by traffic.

The rheological behavior of asphalt binders is highly dependent on parameters such as loading time and temperature. The mutual relationship between temperature and frequency in these materials allows similar mechanical responses to be obtained under different experimental conditions. In other words, the mechanical behavior observed under low temperature and high frequency conditions can also appear similarly under high temperature and low frequency conditions. Especially in bituminous binders, which are considered thermo-rheologically simple, the temperature-time equivalence principle applies, and this allows data obtained at different temperatures to be appropriately shifted to form a continuous main curve at a reference temperature [7]. In principle, it is possible to measure the full modulus-frequency (or loading time) behavior of any polymer or asphalt binder at a specific temperature [8]. In this context, the data is appropriately shifted using the

[1]Corresponding author

concept of reduced frequency (ωr), thereby combining the curves obtained at different temperatures into a single continuous master curve. The resulting master curve represents the rheological behavior of the binder at a specific temperature over a wide frequency range. The rheological properties of bituminous binders are primarily characterized by two limiting behaviors: At high frequencies or low temperatures, the material approaches the glass transition value of the elastic modulus, Gg; at low frequencies or high temperatures, it behaves like a Newtonian fluid, exhibiting viscous flow tendencies. A good main curve is expected to be smooth, continuous, and consistent with physical reality.

According to Chailleux and colleagues [7], the process of creating a main curve is meaningful only when there are no phase changes or large-scale structural rearrangements in the material's structure and the tests are conducted within linear viscoelastic (LVE) limits. Using the master curve and shear factors, it is possible to interpolate the stiffness behavior of the binder over a wider frequency and temperature range beyond the experimentally obtained data range. At this point, if functional forms appropriate to the master curve shape and shear factor relationships are preferred, this interpolation process can be performed quickly and practically in a software environment. Additionally, when functional models based on a certain degree of thermodynamic principles are used, the equations obtained can produce reliable predictions beyond the observed temperature and frequency range [8].

A significant number of studies have been conducted using various models to create smooth and continuous main curves [9]. In this study, the rheological behavior differences between 70/100 pure bitumen and SBS-modified bitumen were analyzed using the Christensen-Anderson (CA) model.

2. Materials and Method

2.1. Materials

In the experimental study, 70/100 penetration grade (PG 70-22) pure bitumen obtained from Batman Petroleum Refineries in Turkey was used. The properties of the pure binder are given in Table 1.

Table 1. Properties of the pure binder

Feature	Standart	B 70/100
Penetration (0.1 mm), 100 g, 5 sec	ASTM D5	85.10
Softening point (°C)	ASTM D36	52.65
Penetration index (PI)		0.87
Viscosity (cP, 135 °C)	ASTM D4402	562.50
Viscosity (cP, 165 °C)	ASTM D4402	175.00
Specific gravity	ASTM D70	1.038

The additive styrene-butadiene-styrene (SBS) copolymer (Kraton D-1101®) was supplied by Shell Chemical Company. The properties of SBS are given in Table 2.

Table 2. Properties of styrene-butadiene-styrene (SBS) additive

Polystyrene Content (%)	Specific gravity	Melt flow rate 200 °C/5kg (g/10min)	Elongation at break (%)	Tensile strength (MPa)	Ash (%)
30-32	0.94	< 1	880	33	< 0.2

The modification was carried out with the SBS ratio fixed at 4% by weight of the bituminous binder. SBS-modified bitumens were prepared by slowly adding the polymer in the specified content after the pure bitumen was melted and

transferred to the mixing tank, and then mixing it at a constant temperature of 180°C for 1 hour in a mixer operating at 1000 rpm. The prepared modified bitumens were poured into small containers, covered with aluminum foil, and stored for use in experiments. The appearance of the mechanical mixer and SBS used is shown in Figure 1.



Figure 1. Appearance of mechanical mixer with SBS and 4 blade impeller

2.2. Dynamic Shear Rheometer (DSR)

The DSR test was performed using a Bohlin DSRII rheometer under controlled stress conditions for unaged binders in order to determine the wheel track (G*/sinδ) parameters and thus determine the high-temperature performance of the binders. The test was performed using a parallel plate with a diameter of 25 mm and a gap width of 1 mm. The frequency sweep test involves oscillating the bituminous binder at different frequencies under constant stress. Since frequency is the inverse of time, it is possible to evaluate the time-dependent viscoelastic characteristics of a material by changing the oscillation frequency. The frequency sweep test is performed using a DSR test device based on the AASHTO T315 standard, at frequencies ranging from 0.1 to 10 Hz and temperatures ranging from 40 to 80°C, with 10°C increments. The test is applied at a minimum strain of 1% within the linear viscoelastic region of the binder.

Main curves were obtained by adhering to the Time Temperature Superposition Principle (TTSP) to examine a wide frequency and temperature range for pure and modified binders. The viscoelastic behavior of the binders was analyzed using the Christensen-Anderson (CA) model with the main curves. The transition frequency (Wc), glassy modulus (Gg), and rheological index (R) are parameters that characterize the viscoelastic behavior of bitumen according to the CA model (Equation 1). The glassy modulus, also known as the shear modulus limit, characterizes the stiffness at low frequencies. This value is generally accepted as 1 GPa [10].

$$|G^*| = G_g \left[1 + \left(\frac{\omega_c}{\omega}\right)^{\frac{log2}{R}}\right]^{\frac{-R}{log2}} \quad (1)$$

3. Results and Discussion

The master curves of pure and SBS modified binders plotted according to the time-temperature superposition principle (TTSP) using the G*-frequency curves obtained from the DSR frequency sweep test are given in Figure 2.

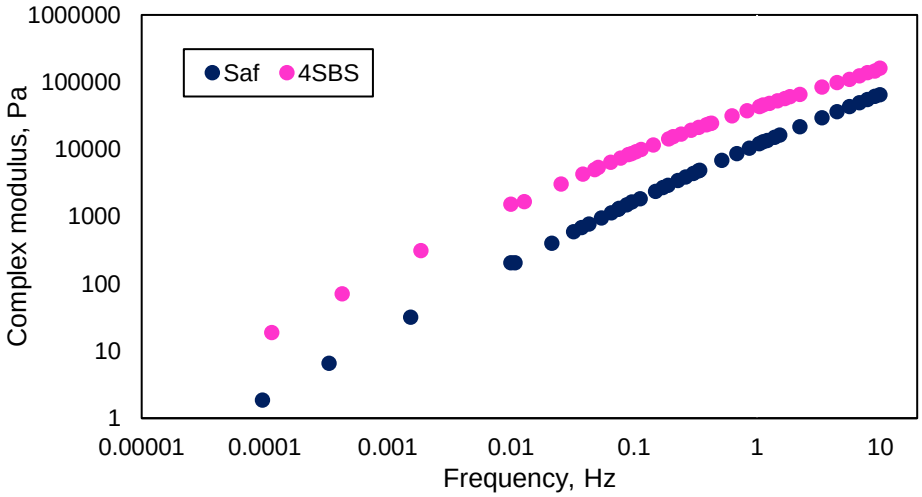


Figure 2. Master curves of pure and SBS modified binders

In the main curve graph presented in Figure 2, the rheological behaviors of pure bitumen and bitumen modified with 4% SBS (Styrene-Butadiene-Styrene) polymer are compared. The horizontal axis shows frequency (Hz) on a logarithmic scale, representing time-dependent loading conditions. The low-frequency range simulates long-term loading conditions (low speed or high temperature conditions), while the high-frequency range simulates short-term loading conditions (high speed or low temperature conditions). The vertical axis displays the complex modulus (Pa) of the adhesives, which is a general stiffness indicator representing both the elastic and viscous behavior of the material. Upon examining the graph, it is observed that the complex modulus values for both adhesives increase with increasing frequency. This confirms classical viscoelastic material behavior, indicating that the material exhibits a stiffer structure at high frequencies and a more fluid structure at low frequencies. However, the performance difference between the two adhesives becomes particularly pronounced in the low-frequency range. In the low-frequency region, the complex modulus values of the SBS-modified binder are significantly higher than those of the pure binder. This indicates that the modified binder is more resistant to long-term loading, such as permanent deformation (rutting). It is understood that the SBS additive imparts elastic characteristics to the binder matrix through its polymer chain structure, and this structure limits the viscous flow of the binder, particularly under low-frequency conditions. In the high-frequency region, the complex modulus values of both binders approach each other, but the modified binder still exhibits higher values. This indicates that the SBS additive not only enhances the binder's rigidity under long-term loading but also under short-term dynamic loading, thereby contributing to crack resistance. ANOVA results obtained after CA Model analysis are given in Table 3

Table 3. ANOVA sonuçları

		DF	Sum of Squares	Mean Square	F Value	Prob>F
Pure	Regression	2	1.93E+10	9.67E+09	418813.2	3.00E-83
	Residual	38	877617.1	23095.19		
	Uncorrected Total	40	1.93E+10			
	Corrected Total	39	1.31E+10			
4SBS	Regression	2	1.37E+11	6.87E+10	52117.25	4.68E-66
	Residual	38	5.01E+07	1318150		
	Uncorrected Total	40	1.37E+11			
	Corrected Total	39	7.76E+10			
			Pure	**4SBS**		
	Number of Points		40	40		
	Degrees of Freedom		38	38		

Reduced Chi-Sqr	23095.19	1318150
Residual Sum of Residual	877617.1	5.01E+07
R-Square(COD)	0.99993	0.99935
Adj. R-Square	0.99993	0.99934

The ANOVA results for the CA model presented in Table 3 strongly demonstrate the statistical significance of the model for both pure binder and 4% SBS modified binder. The F value for the regression model in the pure binder is 418,813.2, while the corresponding p value (Prob>F) is 3.00E-83. This value indicates that the model is statistically significant at a confidence level of over 99.9%.

Similarly, the F value of the model in the SBS-modified binder was calculated as 52.117.25, and the corresponding p value was 4.68E-66, confirming the high significance of the model. The reason for the lower F value in the SBS-modified binder compared to the pure binder may be due to the relatively limited explanatory power of the model compared to the pure binder, resulting from the increased material complexity in the modified binder. Nevertheless, the high F values and low p values obtained indicate that the CA model successfully represents the complex modulus-frequency relationship for both binders. The R2 (COD) and adjusted R^2 (Adj. R^2) values are at the 0.9999 level for both binders. This indicates that the model's explanatory power is nearly flawless and that nearly all of the observed variance is explained by the model. In conclusion, based on the obtained ANOVA outputs, it can be stated that the CA model accurately represents the rheological behavior of both pure and SBS-modified asphalt binders, and the parameters obtained from the model are statistically reliable. Figure 3 shows the main curves of pure and 4SBS binders fitted to the CA model.

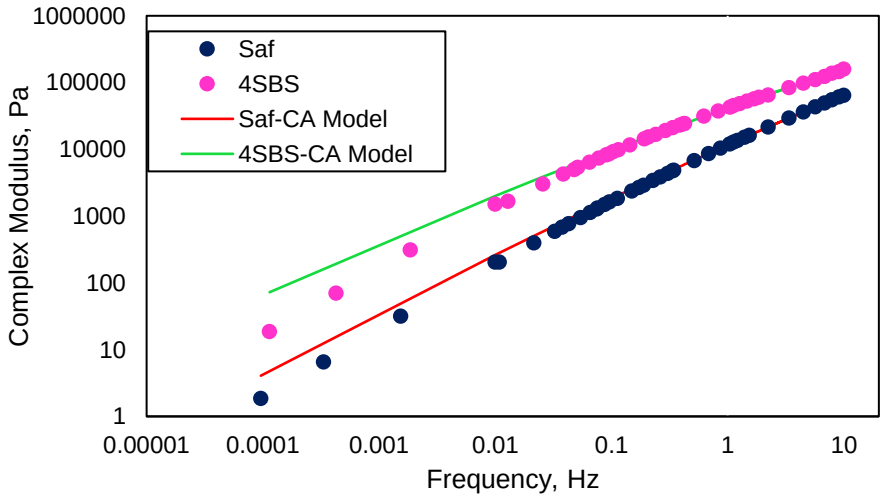


Figure 3. CA model curves

Figure 3 presents a comparative analysis of the main curves obtained experimentally for pure bitumen and 4% SBS modified bitumen, along with the theoretical curves obtained using the Christensen-Anderson (CA) model. The horizontal axis shows frequency (Hz) and the vertical axis shows complex modulus (Pa) on a logarithmic scale. The CA model is an empirical model widely used to predict the viscoelastic behavior of asphalt binders over a wide frequency range, theoretically expressing the complex modulus behavior of the binder. Upon examining the graph, it is observed that the CA model curves (red and green lines) for both pure binder and SBS-modified binder show a high degree of agreement with experimental data. Especially in the medium and high-frequency regions, the model curves and experimental data largely overlap. This indicates that the CA model is successful in predicting the elastic properties and short-term loading behavior of the adhesives in question.

When examining the low-frequency region, it is observed that the CA model yields results that are quite close to the experimental data for the pure binder, while there are partial deviations between the model and experimental data for the SBS-modified binder, particularly at very low frequencies. This indicates that the complex long-term structural effects

arising from the polymeric structure of the SBS additive cannot be fully represented by the classical CA model. It can be said that the effects of polymer modification on the long-term deformation behavior of the binder limit the predictive capacity of the model. Nevertheless, when considering the entire frequency range, it can be stated that the CA model accurately represents the general trends for both the pure and SBS-modified binders, indicating a high level of model usability in this context. Indeed, the high R^2 values and low mean square errors obtained in the previously presented ANOVA analyses also statistically support this situation. The CA Model parameters are presented in Table 4.

Table 4. CA Model parameters

		Value	Standart Hata	t-Value	Prob>|t|	Dependency
Pure	G	1.00E+09	0	--	--	0
	wc	12794.11	379.4157	33.72056	0	0.99745
	R	2.26902	0.01128	201.0977	0	0.99745
4SBS	G	1.00E+09	0	--	--	0
	wc	108.4948	13.8336	7.84285	1.83E-09	0.99817
	R	3.2481	0.03243	100.1612	0	0.99817

The Christensen-Anderson (CA) model parameters presented in Table 4 numerically reveal the rheological characteristics of pure and 4% SBS-modified asphalt binders. The characteristic frequency (wc) and rheological index (R) parameters included in the model play a critical role in defining the viscoelastic behavior of the binder. The wc value is calculated at a high level in the pure binder, indicating that the material reaches the elastic-viscous transition region at higher frequencies. In contrast, the wc value decreases significantly after SBS modification, indicating that the elastic properties of the binder shift to the low-frequency region, meaning it can maintain its elastic behavior even under prolonged loading conditions. When evaluated in terms of the rheological index (R), the R value was 2.269 in the pure binder, while it increased to 3.248 after SBS modification. As is known, the R parameter is a structural indicator that defines the flow and deformation behavior of the binder in the viscoelastic transition region, and an increasing R value indicates that the elastic-viscous transition behavior of the binder is more pronounced and controlled, thereby increasing the deformation resistance of the material. The high t-statistics and very low p-values ($p < 0.05$) calculated for all parameters indicate that the model parameters are statistically significant and reliable. Additionally, the fact that the R^2 and adjusted R^2 (Adj. R^2) values are above 0.997 for both adhesives indicates that the CA model represents the relationship between complex modulus and frequency with extremely high accuracy.

4. Conclusions

The results obtained show that the rheological behavior of asphalt binders can be significantly improved, particularly through polymer modification, and that this effect can be clearly demonstrated using main curves based on the Time-Temperature Superposition (TTSP) principle. Within the scope of the study, it was determined that the SBS additive improves deformation resistance by increasing the complex modulus of the binder in the low-frequency (long-term loading) region, thereby positively affecting high-temperature performance. The main curve results reveal that the modified binder exhibits higher stiffness than the pure binder, especially at low frequencies, indicating that the SBS additive plays an effective role in making the asphalt mixture more resistant to problems such as permanent deformation (rutting).

Analyses conducted using the Christensen-Anderson (CA) model demonstrate that both pure and SBS-modified binders can be accurately represented in terms of their viscoelastic behavior across a wide frequency range. The high R^2 values and low mean square errors obtained from the ANOVA analysis indicate that the model is statistically robust and that the complex modulus-frequency relationship of the binders can be reliably predicted. However, the significant decrease in the characteristic frequency (wc) value after SBS addition indicates that the elastic properties of the modified binder shift to the low-frequency region, meaning that it can exhibit elastic behavior even under long-term loading conditions. Additionally, the increase in the rheological index (R) value indicates that the elastic-viscous transition behavior of the modified binder occurs in a sharper and more controlled manner.

Particularly in the low-frequency region, the inability of the CA model to fully match the experimental data for the SBS-modified binder indicates that the complex long-term structural effects created by polymer modification can only be represented to a limited extent by the classical CA model. This situation indicates the need for additional model development studies to accurately interpret the current model parameters for binders containing polymer modification.

5. References

[1] R. Hunter, A. Self, and J. Read, *The Shell Bitumen Handbook, 6th edition*. 2015.

[2] J. Zhang, G. S. Simate, X. Hu, M. Souliman, and L. F. Walubita, "Impact of recycled asphalt materials on asphalt binder properties and rutting and cracking performance of plant-produced mixtures," *Construction and Building Materials*, vol. 155, pp. 654–663, 2017.

[3] M. Lagos-Varas *et al.*, "Influence of limestone filler on the rheological properties of bituminous mastics through susceptibility master curves," *Construction and Building Materials*, vol. 231, p. 117126, 2020.

[4] M. Lagos-Varas *et al.*, "Study of the permanent deformation of binders and asphalt mixtures using rheological models of fractional viscoelasticity," *Construction and Building Materials*, vol. 260, p. 120438, 2020.

[5] M. Baqersad and H. Ali, "Rheological and chemical characteristics of asphalt binders recycled using different recycling agents," *Construction and Building Materials*, vol. 228, p. 116738, 2019.

[6] Y. Sun, W. Wang, and J. Chen, "Investigating impacts of warm-mix asphalt technologies and high reclaimed asphalt pavement binder content on rutting and fatigue performance of asphalt binder through MSCR and LAS tests," *Journal of Cleaner Production*, vol. 219, pp. 879–893, 2019.

[7] E. Chailleux, G. Ramond, C. Such, and C. de La Roche, "A mathematical-based master-curve construction method applied to complex modulus of bituminous materials," *Road Materials and Pavement Design*, vol. 7, no. sup1, pp. 75–92, 2006.

[8] N. I. M. Yusoff, F. M. Jakarni, V. H. Nguyen, M. R. Hainin, and G. D. Airey, "Modelling the rheological properties of bituminous binders using mathematical equations," *Construction and Building Materials*, vol. 40, pp. 174–188, 2013.

[9] D. W. Christensen Jr. and D. A. Anderson, "Interpretation of dynamic mechanical test data for paving grade asphalt," in *Asphalt Paving Technology: Association of Asphalt Paving Technologists-Proceedings of the Technical Sessions*, 1992, vol. 61, pp. 67–116.

[10] M. Marateanu and D. Anderson, "Time-temperature dependency of asphalt binders-an improved model," in *Asphalt Paving Technology: Association of Asphalt Paving Technologists-Proceedings of the Technical Sessions*, 1996, vol. 65, pp. 408–448.

Effect of Reactive Polymer on High Temperature Performance of Crumb Rubber Modified Bitumen

Tuba BİNGÖL[1], Baha Vural KÖK [2]

[1,2]Civil Engineering Department, Engineering Faculty, Fırat University, Elazığ, Türkiye.
[1]tubabingol2323@gmail.com, [2]bvural@firat.edu.tr
[1](ORCID: 0009-0000-2552-6488), [2](ORCID: 0000-0002-7496-6006)

Abstract
Bituminous mixtures have recently been mostly produced with various additives to resist the increasing volume of traffic and the corrosive effects of the environment. Economically and sustainably advantageous, crumb rubber (CR) recycled from used vehicle tires has been widely used as a modifier in this context. However, in order to achieve high performance, a high amount of CR is required, which causes compatibility and workability problems with bitumen, so a second additive is used in addition to a reasonable amount of CR. In this study, the effects of CR modification combined with the use of an additive based on isocyanate (B2Last), a reactive polymer, on the high temperature performance of bitumen were investigated. Since heat and time are required for the reaction between isocyanate and bitumen to take place, the binders were subjected to short-term aging (RTFO), thus also taking into account the condition to which the bitumen is exposed in real practice. Softening point and rotational viscosity tests were performed on the original and short-term-aged binders. The results showed that short-term aging significantly increased the bitumen-CR-B2last interaction. The addition of B2Last improved the high temperature resistance of the CR modification, resulting in a significant performance improvement.

Keywords: bitumen, modification, crumb rubber, reactive polymer

1. Introduction

The storage of waste materials and their disposal without recycling brings many problems. In order to improve the properties of bituminous mixtures, various additives as well as waste materials are added to the bitumen or bituminous mixture. In this context, Crumb Rubber (CR) obtained from waste vehicle tires containing valuable polymers is widely used in bituminous mixtures. CR is added to bituminous mixtures as a wet or dry process.

In the dry process, which is less preferred due to its limited advantages, CR particles are added directly to the mix as part of the aggregate. It is stated that the void ratio increases with the increase in CR ratio in the dry process, this increase causes a rebound effect during compaction due to the flexibility of CR particles during compaction, causing the mixture to be less dense and that the optimum CR ratio by weight of the mixture is 1% [1] [2]. Tahami et.al. stated that dry processing has more CR consumption and thus environmental benefits, but it has not attracted much attention due to its lower performance compared to wet processing [3]. In the study investigating the modification of bituminous mixture with CR using dry process by surface morphology, it was found that microscopically, in the rubber-bitumen interaction, carbon black passes from the ground rubber to the bitumen, giving the bitumen a darker color and yellow fluorescent areas appear around the rubber particles, and the most suitable particle size for dry process is less than 200 μm. [4]. Junan Shen et.al. investigated the interaction between CRM and asphalt binder in asphalt mixtures to which CRM was added by dry process during the time period after mixing and before paving, including storage in the silo. As a result, they observed that the large molecular size (LMS) value increased with increasing storage time. They observed that the LMS value obtained from the wet process was much higher than the LMS value obtained from the dry process without storage. They explained that this was due to the fact that the CRM-modified binder was fully mixed for 45 minutes in the wet process and the binder experienced some aging as well as CRM and binder interaction [5].

It has been recommended to limit the CR content to no more than 10% of the binder weight for adequate workability in wet processing [6]. It has been reported that the addition of rubber improves the asphalt-aggregate interface, is suitable for countries with tropical climates to combat moisture damage, improves the skid resistance of bituminous mixtures, offers better resistance to permanent deformation at high temperature and cracking at moderate temperature [7] [8] [9].

[1]Corresponding author

Studies show that CR is a passive admixture and has a physical interaction with bitumen. Recently, additives that react with bitumen and are called “active additives” have come to the agenda. Polyurethanes (PU) are the most important of these. Polyurethanes are polymeric materials formed by the combination of various chemical groups and generally have elastomeric properties. Polyurethane consists of soft and hard segments. This helps to determine the mechanical properties of the polymer and its application areas. The soft segment usually refers to a part that provides flexibility and elasticity, while the hard segment represents the more durable and rigid part. The soft segments come from polyols, while the hard segments are composed of isocyanates (NCO) or small molecule chain extenders. The combination of soft and hard segments changes the structure within the polyurethane, the microstructure between them and the interaction between polyurethane and asphalt. [10].

Liu H. et.al investigated the effect of soft segment length on the performance of polyurethane modified asphalt (PUMA) using different polytetramethylene ether glycol (PTMEG) prepolymers and observed that high and low temperature properties improved with increasing soft segment length [11]. PU/epoxy resin composite modified asphalt has sufficient storage stability, high temperature properties and mechanical strength, and the C-O-C group in PU improves the flexibility of the binder [12]. The addition of PU and modification with a higher content of hard segments significantly increases the fatigue life of asphalt and improves its recovery properties, i.e. its flexibility, fatigue resistance and complex modulus. Gong X. et.al reported that environmentally curable CO_2-based polyurethane (PPC-PU) shows good thermal stability and gradual curing ability and can be synthesized to improve asphalt properties [13]. Jiang W. et.al emphasized that polyurethane modified asphalt (PA) prepared using poly-tetra methylene ethylene ether glycol (PTEMG) and 4,4'-diphenylmethane diisocyanate (MDI) has become a sustainable road material, showing improved thermal performance and cryogenic stability, PA has the advantages of high temperature stability, low temperature cracking resistance and long fatigue life [14]. Jin X. et.al focused on analyzing the fatigue self-healing properties of thermoplastic polyurethane (TPU) modified asphalt using dynamic shear rheometer tests and showed that increasing hard segment content and isocyanate radical index can improve the shear deformation resistance and self-healing degree [15].
In this study, it was determined how the effectiveness of the CR modification changes with the addition of a reactive polymer additive (B2Last) with MDI as the main component. The high temperature and flow properties of CR and CR+B2Last modifications were evaluated.

2. Materials and Method

2.1 Materials

B 70/100 grade bitumen obtained from Tüpraş Batman refinery was used as pure bitumen in the study. The properties of the pure binder are presented in Table 1. Crumb rubber (CR) was obtained from Samsun Akın Plastik Company and has a granular structure with typical diameters ranging from 0.3 to 0.6 mm.

Another additive used in the study is B2Last, a reactive polymer from BASF. At ambient temperature, B2Last appears as a low viscosity, black liquid. This additive is both water and solvent free. The density of B2Last was measured to be 1.2 g/cm³ and its boiling point exceeds 200°C, ignition temperature exceeds 530°C and thermal decomposition temperature exceeds 230°C. The additives used in the study are shown in Figure 1.

Table 1. Properties of pure binder

Properties	Standard	B 70/100
Penetration (0.1 mm), 100 g, 5 s	ASTM D5	85.10
Softening point (°C)	ASTM D36	52.65
Penetration index (PI)		0.87
Viscosity (cP, 135 °C)	ASTM D4402	562.50
Viscosity (cP, 165 °C)	ASTM D4402	175.00
Specific gravity	ASTM D70	1.038

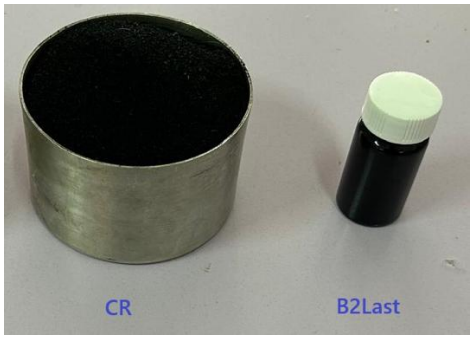


Figure.1 The additives used in the study

2.2 Preparation of Modified Asphalt Binders

The CR and B2Last modifications were performed using a high shear rate mixer housed in a cylindrical insulated chamber (Figure 2). The CR modification process requires high mixing speed and shear action for effective particle disintegration. To ensure that pure, CR and CR+B2Last modified binders were subjected to similar levels of oxidation during the mixing process, the mixing parameters set for CR modification were used for all binders. The mixtures were obtained at a speed of 2,000 rpm and a temperature of 180℃ during a mixing process lasting 90 minutes. Given that B2Last is a liquid additive, it was mixed with pure bitumen and CR modified bitumen for as little as 2 minutes to ensure proper homogenization.

The pure binder was modified by the addition of CR and B2Last. Specifically, the CR modification included addition rates of 6% and 8% by weight of bitumen, indicated by 6CR and 8CR respectively. In cases where CR and B2last were combined, the B2Last content was kept at a constant 2%, while the CR content varied between 6% and 8%. As a result, the binder containing 6% CR and 2% B2last is designated 6CR+2B, while the binder containing 8% CR and 2% B2last is designated 8CR+2B.

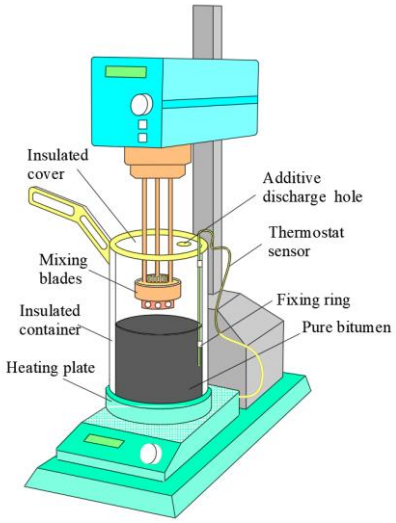


Figure 2. High shear mixer

In order for the B2Last additive to interact effectively with bitumen, it must be exposed to high temperatures for a certain period of time. In polyurethane modification, the curing time under a certain temperature is important in the consumption of isocyanates, and the temperature leads to the gradual formation of polyurethane cross-linked network structures, resulting in improved performance [16].
In this research, a short-term aging procedure was applied to facilitate the binders' exposure to heat. This approach ensures the interaction between B2Last and bitumen while at the same time taking into account the practical situation of the

bitumen itself. Short-term aging of the binders was carried out by rotary thin film oven (RTFO) testing at 163°C for 85 minutes in accordance with ASTM D2872. The study examined the effectiveness of both pure and modified binders in their original unaged state and after RTFO aging.

2.3. Softening Point Test

It is an experiment performed to determine the softening temperature of the bituminous binder. This experiment aims to measure the sensitivity of bitumen to temperature changes. Bitumen samples are filled into rings placed on a pre-lubricated plate. After cooling for one hour, the sample is cooled for one hour, the overflow is smoothed and placed in the experimental apparatus (Figure 3). When water is used, the starting temperature is 5°C. If the softening point exceeds 80°C, glycerin is used and the initial temperature is set at 32°C. Balls are dropped on the bitumen samples from a height of 2.5 cm and the temperature is increased by 5°C every minute during the experiment. The value seen on the thermometer when the softening bitumen sample touches the bottom of the container is considered as the softening point temperature.

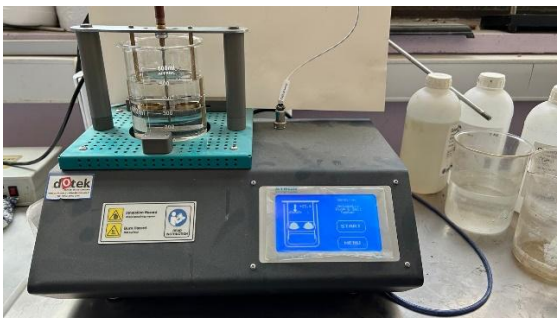


Figure 3. Softening point test device

2.4. Rotational Viscosity Test

The rotational viscosity test is performed to evaluate the high temperature viscosity properties of bituminous binders. In this study, the viscosities of pure and modified binders were determined using a Brookfield Rotational Viscometer at 135 and 165°C in accordance with ASTM D4402. High viscosity is associated with improved resistance to permanent deformation; however, the viscosity properties of bituminous binders also play an important role during operations such as pumping, mixing and compaction. In this context, it is important that the viscosity of the binders at 135°C does not exceed 3 Pa.s (3000 cP) to ensure adequate workability. In the experiment, 11g of binder was placed in the tube shown in Figure 4 and viscosities were measured with a rotating head at 20 rpm.



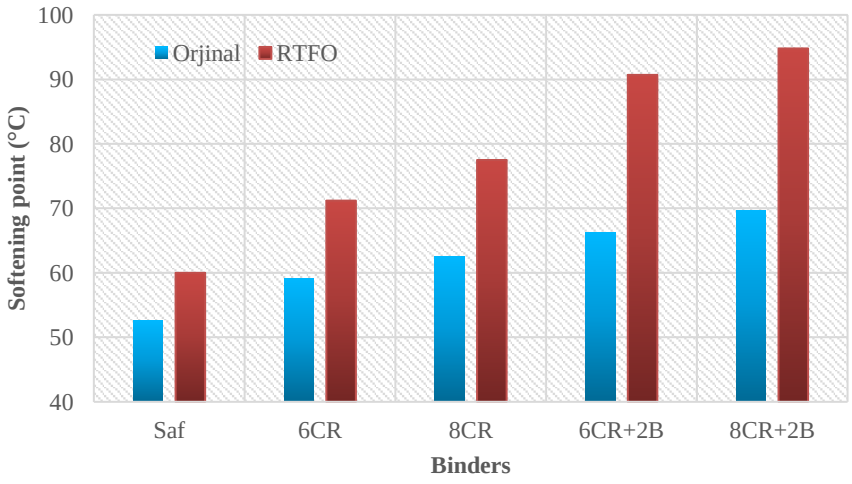
Figure 4. Rotational viscosity test device

3. Experimental Results

The softening point values of the binders in the original state and after aging are given in Figure 5. The binders containing B2Last after RTFO were tested in glycerin. Increasing the CR content increased the softening point both in the original state and after aging. Increasing the CR content from 6% to 8% increases the softening point by 3.4°C. The high softening point of the B2Last modified samples indicates that there is some bitumen-isocyanate interaction due to the experiment in water and the temperature increase. Although there is not a complete reaction here, the increase in softening point confirms the effect of water and temperature. In the original case, the 6CR+2B binder gives a softening point value 7.1°C higher than 6CR and the 8CR+2B binder 7.2°C higher than 8CR.

Compared to the pure binder in the original state, 6CR+2B and 8CR+2B binders have 25% and 32% higher softening point values, respectively. The softening point values of all binders increase after aging. This increase is significantly higher in binders containing B2Last. While the softening point of the pure binder increases by 14% after aging, this increase is 37% in the 6CR+2B binder. Considering that all of the softening point increase in the pure binder is due to oxidation, it can be said that the 33% increase obtained by subtracting this rate (14%) from the 6CR+2B binder is due to the B2Last-CR-bitumen reaction.

The B2Last-bitumen interaction increases the consistency of the bitumen and increases the softening tendency of the bitumen at higher temperatures, thus improving the high temperature resistance. With a softening point value of 90.7°C after aging, the 6CR+2B binder is 19.5°C higher than 6CR. The softening point of 8CR+2B binder after aging is 17.3°C higher than that of 8CR. It was found that the high temperature resistance of binders containing B2Last improved significantly after aging.



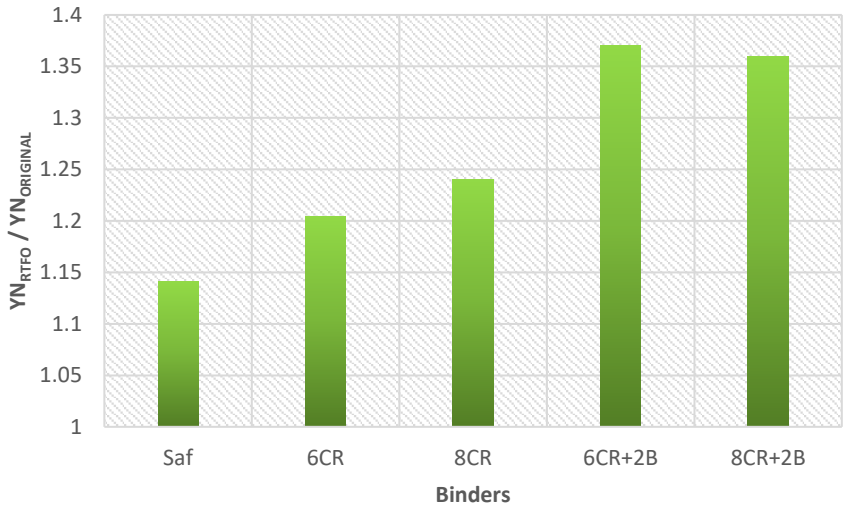


Figure 5. Softening point values of binders in original condition and after aging

In order to determine the viscosity characteristics of the binders, rotational viscosity tests were first performed on the unaged binders at 135 and 165°C. The results obtained are given in Figure 6. In the original case, 6% and 8% CR modification results in 2.82 and 3.07 times viscosity increase at 135°C compared to the pure binder. Here, the hydrocarbons of the polymeric chains of CR interact with the aromatic component of bitumen and the CR particles absorb the aromatic oil, swell and reduce the oil content in bitumen. As the oil content in the bitumen decreases, the viscosity increases and a thick layer of bitumen forms around the CR particles [17]. The modified binders did not pose a problem in terms of workability by not exceeding the limit value of 3000 cP at 135°C.

Addition of 2% B2Last to 6CR binder causes 25% and 12% viscosity increase at 135 and 165°C, respectively. Addition of 2% B2Last to 8CR binder causes 21% and 11% viscosity increase at 135 and 165°C, respectively. These increases are thought to be caused by some reaction of the -NCO component with the hydroxyl groups of bitumen and CR due to the effect of temperature and duration of the experiment.

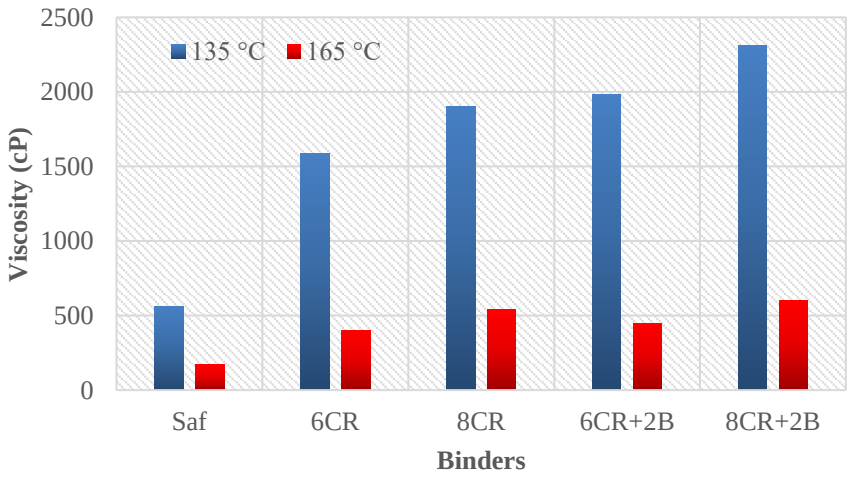


Figure 6. The viscosity values of unaged asphalt binders

The viscosity values of the binders after short-term aging are given in Figure 7. The maximum viscosity that the test equipment can measure is 12500 cP. Since the viscosity values of the binders containing B2Last after aging at 135 °C were above this limit value, the experiment was carried out at 140 and 165 °C. Here, a significant viscosity increase in B2Last

added binders is observed with the curing after aging. At 140°C, 6C2B and 8C2B binders have 14 and 15.6 times higher viscosity values than pure binder, respectively. These ratios become 11.6 and 11.8 at 165°C.

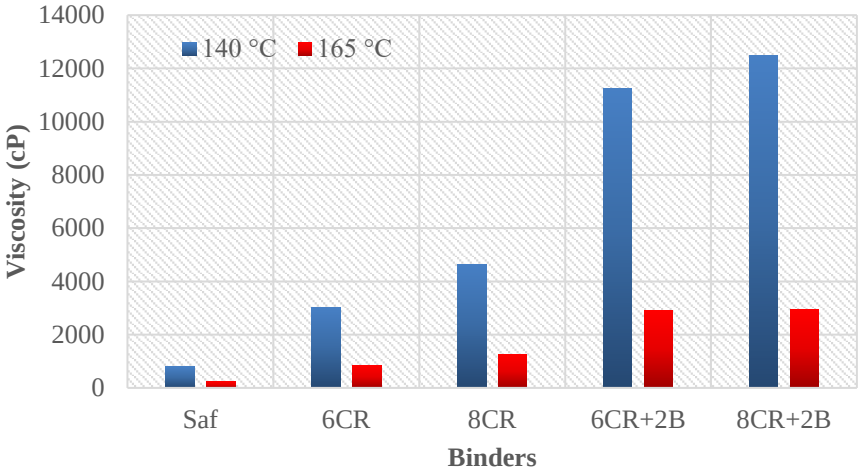
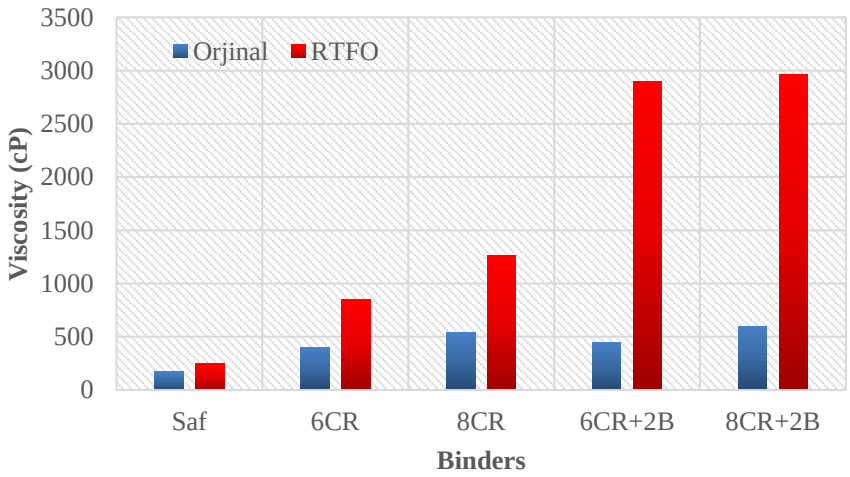


Figure 7. The viscosity values of RTFO-aged asphalt binders

Figure 8 shows the original and post-RTFO viscosity values of the binders at 165°C and the increase in viscosity with aging. Significant change occurs in B2Last-added binders with aging. It can be said that the B2Last content does not lack the hydroxyl group required for the NCO bitumen reaction, in other words, there are enough hydroxyl groups in the environment to fully react with NCO, thus realizing an advanced polymer network. This increase is 6.4 and 5 times in 6CR+2B and 8CR+2B binders after aging. The increase in 6CR+2B and 8CR+2B binders indicates that NCO interacts with CR as well as bitumen to develop a stronger polymer network.



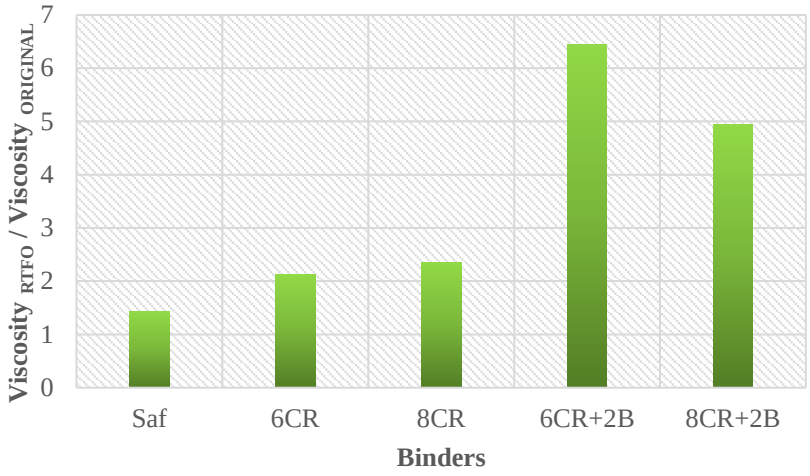


Figure 8. Change in viscosity values of binders at 165℃ after RTFO-aging

4. Conclusion

This study investigated how the effectiveness of CR modification changes with the addition of isocyanates. The high temperature performance of these combined binders was evaluated by adding 2% B2last to bitumen modified with 6% and 8% CR by weight.

The addition of 6% and 8% CR to the bitumen resulted in a 12% and 18% increase in softening point compared to the pure binder, respectively. The use of 2% B2last with 6% CR resulted in a 25% increase compared to the pure binder and a 12% increase compared to the 6%CR modification. The use of 2% B2last with 8% CR resulted in a softening point value 32% higher than the pure binder and 11% higher than the binder containing 8%CR. The use of B2last in combination with CR modification significantly increased the temperature level of the viscoelastic behavior of the binder. The reaction between the isocyanate and the polar functional compounds of bitumen and CR at the end of the short-term aging procedure resulted in significant increases in the softening point values. The combination of 6%CR+%2B2Last after RTFO gave a softening point value 35% higher than that of the pure binder, offering significant resistance to permanent deformations.

CR modification causes increased viscosity due to the swelling of CR particles by absorbing aromatic oils in the bitumen. 6% CR and 8% CR modifications increased the viscosity of the pure binder at 135 ºC by 2.82 and 3.07 times, respectively. The use of B2last in combination with CR did not significantly affect the viscosity of binders doped with CR only in the original case. The viscosity of the 6%CR+2B2last combination was 25% higher than that of the 6% CR modification in the original condition. The viscosity of the 8CR+2B2last combination is 21% higher than that of the 8% CR modification in the original condition. However, after RTFO, the viscosities of the CR modified binders containing B2Last increased significantly as a result of the reaction. The 6%CR+2B2last modified binder gave a viscosity value 11.6 times higher than the pure binder and 3.4 times higher than the 6% CR modified binder after RTFO. The use of B2Last in combination with CR offered significant temperature resistance. According to the viscosity values at 135℃, none of the modified binders exceeded the 3000 cP limit value and did not pose a problem in terms of workability.

The use of CR in bitumen modification provides an economical increase in performance, while also contributing to the conservation of natural resources and ensuring significant sustainability. The use of as low as 2% B2Last in combination with CR increases the high temperature performance of bitumen to very high levels.

6. Acknowledgments

This article is extracted from the first author's master thesis entitled "Investigation of Structures Produced From Historical Mortars Using Fracture Mechanics Principles", supervised by Baha Vural KÖK. This study was carried out within the

scope of the Scientific and Technological Research Council of Türkiye (TUBITAK). We gratefully acknowledge the financial support provided by TUBITAK to Research Project 124M464.

5. References

[1] Tan Q, Zhu H, Yang S, Zhao H, Yang X, Wang Y, et al. Tannic acid/γ-aminopropyltriethoxysilane/molybdenum disulfide composite prepared by co-deposition as an environment-friendly modifier for crumb-rubber-modified asphalt. Constr Build Mater 2024;456:139191. https://doi.org/10.1016/j.conbuildmat.2024.139191.
[2] Moreno F, Rubio MC, Martinez-Echevarria MJ. Analysis of digestion time and the crumb rubber percentage in dry-process crumb rubber modified hot bituminous mixes. Constr Build Mater 2011;25:2323–34. https://doi.org/10.1016/j.conbuildmat.2010.11.029.
[3] Tahami SA, Mirhosseini AF, Dessouky S, Mork H, Kavussi A. The use of high content of fine crumb rubber in asphalt mixes using dry process. Constr Build Mater 2019;222:643–53. https://doi.org/10.1016/j.conbuildmat.2019.06.180.
[4] López-Moro FJ, Moro MC, Hernández-Olivares F, Witoszek-Schultz B, Alonso-Fernández M. Microscopic analysis of the interaction between crumb rubber and bitumen in asphalt mixtures using the dry process. Constr Build Mater 2013;48:691–9. https://doi.org/10.1016/j.conbuildmat.2013.07.041.
[5] Sheng X, Wang M, Xu T, Chen J. Preparation, properties and modification mechanism of polyurethane modified emulsified asphalt. Constr Build Mater 2018;189:375–83. https://doi.org/10.1016/j.conbuildmat.2018.08.177.
[6] Khalid HA, Artamendi I. Performance Based Characterisation of Crumb Rubber Asphalt Modified using the Wet Process. Road Mater Pavement Des 2003;4:385–99. https://doi.org/10.1080/14680629.2003.9689955.
[7] Poovaneshvaran S, Zheng LW, Hasan MRM, Yang X, Diab A. Workability, compactibility and engineering properties of rubber-modified asphalt mixtures prepared via wet process. Int J Pavement Res Technol 2021;14:560–9. https://doi.org/10.1007/s42947-020-1006-z.
[8] Chen R, Zhu H, Kong L, Xu Y, Ou L. Stage-aging characteristics and stages division of crumb rubber modified asphalt binder. Constr Build Mater 2023;367:129712. https://doi.org/10.1016/j.conbuildmat.2022.129712.
[9] Kök BV, Çolak H. Laboratory comparison of the crumb-rubber and SBS modified bitumen and hot mix asphalt. Constr Build Mater 2011;25:3204–12. https://doi.org/10.1016/j.conbuildmat.2011.03.005.
[10] Shen J, Amirkhanian S, Xiao F, Tang B. Influence of surface area and size of crumb rubber on high temperature properties of crumb rubber modified binders. Constr Build Mater 2009;23:304–10. https://doi.org/10.1016/j.conbuildmat.2007.12.005.
[11] Liu H, Zhang Z, Zhang S, Chang P, Liang Y, Wang Z, et al. Study on the effect of soft segment length on the performance of polyether-based polyurethane modified asphalt. Int J Adhes Adhes 2024;130:103642. https://doi.org/10.1016/j.ijadhadh.2024.103642.
[12] Wang C, Huang S, Chen Q, Ji X, Duan K. Materials, preparation, performances and mechanism of polyurethane modified asphalt and its mixture: A systematic review. J Road Eng 2023;3:16–34. https://doi.org/10.1016/j.jreng.2023.01.002.
[13] Gong X, Liu Q, Wang H, Wan P, Chen S, Wu J, et al. Synthesis of environmental-curable CO2-based polyurethane and its enhancement on properties of asphalt binder. J Clean Prod 2023;384:135576. https://doi.org/10.1016/j.jclepro.2022.135576.
[14] Jiang W, Zhang M, Ren P, Xing C, Yuan D, Wu W. Development of porous asphalt mixture based on the synthesis of PTEMG and MDI polyurethane asphalt. Constr Build Mater 2024;411:134537. https://doi.org/10.1016/j.conbuildmat.2023.134537.
[15] Jin X, Li D, Fu H, Sun B, Yang Y, Guo N, et al. Research on fatigue self-healing of asphalts modified by thermoplastic polyurethane (TPU) with different soft segments. Constr Build Mater 2024;420:135574. https://doi.org/10.1016/j.conbuildmat.2024.135574.
[16] Luo Z, Duan J, Cao Z, Yi J, Xu S, Wang R. Environmentally friendly high-content polyurethane modified asphalt: Workability evaluation and early performance evolution. Constr Build Mater 2024;453:139093. https://doi.org/10.1016/j.conbuildmat.2024.139093.
[17] Karakurt C. Microstructure properties of waste tire rubber composites: an overview. J Mater Cycles Waste Manag 2015;17:422–33. https://doi.org/10.1007/s10163-014-0263-9.

Effect of Isocyanate and Crumb Rubber on Stability Properties of Bituminous Mixtures

Furkan YILDIRIM[1], Baha Vural KÖK[21]

[1]Civil Engineering, Engineering Faculty, Bursa Technical University, Bursa, Türkiye.
[2]Civil Engineering, Engineering Faculty, Firat University, Elazig, Türkiye.
[1]furkan.yildirim@btu.edu.tr, [2]bvural@firat.edu.tr
[1](ORCID: 0009-0004-6534-7416), [2](ORCID: 0000-0002-7496-6006)

Abstract
This study investigates the effects of styrene-butadiene-styrene (SBS), crumb rubber (CR), and a reactive polymer (MDI-based) on the mechanical and stability properties of bituminous mixtures. Modified binders containing 4% SBS, 8% CR, and a combination of 8% CR + 2% B2Last (reactive polymer) were prepared and used to produce hot mix asphalt specimens. The volumetric properties and Marshall stability of the mixtures were evaluated. Results indicate that all additives improved the stability of the mixtures compared to pure asphalt, with the combination of CR and reactive polymer achieving the highest stability and stiffness. Specifically, the CR + B2Last mixture exhibited 67%, 36%, and 47% higher stability than the pure, CR, and SBS mixtures, respectively. The findings demonstrate that incorporating a reactive polymer with CR significantly enhances the mechanical performance and deformation resistance of bituminous mixtures.

Keywords: bituminous mixtures, crumb rubber, sbs, reactive polymer, stability.

1. Introduction

Although bitumen is present in bituminous mixtures in relatively low volumetric proportions, it has a critical impact on the overall performance of pavements due to its rheological, physical, and chemical characteristics. Bitumen, which exhibits viscoelastic and thermoplastic properties, is inherently a temperature-sensitive material [1,2]. At high temperatures, bitumen behaves like a Newtonian fluid; in the intermediate temperature range, it acts as a viscoelastic liquid; and at low temperatures, it exhibits brittle, viscoelastic solid behavior. This increases the risk of cracking at low temperatures [3]. The performance of bitumen can be adversely affected by external factors such as temperature variations, aging over time, traffic loads, environmental and climatic conditions, ultraviolet (UV) radiation, and oxidative processes [4,5]. Such effects lead to structural damage in flexible pavements over time and shorten their service life.

One of the most commonly applied methods to enhance pavement strength and extend its service life is the modification of bitumen. In this context, polymer additives, particularly thermoplastic elastomers such as styrene-butadiene-styrene (SBS), have been widely preferred in recent years. Studies have shown that the interaction between SBS and asphalt occurs at a physical rather than a chemical level. However, since high molecular weight thermoplastic rubbers, similar to asphaltenes, compete for solubility in the maltene phase, phase separation may occur when the maltene content is low [6][7]. While SBS improves the high-temperature resistance [8] and fatigue life [9] of mixtures, increasing the additive content also leads to compatibility issues with bitumen and higher costs. To overcome these disadvantages, research has focused on alternative additives that are less expensive, more compatible with bitumen, and offer higher aging resistance. One such alternative is crumb rubber (CR), obtained from waste vehicle tires. [10]. To overcome these disadvantages, research has focused on alternative additives that are more cost-effective, more compatible with bitumen, and offer high aging resistance. One such alternative is crumb rubber (CR), obtained from waste vehicle tires.

The addition of CR extends the service life of pavements and enhances their resistance to cracking, offering the potential to reduce maintenance and repair costs [11] [12]. Moreover, the use of waste tires not only contributes to environmental sustainability but also provides benefits such as improved surface durability and reduced tire pavement noise [13][14]. Studies have shown that CR contents above 8% can improve the performance of bitumen, both before and after aging, by up to 1–2 degrees [15].

Additives that chemically interact with specific fractions of bitumen and prevent phase separation are defined in the literature as “active additives” [16]. The growing global interest in conserving natural resources has brought the use of methylene diphenyl diisocyanate (MDI) based prepolymers, classified as active additives, into focus as an alternative to

[1]Corresponding author

conventional petrochemical derivatives [17]. The highly reactive functional groups (isocyanate, NCO) of MDI based prepolymers form chemical bonds by reacting with multifunctional active hydrogen compounds present in bitumen, particularly the –OH groups found in the asphaltene structure [18–20]. To produce polyurethane (PU), MDI has been used as a curing agent and combined with different types of polyols.

Zhang et al. reported that the low-temperature properties and mechanical durability of bituminous binder were significantly improved when modified with polyurethane (PU) [21]. Liu et al. developed a waste polyethylene terephthalate (PET) enhanced PU modified asphalt to reduce both PET accumulation and the dependency of PU-modified asphalt on petrochemicals. They concluded that polyols obtained through the depolymerization of waste PET with polyols can be used in the production of PU-modified asphalt and improve both the high- and low-temperature performance compared to pure asphalt. They also noted that the binder prepared with waste PET, diethylene glycol, and MDI exhibited the best high-temperature stability, while the binder prepared with waste PET, polyethylene glycol 400, and MDI demonstrated the best low-temperature performance [22].

Wu et al. used a composite polymer material synthesized with polyurethane and polyurea (PU-PUa) to modify asphalt binder. Their results showed that asphalt binders modified with polyurethane and polyurea exhibited significant improvements in strength, ductility, viscoelastic, and rheological properties [23].

Dong et al. employed a composite modifier containing butyl rubber (BR) to enhance the low-temperature performance of asphalt modified with waste polyurethane (WPU). They reported that butyl rubber and waste polyurethane effectively improved both the high and low temperature performance of asphalt, particularly enhancing its low temperature behavior; however, excessive butyl rubber could reduce the colloidal stability of the asphalt. Therefore, a combination of 15% waste polyurethane and 3% butyl rubber was identified as the optimal modification ratio [24]. Li et al. aimed to develop an asphalt sealant modified with crumb rubber (CR) and polyurethane (PU). Their results indicated that a sealant containing 30% PU and 15% CR exhibited the best thermal stability and mechanical properties, making it the optimal composition [25].

The main objective of this study is to investigate the effects of SBS (styrene-butadiene-styrene), CR, and MDI additives on the mechanical properties of bituminous mixtures. In particular, the effects of MDI, a reactive polymer used together with CR, on the stability of the mixture were examined in this study.

2. Experimental Study

2.1. Materials

In the production of hot mix asphalt specimens, crushed limestone obtained from the Karayazı region of Elazığ was used as aggregate. The physical properties of the aggregate are presented in Table 1. In the study, B 70/100 grade bitumen, supplied from the TÜPRAŞ Batman Refinery, was used as the base bitumen. The technical properties of the base bitumen are presented in Table 2.

Table 1. Physical properties of the aggregate

Properties	Standard	Limits	Coarse	Fine	Filler
Abrasion loss (%) (Los Angeles)	ASTM C131	Max. 30	29	-	-
Freeze effect (%) (Na_2SO_4)	ASTM C88	Max. 10	4.5	-	-
Flat and elongated particles (%)	ASTM D4791	Max. 10	4		
Water absorption (%)	ASTM C127	Max. 2	1.37		
Specific gravity (g/cm^3)	ASTM C127		2.533	-	-
Specific gravity (g/cm^3)	ASTM C128		-	2.619	-
Specific gravity (g/cm^3)	ASTM D854		-	-	2.732

Table 2. Properties of the pure binder

Properties	Standard	B 70/100
Penetration (0.1 mm), 100 g, 5 s	ASTM D5	77
Softening point (°C)	ASTM D36	52.6
Penetration index (PI)		0.55
Specific gravity	ASTM D70	1.024
Viscosity (cP, 135 °C)	ASTM D4402	562.5
Viscosity (cP, 165 °C)	ASTM D4402	175

In the study, SBS, CR, and B2Last additives were used as modifiers. The SBS in powder form (Kraton D1101) used in the study was supplied by the chemical company Shell. The CR, which was passed through a No. 40 sieve before use, was obtained from Samsun Akın Plastik. The B2Last additive, with an isocyanic acid and polymethylene polyphenyl ester chemical structure and reactive properties, was supplied by İstanbul Teknik. B2Last is in liquid form, black in color, and has an earthy odor.

2.2. Preparation of Modified Bitumen

In the study, three different types of modified bitumen were prepared, containing SBS, CR, and CR+B2Last additives. For this purpose, a high-shear mixer, shown in Figure 1, was used. The modified binders were prepared by mixing at a speed of 2000 rpm, at a temperature of 180 °C, for 90 minutes. SBS was used at 4% by weight of bitumen, and the bituminous mixture produced with this modified binder was designated as 4SBS. CR was used at 8%, and the bituminous mixture produced with this binder was designated as 8CR. B2Last was added at 2% to the 8% CR-modified bitumen, and the mixture prepared with this modified binder was designated as **8C2B**.



Figure 1. High-shear mixer

2.3. Preparation of Hot Mix Asphalt Specimens

In the study, a total of four different types of HMA specimens, including one with unmodified (pure) bitumen, were prepared. For the HMA specimens, the aggregate gradation was selected as the average of the lower and upper limits of the Type-1 wearing course gradation specified in the General Directorate of Highways (KGM) technical specifications. The gradation values are presented in Table 3.

Table 3. Aggregate gradation

Sieve opening (mm)	19.00	12.50	9.50	4.75	2.00	0.425	0.180	0.075
Passing (%)	100	94	81	47	30	15	10	5

To determine the optimum bitumen content of the mixtures prepared with the pure binder, Marshall specimens were prepared by applying 75 blows on each side. The bitumen content was increased from 4.5% to 6.5% in increments of 0.5%. The volumetric properties as well as the stability and flow values of the specimens, three prepared at each bitumen content, were determined. The results obtained are presented in Table 4. Figure 2 graphically shows the variation of these properties with bitumen content.

Table 4. Marshall design test results

Coarse aggregate ratio			%53			**Bitumen specific gravity**			1.024	
Fine aggregate ratio			%42			**Aggregate blend specific gravity**			2.576	
Filler ratio			%5			**Effective specific gravity (Gse)**			2.610	
Wa (%)	**W_{dry} (gr)**	**$W_{saturated}$ (gr)**	**W_{wet} (gr)**	**Gmb**	**Gmm**	**Va(%)**	**Vma (%)**	**Vfa (%)**	**Stability (kN)**	**Flow(mm)**
	1142.60	1151.89	651.40	2.283		6.70	15.22	56.01	11.25	2.05
4.5	1137.94	1146.41	649.15	2.288	2.447	6.47	15.02	56.91	11.52	2.56
	1140.20	1147.1	646.00	2.275		7.00	15.50	54.81	10.92	2.78
Average				**2.282**		**6.72**	**15.25**	**55.91**	**11.23**	**2.46**
	1150.50	1154.07	656.29	2.311		4.91	14.58	66.29	12.37	2.94
5.0	1148.92	1153.22	655.46	2.308	2.431	5.04	14.69	65.69	11.81	3.15
	1153.90	1157.91	659.26	2.314		4.80	14.47	66.85	11.84	3.08
Average				**2.311**		**4.91**	**14.58**	**66.27**	**12.01**	**3.05**
	1154.86	1158.70	663.71	2.333		3.39	14.18	76.09	12.12	3.54
5.5	1158.06	1161.00	664.70	2.333	2.420	3.38	14.17	76.16	12.52	3.11
	1155.59	1159.70	663.90	2.331		3.49	14.27	75.56	12.29	3.65
Average				**2.332**		**3.42**	**14.20**	**75.94**	**12.31**	**3.43**
	1162.17	1165.36	666.53	2.330		2.91	14.71	80.22	11.26	4.07
6.0	1160.42	1163.74	665.22	2.328	2.400	2.99	14.78	79.74	11.14	3.98
	1160.79	1164.08	666.20	2.331		2.84	14.64	80.61	11.52	4.22
Average				**2.330**		**2.91**	**14.711**	**80.19**	**11.31**	**4.09**
	1164.35	1168.00	666.56	2.322		2.62	15.39	82.95	9.77	4.83
6.5	1169.58	1173.00	667.96	2.316	2.380	2.88	15.62	81.53	10.03	5.16
	1165.22	1169.52	665.71	2.313		3.01	15.72	80.87	9.09	5.2
Average				**2.317**		**2.84**	**15.57**	**81.79**	**9.63**	**5.06**

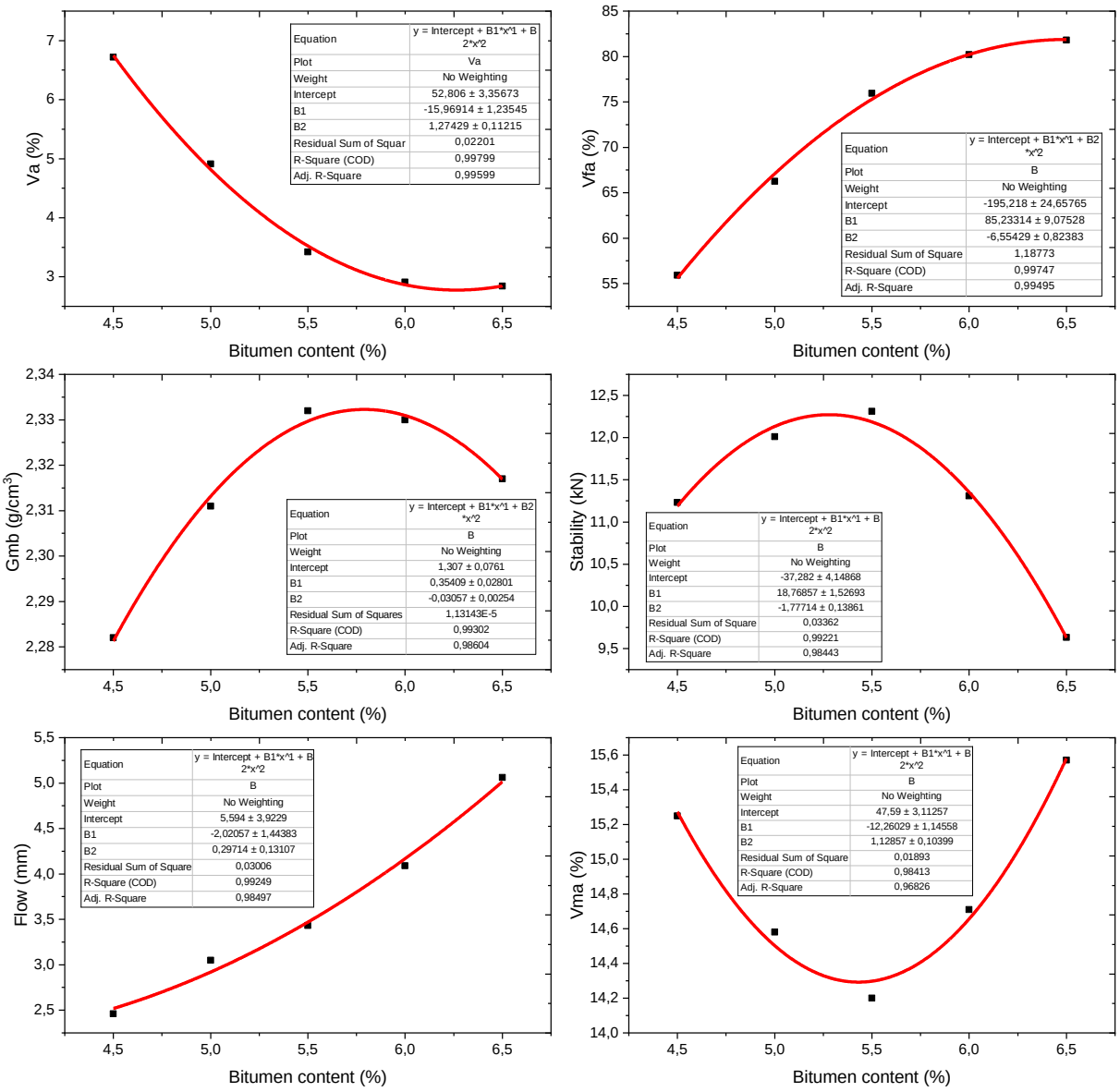


Figure 2. Relationship between bitumen content and physical properties

According to the highways technical specifications, the bitumen contents that satisfy the criteria of maximum stability and Gmb, 4% air voids, and 70% voids filled with bitumen are presented in Table 5.

Table 5. Bitumen contents meeting the specification criteria

Bitumen content providing maximum stability	5.28
Bitumen content providing maximum Gmb	5.78
Bitumen content providing 4% Va	5.27
Bitumen content providing 70% VFA	5.19
Average	**5.38**
Flow value at 5.38% bitumen content (mm)	3.32
VMA value at 5.38% bitumen content (%)	14.29

At the bitumen content determined according to the Marshall method, three additional specimens were prepared, and their volumetric properties, stability, and flow were determined. The results obtained are presented in Table 6.

Table 6. Properties of specimens prepared with pure binder at optimum bitumen content

Wa (%)	W_{dry} (gr)	$W_{saturated}$ (gr)	W_{wet} (gr)	Gmb	Gmm	Va(%)	Vma (%)	Vfa (%)	Stability (kN)	Flow (mm)
5.38	1153.24	1157.65	662.04	2.327	2.419	3.80	14.31	73.47	12.17	3.19
	1152.87	1157.13	661.12	2.324		3.91	14.41	72.90	12.54	3.57
	1152.15	1156.66	660.90	2.324		3.92	14.42	72.83	12.28	3.30
		Average				**3.87**	**14.38**	**73.07**	**12.33**	**3.35**

It was determined that the specimens prepared with 5.38% bitumen content had void ratios between 3% and 5%, VFA values between 65% and 75%, and flow values between 2 mm and 4 mm, and that they met the minimum VMA requirement of 14% and the minimum stability requirement of 900 kg.

2.4. Marshall Stability and Flow Test Results

The stability of the mixture represents its capacity to resist deformation under heavy vehicle loads. Therefore, sufficient stability is required to construct an effective pavement that meets specific traffic demands. As seen in Table 7, the stabilities of the mixtures meet the minimum specification criteria. However, the use of additives provides significant improvements in both stability and flow. The increase in the Marshall stability value indicates an improvement in the asphalt mixture's ability to resist rutting and friction under heavy traffic loads [26]. Figure 3 presents the stability and flow values of the mixtures, as well as the stability-to-flow ratio, which is an indicator of stiffness. The 8C2B mixture exhibits stability values that are 67%, 36%, and 47% higher than the pure, 8CR, and 4SBS mixtures, respectively. Considering that all mixture types have the same gradation and similar air void content, it can be stated that the increase in stability with the use of additives results from the increased cohesion of the binder in the mixtures. It was observed that the 8CR and 4SBS mixtures showed similar performance, whereas the addition of 2% B2Last significantly improved the strength of the 8CR-modified mixture and also reduced its plastic deformation capacity. Based on the stability-to-flow ratio, which indicates stiffness, the 8CR and 4SBS mixtures again showed similar performance, while the 8C2B mixture exhibited a stiffness value 63% higher than those of these mixtures.

Table 7. Stability values of the mixtures

	Stability (kN)				
	1	2	3	Std	Ortalama
Pure	12.5	12.7	12.1	0.305	12.43
8CR	14.9	15.9	15.3	0.503	15.36
8C2B	21.5	20.3	20.7	0.611	20.83
4SBS	14.0	14.5	14.1	0.264	14.20
	Flow (mm)				
Pure	3.59	3.62	3.76	0.090	3.65
8CR	3.38	3.49	3.10	0.201	3.32
8C2B	2.67	2.80	2.75	0.065	2.74
4SBS	3.05	2.94	3.12	0.091	3.03

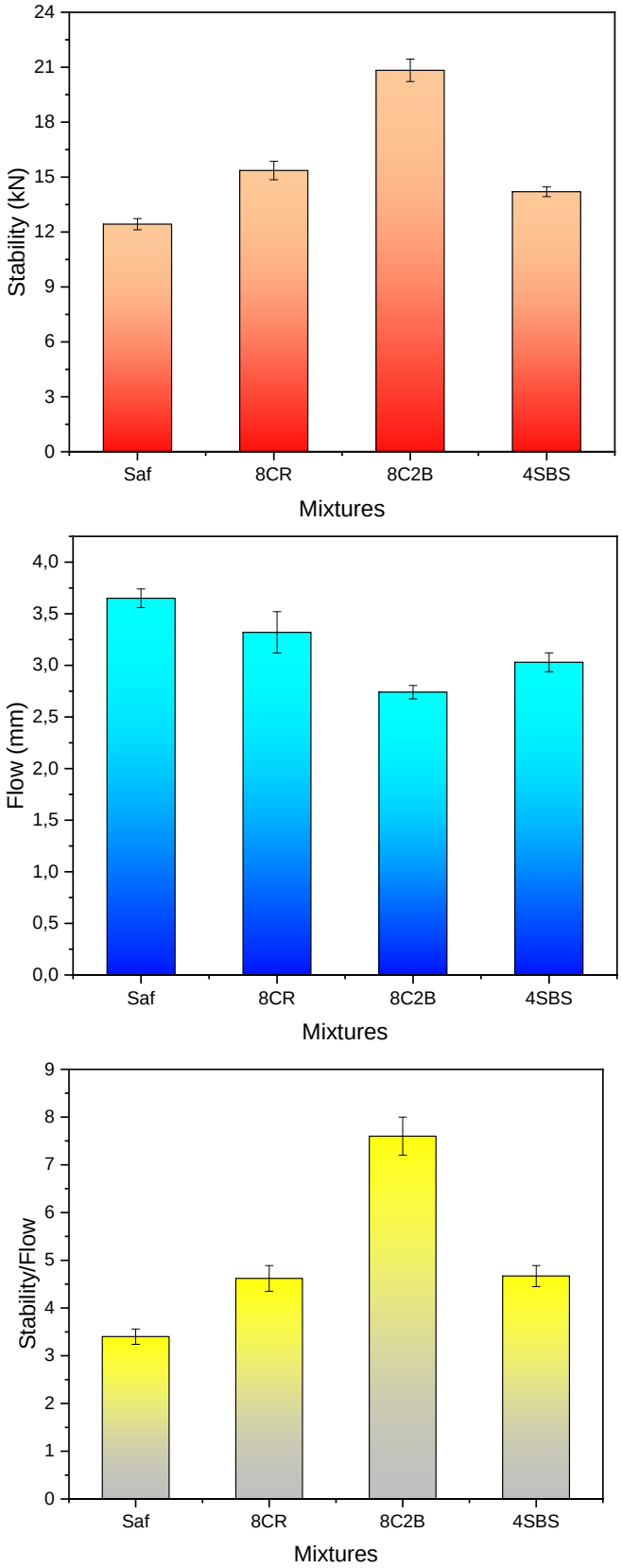


Figure 3. Stability, flow and stability/flow values of the mixtures

3. Conclusion

In this study. previous research on bitumen and bituminous mixtures modified with crumb rubber (CR). styrene-butadiene-styrene (SBS). and reactive polymers was reviewed. Additionally, the effects of CR. SBS. and the combined use of CR and a reactive polymer on the volumetric and stability properties of bituminous mixtures were investigated experimentally. For this purpose, binders modified with 8% CR. 4% SBS. and 8% CR + 2% B2Last were first prepared. Subsequently. bituminous mixture specimens were produced using these modified binders. and their volumetric and stability properties were compared.

According to the literature review. it has been observed that in recent years. additives that chemically interact with bitumen. rather than those that interact physically. such as SBS or CR. have become widely used. In this context. the use of methylene diphenyl diisocyanate (MDI). a reactive polymer. has shown positive results. Studies on the use of reactive polymers have reported that an MDI content of 0.5–2% in bitumen is sufficient to improve its properties. It has also been noted that an MDI content of 2–5% can enhance the polymerization effect; however, excessive MDI can increase the viscosity of the mixture. making it more difficult to work with. Recommended mixing conditions include a temperature of 160 °C–180 °C and a mixing time of 30–60 minutes.

The bituminous mixtures with 8% CR. 4% SBS. and 8% CR + 2% B2Last additives. prepared using the average of the lower and upper limits of the Type-1 wearing course gradation specified in the General Directorate of Highways (KGM) technical specifications. were found. at a bitumen content of 5.38%. to have void ratios between 3% and 5%. VFA values between 65% and 75%. and flow values between 2 mm and 4 mm. while also meeting the minimum VMA requirement of 14%. The experimental study demonstrated that the use of 2% B2Last in combination with 8% CR significantly increased stability. substantially improved the strength of the CR-modified mixture. enhanced resistance to plastic deformation. and provided better stiffness properties compared to the 4% SBS-modified mixture.

4. Acknowledgments

This article is extracted from the first author's master thesis entitled "Investigation of the Effects of Using Rubber Obtained from Waste Vehicle Tires with a Reactive Polymer on the Mechanical Properties of Bituminous Mixtures", supervised by Baha Vural KÖK. This study was carried out within the scope of the Scientific and Technological Research Council of Türkiye (TUBITAK). We gratefully acknowledge the financial support provided by TUBITAK to Research Project 124M464.

5. References

[1] Chen S, Xu L, Jia S, Wang J. Characterization of the nonlinear viscoelastic constitutive model of asphalt mixture. Case Stud Constr Mater 2023;18:e01902. https://doi.org/10.1016/j.cscm.2023.e01902.
[2] Du C, Zhang N, Wu J, Ma C, Zhang H, Tian Y, et al. Viscoelastic-viscoplastic characterization of asphalt mixtures based on the locally homogenous model simulations. Constr Build Mater 2024;419:135419. https://doi.org/10.1016/j.conbuildmat.2024.135419.
[3] Carreaul PJ, Bousmina M, Bonniotl F. The Viscoelastic Properties of Polymer-Modified Asphalts 2000:495–503.
[4] Fu H, Chai J, Xu Z, Qin Y, Liu Y, Wang Y, et al. Research on material selection and low-temperature anti-cracking mechanism of hydraulic asphalt concrete panels in the alpine region. Constr Build Mater 2024;423:135830. https://doi.org/10.1016/j.conbuildmat.2024.135830.
[5] Li H, Sun J, Hao G, Zhao Q, Hu Y, Zou X, et al. Effect of styrene butadiene styrene and desulfurized rubber powder on asphalt modification: Preparation, performance enhancement, mechanism analysis. Sci Total Environ 2024;912:169077. https://doi.org/10.1016/j.scitotenv.2023.169077.
[6] Zeng Y, Liu Q, Zeng Q, He Y, Xu Z. Research on the Mechanism of Interaction between Styrene–Butadiene–Styrene (SBS) and Asphalt Based on Molecular Vibration Frequency. Materials (Basel) 2021;14:358. https://doi.org/10.3390/ma14020358.
[7] Airey GD. Rheological properties of styrene butadiene styrene polymer modified road bitumens. Fuel 2003;82:1709–19. https://doi.org/10.1016/S0016-2361(03)00146-7.

[8] Zhang M, Yuan D, Jiang W. Effects of TLA on rheological, aging, and chemical performance of SBS modified asphalt. Constr Build Mater 2024;438:137283. https://doi.org/10.1016/j.conbuildmat.2024.137283.
[9] Nan H, Sun Y, Chen J, Gong M. Investigation of fatigue performance of asphalt binders containing SBS and CR through TS and LAS tests. Constr Build Mater 2022;361:129651. https://doi.org/10.1016/j.conbuildmat.2022.129651.
[10] Sengoz B, Isikyakar G. Evaluation of the properties and microstructure of SBS and EVA polymer modified bitumen. Constr Build Mater 2008;22:1897–905. https://doi.org/10.1016/j.conbuildmat.2007.07.013.
[11] Arabani M, Mirabdolazimi SM, Sasani AR. The effect of waste tire thread mesh on the dynamic behaviour of asphalt mixtures. Constr Build Mater 2010;24:1060–8. https://doi.org/10.1016/j.conbuildmat.2009.11.011.
[12] Huang B, Mohammad LN, Graves PS, Abadie C. Louisiana Experience with Crumb Rubber-Modified Hot-Mix Asphalt Pavement. Transp Res Rec J Transp Res Board 2002;1789:1–13. https://doi.org/10.3141/1789-01.
[13] Bartolozzi I, Mavridou S, Rizzi F, Frey M. Life cycle thinking in sustainable supply chains: the case of rubberized asphalt pavement. Environ Eng Manag J 2015;14:1203–15. https://doi.org/10.30638/eemj.2015.131.
[14] Cao R, Leng Z, Yu H, Hsu S-C. Comparative life cycle assessment of warm mix technologies in asphalt rubber pavements with uncertainty analysis. Resour Conserv Recycl 2019;147:137–44. https://doi.org/10.1016/j.resconrec.2019.04.031.
[15] Yousefi Kebria D, Moafimadani SR, Goli Y. Laboratory investigation of the effect of crumb rubber on the characteristics and rheological behaviour of asphalt binder. Road Mater Pavement Des 2015;16:946–56. https://doi.org/10.1080/14680629.2015.1042015.
[16] Pérez-Lepe A, Martínez-Boza FJ, Attané P, Gallegos C. Destabilization mechanism of polyethylene-modified bitumen. J Appl Polym Sci 2006;100:260–7. https://doi.org/10.1002/app.23091.
[17] Swamy BKK, Siddaramaiah, Somashekar R. Structure-property relationship of castor oil based diol chain extended polyurethanes (PUs). J Mater Sci 2003;38:451–60. https://doi.org/10.1023/A:1021815514141.
[18] Segura DM, Nurse AD, McCourt A, Phelps R, Segura A. Chapter 3 Chemistry of polyurethane adhesives and sealants. Handb. Adhes. Sealants, vol. 1, Elsevier Science Ltd; 2005, p. 101–62. https://doi.org/10.1016/S1874-5695(02)80004-5.
[19] Topal A. Evaluation of the properties and microstructure of plastomeric polymer modified bitumens. Fuel Process Technol 2010;91:45–51. https://doi.org/10.1016/j.fuproc.2009.08.007.
[20] Singh B, Tarannum H, Gupta M. Use of isocyanate production waste in the preparation of improved waterproofing bitumen. J Appl Polym Sci 2003;90:1365–77. https://doi.org/10.1002/app.12808.

Effect of Ethylene-Propylene-Diene-Monomer Addition on Rheological Properties of Bitumen

Kader YILDIRIM[1], Baha Vural KÖK[2], Ahmet Münir ÖZDEMIR[3]

[1,]Civil Engineering Department, Engineering Faculty, Eskişehir Technical University, Eskişehir, Türkiye.
[2] Civil Engineering Department, Engineering Faculty, Fırat University, Elazığ, Türkiye.
[3] Civil Engineering Department, Engineering and Natural Sciences Faculty, Bursa Technical University, Bursa, Türkiye
[1] kaderyildirim@eskisehir.edu.tr, [2]bvural@firat.edu.tr, [3] ahmet.ozdemir@btu.edu.tr
[1](ORCID: 0000-0003-4329-6158), [2](ORCID: 0000-0002-7496-6006), [3](ORCID: 0000-0002-4872-154X)

Abstract
Since thermoplastic bitumen is prone to permanent deformation at high temperatures and brittleness at low temperatures, the search for sustainable additives to overcome these mechanical weaknesses has gained importance. This study aims to investigate the rheological behavior of bitumen modified with recycled ethylene-propylene-diene monomer (EPDM) to improve the high temperature performance of asphalt binders. Pure bitumen of 70/100 penetration class from Batman Petroleum Refineries in Turkey and waste EPDM from automobile glass gaskets and washing machine lids were used in the study. Rheological characterization was carried out by frequency sweep tests using Dynamic Shear Rheometer (DSR) and master curves were generated by Time-Temperature Superposition Principle (TTSP) at a reference temperature of 40°C. The obtained shift factors (aT) were fitted with Arrhenius equation and activation energies (Ea) were calculated. According to the findings, 8% and 10% EPDM modified binders showed higher complex modulus (G*) values compared to pure binder, especially in low frequency regions. This indicates that EPDM admixture increases the resistance to permanent deformation under long-term loading. The activation energy values increased in parallel with the EPDM additive ratio to 143.0, 144.9 and 145.2 kJ/mol, respectively, indicating that EPDM resists temperature-dependent flow behavior by limiting molecular motion within the binder. In conclusion, recycled EPDM additive is an effective additive for enhancing the rheological performance of asphalt binders at elevated temperature and significantly improves the resistance to deformation, especially under low-frequency loading conditions.

Keywords: Bitumen, rheology, modification, EPDM

1. Introduction

Bitumen exhibits different mechanical responses depending on temperature changes due to its thermoplastic properties. Bitumen, which is exposed to permanent deformations due to its low strength under high temperature conditions, gains a brittle structure at low temperatures and prepares the ground for crack formation. This situation, together with the continuity of traffic loads and environmental impacts, causes the pavement performance to deteriorate over time. In order to minimize these mechanical and environmental problems and to increase both the strength and service life of asphalt pavements, the addition of various additives to bitumen binders has become a common method [1].

Bitumen modification has become a strategic application in the context of modern highway engineering, not only to meet technical performance requirements but also to meet environmental sustainability goals. In this context, Styrene-Butadiene-Styrene (SBS) copolymer is one of the most preferred modifiers due to its positive effects such as increasing elasticity, providing resistance to low temperature cracking and increasing resistance to aging. However, the high cost of SBS and its production process based on non-renewable resources necessitates the search for alternative additives with a focus on sustainability. Furthermore, the use of SBS, especially in high doses, has several limitations in terms of both economic efficiency and compatibility with bitumen [2][3][4]. In their study on SBS modification, Ren et al. emphasized that although general properties improved with an increase in SBS content, compatibility and storage stability deteriorated [5]. Wang et al. also stated in their study that the differences in density, molecular weight and polarity between SBS and bitumen create an unstable system, resulting in phase separation and segregation [6]. Zhao et al. reported that the direct cost of SBS is high and this is reflected in asphalt preparation costs, straining project budgets [7]. Huang et al. reported in their study that even the use of only 2-3% SBS results in approximately 10.9% higher road user costs than standard polymer modification [8].

[1]Corresponding author

Conventional bituminous mixtures are subject to various types of deterioration over time, such as crack formation, rutting deformation and surface deterioration. The use of waste materials in asphalt pavements is becoming increasingly widespread to address such structural problems and at the same time support economic and environmental sustainability. Waste-based additives are considered as an environmentally friendly and functional solution for both improving rheological parameters and integrating into waste management processes. In this respect, bitumen modification with recycled additives has gained an important place in the development of current asphalt technology [9].

One of the waste-based synthetic rubbers used as additives within the scope of modification with waste materials is ethylene-propylene-diene-monomer (EPDM). EPDM has high elasticity due to its chemical composition, maintains its thermal stability in the temperature range of -40°C to +150°C, resists ozone and UV rays and offers resistance to chemicals. In addition, thanks to its electrical insulating properties, it has a wide range of applications in different engineering applications. In the literature, it has been found that EPDM-added bitumen improves deformation resistance at high temperatures, reduces rutting and decreases cracking tendency at low temperatures. In addition, various scientific studies have shown that the addition of EPDM increases the aging resistance, increases the coefficient of elasticity and significantly improves the fatigue life of asphalt pavements [10][11].

Chegenizadeh et al. stated in their study that the sample containing 4% EPDM had the lowest rut depth with a reduction of approximately 75% compared to the control sample. In general, it was emphasized that EPDM additive significantly reduced the rutting value. 4% EPDM content was determined as optimum and it was found that this ratio provided improvement in both rutting and fatigue performance values [12]. Kumar et al. stated that the addition of EPDM increases the viscosity of the binder. It was observed that the failure temperature of both unaged and RTFOT aged binders increased with increasing EPDM content. It was emphasized that EPDM increases the stiffness and elastic properties of the binder, thus improving resistance to rutting. The MSCR test also indicated an increase in the binder's resistance to permanent deformation [13]. Niu et al. reported that EPDM improved the fatigue resistance of asphalt at moderate temperature. It was also found that EPDM improved the high and low temperature performance of asphalt [14].

In this study, the rheological properties of EPDM blended bituminous binders prepared at two different ratios were investigated using frequency sweep test results obtained from dynamic shear rheometer (DSR) test.

2. Materials and Method

2.1 Materials

In the experimental study, 70/100 penetration grade (PG 70-22) pure bitumen obtained from Batman Petroleum Refineries in Turkey was used. The properties of the pure binder are given in Table 1.

Table 1. Properties of the pure binder

Properties	Standart	B 70/100
Penetration (0.1 mm), 100 g, 5 sec	ASTM D5	85.10
Softening point (°C)	ASTM D36	52.65
Penetration index (PI)		0.87
Viscosity (cP, 135°C)	ASTM D4402	562.50
Viscosity (cP, 165°C)	ASTM D4402	175.00
Specific gravity	ASTM D70	1.038

Ethylene Propylene Diene Monomer (EPDM) is a synthetic elastomer with a wide range of applications from asphalt modification to the automotive industry, from construction applications to the electrical industry. EPDM has a terpolymer structure obtained through controlled chemical polymerization and formed by the combination of three different monomers (ethylene, propylene and diene). The EPDM produced can be black, white or light colored, while its density is usually in the range of 0.86-0.90 g/cm³. It is known for its high water resistance and superior insulating properties [15][16]. The

EPDM used in this study is recycled waste material from car window seals and washing machine lids. This material (Figure 1) was sieved through sieve number 40 and used in bitumen modification.



Figure 1. Waste EPDM used in the study

The modification of EPDM was conducted utilizing a high shear rate mixer situated within a cylindrical insulated chamber (refer to Figure 2). This process necessitates a high mixing speed and shear force to achieve effective particle disintegration. The mixtures were produced at a speed of 2,000 rpm and a temperature of 180°C over a mixing duration of 90 minutes. In order to mitigate the oxidation effect during the mixing process, the pure binder underwent the identical mixing procedure.



Figure 2. High shear mixer

2.2. Method

Dynamic Shear Rheometer (DSR) analysis was conducted on unaged asphalt binders using a Bohlin DSRII system operating in controlled-stress mode. This procedure aimed to assess the rutting resistance of the material, represented by the G*/sinδ parameter, thereby providing insight into the binder's performance at elevated temperatures. The test setup included 25 mm parallel plates with a 1 mm gap, operating at a frequency of 1.59 Hz, and was performed at temperatures ranging from 58°C to 82°C in 6°C increments under single-point testing conditions.

To characterize the time-dependent viscoelastic properties of the binders, a frequency sweep test was also performed. This test applies oscillatory loading at various frequencies while maintaining constant strain, effectively capturing the material's behavior across different loading times. Conducted in line with AASHTO T315 guidelines, the test used the DSR to evaluate binder response over a frequency range of 0.1 to 10 Hz and at temperatures from 40°C to 80°C in 10°C intervals. The measurements were carried out within the binder's linear viscoelastic range, ensuring strain did not exceed the lower limit of 1%.
The G* curves obtained from the frequency sweep test were converted to master curve format in accordance with TTSP. This made it possible to perform rheological evaluation over a wide frequency range. The master curves were obtained at a reference temperature of 40°C, and the other curves were shifted to this degree. Subsequently, the shift

factors (aT) were fitted to the Arrhenius equation as a function of temperature, and the activation energy (Ea) was obtained. The Arrhenius equation is given in Equation 1 [17].

$$\log\big(aT(T)\big) = \log\left(\frac{\tau(T)}{\tau(T_g)}\right) = \frac{Ea}{c1R}\left(\frac{1}{T} - \frac{1}{Tg}\right) \quad (1)$$

3. Results and Discussion

Research has found that there is a reciprocal relationship between temperature and frequency (or temperature and loading time) through time-temperature shift factors, which allows measurements taken at different temperatures to conform to a single general continuous curve at a reduced frequency. This curve is called the master curve and is preferred for describing the behavior of bituminous binders at a specific temperature over a wide frequency range. The test setup (DSR) may not always be able to measure at very low or very high frequencies, and maintaining a wide frequency range for evaluation is necessary to determine the rheological properties of bitumen. The time-temperature superposition principle (TTSP) can be used to relate the equivalence between temperature and frequency and thus produce a master curve [18]. Figure 3 shows the Master curves of pure, 8% and 10% EPDM modified binders obtained at 40°C reference temperature. These curves are highly valuable in representing the viscoelastic behavior of the binders over a wide temperature and loading time range at a single reference temperature.

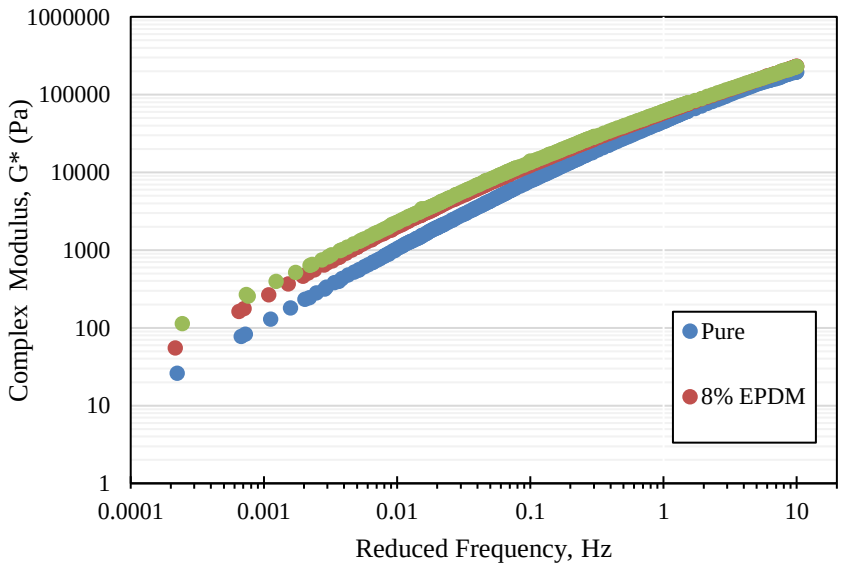


Figure 3. Master curves of binders containing pure, 8% and 10% EPDM

Pure binder has provided the lowest complex modulus values in the low frequency region (i.e., under long-term loading conditions). This indicates that pure asphalt is less resistant to permanent deformation and exhibits a more viscous structure, and it can be said that this curve, reduced to a single temperature, generally provides the weakest permanent deformation resistance. In contrast, the G* values of binders containing both 8% and 10% EPDM are higher than those of pure binder at each frequency value. As the frequency value increases, it becomes more difficult to observe differences in the G* values of the binders. The role of the additive is more clearly evident in the low-frequency region. Considering the viscoelastic structure of asphalt, low frequency (long loading time) presents a more unfavorable condition, necessitating a greater need for the additive in this region. As shown in Figure 3, the highest G* values in the low-frequency region are observed in the 10% EPDM binder. For viscoelastic materials like bitumen, TTSP master curves are desired to have a smooth and continuous structure and to be thermoreologically simple. Upon examining Figure 3, it is observed that quite smooth and continuous master curves were obtained for all three binder types. Table 2 and Figure 4 show the G* values corresponding to some characteristic frequency values.

Table 2. G* values at different frequency values

	Binder type		
Frequency	**Pure**	**8E**	**10E**
0.0002	25.98	54.84	112.8
0.002	231	459.7	634.1
0.02	1918	3444	4154
0.1	8019	12940	13440
1	45250	58830	61310
5	134900	153000	153800
10	194600	229300	227800

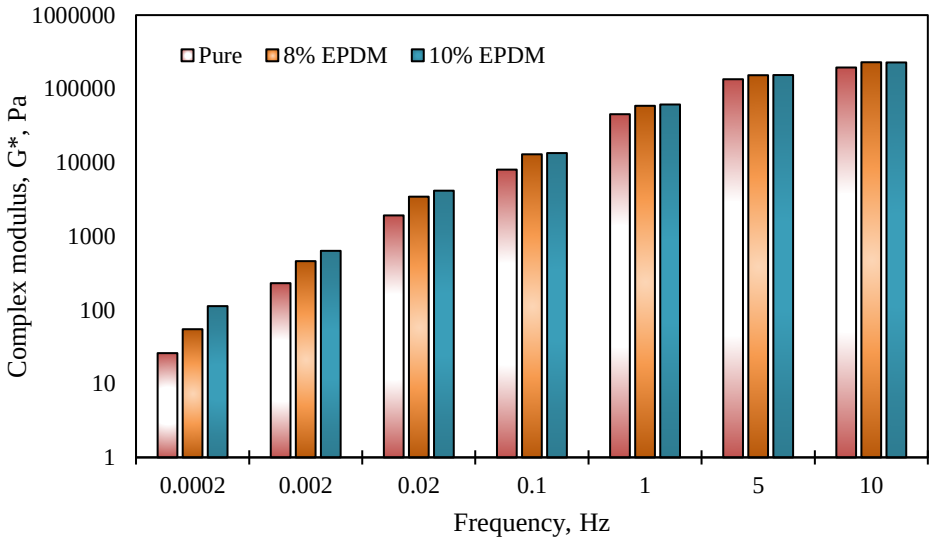


Figure 4. Display of G* values between 0.0002-10 Hz

In Figure 4, the complex modulus (G*) values obtained in the 0.0002–10 Hz range for asphalt binders modified with different EPDM ratios are shown using a logarithmic Y axis. In general, it was observed that G* values increased significantly in all binders as frequency increased. There is also an increasing trend in G* values with increasing modification ratio; this increase is particularly pronounced in the low-frequency range. This indicates that the EPDM additive increases the stiffness of the binder, thereby enhancing its resistance to deformation. Additionally, the greater influence on viscoelastic behavior at low frequencies (e.g., 0.0002–0.02 Hz) as the additive ratio increases indicates the additive's effectiveness in structural modification.

The shift factor (aT) values calculated while obtaining the master curves are analyzed using the Arrhenius equation, and the activation energy values are found. Figure 5 shows the Arrhenius fit curves of the aT values against temperature. When Figure 3 is examined, it can be seen that a very successful fitting process has been carried out for all three binders.

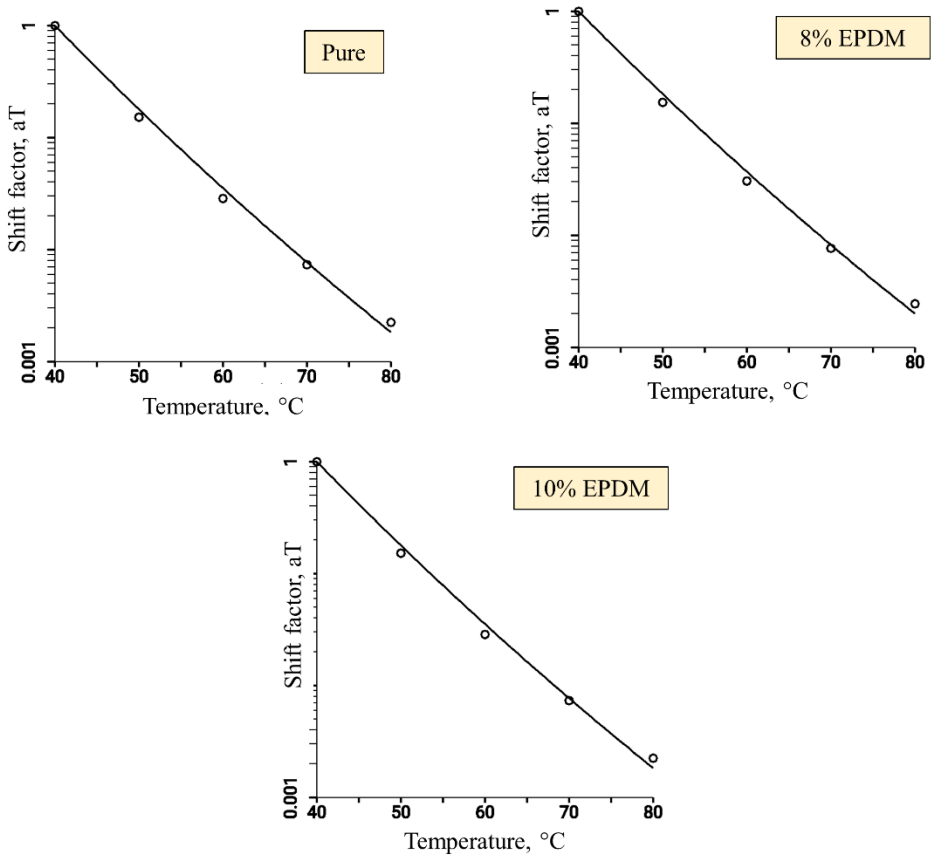


Figure 5. Arrhenius fit of shift factors

The activation energy values obtained as a result of the fitting process given in Figure 5 are given in Figure 6.

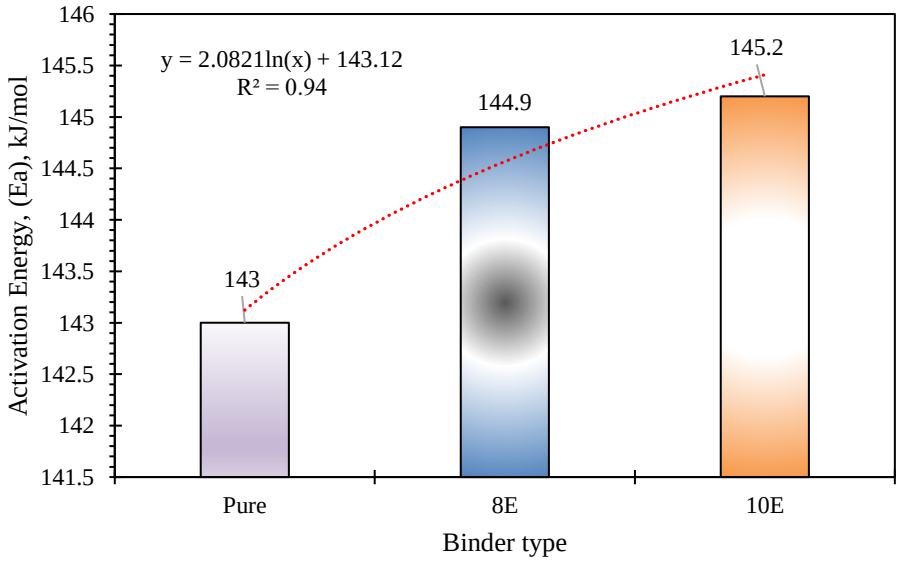


Figure 6. Activation energy values for pure and modified binders

Activation energy (Ea) refers to the energy barrier that the asphalt binder must overcome in order to transition to a flow state under the influence of temperature. The higher the Ea value, the more difficult it becomes for the binder to undergo molecular movement, which in turn causes the viscosity to decrease more slowly as temperature increases,

meaning the binder remains more rigid. In this respect, high Ea is a strong indicator of the asphalt binder's high temperature resistance. Especially in polymer-modified systems, the activation energy increases with the increase in interchain interactions, and the binder's resistance to deformation increases. Therefore, Ea is of critical importance in evaluating the binder's temperature-related flow behavior.

When examining Figure 6, the lowest Ea value is observed in the pure binder, and this value increases regularly as the additive ratio increases. While the activation energy of the pure binder was 143 kJ/mol, this value increased to 144.9 kJ/mol in the binder with 8% EPDM additive and to 145.2 kJ/mol in the binder with 10% additive. This increasing trend indicates a logarithmic relationship between the additive ratio and Ea, and this relationship is supported by the regression curve added to the graph with a red dashed line. The high correlation coefficient indicates that increasing the additive ratio has a strong and consistent effect on the activation energy. The physical barrier and interchain interactions created by the EPDM additive in the binder structure restrict molecular mobility and make it difficult for the binder to transition to a flow state. Therefore, as the contribution ratio increases, the tendency for the binder's viscosity to decrease with temperature slows down; thus, the binder remains more rigid at high temperatures and resists deformation. In other words, the increase in activation energy observed here is a positive indicator not only in terms of temperature sensitivity but also in terms of high-temperature resistance.

4. Acknowledgments

This article is extracted from the first author's master thesis entitled "Investigation of the Effects of Using Waste EPDM with SBS on Bitumen Properties", supervised by Baha Vural KÖK.

5. Conclusions

In this study, the rheological behavior of asphalt binders modified with EPDM at different concentrations (8% and 10%) was evaluated through frequency sweep tests conducted with a Dynamic Shear Rheometer (DSR). Master curves were constructed using the Time–Temperature Superposition Principle (TTSP), and shift factors (aT) obtained from this process were fitted to the Arrhenius model to calculate activation energy (Ea) values. Based on the experimental results and model interpretations, the following conclusions can be drawn:

- The master curves obtained at a reference temperature of 40°C successfully captured the viscoelastic behavior of the binders across a wide frequency range, confirming that all tested binders exhibit thermorheologically simple behavior.
- Both 8% and 10% EPDM-modified binders exhibited higher complex modulus (G*) values than the pure binder, particularly in the low-frequency region, indicating enhanced resistance to permanent deformation under long loading times.
- The shift factor (aT) values obtained from TTSP fitting followed the Arrhenius model with high accuracy ($R^2 > 0.98$ for all binders), demonstrating the validity of the time–temperature equivalence for the binders tested.
- Activation energy (Ea) values increased systematically with EPDM content: 143.0 kJ/mol for the pure binder, 144.9 kJ/mol for 8% EPDM, and 145.2 kJ/mol for 10% EPDM binders. This trend implies that EPDM addition increases the energy barrier for flow initiation, enhancing the high-temperature stiffness of the binder.
- The increase in Ea with higher polymer content confirms that EPDM incorporation limits molecular mobility within the bitumen matrix and improves resistance to temperature-induced deformation.
- The log-linear relationship between EPDM content and Ea, with a high coefficient of determination ($R^2 = 0.94$), suggests a predictable and positive influence of EPDM on the thermal activation behavior of the binder.

Overall, the findings of this study demonstrate that EPDM is an effective polymer additive for enhancing the high-temperature rheological properties of bituminous binders. Increasing EPDM content improves the deformation resistance of the binder by raising its activation energy and complex modulus, especially under low frequencies.

5. References

[1] Hunter R.N., Self A., Read J., The Shell Bitumen... - Google Akademik n.d.

[2] Laukkanen O-V, Soenen H, Winter HH, Seppälä J. Low-temperature rheological and morphological characterization of SBS modified bitumen. Constr Build Mater 2018;179:348–59. https://doi.org/10.1016/j.conbuildmat.2018.05.160.

[3] Yalikun N, Yu S, Yang H, Liu C, Zhang H, Wang Q, et al. Preparation and performance study of oxidized CR/SBS composite modified asphalt. J Environ Chem Eng 2025;13:117094. https://doi.org/10.1016/j.jece.2025.117094.

[4] Airey G. Rheological properties of styrene butadiene styrene polymer modified road bitumens⋆ . Fuel 2003;82:1709–19. https://doi.org/10.1016/S0016-2361(03)00146-7.

[5] Ren S, Liu X, Fan W, Wang H, Erkens S. Rheological Properties, Compatibility, and Storage Stability of SBS Latex-Modified Asphalt. Materials (Basel) 2019;12:3683. https://doi.org/10.3390/ma12223683.

[6] Wang P, Wei H-R, Liu X-Y, Ren R-B, Wang L-Z. Identifying the Long-Term Thermal Storage Stability of SBS-Polymer-Modified Asphalt, including Physical Indexes, Rheological Properties, and Micro-Structures Characteristics. Sustainability 2021;13:10582. https://doi.org/10.3390/su131910582.

[7] Zhao Z, Wang Z, Wu S, Xie J, Yang C, Li N, et al. Road performance, VOCs emission and economic benefit evaluation of asphalt mixture by incorporating steel slag and SBS/CR composite modified asphalt. Case Stud Constr Mater 2023;18:e01929. https://doi.org/10.1016/j.cscm.2023.e01929.

[8] Huang J, Yan K, Wang M, Shi K, Li Y, Zhang Y. Performance evaluation of SBS-modified asphalt mixtures incorporating waste tire rubber and HDPE. Constr Build Mater 2024;430:136423. https://doi.org/10.1016/j.conbuildmat.2024.136423.

[9] Yüknü K, Öztürk T. Bitümlü Bağlayıcılar Laboratuvar El Kitabı. Ankara: Araştırma Ve Geliştirme Dairesi Başkanlığı Üstyapı Geliştirme Şubesi Müdürlüğü; 2021.

[10] Kumar A, Choudhary R, Kumar A. Characterization of Storage Stability of EPDM Rubberized Asphalt with Tire Pyrolytic Oil and Sulfur. J Mater Civ Eng 2023;35. https://doi.org/10.1061/(ASCE)MT.1943-5533.0004716.

[11] Aldagari S, Kabir SF, Fini EH. A comparative study on efficacy of waste plastic and waste Rubber in bitumen. Constr Build Mater 2022;325:126724. https://doi.org/10.1016/j.conbuildmat.2022.126724.

[12] Chegenizadeh A, Aung M-O, Nikraz H. Ethylene Propylene Diene Monomer (EPDM) Effect on Asphalt Performance. Buildings 2021;11:315. https://doi.org/10.3390/buildings11080315.

[13] Kumar A, Choudhary R, Kumar A. Characterisation of asphalt binder modified with ethylene–propylene–diene–monomer (EPDM) rubber waste from automobile industry. Road Mater Pavement Des 2021;22:2044–68. https://doi.org/10.1080/14680629.2020.1740772.

[14] Niu Y, Wang W, Geng J, Zhao W, Zhang Z, Li X. Study on rheological properties and index of Ethylene-Propylene-Diene Monomer (EPDM) sulfur modified asphalt based on principal component analysis (PCA). Constr Build Mater 2024;449:138546. https://doi.org/10.1016/j.conbuildmat.2024.138546.

[15] Wei H, Ma Z, He X, Han S, Jiang B. Structural response analysis of conductive ethylene–propylene–diene monomer rubber composite pavement under validated temperature field. Constr Build Mater 2022;328:127094. https://doi.org/10.1016/j.conbuildmat.2022.127094.

[16] Zirnstein B, Schulze D, Schartel B. Mechanical and Fire Properties of Multicomponent Flame Retardant EPDM Rubbers Using Aluminum Trihydroxide, Ammonium Polyphosphate, and Polyaniline. Materials (Basel) 2019;12:1932. https://doi.org/10.3390/ma12121932.

[17] Liang B, Chen Y, Lan F, Zheng J. Evaluation of rheological and aging behavior of modified asphalt based on activation energy of viscous flow. Constr Build Mater 2022;321:126347. https://doi.org/10.1016/j.conbuildmat.2022.126347.

[18] Yusoff NIM, Shaw MT, Airey GD. Modelling the linear viscoelastic rheological properties of bituminous binders. Constr Build Mater 2011;25:2171–89. https://doi.org/10.1016/j.conbuildmat.2010.11.086.

International Conference on Advances and Innovations in Engineering

3.ICAIE

Effect of Damping Ratio on Seismic Response of a Reinforced Concrete Building Using Nonlinear Time History Analysis

Burak ÇAKIL[1], Ozan İNCE[2], Ömer Faruk TAŞ[3], Ömer Faruk OSMANLI[4], Muhammet KARATON[5], Erkut SAYIN[6]

[1,2,3,4,5,6]Civil Engineering, Engineering Faculty, Firat University, Elazığ, Türkiye.
[1]burakk.cakill@gmail.com, [2]o.ince@firat.edu.tr, [3]oftas@firat.edu.tr, [4]ofosmanli@firat.edu.tr, [5]mkaraton@firat.edu.tr, [6]esayin@firat.edu.tr
[1](ORCID: 0000-0002-4972-8874), [2](ORCID: 0000-0002-6319-1129), [3](ORCID: 0000-0002-1431-5316), [4](ORCID: 0000-0002-9799-3103) [5](ORCID: 0000-0002-1498-4659), [6](ORCID: 0000-0003-0266-759X)

Abstract
Earthquakes are among the natural disasters that cause abrupt and devastating impacts on structures, frequently leading to substantial casualties and property damage. Accurate modeling of a structure's seismic behavior is critically important for ensuring the safety of engineering designs. In this context, the damping ratio used in seismic analysis is a key parameter that determines the extent to which the system dissipates vibrational energy. Although a constant value—typically 5%—is commonly assumed in engineering practice, the impact of this assumption on structural performance is often not explicitly addressed. However, the damping ratio has a decisive influence on how much of the vibrational energy is absorbed by the structure.
In this study, a three-story residential reinforced concrete building was modelled using SAP2000, and nonlinear time history analyses were conducted for different damping ratios using the ground motion record of the Kahramanmaraş–Elbistan earthquake(Mw:7.6) that occurred on February 6, 2023. For this purpose, three different damping ratios—2%, 5%, and 10%—were used, and the structural responses of the building were compared. In particular, the relationship between base shear force and lateral displacement was examined, and the hysteresis curves obtained from the dynamic analyses were evaluated to assess the energy dissipation capacities of the structures with different damping ratio.

Keywords: damping ratio, nonlinear time history analysis, reinforced concrete buildings

1. Introduction

Earthquakes are natural disasters that generate sudden and destructive effects on structures, leading to significant loss of life and property. In regions located on active fault lines, seismic ground motions can cause severe damage to the load-bearing systems of buildings. Türkiye is situated in a highly seismically active zone, making accurate assessment of earthquake effects essential for structural safety. The earthquakes that occurred on February 6, 2023, with epicenters in Kahramanmaraş, caused widespread destruction across a vast region and brought renewed attention to the seismic resistance and energy dissipation capacities of structures. In order to understand the behavior of structures under seismic loading and to develop appropriate engineering solutions, nonlinear time history analyses offer a powerful tool. In these analyses, the damping ratio is a fundamental parameter that defines the extent to which the system dissipates vibrational energy, and it can directly influence the seismic performance of the structure. Although a damping ratio of 5% is commonly assumed in engineering practice, the effect of this assumption on structural response is often not thoroughly examined. However, different damping ratios can lead to significant variations in lateral displacements, base shear forces, and energy dissipation capacity of structures. The damping ratio may vary depending on the type of structure, characteristics of the load-bearing system, and material behavior. This has been supported by various academic studies. Gilles and McClure (2011), in an experimental study on reinforced concrete shear wall structures, reported that the damping ratios of 27 buildings ranged from 1% to 4%, with an average value of 2% [1]. Karaton et al (2021), investigated the effects of different damping ratios and damping types on reinforced concrete shear wall systems, and found that Rayleigh and stiffness-proportional damping models with damping ratios between 2% and 3% produced more consistent results compared to mass-proportional damping [2]. Finally, in the study conducted by Cakil et al (2025), experimental data of seismically isolated steel structures were numerically modeled, and nonlinear time history analyses were performed using damping ratios ranging from 1% to 8%. The best agreement with experimental results was obtained with a 6% damping ratio for mass-proportional damping, and a 1% damping ratio for Rayleigh and stiffness-proportional damping types [3].
In this study, the seismic response of a three-story reinforced concrete residential building was investigated under different damping ratios using the Elbistan earthquake (Mw: 7.6) that occurred on February 6, 2023. The structural model was developed using SAP2000 software, and nonlinear time history analyses were performed. Within this scope, structural responses were compared for three different damping ratios: 2%, 5%, and 10%. In particular, the relationship between

[1]Corresponding author

lateral displacement and base shear force was evaluated, and the energy dissipation capacities of the structures were analyzed based on the hysteresis curves obtained for each damping ratio.

2. Material Method

In typical reinforced concrete structures, the damping ratio generally ranges between 1% and 10%, except for special cases [4]. In numerical analyses of structures, a damping ratio of 5% is commonly assumed. Most building codes define seismic design spectra based on a 5% damping ratio. For the dynamic analyses, several damping models have been proposed by researchers. These models include viscous damping, mass- and stiffness-proportional (Rayleigh) viscous damping, Coulomb damping, structural damping, and others. To formulate the classical damping matrix, two types of damping are commonly used: Rayleigh and Caughey damping. Rayleigh damping is a classical damping model that combines mass-proportional and stiffness-proportional damping components. Stiffness-proportional damping mainly accounts for energy dissipation due to structural deformations. However, expressing mass-proportional damping is challenging because air damping, which is typically included in the model, is usually negligible. Rayleigh damping corresponds to two damping ratios and two effective modal masses. On the other hand, Caughey damping incorporates the contribution of higher modes by applying a series of coefficients to the mass and stiffness matrices. Rayleigh damping is frequently preferred due to its computational simplicity and ease of implementation in software. In this study, Rayleigh damping models were used in the dynamic analyses among the available damping models [5-7].

In the numerical analyses, structural damping is expressed through the damping matrix. The damping matrix can be obtained using the following equation

$$[C] = \alpha_c . [M] + \beta_c . [K] \tag{1}$$

in this equation, α_c and β_c are denoted coefficients of the mass matrix and the stiffness matrix in the Rayleigh damping, respectively [5]. Two natural frequencies and damping ratios with significant modal participation are used to calculate the two Rayleigh damping coefficients. These coefficients are determined using the following equations.

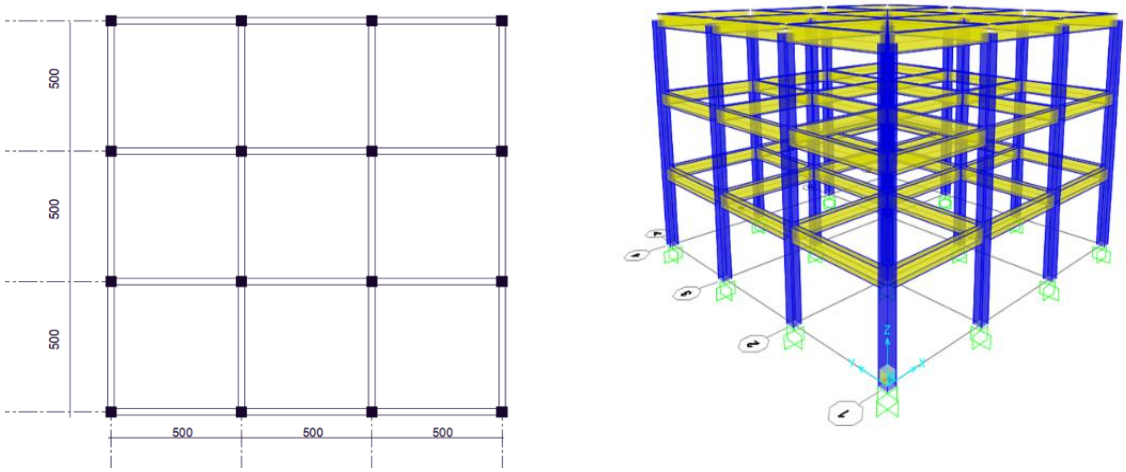
$$\alpha_c = (2.\xi.\omega_i.w_j)/(\omega_i + \omega_j) \tag{2}$$

$$\beta_c = 2.\xi/(\omega_i + \omega_j) \tag{3}$$

where, ξ and ω define damping ratio and natural frequency of the structure, respectively. Subscripts i and j indicate modal numbers of the structure. In this study, Rayleigh damping is selected and the damping ratio is changed as 2%, 5% and 10%.

3. Numerical Modelling

In this study, a nonlinear analysis was conducted on a three-story reinforced concrete building using various analytical methods. The structural model developed for the analysis is given in Figure 1. The building was designed with concrete of 25 MPa compressive strength and reinforcing steel with a yield strength of 420 MPa. Each story has a uniform height of 3 meters. Columns were modeled with cross-sectional dimensions of 40×40 cm, while beams were designed as 25×50 cm. Rigid diaphragm assumptions were applied at all floor levels. Floor and wall loads were distributed onto the beams.



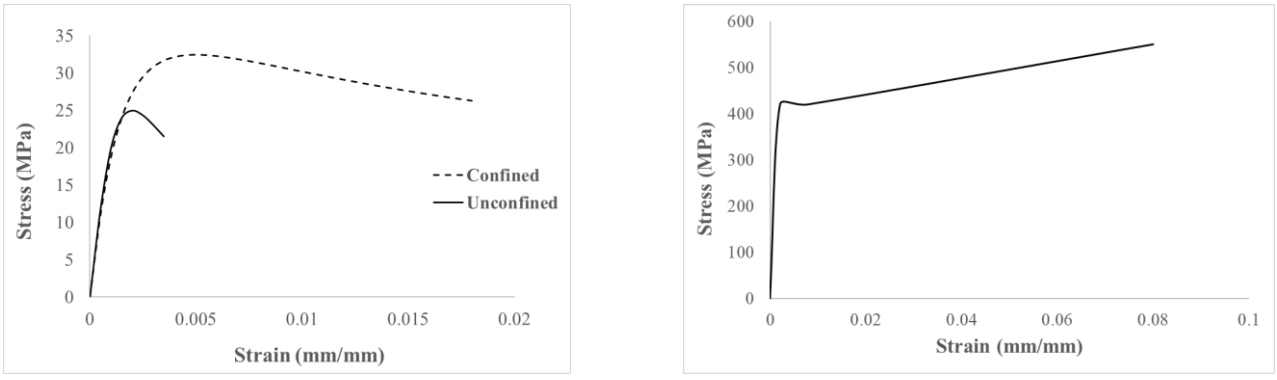
(a) (b)

Figure 1. (a) Floor plan (cm) (b) Three-dimensional finite element model

To accurately capture the nonlinear behavior of the reinforced concrete structure, the fiber element method was performed. Plastic hinges were assigned to the end regions of all column and beam elements. For concrete modeling, the stress–strain relationships of both confined and unconfined concrete were incorporated using the Mander model [8]. Material damage limits for confined concrete, unconfined concrete, and reinforcement steel were determined in accordance with the Türkiye Building Earthquake Code (TBEC – 2018) and implemented within the SAP2000 [9–10]. These damage thresholds are summarized in Table 1. The stress–strain curve of the Mander model applied to the concrete material is given in Figure 2a, while the constitutive model adopted for reinforcement steel is presented in Figure 2b. Additionally, effective section stiffness coefficients were considered in the modeling process to ensure a more realistic representation of element stiffness.

Table 1. Damage limits of materials

Damage limits	Unconfined concrete	Confined concrete	Steel
Limited damage	0.0025	0.0025	0.0075
Controlled damage	0.0026	0.014	0.024
Collapse prevention	0.0035	0.018	0.032



(a) Mander concrete model (b) Reinforcement steel model

Figure 2. Nonlinear behaviors for concrete and reinforcement steel materials

In order to determine the dynamic characteristics of the building and calculate damping coefficient, modal analyses were conducted. The results of this analysis include the vibration modes of the structural systems, as well as the

corresponding natural periods and mass participation ratios. The natural periods and mass participation ratios of the model are presented in Table 2. The damping coefficients that vary depending on the damping ratio are given in Table 3.

Table 2. Modal analysis results

Mode	Period (s)	Mass participation ratio	
		U(x)	U(y)
1	**0.53**	0.00	**0.87**
2	**0.53**	**0.87**	0.00
3	0.46	0.00	0.00
4	0.16	0.08	0.03

Table 3. Damping coefficient

Damping ratio	α	β
2%	0.2370	0.0017
5%	0.5924	0.0042
10%	1.1848	0.0084

A nonlinear time history analysis was carried out for the designed reinforced concrete building. For this purpose, the ground acceleration record from station number 4406, recorded during the 2023 Elbistan earthquake, was utilized. The selected ground motion was scaled in accordance with the TBEC-2018 to match the target design spectrum [9]. Although the original acceleration record spans a longer duration, only the most significant 35-second segment—corresponding to the period of peak ground motion—was considered in the analysis to capture the most critical structural response (Figure 3).

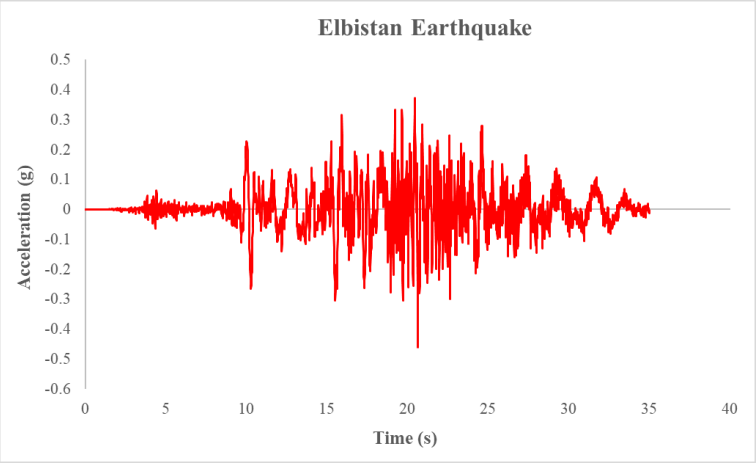


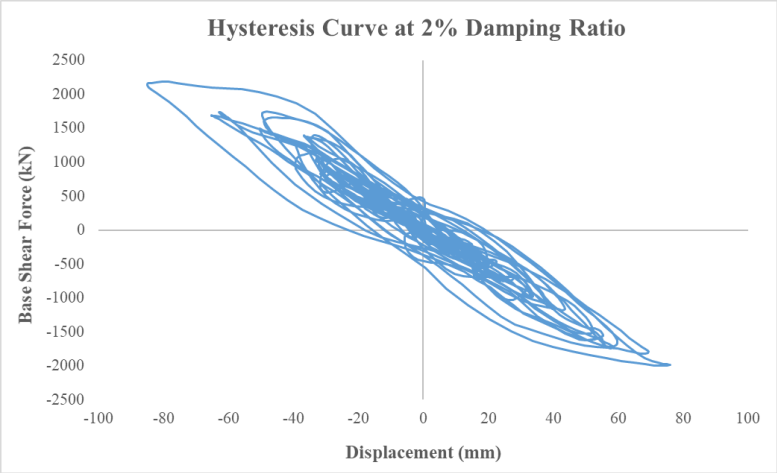
Figure 3. Elbistan earthquake, scaled acceleration-time graph of 4406 station

To evaluate the effect of damping on structural response, analyses were conducted for three different damping ratios: 2%, 5%, and 10%. As shown in Table 4 and Figure 4, an increase in the damping ratio leads to a noticeable reduction in both base shear force and maximum displacement. For instance, when the damping ratio increased from 2% to 10%, the base shear force decreased from 2188 kN to 1972 kN, while the maximum displacement was reduced from 85 mm to 48 mm. The hysteresis curves in Figure 4 clearly illustrate the reduction in the width of the force-displacement response with increasing damping, indicating a lower energy dissipation demand on the structure.

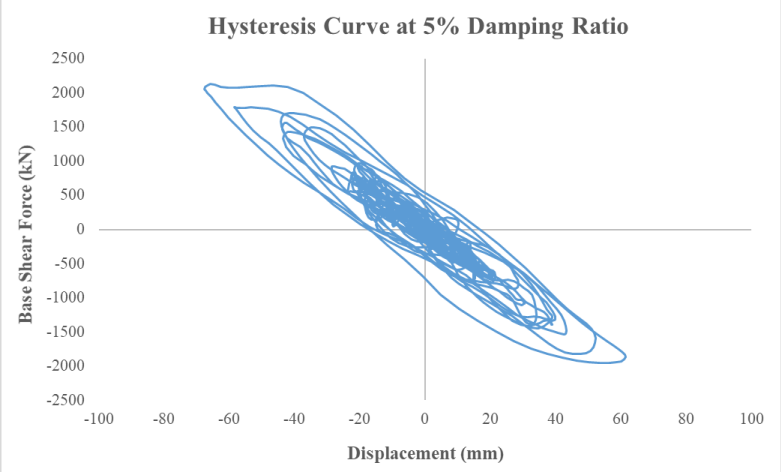
Table 4. Base shear force and maximum displacement values according to damping ratios

Damping ratio	Base shear force(kN)	Maximum displacement(mm)
2%	2188	85
5%	2129	68

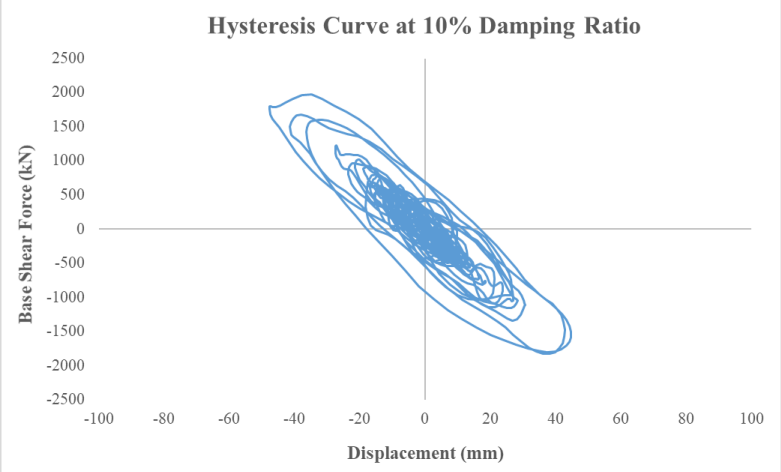
10%	1972	48



a) 2% damping ratio



b) 5% damping ratio



c) 10% damping ratio

Figure 4. Hysteresis curves

4. Results

In this study, the effect of different damping ratios on the seismic behavior of a three-story reinforced concrete residential building was investigated through nonlinear time history analyses using the ground motion record of the 2023 Elbistan earthquake. The structural responses for damping ratios of 2%, 5%, and 10% were compared in terms of base shear force, maximum lateral displacement, and energy dissipation characteristics derived from hysteresis curves.

The analysis results are summarized below:

- An increase in the damping ratio from 2% to 10% resulted in a reduction of the base shear force from 2188 kN to 1972 kN, corresponding to an approximate decrease of 10%.
- The maximum horizontal displacement decreased from 85mm to 48mm in the same increment range, which showed a decrease of more than 40%.
- Examination of the hysteresis curves revealed that as the damping ratio increased, the force–displacement cycles became smaller, and the amount of energy that the structure needed to absorb decreased.
- Structural responses are highly sensitive to the damping ratio; damping values other than the commonly used fixed 5% can lead to significant differences in structural behavior. Therefore, to achieve more realistic results in performance-based analyses, the damping ratio should be carefully selected according to the characteristics of the structure and the desired safety level.

5. References

[1] Gilles D, McClure G. "In situ dynamic characteristics of reinforced concrete shear wall buildings." In Structures Congress, pp. 2235-2245. 2012. https://doi.org/10.1061/9780784412367.196

[2] Karaton M, Osmanlı ÖF, Gülşan ME. “Numerical simulation of reinforced concrete shear walls using force-based fiber element method: effect of damping type and damping ratio.” Bulletin of Earthquake Engineering, 19(14), 6129-6156, 2021. https://doi.org/10.1007/s10518-021-01221-x

[3] Cakil B, Osmanlı OF, Karaton M. “Investigation of damping type and damping ratio for nonlinear seismic analysis of base-isolated steel structures.” Structural Engineering and Mechanics, 95(1), 1-13, 2025. https://doi.org/10.12989/sem.2025.95.1.001

[4] Celep Z, Kumbasar N. "Deprem muhendisligine giris ve depreme dayanikli yapi tasarimi." Istanbul: Beta Dagitim, 2004.

[5] Chopra AK. “Dynamics of Structures: Theory and Applications to Earthquake Engineering.” Prentice Hall, Pearson Education, United States of America, 2012.

[6] Puthanpurayil AM, Lavan O, Carr AJ, Dhakal RP. “Elemental damping formulation: an alternative modelling of inherent damping in nonlinear dynamic analysis.” Bulletin of Earthquake Engineering, 14(8), 2405-2434, 2016. https://doi.org/10.1007/s10518-016-9904-9

[7] Karaton M, Osmanlı ÖF, Gülşan ME. “Investigation of uncertainties in nonlinear seismic analysis of the reinforced concrete shear walls.” International Journal of Civil Engineering, 19(3), 301-318, 2021. https://doi.org/10.1007/s40999-020-00567-8

[8] Mander, J. B., Priestley, M. J., & Park, R. (1988). Theoretical stress-strain model for confined concrete. Journal of structural engineering, 114(8), 1804-1826.

[9] TBEC-2018, Türkiye Building Earthquake Specifications, Disaster and Emergency Management Authority, Ankara, Türkiye, 2018.

[10] SAP2000, Integrated Finite Element Analysis and Design of Structures, Computers and Structures Inc., Berkeley, California, USA.

Evaluation of Ground Motion Attenuation Relationships: Case Study of the 2023 Kahramanmaraş Earthquakes in Elazığ, Türkiye

Ömer Faruk OSMANLI[1] , Muhammet KARATON [2], Burak ÇAKIL[3], Ömer Faruk TAŞ[4], Erkut SAYIN [5], Ozan İNCE [6]

[1,2,3,4,5,6]Civil Engineering, Engineering Faculty, Firat University, Elazığ, Türkiye.
[1]ofosmanli@firat.edu.tr, [2] mkaraton@firat.edu.tr, [3] burakk.cakill@gmail.com, [4] oftas@firat.edu.tr, [5] esayin@firat.edu.tr, [6] o.ince@firat.edu.tr
[1](ORCID: 0000-0002-9799-3103), [2](ORCID: 0000-0002-1498-4659), [3](ORCID: 0000-0002-4972-8874), [4](ORCID: 0000-0002-1431-5316) [5](ORCID: 0000-0003-0266-759X), [6](ORCID: 0000-0002-6319-1129)

Abstract
As earthquakes propagate away from the epicenter, the intensity attenuates depending on various parameters such as local site conditions. The seismic effects generally decrease with increasing distance from the fault zone. However, the extent of this attenuation and the precise impact of an earthquake on distant structures remain uncertain. In the literature, various studies have been conducted to model ground motion attenuation as a function of distance and the magnitude of the earthquake. Nevertheless, the validity of these attenuation relationships must be verified with actual earthquake data. Ground motion attenuation functions are significant for estimating the level of seismic effect imposed on a structure located at a certain distance from the earthquake epicenter, and for assessing the potential extent of structural damage resulting from this effect. In this study, the attenuation characteristics of the 2023 Kahramanmaraş earthquakes were examined by considering several ground motion attenuation functions proposed in the literature. For this purpose, acceleration records from various seismic stations during the 2023 Kahramanmaraş earthquakes were utilized. These acceleration values were compared with the attenuation models' values for Elazığ province.

Keywords: earthquake effects, ground motion attenuation functions, acceleration records.

1. Introduction

Earthquakes are natural disasters that result from the sudden release of energy within the Earth's crust, leading to intense ground shaking and causing significant damage in residential areas. Parameters such as the earthquake's intensity, duration, wave characteristics, and distance to the epicenter directly determine how a structure will be affected by an earthquake. In particular, local soil conditions can significantly amplify or attenuate the seismic forces acting on a structure [1]. The forces induced by earthquakes generate both horizontal and vertical accelerations within the structural system, which can lead to displacements and, consequently, damages. Structural damage caused by earthquakes is often attributed to deficiencies in the design and construction processes [2, 3]. However, in seismic-resistant design, it is essential to consider not only the structure-specific parameters but also the general characteristics of ground motion and the variation of these ground motions depending on the location.

Turkey is a seismically active country where seismic activity is frequently observed. Major earthquakes such as the 2020 Elazığ earthquake, the 2020 Aegean Sea earthquake, and the 2023 Kahramanmaraş earthquakes have resulted in significant loss of life and the collapse or severe damage of many buildings. In particular, the impacts of the 2023 Kahramanmaraş earthquakes have demonstrated that loss of lives and structural damage can occur not only in areas near the epicenter and fault rupture but also in regions located at considerable distances from the epicenter [4–7]. To address this uncertainty, mathematical models have been developed to describe how ground motion parameters attenuate with increasing distance from the epicenter. These models, known as ground motion attenuation functions, play a critical role in estimating the potential impact of an earthquake on structures within a given region. By considering various factors, such as earthquake magnitude, focal depth, wave propagation direction, and local site conditions, these functions help predict the expected levels of ground motion at different distances from the epicenter. [8–10]. Furthermore, the use of a reliable attenuation function is crucial not only for the seismic design of new structures but also for the evaluation of seismic performance and the identification of retrofitting needs in existing buildings, thereby ensuring the accuracy of design and assessment procedures.

[1]Corresponding author

In this study, the impact of ground motion attenuation relationships available in the literature was examined for the 2023 Kahramanmaraş Pazarcık and Elbistan earthquakes, specifically with respect to the city center of Elazığ. For this purpose, the distances between the earthquake epicenter and the Elazığ city center were determined, and attenuation functions based on moment magnitude (Mw) were considered. To evaluate the effectiveness of these attenuation models and to identify the attenuation model closest to actual ground motion, acceleration data recorded at 11 seismic stations located around Elazığ during the 2023 Kahramanmaraş earthquakes were obtained from the Turkish Strong Motion Database System (TADAS). The observed peak ground acceleration (PGA) values were compared with those estimated using selected attenuation functions. The model that most accurately approximated the recorded ground motions for this earthquake event was identified.

2. Ground Motion Attenuation Functions

Ground motion parameters in earthquakes, particularly peak ground acceleration (PGA), have been modeled through attenuation equations that describe their reduction with increasing distance from the epicenter as a function of moment magnitude. These equations provide the opportunity to determine the earthquake effects that the structure will be exposed to by modeling the damping of the ground motion as it moves away from the epicenter and its variation depending on the moment magnitude. They are developed through statistical analysis of a large number of accelerometer records and are regionally calibrated by considering parameters such as soil types, fault mechanisms, and local geological conditions. In this study, the effectiveness of several attenuation equations available in the literature was evaluated for the city center of Elazığ by comparing their predictions with acceleration data recorded at selected stations for the 2023 Kahramanmaraş earthquakes.

In the literature, ground motion attenuation functions are generally derived as a function of distance from the epicenter. These functions, although distance-dependent, have been developed based on local magnitude (M_L) [11-15], surface-wave magnitude (M_s) [16-20], and moment magnitude (M_w) [21-27] of the earthquake [28, 29]. Within the scope of this study, only attenuation relationships that incorporate moment magnitude (Mw) were considered. The attenuation functions used in this study, which are based on distance (R) and moment magnitude (Mw), are presented in Table 1. In the table, R represents the distance of the earthquake epicenter to the considered location, M represents the moment magnitude of the earthquake, and A refers to the attenuated ground acceleration value at the considered location.

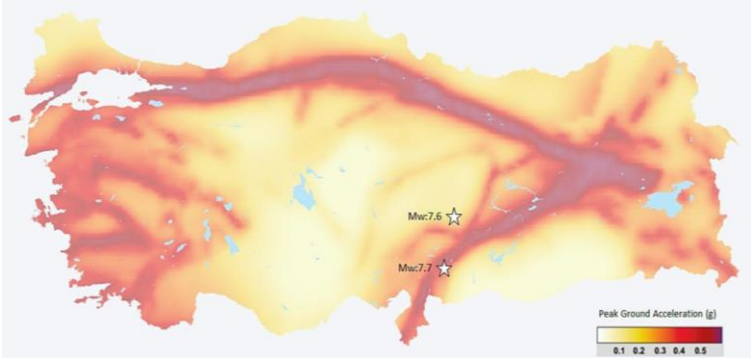
Table 1. Ground motion attenuation functions

$A = 1300\ e^{0.67M}(R+25)^{-1.6}$	Donovan (1973) [21]	(Eq.1)
$A = 1230\ e^{0.58M}(R+25)^{-1.32}$	Donovan (1973) [22]	(Eq.2)
$A = 472.3\ e^{0.64M}(R+25)^{-1.1031}$	McGuire (1974) [23]	(Eq.3)
$A = 69\ e^{0.92M}(R)^{-1.30}$	Orphal and Lahoud (1974) [24]	(Eq.4)
$A = 5000\ e^{0.8M}(R+40)^{-2}$	Shah and Movassate (1975) [25]	(Eq.5)
$LogA = 2.308 + 0.411M - 1.637\log(R+30)$	Katayama (1974) [26]	(Eq.6)
$LogA = 0.505 + 0.538M - \log(R + 0.00810^{0.5M}) - 0.002R$	Ulutas and Ozer (2010) [27]	(Eq.7)

3. 2023 Kahramanmaraş Earthquakes

Türkiye is located in a highly seismically active region, characterized by major active fault zones capable of generating destructive earthquakes. The principal fault systems include the North Anatolian Fault (NAF) and East Anatolian Fault (EAF) zones. The national seismic hazard map of Türkiye has been developed based on the activity and characteristics of these fault systems. The Pazarcık (Mw = 7.7) and Elbistan (Mw = 7.6) earthquakes that occurred on February 6, 2023, resulted in the loss of thousands of lives and the collapse or severe damage of hundreds of thousands of buildings. The Pazarcık earthquake, recorded on February 6 at 04:17 local time at latitude 37.288 and longitude 37.043, occurred at a depth of 8.6 km. On the same day, the Elbistan earthquake occurred at 13:24

local time at latitude 38.089 and longitude 37.239, at a depth of 7.0 kilometers. These two major events, occurring approximately nine hours apart, significantly affected 11 provinces. The seismic hazard map of Türkiye, including the epicenters of the 2023 Kahramanmaraş earthquakes, and the provinces influenced by the consequent earthquakes are illustrated in Figure 1 [2]. Although the number of collapsed buildings in the city center of Elazığ was relatively lower compared to other provinces, a considerable number of buildings suffered severe damage.



(a)



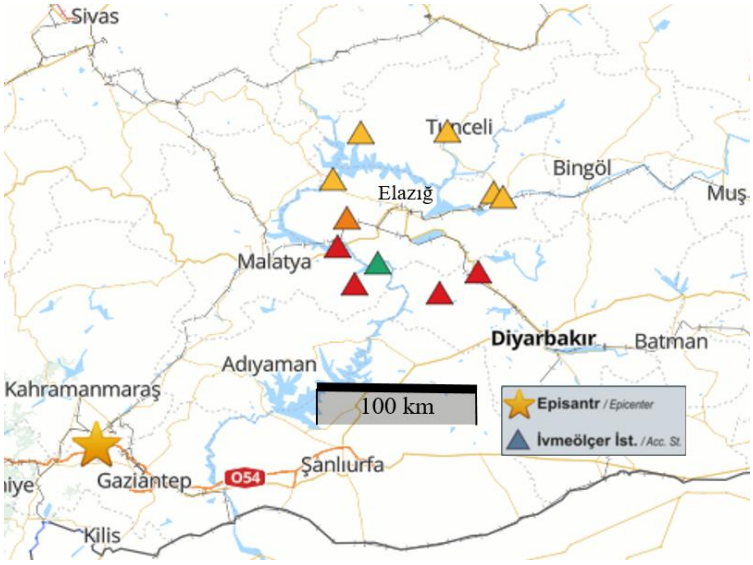
(b)

Figure 1. Epicenters of the Kahramanmaraş Earthquakes and influenced provinces [2]

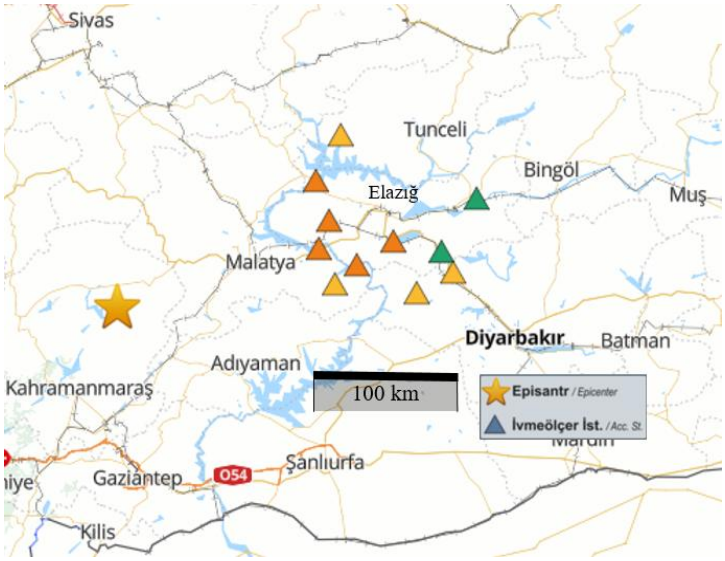
4. Case Study

In this study, to evaluate the effectiveness of ground motion attenuation functions, acceleration records obtained from 11 different accelerometers located around Elazığ were evaulated for the 2023 Kahramanmaraş earthquakes. For each strong motion station, the peak ground acceleration (PGA) values estimated using the attenuation functions described above were compared with the corresponding observed values recorded at the stations. The locations of the strong motion stations considered in this study, corresponding to both the Pazarcık and Elbistan earthquake events, are presented on the map in Figure 2. Considered strong motion records used in this study were obtained from the Turkish Strong Motion Database and Analysis System (TADAS), operated by the Earthquake Department of the Disaster and Emergency Management Authority (AFAD) [30]. Since the strong motion station records taken from the TADAS system are different for each earthquake, the 11 station records closest to the Elazığ city center were used for each earthquake. Station codes, locations, recorded acceleration values and the distance from the epicenter of the earthquake to the location of the station (R) were given in Table 2 and Table 3 for the Pazarcık and Elbistan earthquakes, respectively. Since the attenuation functions available in the literature are typically derived for the horizontal components (North-South (NS) and East-West (EW) directions) of ground motion, the vertical acceleration components (Up-Down (UD) direction) of the earthquake were not taken into consideration in this study.

The distance between the earthquake epicenter and the accelerometer stations ranges from 190 to 295 km for the Pazarcık earthquake, and from 137 to 244 km for the Elbistan earthquake.



(a) Pazarcık Earthquake



(b) Elbistan Earthquake

Figure 2. Considered strong motion stations for the 2023 Kahramanmaraş earthquakes

Table 2. Details of the considered strong motion stations for Pazarcık earthquake

	Station Code	Longitude	Latitude	Province	District	PGA_NS (cm/sec^2)	PGA_EW (cm/sec^2)	R (km)
1	2310	38.82453	38.57266	Elazığ	Baskil	**60.46**	51.20	211.70
2	4414	38.75413	38.40695	Malatya	Kale	106.62	**163.84**	195.07
3	4413	39.03934	38.31353	Malatya	Doğanyol	**13.15**	10.19	209.20
4	4404	38.87385	38.19588	Malatya	Pütürge	136.24	**137.42**	190.02
5	2107	39.48379	38.14594	Diyarbakır	Çermik	74.76	**112.27**	234.92
6	2104	39.759	38.2644	Diyarbakır	Ergani	72.84	**116.47**	262.22
7	2307	39.93199	38.69583	Elazığ	Palu	33.63	**38.11**	297.63
8	2304	39.86293	38.72096	Elazığ	Kovancılar	32.68	**49.40**	293.97
9	6201	39.53469	39.07473	Tunceli	Merkez	15.29	**20.66**	294.77
10	6203	38.91709	39.06302	Tunceli	Çemişgezek	31.83	**36.59**	256.49
11	2309	38.72728	38.79913	Elazığ	Keban	**38.26**	35.34	223.57

Table 3. Details of the considered strong motion stations for Elbistan earthquake

	Station Code	Longitude	Latitude	Province	District	PGA_NS (cm/sec^2)	PGA_EW (cm/sec^2)	R (km)
1	2310	38.82453	38.57266	Elazığ	Baskil	41.47	**55.24**	148.39
2	4414	38.75413	38.40695	Malatya	Kale	**81.41**	63.01	136.95
3	4413	39.03934	38.31353	Malatya	Doğanyol	36.79	**50.93**	159.28
4	4404	38.87385	38.19588	Malatya	Pütürge	45.36	**48.54**	143.46
5	2308	39.3102	38.45063	Elazığ	Sivrice	**69.80**	48.52	185.23
6	2107	39.48379	38.14594	Diyarbakır	Çermik	28.64	**47.61**	196.48
7	2104	39.759	38.2644	Diyarbakır	Ergani	**27.52**	21.22	221.13
8	2302	39.67541	38.39231	Elazığ	Maden	12.36	**16.34**	215.43
9	2307	39.93199	38.69583	Elazığ	Palu	12.96	**15.56**	244.20
10	6203	38.91709	39.06302	Tunceli	Çemişgezek	26.44	**37.09**	181.68
11	2309	38.72728	38.79913	Elazığ	Keban	**62.65**	38.65	151.77

For the comparison of the attenuation functions with the recorded earthquake PGA values, the maximum PGA values of horizontal directions of the recorded earthquake data were taken into consideration. These values are indicated as highlight in Table 2 and 3. The recorded maximum PGA values from the strong motion stations were compared with the calculated acceleration values using attenuation functions given in Equations (1) - (7). The differences between recorded data and calculated values for each attenuation function were given in Table 4 for Pazarcık earthquake, in Table 5 for Elbistan earthquake. The average differences were calculated for each attenuation function to observe the function that provides the best approximation to the measured maximum horizontal PGA values. The minimum differences were obtained as 53% with Eqs. (2), (4), and (7) for the Pazarcık earthquake. Also, the minimum differences for the Elbistan earthquake were obtained as 25% with Eq. (3), as 26% for Eqs. (1) and (2), and as 31% with Eq. (4). It was observed that attenuation functions provide closer results to the data obtained from strong motion stations for the Elbistan earthquake (in 137-244 km range) compared to the Pazarcık earthquake (in 190-295 km range). Evaluation of the recordings from both earthquakes indicates that among the attenuation functions considered, Eq. (2) and Eq. (4) exhibit a more acceptable agreement with the observed ground motion data at the stations.

Among the attenuation functions considered in this study, the acceleration values obtained using Eq. (2) showed the closest agreement with the peak ground acceleration (PGA) values recorded at the measurement stations around Elazığ. Therefore, Eq. (2) was selected as the attenuation function to be used for the Elazığ city center. For the calculation of ground acceleration at the coordinates of Elazığ city center (38.674874299366905, 39.2167341392772), the epicentral distance from the Pazarcık earthquake was measured as 245.27 km, while the distance from the Elbistan earthquake was measured as 184.63 km. Using Eq. (2), the calculated PGA at this location was 65.99 cm/s² for the Pazarcık earthquake and 41.17 cm/s² for the Elbistan earthquake. If a structure located in the Elazığ city center is to be analyzed under the effects of the Kahramanmaraş earthquakes, it is recommended to use Eq. (2) and the corresponding PGA values provided in this study.

Table 4. Comparison of recorded acceleration values with attenuation functions for Pazarcık earthquake

Station Code		2310	4414	4413	4404	2107	2104	2307	2304	6201	6203	2309	Average Difference (%)
Max Horizontal PGA (cm/sec^2)		60.46	163.84	13.15	137.42	112.27	116.47	38.11	49.40	20.66	36.59	38.26	
Distance (R) (km)		211.70	195.07	209.20	190.02	234.92	262.22	297.63	293.97	294.77	256.49	223.57	
Eq.(1)	A (cm/sec^2)	35.95	40.40	36.57	41.93	30.95	26.38	21.90	22.31	22.22	27.25	33.25	
	Difference (%)	68.1	305.6	64.0	227.7	262.7	341.4	74.0	121.5	7.0	34.3	15.1	**138**
Eq.(2)	A (cm/sec^2)	78.62	86.55	79.73	89.25	69.48	60.90	52.24	53.03	52.85	62.54	73.70	
	Difference (%)	23.1	89.3	83.5	54.0	61.6	91.2	27.0	6.8	60.9	41.5	48.1	**53**
Eq.(3)	A (cm/sec^2)	53.16	58.45	53.90	60.24	47.07	41.33	35.53	36.06	35.94	42.43	49.88	
	Difference (%)	13.7	180.3	75.6	128.1	138.5	181.8	7.3	37.0	42.5	13.8	23.3	**77**
Eq.(4)	A (cm/sec^2)	77.98	86.73	79.19	89.74	68.11	59.04	50.07	50.88	50.71	60.76	72.64	
	Difference (%)	22.5	88.9	83.4	53.1	64.8	97.3	23.9	2.9	59.3	39.8	47.3	**53**
Eq.(5)	A (cm/sec^2)	37.36	42.84	38.12	44.74	31.32	25.92	20.76	21.22	21.12	26.93	34.07	
	Difference (%)	61.8	282.5	65.5	207.1	258.5	349.4	83.5	132.8	2.2	35.9	12.3	**136**
Eq.(6)	A (cm/sec^2)	37.26	41.88	37.90	43.46	32.07	27.31	22.65	23.07	22.97	28.21	34.45	
	Difference (%)	62.2	291.3	65.3	216.2	250.1	326.5	68.3	114.2	10.1	29.7	11.1	**131**
Eq.(7)	A (cm/sec^2)	79.16	92.74	81.03	97.45	64.10	50.64	37.90	39.02	38.78	53.16	70.96	
	Difference (%)	23.6	76.7	83.8	41.0	75.2	130.0	0.5	26.6	46.7	31.2	46.1	**53**

Table 5. Comparison of recorded acceleration values with attenuation functions for Elbistan earthquake

Station Code	2310	4414	4413	4404	2308	2107	2104	2302	2307	6203	2309	Aver
Max Horizontal PGA (cm/sec^2)	55.24	81.41	50.93	48.54	69.80	47.61	27.52	16.34	15.56	37.09	62.65	

													Average Difference (%)
Distance R (km)		148.39	136.95	159.28	143.46	185.23	196.48	221.13	215.43	244.20	181.68	151.77	
Eq.(1)	A (cm/sec²)	55.33	61.71	50.19	57.94	40.65	37.40	31.59	32.79	27.37	41.77	53.64	
	Difference(%)	0.2	31.9	1.5	16.2	71.7	27.3	12.9	50.2	43.1	11.2	16.8	**26**
Eq.(2)	A (cm/sec²)	55.73	62.12	50.58	58.35	40.99	37.71	31.84	33.06	27.57	42.12	54.05	
	Difference(%)	0.9	31.1	0.7	16.8	70.3	26.3	13.6	50.6	43.6	11.9	15.9	**26**
Eq.(3)	A (cm/sec²)	61.57	69.79	55.02	64.92	43.08	39.08	32.05	33.49	27.05	44.47	59.42	
	Difference(%)	10.3	16.7	7.4	25.2	62.0	21.9	14.1	51.2	42.5	16.6	5.4	**25**
Eq.(4)	A (cm/sec²)	74.76	81.70	69.06	77.61	58.18	54.37	47.39	48.86	42.18	59.49	72.90	
	Difference(%)	26.1	0.4	26.2	37.5	20.0	12.4	41.9	66.6	63.1	37.7	14.1	**31**
Eq.(5)	A (cm/sec²)	111.89	122.43	103.24	116.22	86.76	80.99	70.46	72.67	62.60	88.73	109.07	
	Difference(%)	50.6	33.5	50.7	58.2	19.6	41.2	60.9	77.5	75.1	58.2	42.6	**52**
Eq.(6)	A (cm/sec²)	112.88	125.29	102.95	117.94	84.61	78.37	67.20	69.52	59.07	86.76	109.62	
	Difference(%)	51.1	35.0	50.5	58.8	17.5	39.2	59.0	76.5	73.7	57.3	42.8	**51**
Eq.(7)	A (cm/sec²)	133.55	152.52	118.32	141.30	90.29	80.82	64.10	67.55	52.20	93.57	128.55	
	Difference(%)	58.6	46.6	57.0	65.6	22.7	41.1	57.1	75.8	70.2	60.4	51.3	**55**

5. Conclusions

In this study, the effects of the 2023 Kahramanmaraş earthquakes on Elazığ province were investigated using various ground motion attenuation functions based on the moment magnitude (M_w) and available in the literature. Peak ground acceleration (PGA) values recorded at 11 different seismic stations were compared with the values calculated using 10 attenuation models. The comparison of several attenuation functions with strong motion recordings from the 2023 Kahramanmaraş earthquakes demonstrated that Equations (2) and (4) yielded the closest agreement with the observed PGA values around Elazığ. In particular, Equation (2) showed the lowest average difference and was found to be most suitable for estimating ground motion at Elazığ city center. The attenuation functions provide closer results to the data obtained from strong motion stations for the Elbistan earthquake compared to the Pazarcık earthquake. The distance from the epicenter of the earthquake to the Elazığ city center were 245.27 km for the Pazarcık earthquake (M_w=7.7) and 184.63 km for the Elbistan earthquake (M_w=7.6), resulting in estimated PGAs of 65.99 cm/s² and 41.17 cm/s², respectively. Therefore, Equation (2) and the corresponding PGA values to be adopted for structural analyses of buildings located in Elazığ under the seismic effects of the Kahramanmaraş earthquakes are recommended.

6. References

[1] Miari, M., Jankowski, R. (2022). "Incremental dynamic analysis and fragility assessment of buildings founded on different soil types experiencing structural pounding during earthquakes." Engineering Structures, 252, 113118. https://doi.org/10.1016/j.engstruct.2021.113118

[2] Ince, O. (2024). "Structural damage assessment of reinforced concrete buildings in Adıyaman after Kahramanmaraş (Türkiye) Earthquakes on 6 February 2023." Engineering Failure Analysis, 156, 107799. https://doi.org/10.1016/j.engfailanal.2023.107799

[3] Özmen, A., İnce, O., Taş, Ö. F., Atar, M., Ozbulut, O. E., Sayın, E. (2024). "Seismic performance assessment of structures in Malatya Province after 6 February 2023 Turkey earthquake sequences." Natural Hazards, 1-34. https://doi.org/10.1007/s11069-024-07017-x

[4] Altunsu, E., Güneş, O., Öztürk, S., Sorosh, S., Sarı, A., Beeson, S. T. (2024). "Investigating the structural damage in Hatay province after Kahramanmaraş-Türkiye earthquake sequences." Engineering Failure Analysis, 157, 107857. https://doi.org/10.1016/j.engfailanal.2023.107857

[5] Kazaz, İ., Avşar, Ö., Dilsiz, A. (2024). "Importance of building inspection on the seismic response of a severely damaged RC structure during the February 6, 2023 Kahramanmaraş earthquake sequence." Engineering Failure Analysis, 162, 108410. https://doi.org/10.1016/j.engfailanal.2024.108410

[6] Yuzbasi, J. (2024). "Post-earthquake damage assessment: field observations and recent developments with recommendations from the Kahramanmaraş earthquakes in Türkiye on February 6th, 2023 (Pazarcık M7. 8 and Elbistan M7. 6)." Journal of Earthquake Engineering, 1-26. https://doi.org/10.1080/13632469.2024.2353864

[7] Vuran, E., Serhatoğlu, C., Timurağaoğlu, M. Ö., Smyrou, E., Bal, İ. E., Livaoğlu, R. (2025). "Damage observations of RC buildings from 2023 Kahramanmaraş earthquake sequence and discussion on the seismic code regulations." Bulletin of Earthquake Engineering, 23(3), 1153-1182. https://doi.org/10.1007/s10518-023-01843-3

[8] Boore, D. M., Joyner, W. B., Fumal, T. E. (1997). "Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes: A summary of recent work." Seismological research letters, 68(1), 128-153.

[9] Ulusay, R., Tuncay, E., Sonmez, H., Gokceoglu, C. (2004). "An attenuation relationship based on Turkish strong motion data and iso-acceleration map of Turkey." Engineering Geology, 74(3-4), 265-291. https://doi.org/10.1016/j.enggeo.2004.04.002

[10] Fahjan, Y., Pakdamar, F., Eryılmaz, Y., Kara, F. (2015). "Afet planlamasında deprem riski belirsizliklerinin değerlendirilmesi." Doğal Afetler ve Çevre Dergisi, 1(1-2), 21-39.

[11] Denham, D., Small, G.R., Everingham, I.B. (1973). "Some strong-motion results from Papua New Guinea 1967-1972, In Proceedings of Fifth World Conference on Earthquake Engineering, Cilt. 2, s. 2324-2327.

[12] Cornell, C.A., Banon, H., Shakal, A.F. (1979). "Seismic motion and response prediction alternatives," Earthquake Engineering and Structural Dynamics, Cilt. 7(4), s. 295-315.

[13] Gaull, B.A. (1988). "Attenuation of strong ground motion in space and time in southwest western Australia," In Proceedings of Ninth World Conference on Earthquake Engineering, Cilt. 2, s. 361-366.

[14] Marin, S., Avouac, J.P., Nicolas, M., Schlupp, A. (2004). "A probabilistic approach to seismic hazard in metropolitan France," Bulletin of the Seismological Society of America, Cilt. 94(6), s. 2137-2163.

[15] Zuccolo, E., Bozzoni, F., Lai, C.G. (2017). "Regional low-magnitude GMPE to estimate spectral accelerations for earthquake early warning applications in southern Italy," Seismological Research Letters, Cilt. 88(1), s. 61-71.

[16] Castillo, W.T., Lopez, P.S., Dahle, A., Bungum, H. (1992). "Digitization of strong motion data and estimation of PGA attenuation." Technical report no: 2-4, NORSAR, Kjeller, Norway, Reduction of Natural Disasters in central America Earthquake Preparedness and Hazard Mitigation Seismic Zonation and Earthquake Hazard Assessment.

[17] Wang, B.Q., Wu, F.T., Bian, Y.J. (1999). "Attenuation characteristics of peak acceleration in north China and comparison with those in the eastern part of North America," Acta Seismologica Sinica, Cilt. 12(1), s. 26-34.

[18] Aydan, Ö. (2001). "İstanbul Boğazı denizaltı geçişi için tüp tünel ile kalkan tünelin uygunluğunun karşılaştırılması," Jeoloji Mühendisliği Dergisi, Cilt. 25(1), s. 1-17.

[19] Shi, S., Shen, J. (2003). "A study on attenuation relations of strong earth movements in Shanghai and its adjacent area," Earthquake Research in China, Chinese, Cilt. 19, s. 315-323.

[20] Cui, J.W., Zhang, J.G., Gao, D., Duan, J.X., Wang, T. (2012). "The ground motion attenuation relation for the mountainous area in Sichuan and Yunnan." In Proceedings of Fifteenth World Conference on Earthquake Engineering, 149, Lisbon, Portugal.
[21] Donovan, N. C. (1973). "A statistical evaluation of strong motion data including the Feb. 9 1971, San Fernando earthquake." In Proc. 5th World Conf. on Earthq. Eng., Rome (Vol. 2).
[22] Donovan, N. C. (1973). "Earthquake hazards for buildings, in Building Practice for Disaster Mitigation," National Bureau of Standards Building Science Series 46, 82-111.
[23] McGuire, R. K. (1974). "Seismic structural response risk analysis, incorporating peak response regressions on earthquake magnitude and distance." Report R74-51, Structures Publication, (399).
[24] Orphal, D. L., Lahoud, J. A. (1974). "Prediction of peak ground motion from earthquakes." Bulletin of the Seismological Society of America, 64(5), 1563-1574. https://doi.org/10.1785/BSSA0640051563
[25] Shah, H. C., Movassate, M. (1975). "Seismic risk analysis of California state water project." In Proc. 5th Europ. Conf. Earth. Eng., Istanbul.
[26] Katayama, T., (1974) "Statistical analysis of peak accelerations of recorded earthquake ground motions," SEISAN-KENKYU, Univ. of Tokyo, 26-1, 18-20.
[27] Ulutas, E., Ozer, M.F. (2010) "Empirical attenuation relationship of peak ground acceleration for Eastern Marmara region in Turkey," Arabian Journal of Science and Engineering, Cilt. 35, s. 187−203.
[28] Uyanık, O., Ekin, N., Çoşkun, O. (2021). "Zemin ve deprem parametrelerinden ivme azalım ilişkisi." Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi, 23(68), 575-593. https://doi.org/10.21205/deufmd.2021236820
[29] Güzel, M. (2009). "Mikrobölgeleme çalışmalarında jeolojik, jeofizik, jeoteknik verilerin birlikte kullanımı (Kuzey Adana örneği)." Çukurova Üniversitesi Fen Bilimleri Enstitüsü, Doktora Tezi.
[30] Ministry of Interior of Türkiye, Disaster and Emergency Management Authority (AFAD), Earthquake Department (under AFAD), "Turkish Strong Motion Database and Analysis System (TADAS)" https://tadas.afad.gov.tr

Road Surface Crack Detection Using Deep Learning in Smart Cities

Fatmanur Polat[1], Muhammed Emre Çolak[2], İrfan Kılıç [3]

[1,3]Department of Software Engineering, Faculty of Engineering, Fırat University, Elazığ, Türkiye.
[2]Department of Artificial Intelligence and Data Science Engineering, Faculty of Engineering, Fırat University, Elazığ, Türkiye.
[1]fatmannurpolat@gmail.com, [2]memrecolak@firat.edu.tr, [3]irfankilic@firat.edu.tr
[1](ORCID: 0009-0005-5374-9933), [2](ORCID: 0000-0003-2317-2385),[3](ORCID: 0000-0001-5079-2825)

Abstract
This study aims to develop an automatic and high-accuracy system for detecting cracks on asphalt road surfaces as part of smart city applications. The proposed method employs deep learning techniques, specifically a MobileNetV2-based convolutional neural network, trained on a large, balanced dataset of 40,000 images. Preprocessing steps such as resizing, normalization, and data augmentation were applied to improve model generalization. The system achieved 99.78% accuracy on the training set and 99.73% on the test set, demonstrating its capability for real-time operation on low-cost devices. The results indicate the proposed approach is a feasible and practical solution for urban infrastructure maintenance planning, with potential integration into mobile and embedded systems for real-world deployment.

Keywords: asphalt crack, deep learning, image processing, smart city, road maintenance

1. Introduction

Rapid urbanization has increased the load on transportation infrastructure, causing various types of road surface degradation. Cracks on asphalt surfaces pose safety hazards and lead to higher maintenance costs. Traditional detection methods are labor-intensive, time-consuming, and prone to human error. Smart city technologies aim to address these challenges by integrating automated monitoring systems and data-driven decision-making into urban management. Early, accurate crack detection enables more efficient maintenance planning, offering both economic and environmental benefits. Recent studies highlight the success of deep learning approaches, particularly convolutional neural networks (CNN), YOLO, EfficientNet, and MobileNet models, achieving high accuracy (99%) while being lightweight enough for real-time and mobile applications [1], [2], [3], [4], [5].

This paper presents a MobileNetV2-based deep learning system trained on the public “Surface Crack Detection” dataset from Kaggle. The model is optimized for low-cost device integration and evaluated with metrics such as accuracy, F1-score, and Dice coefficient to assess its suitability for smart city applications.

1.1. Related Works

The timely and accurate detection of cracks on asphalt surfaces is of great importance for ensuring the sustainability of urban transportation infrastructure, increasing road safety, and planning maintenance and repair processes more effectively [6], [7]. Traditional crack detection methods generally rely on field observations and manual measurements, resulting in high labor requirements, time loss, and vulnerability to human error [8]. Especially in regions where urbanization is rapidly increasing, the applicability of these methods has decreased, and the need for automatic, fast, and reliable solutions has become evident.In this context, in recent years, artificial intelligence (AI)-based methods have offered groundbreaking advances in the detection and classification of asphalt cracks, particularly through image processing, machine learning (ML), and deep learning (DL) techniques [9], [10]. Deep learning models are frequently preferred in this field due to their ability to perform automatic feature extraction, generalization learning from large datasets, and high accuracy rates [11]. Convolutional neural network (CNN)-based architectures, such as EfficientNet, YOLOv8s, Faster R-CNN, and ResNet; It attracts attention with both its real-time detection capacity and mobile compatibility [12], [13]. For example, Chen et al. [14] achieved over 94% accuracy by analyzing preprocessed crack images with an EfficientNet-based model; the study reported that data augmentation strategies significantly increased model performance. Similarly, the Faster R-CNN-based system proposed by Zou et al. [15] successfully distinguished various crack types (e.g., block, linear, alligator cracks) and provided 30% faster and more accurate results than traditional methods. An improved lightweight semantic segmentation model based on

[1]Corresponding author

BiSeNetv2 detected road surface cracks with high accuracy by effectively combining low- and high-level features, and achieved an improvement of up to 10.14% in F1 score compared to conventional models [16]. Another study proposed CP-YOLOX and SViT based deep learning models for the detection of structural deteriorations (cracks, voids, weak bonds, etc.) on the road surface using 3D GPR horizontal plane images [17]. In particular, methods that use manually extracted features, such as HOG and LBP, in conjunction with classifiers, such as SVM and KNN; It has achieved high success with low parameter requirements. To overcome the limitations of traditional laser scanning-based methods, Jin et al. combined the oriented gradient histogram with the improved watershed algorithm to effectively detect road cracks based on direction and edge features on range images [18]. Li et al. aimed to increase both accuracy (95.2%) and efficiency (95.8% recall) in visual safety inspections of dams by proposing a method that enables automatic, rapid, and high-accuracy detection of cracks on the dam surface using the YOLOv8 algorithm on high-resolution images collected by an UAV [19]. Hu et al. proposed MobiLiteNet, a lightweight DL-based road nuisance detection method, to improve road infrastructure management and intelligent transportation systems. This method combines the techniques of channel attention, KD, and model compression to achieve high accuracy and low computational cost. It can be integrated into mobile devices to overcome traditional inspection challenges [20].

This study aims to develop a crack detection system that simultaneously meets the key criteria of high accuracy, low hardware requirements, and real-time performance, within the framework of the literature summarized above. To this end, an optimized model is proposed by integrating the deep learning-based MobileNetV2 architecture, data augmentation strategies, and traditional classifiers. In terms of accuracy and field applicability, the proposed model is competitive with the current approaches in the literature.

2. Material and Methods

2.1. Dataset

We used the Surface Crack Detection dataset shared on Kaggle [21], containing 40,000 color images (227×227 pixel) evenly split between positive (crack) and negative (no crack) classes. The data was divided into training (80%), validation (10%), and test (10%) sets.

- Positive: Images containing cracks (20,000 ima
- Negative: Crack-free images (20.000 images)

The balanced structure of the dataset ensures equal distribution across classes. The images were obtained from various surface types (asphalt, concrete, etc.) and represent a single surface segment. The dataset was split into training, validation, and test segments at a ratio of 80% / 10% / 10%. Accordingly, the distribution is as shown in Table 1.

Table 1. Class distribution of road crack detection dataset

Dataset	Number of images
Training Set	~32.000
Validation Set	~4.000
Test Set	~4.000

The images were loaded by opening the .zip file in Google Colab. The directory structure was reorganized to suit the ImageDataGenerator function.

2.2. Preprocessing and Data Augmentation

Various preprocessing and data augmentation techniques were applied to the images to enable the model to generalize better:

- Resizing to 224×224 pixels
- Normalization to [0–1]
- Augmentation techniques:
  - Horizontal flipping
  - Random rotation (±20°)
  - Zoom (20%)
  - Brightness adjustments

These were applied via ImageDataGenerator in Keras, with only resizing and normalization used for validation and test sets.

2.3. The Proposed Deep Learning Model

We adopted a transfer learning approach using MobileNetV2 [22] pre-trained on ImageNet. The base network was frozen, and only the final layers were retrained for binary classification (crack/no crack) using a sigmoid activation. The architecture uses "inverted residual blocks" with expansion, depthwise convolution, and projection layers to reduce computational cost while maintaining accuracy. Dice Loss was used as the primary loss function to handle class imbalance effectively.

A transfer learning approach was adopted in this study, and the MobileNetV2 architecture was chosen as the base model. This architecture, which was previously trained on ImageNet, was chosen for its low parameter count and high accuracy. MobileNetV2 is a lightweight and fast CNN architecture suitable for integration into mobile devices. A key feature of the model is its bottleneck blocks, which minimize information loss by preserving features from input to output.

During model training, the body of the pre-trained MobileNetV2 network was left frozen, and only the final layers were retrained. This allows the model to learn the required outputs for the new task and progress faster using the previously trained weights' information (Figure 1).

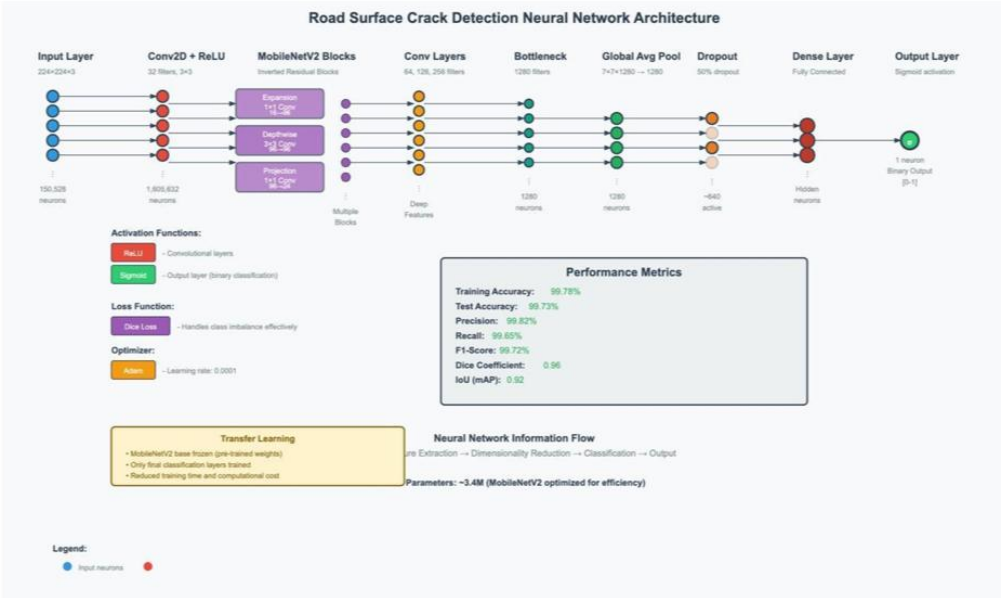


Figure 1. The proposed deep learning model architect

The model's output performs binary classification using a sigmoid activation function that generates values between 0 and 1. This structure allows the model to probabilistically distinguish between the presence and absence of cracks. The model structure is schematically illustrated in Figure 1. This structure reduces both training time and memory usage, allowing the model to be used in real-time applications. An image classification-based approach was used to detect asphalt cracks in this study. This study aims to design an optimized and lightweight DL architecture that determines whether the input surface images contain cracks.

Training Parameters;

- Optimizer: Adam
- Learning rate: 0.0001
- Batch size: 32
- Epochs: 50 (with EarlyStopping)
- Callbacks: ReduceLROnPlateau, EarlyStopping
- Dropout: 50%
- Environment: Google Colab with NVIDIA Tesla P100 GPU

The following steps were taken to combat overfitting:

- ReduceLROnPlateau: The learning rate was reduced when the validation loss became stationary.
- EarlyStopping: Training was stopped when validation performance did not improve.
- Dropout: 50% was applied.
- Data augmentation: The model was allowed to see various variations.

3. Results and Discussion

After our model was trained according to the applications and libraries given in Table 2, the performance results given in Table 3 were obtained.

Table 2. Hardware and software components used in training environment

Component	Version
Python	3.8
TensorFlow	2.x
Keras	Integrated (TF 2.x)

The Dice score and the intersection over union (IoU) or mean average precision (mAP) value were also calculated. The Dice score and mAP value yielded state-of-the-art results.

- Training accuracy: **99.78%**
- Test accuracy: **99.73%**
- Dice coefficient: **0.96**
- IoU (mAP): **0.92**

Precision, recall, and F1-score were also above 99%, demonstrating strong performance.

Table 3. Model training performance metrics

	Precision	Recall	F1-Score	Accuracy
Train	%99.87	%99.70	%99.78	%99.78
Test	%99.82	%99.65	%99.72	%99.73

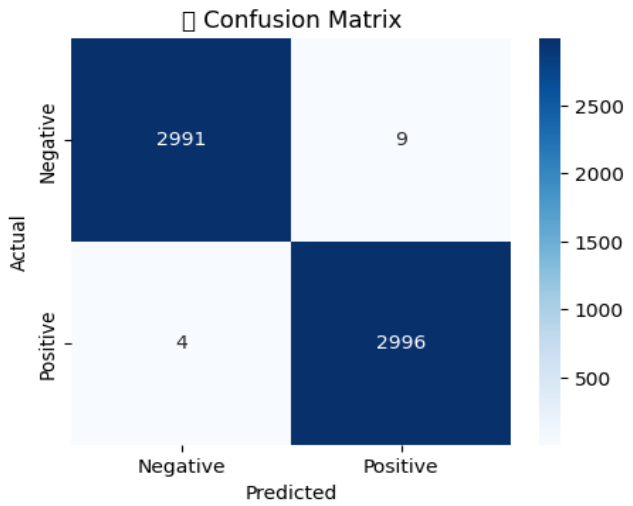


Figure 2. Confusion Matrix

The confusion matrix showed balanced and high-quality predictions across classes (Figure 2). However, false negatives were observed primarily in images with very small or low-contrast cracks.

Accuracy, precision, recall and F1-score were calculated according to equation (3-6).

$$ccuracy = \frac{TP+TN}{TP+FN+TN+FP} \quad (3)$$

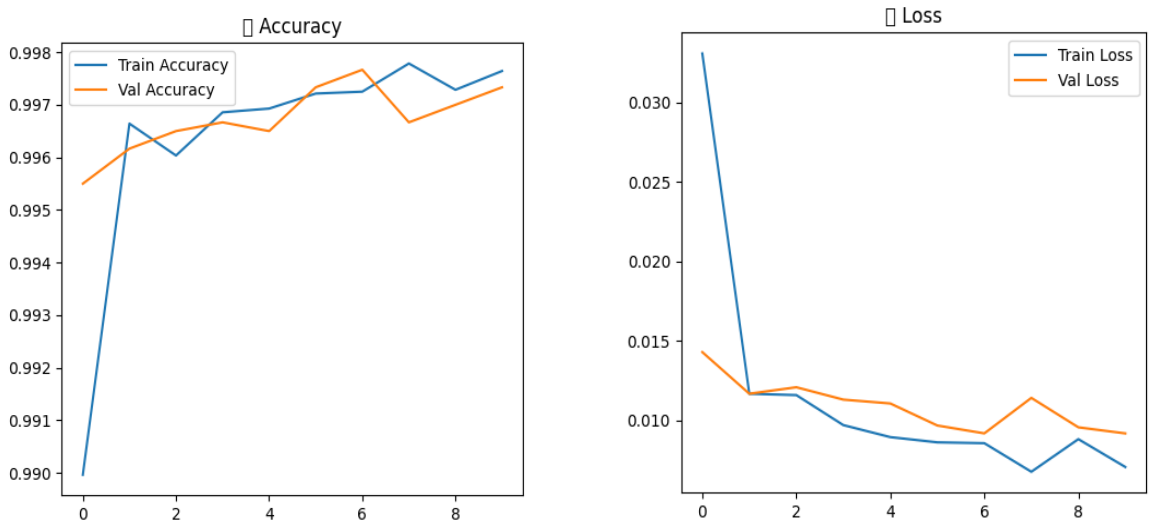
$$recall = \frac{TP}{TP+FP} \quad (4)$$

$$precision = \frac{TP}{TP+FN} \quad (5)$$

$$F1 - score = \frac{recall\ x\ precision}{recall + precision} \quad (6)$$

TP: True-Positive, FN= False-Negative, TN: True-Negative, FP: False-Positive

Figure 3 shows the accuracy and loss graphs for training and testing.



Samples x 1000
a) Accuracy

Sample x 1000
b) Loss

Figure 3. Accuracy and loss plots a) Accuracy b) Loss

The accuracy graph shows the model's correct classification rate on the training and validation data at the end of each epoch. The loss graph visualizes the change in the loss values during the training process, representing the model's error. Figure 4 presents the predictions of the proposed model using the corresponding sample images. Figure 5 shows the sample images that the model incorrectly predicted.

Based on these values, the following evaluation metrics were calculated:

- Accuracy: The proportion of all samples that the model correctly predicted.
- Precision: The fraction of samples predicted as belonging to the positive class that were actually positive.
- Recall: The fraction of true positives that were correctly predicted.
- F1-Score: The harmonic mean of Precision and Recall is important for imbalanced classes.
- Dice Coefficient: A similarity measure that shows the extent to which the true and predicted classes overlap.
- IoU: The ratio of the intersection of the predicted and true crack regions to the union.

According to the results, the model generally demonstrated high performance on the test set. However, the FN rate increased on some surfaces, particularly those with small cracks. This suggests that some fine cracks may be overlooked due to their blending with the surface's natural patterns. Figure 4 shows the images that our model predicted correctly.



Figure 4. Examples correctly classified by the proposed model

Figure 5 shows the images in which the model incorrectly predicted that the objects were outside the road. The accuracy of the model was evaluated visually using examples of correct and incorrect predictions. Correct predictions reveal the model's strengths, whereas incorrect predictions reveal areas for improvement. Figure 6 shows where the model focused on its predictions by extracting heat memories using the Grad-CAM method.

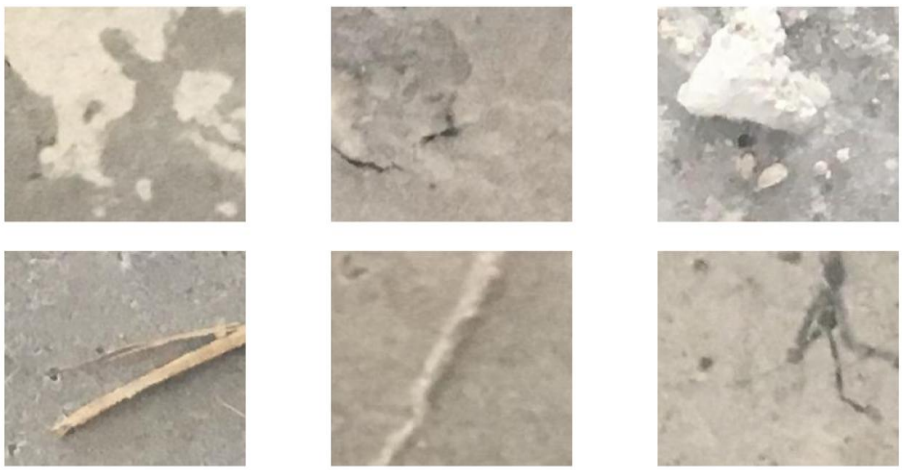


Figure 5. Examples incorrectly classified by the proposed model

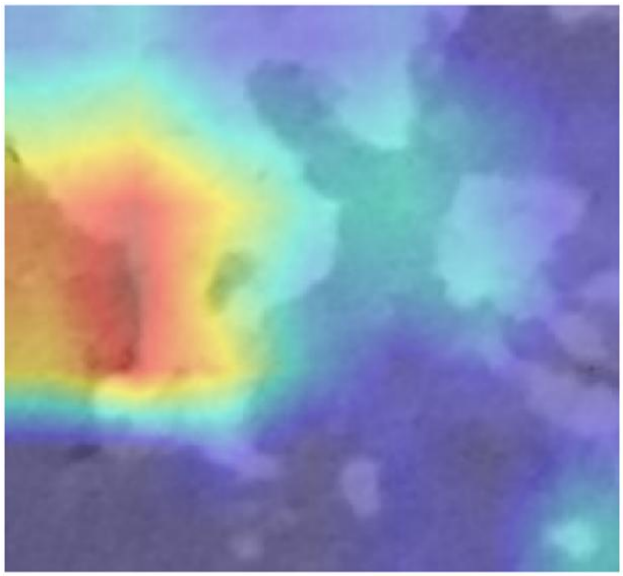


Figure 6. Heat map created with the Grad-CAM method

Grad-CAM displays the regions in the image to which the model pays the most attention during classification as a heat map. This visual aids in interpreting the decision-making process of the model and understanding the areas it focuses on. Figure 7 shows the number of broken/unbroken data points in the training dataset after model training.

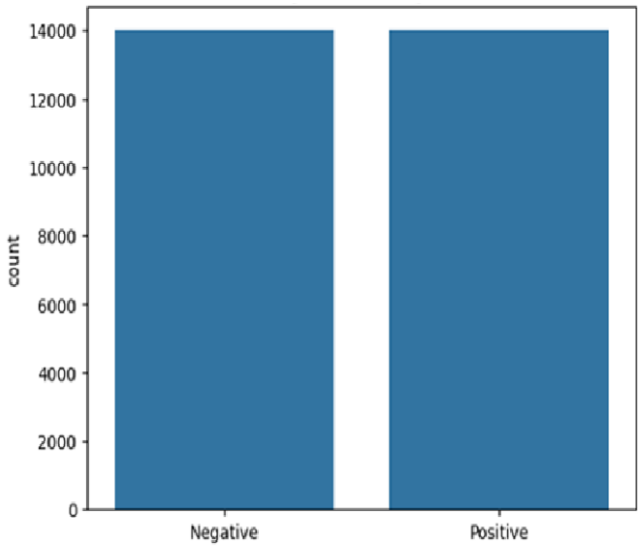


Figure 7. Class distribution plot in the training dataset

Grad-CAM displays the regions in the image to which the model pays the most attention during classification as a heat map. This visual aids in interpreting the decision-making process of the model and understanding the areas it focuses on. Figure 7 shows the number of broken/unbroken data points in the training dataset after model training.

Correctly classified examples confirmed robust learning. Misclassifications often included images with non-road textures or very subtle cracks. Grad-CAM heatmaps revealed the model focused on crack regions, validating its interpretability. Slight overfitting was detected, with training performance slightly surpassing validation. Strategies to mitigate this included aggressive data augmentation, increased dropout, EarlyStopping, testing with simpler architectures, and cross-validation.

4. Conclusions

This study demonstrates the feasibility of using MobileNetV2-based deep learning models for automated, high-accuracy asphalt crack detection in smart city contexts. The system achieved over 99% accuracy with low computational requirements, making it suitable for mobile and embedded deployment. Future work should focus on expanding datasets to include diverse environmental conditions, implementing multi-class classification to identify crack types, enabling mobile deployment using TensorFlow Lite or ONNX, incorporating continual learning and online updates, collaborating with municipalities for field testing, and openly sharing models and code for community advancement.

Through extensive experimental testing, the deep learning-based asphalt crack detection model developed in this study has demonstrated high accuracy and reliability. Built on the MobileNetV2 architecture, the model achieved an accuracy of 99.70% and above, achieved an F1 score, and demonstrated superior performance with a Dice coefficient (0.96) in both pixel-based segmentation and image-based classification tasks. These results support the practical application of the proposed method. The model successfully detected both fine and complex cracks as well as prominent and wide cracks. Effective data augmentation strategies, a transfer learning approach, and a carefully structured model architecture drive the success of this approach. Furthermore, the low hardware requirements and fast processing time of the model enhance its usability in real-time systems and enable its integration with mobile platforms.

However, some environmental challenges could negatively impact the model's performance. The detection accuracy was particularly low under environmental influences such as low-contrast asphalt surfaces, shadows, and dirt. Furthermore, the model's high adaptation to the training data introduced a tendency for overfitting, leading to performance degradation on data types it has not yet encountered. Although the dataset encompasses various surface and lighting conditions, it is insufficient to cover all real-world possibilities. This demonstrates the need for larger and more representative datasets to enhance the model's generalizability. Overall, this study presents a deep learning-based, lightweight, and mobile-friendly solution for asphalt crack detection, offering a compelling alternative to existing methods in terms of accuracy, efficiency, and practical application.

5. References

[1] R. Xie, M. Chen, M. Tao, K. Ding, and H. Chen, "Gradient and self-attention enabled convolutional neural network for crack detection in smart cities," in *2023 IEEE 29th International Conference on Parallel and Distributed Systems (ICPADS)*, 2023, pp. 2110–2117. doi: 10.1109/ICPADS60453.2023.00286.

[2] Z. Topbaş, Ö. Erdaş Çiçek, and Ş. Gülcü, "DERİN ÖĞRENME İLE ASFALT ÇATLAKLARININ TESPİTİNDE VERİ ARTIRIMI VE EVRİŞİMSEL BLOK SEÇİMİNİN ETKİSİ," *Adıyaman Üniversitesi Mühendislik Bilimleri Dergisi*, vol. 11, no. 23, pp. 172–189, 2024, doi: 10.54365/adyumbd.1453460.

[3] L. S. Furtado, I. S. Bessa, N. de O. Gurjão, and J. B. Soares, "Integrating smart city technologies for sustainable pavement infrastructure," *Canadian Journal of Civil Engineering*, vol. 52, no. 5, pp. 569–582, 2025, doi: 10.1139/cjce-2024-0217.

[4] H. Chu *et al.*, "Deep Learning Method to Detect the Road Cracks and Potholes for Smart Cities," *Computers, Materials & Continua*, vol. 75, pp. 1863–1881, Jul. 2023, doi: 10.32604/cmc.2023.035287.

[5] M. Alkhedher, A. Alsit, M. Alhalabi, S. AlKheder, A. Gad, and M. Ghazal, "Novel pavement crack detection sensor using coordinated mobile robots," *Transp Res Part C Emerg Technol*, vol. 172, p. 105021, 2025, doi: https://doi.org/10.1016/j.trc.2025.105021.

[6] L. Zhang, F. Yang, Y. Daniel Zhang, and Y. J. Zhu, "Road crack detection using deep convolutional neural network," in *2016 IEEE International Conference on Image Processing (ICIP)*, 2016, pp. 3708–3712. doi: 10.1109/ICIP.2016.7533052.

[7] H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama, and H. Omata, "Road Damage Detection Using Deep Neural Networks with Images Captured Through a Smartphone," *ArXiv*, Jul. 2018, doi: 10.48550/arXiv.1801.09454.

[8] H. Oliveira and P. Correia, "CrackIT - An Image Processing Toolbox for Crack Detection and Characterization," in *2014 IEEE International Conference on Image Processing, ICIP 2014*, Jul. 2014. doi: 10.1109/ICIP.2014.7025160.

[9] J. Xing, Y. Liu, and G. Zhang, "Improved YOLOV5-Based UAV Pavement Crack Detection," *IEEE Sens J*, vol. PP, p. 1, Jul. 2023, doi: 10.1109/JSEN.2023.3281585.
[10] Y. Wu, S. Li, J. Li, Y. Yu, J. Li, and Y. Li, "Deep learning in crack detection: A comprehensive scientometric review," *Journal of Infrastructure Intelligence and Resilience*, vol. 4, no. 3, p. 100144, 2025, doi: https://doi.org/10.1016/j.iintel.2025.100144.
[11] M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," Jul. 2019. doi: 10.48550/arXiv.1905.11946.
[12] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," Jul. 2018, doi: 10.48550/arXiv.1804.02767.
[13] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Trans Pattern Anal Mach Intell*, 2017, doi: 10.1109/TPAMI.2016.2577031.
[14] X. Chen *et al.*, "Autonomous Crack Detection for Mountainous Roads Using UAV Inspection System," *Sensors*, vol. 24, no. 14, 2024, doi: 10.3390/s24144751.
[15] Q. Zou, Y. Cao, Q. Li, Q. Mao, and S. Wang, "CrackTree: Automatic crack detection from pavement images," *Pattern Recognit Lett*, vol. 33, no. 3, pp. 227–238, 2012, doi: https://doi.org/10.1016/j.patrec.2011.11.004.
[16] H. Yu, Y. Deng, and F. Guo, "Real-time pavement surface crack detection based on lightweight semantic segmentation model," *Transportation Geotechnics*, vol. 48, p. 101335, 2024, doi: https://doi.org/10.1016/j.trgeo.2024.101335.
[17] J. Yang, K. Ruan, J. Gao, S. Yang, and L. Zhang, "Pavement Distress Detection Using Three-Dimension Ground Penetrating Radar and Deep Learning," *Applied Sciences*, vol. 12, no. 11, 2022, doi: 10.3390/app12115738.
[18] H. Jin, F. Wan, and O. Ruan, "Pavement Crack Detection Fused HOG and Watershed Algorithm of Range Image," Jul. 2018, pp. 475–488. doi: 10.1007/978-3-319-59463-7_47.
[19] X. Li *et al.*, "Dam Crack Detection Studies by UAV Based on YOLO Algorithm," in *2023 2nd International Conference on Robotics, Artificial Intelligence and Intelligent Control (RAIIC)*, 2023, pp. 104–108. doi: 10.1109/RAIIC59453.2023.10281120.
[20] Y. Hu, N. Chen, Y. Hou, X. Lin, B. Jing, and P. Liu, "Lightweight deep learning for real-time road distress detection on mobile devices," *Nat Commun*, vol. 16, no. 1, p. 4212, 2025, doi: 10.1038/s41467-025-59516-5.
[21] "Surface Crack Detection." Accessed: Jul. 12, 2025. [Online]. Available: https://www.kaggle.com/datasets/arunrk7/surface-crack-detection/data
[22] [A. G. Howard *et al.*, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *ArXiv*, vol. abs/1704.0, 2017, [Online]. Available: https://api.semanticscholar.org/CorpusID:12670695

International Conference on Advances and Innovations in Engineering

3.ICAIE

Barrette Pile Foundations and Application Principles: A Case Study of the Çeşmeli–Kızkalesi Highway

Ahmet ALTUNBEY[1], Mesut GÖR[2], Rahim Khan NASERY [3], Kübra SERTKAYA[4]

[1,3] Graduate School of Natural and Applied Sciences, Fırat University, Elazığ, Türkiye.
[2] Department of Civil Engineering, Engineering Faculty, Fırat University, Elazığ, Türkiye.
[4] Civil and Structural Engineering, University of Sheffield,Sheffield, UK.
[1] ahmetaltunbey@gmail.com, [2] mgor@firat.edu.tr, [3] rahim.k.nasery@gmail.com, [4] ksertkaya1@sheffield.ac.uk
[1](ORCID: 0009-0007-8350-304X), [2](ORCID: 0000-0002-5463-9278),[3](ORCID: 0000-0001-8835-4783) ,[4](ORCID: 0000-0001-9723-6674)

Abstract
Barrette piles are deep foundation elements frequently preferred in engineering structures requiring high load-bearing capacity and lateral stiffness. The aim of this study is to present the application principles of barrette piles and to evaluate the main points that should be considered on site. As a field sample, four barrette pile foundations constructed at Viaduct 4 site, located in the Çeşmeli–Kızkalesi section of the Çeşmeli–Taşucu Highway Project within the borders of Mersin province, were examined. The research emphasized the critical role of controlling the viscosity and density parameters of the bentonite suspension in ensuring trench stability, and demonstrated that the use of guide walls prevents excavation deviations, thus preserving the designed cross-sectional geometry. Furthermore, the effects on pile integrity and strength of producing reinforcement cages with the required rigidity and dimensions in accordance with the project and placing them precisely, as well as conducting continuous and controlled concreting using the tremie method, were addressed. Field observations revealed that negligence in these parameters could negatively affect pile performance and structural safety. Consequently, the findings indicate that strict adherence to these aspects is of critical importance for cost-effectiveness and long-term structural performance.

Keywords: barrette piles, pile design, soil stability, deep foundation applications.

1. Introduction

Barrette piles are defined as cast-in-place concrete piles with a rectangular cross-section. Barrette piles are a relatively new deep foundation solution that has recently begun to be implemented in Turkey, and are preferred especially in engineering structures where high load-bearing capacities are required [1]. Unlike conventional circular piles, they have rectangular, elliptical, or specially shaped cross-sections. Their dimensions generally vary between 60–120 cm in width and 200–400 cm in length. They are often preferred due to their high lateral stiffness, increased section moment, and excavation support characteristics [2]. Moreover, compared to conventional circular bored piles, they offer significant advantages by providing a larger peripheral surface area for the same cross-sectional area [3]. Typically, they are constructed by placing a reinforcement cage into a trench excavated using diaphragm wall equipment and then concreting. In structural engineering, the proper selection of foundation systems is critically important for structural safety and economic efficiency. Particularly in projects involving heavy loads and complex ground conditions, such as high-rise buildings, bridge piers, and metro stations, foundation systems that provide high rigidity and large load-bearing capacity are needed. At this point, barrette piles stand out as a system offering significant advantages over conventional circular piles. Barrette piles are generally rectangular cast-in-place reinforced concrete piles formed by placing reinforcement cages into trenches excavated with diaphragm wall equipment and then pouring concrete. Depending on the project, their cross-sections can also be designed in special geometries such as H, T, L, Y or I shape [4]. In general, foundations must safely transfer superstructure loads to the ground. However, when ground conditions are inadequate or the load-bearing capacity requirements are very high, deep foundation systems are preferred to transfer the loads to deeper and stronger soil layers. Pile foundations are the most commonly used type of deep foundation for this purpose, preventing excessive settlements by transferring the weight and service loads of the structure to sublayers with higher bearing capacity [5]. Pile foundations are of vital importance, especially in heavy engineering structures such as bridges and high-rise buildings. For example, pile applications reaching depths of hundreds of meters have been observed in skyscraper projects in Hong Kong, Malaysia, and the United Arab Emirates [6]. However, conventional circular piles sometimes prove insufficient in carrying heavy loads, and in order to increase their load-bearing capacity, either their diameters must be enlarged up to certain limits or multiple piles must be used in groups. This leads to increased costs and construction difficulties. In addition, the limited peripheral surface area

[1]Corresponding author

of circular piles relative to their cross-sectional area emerges as a significant factor restricting load capacity in friction pile systems [5]. For these reasons, the idea of achieving a greater peripheral friction area with the same volume of concrete has paved the way for the development of a new type of pile in the literature, and over the last few decades, the use of rectangular piles called "barrette piles" has become widespread [7]. Particularly in Far Eastern and Middle Eastern countries, it is often preferred to meet the function of multiple circular piles with a single large-section barrette pile in high-rise building foundations [5]. With advancements in construction technologies, their use has become increasingly widespread, as they offer higher shaft friction capacity than circular piles with the same cross-sectional area, thanks to their larger friction surface [1].

2. Material and Method

This study covers the barrette pile applications at Viaduct 4, located in the Çeşmeli–Kızkalesi section of the Çeşmeli–Erdemli–Silifke–Taşucu Highway Project in Mersin, Turkey (Figure 1). At the examined site, four barrette pile foundations were constructed in regions where ground conditions varied. In the initial stage of the study, detailed geological surveys were carried out, and borehole investigations aimed at profiling the subsurface were performed to determine parameters such as layer thicknesses, unit weights, cohesion, and internal friction angles. This enabled the foundation design to be optimized based on site-specific engineering data. Viaduct 4 features a balanced cantilever system and consists of a total of 21 piers. At four of these piers, 144 barrette piles were constructed, with 36 piles under each pier. The pile drilling operations were conducted with the aid of pre-constructed guide walls to maintain the cross-sectional dimensions and alignment of the barrette piles. Thanks to the precise placement of the guide walls, the drilling advanced along the desired line and geometry, preventing directional deviations. To maintain trench stability during drilling, bentonite slurry was used; its viscosity and density were regularly measured on site to ensure they remained within ideal ranges, thereby preventing trench wall collapse or deformation. Reinforcement cages, prepared to project-specified dimensions and rigidity, were carefully lowered into the boreholes with cranes, maintaining their structural integrity. Special attention was given to preserving geometric accuracy during reinforcement placement. Concreting was performed using the tremie method to prevent segregation and void formation. Through tremie pipes, concrete was placed continuously and in a controlled manner, resulting in a homogeneous and durable reinforced concrete structure along the pile. Throughout the construction process, quality control tests and site inspections were carried out regularly. The findings indicated that the process was successful both technically and economically. This comprehensive material and method approach provides an important field example for the construction of barrette piles in accordance with high engineering standards.



Figure 1. Google Earth image of the relevant viaduct

2.1 Manufacturing Stages and Application Principles of Barrette Piles

Barrette piles are constructed using different types of equipment depending on the excavation depth (headroom) and geotechnical properties of the area where the excavation will take place. In this process, excavation operations are generally carried out using either a mechanical grab or a more advanced technology such as a hydrofraise (Figure 2). Mechanical

grab equipment provides high efficiency, particularly in less hard soils, while hydrofraise systems are advanced technology machines capable of reaching rock formations and allowing a more uniform cross-sectional geometry to be achieved. Hydrofraise machines perform cutting with their rotating cutter drums and stabilize the excavation by keeping the excavated material suspended in the bentonite slurry through a continuous circulation system. Thanks to these methods, it becomes possible to excavate deep and large-section barrette piles in a homogeneous and stable manner, ensuring the safety of the trench walls and creating a suitable foundation cavity for subsequent reinforcement and concreting operations.



Figure 2. Grab and hydrofraise

2.1.1 Structural and Functional Importance of Guide Walls in Barrette Pile Construction

Guide walls used in the construction of barrette piles are critically important for ensuring the correct positioning and verticality of the piles. These walls serve as primary reference points for determining the pile axis and accurately guiding the excavation equipment during the digging process (Figure 3). Thus, they guarantee that the piles are constructed along the lines and orientations anticipated in the design projects. Guide walls are generally built as reinforced concrete structures, providing a reliable working platform that can withstand long-term site loads and possible external effects thanks to their strength and geometric precision. By clearly defining the starting and ending points of the pile axes, these walls enable precise control over the starting and ending coordinates of the excavation equipment. Moreover, guide walls play an important role not only in guiding the excavation but also in limiting horizontal deformations that may occur during digging and enhancing trench stability. In this way, axis deviations and cross-sectional deformations, which may be particularly observed in large-section barrette piles, are prevented. In the example application, since barrette piles were present at four foundations (P5L, P6R, P6L, and P7R) of Viaduct 4 located in the Çeşmeli–Kızkalesi section, guide wall construction was carried out only at these piers (Figure 4). This ensured that the excavation operations adhered strictly to the axes defined in the project, achieving the desired criteria in terms of both geometric precision and structural safety during construction.



Figure 3. Guide wall (guide trench) construction

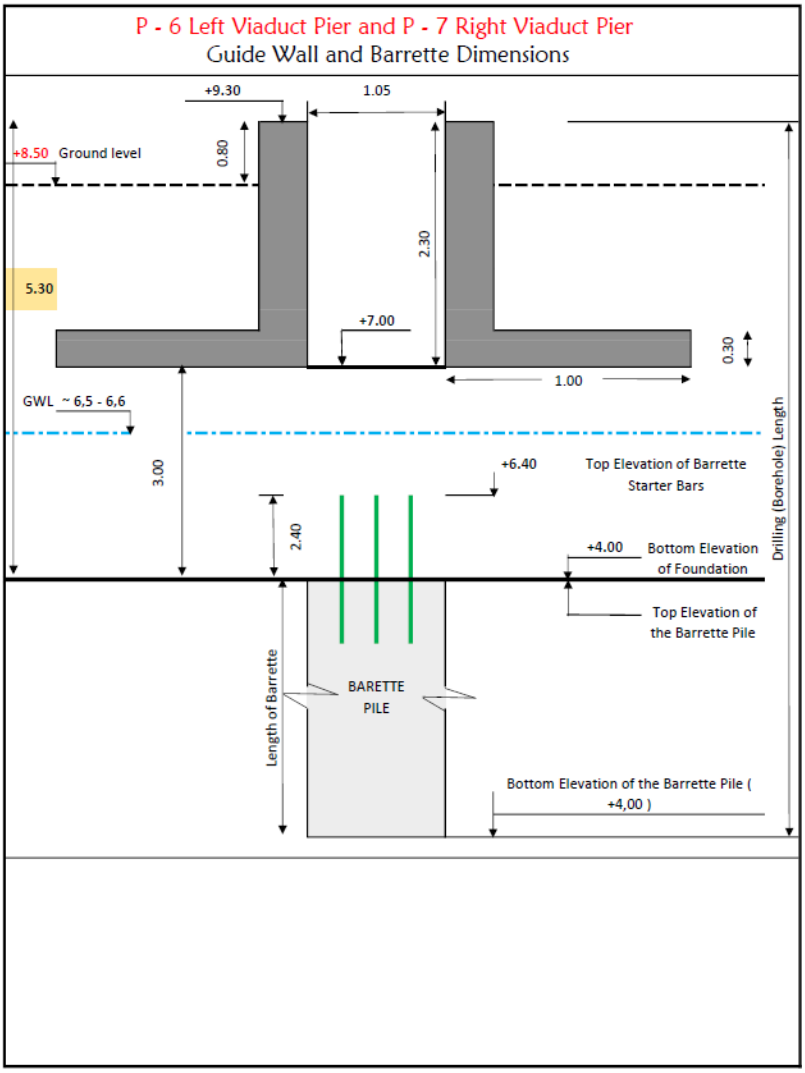


Figure 4. Guide wall details

2.1.2. Preparation of Supporting Bentonite Slurry

One of the most important stages in barrette pile excavation is the stabilization of the excavation walls. To prevent the collapse of the walls during excavation, bentonite- or polymer-based support fluids (slurries) are used [8]. These silt or polymer suspensions circulated along the excavation walls keep the interior of the excavation stable by excluding groundwater. In the literature, these materials are referred to as supporting fluids and are composed of clay minerals (bentonite) or swellable polymers. By carefully controlling the consistency of the suspension, constant support is maintained along the excavation depth; thus, when the fluid is slightly displaced during subsequent concrete pouring, the risk of wall collapse is eliminated [9]. In the sample project, the pH value of the bentonite suspension was regularly measured, ensuring that it remained within the range of 8.5 to 10.5 (Figure 5). During excavation, the bentonite suspension forms a thin, low-permeability filter cake on the borehole walls, preventing the soil from collapsing. This protective barrier plays a crucial role in keeping the excavation stable and securely maintaining the walls.



Figure 5. pH control of bentonite suspension

At the initial stage of the pile construction process, a bentonite suspension is prepared to ensure stability. As the excavation progresses, this suspension is continuously pumped into the borehole to prevent the collapse of the walls. Once the excavation is complete, a reinforcement cage prepared according to the project specifications is placed into the borehole, followed by the careful lowering of tremie pipes. Concreting is carried out through the tremie pipes, starting from the bottom of the borehole and progressing upward. During this process, the poured concrete displaces the bentonite suspension, pushing it out of the borehole. The bentonite, which has partially lost its properties due to interaction with the soil, is then removed and sent to a system called a “desander unit,” where it is separated from coarse particles such as sand and silt. The cleaned bentonite suspension is subsequently transferred to preparation tanks for reuse. This method allows the repeated use of bentonite, a high-cost material, thereby supporting both environmental sustainability and providing economic benefits [1]. The bentonite plant and desander unit are shown (Figure 6).



Figure 6. Bentonite plant and desander plant

2.1.3. Reinforcement Fabrication and Installation

In barrette pile applications, the reinforcement cage is a fundamental component that ensures the load-bearing capacity and structural integrity of the pile. It not only increases the stiffness of the pile, enabling the safe transfer of axial and lateral loads from the structure to the ground, but also contributes to crack control and ductility behavior of the reinforced concrete system. Therefore, both the correct design of the reinforcement cage and its precise fabrication and placement on site are critically important. The reinforcement cage consists of three main components: longitudinal rebars, stirrups, and diagonal bars. The diameters and spacing of these rebars are determined through engineering calculations that consider the dimensions of the pile, the soil profile, and the loads it will carry. In the example project, longitudinal rebars with a diameter of Ø32 mm and stirrups ranging from Ø16 to Ø20 mm were used. These reinforcement ratios provide sufficient stiffness in terms of both sectional strength and reinforced concrete integrity. Reinforcement cages are generally fabricated on site by welding them in a modular form and then brought to the excavation area. Especially for deep piles, instead of lowering the cage into the borehole in a single piece, it is preferred to assemble and place it segment by segment. This method not only facilitates assembly but also reduces the risk of buckling or deformation of the cage. In our example project, since the boreholes are empty, the cages are suspended. The placement of the cage into the borehole is typically carried out with the aid of a crane (Figure 7). In conclusion, the fabrication and installation of reinforcement cages in barrette piles directly affect not only the structural performance but also the long-term reliability of the system. Reinforcement cages that are properly detailed according to the project and meticulously implemented on site play a decisive role in ensuring that the foundation system meets strength, ductility, and continuity criteria. In this study, three reinforcement cages were used in each barrette pile constructed.

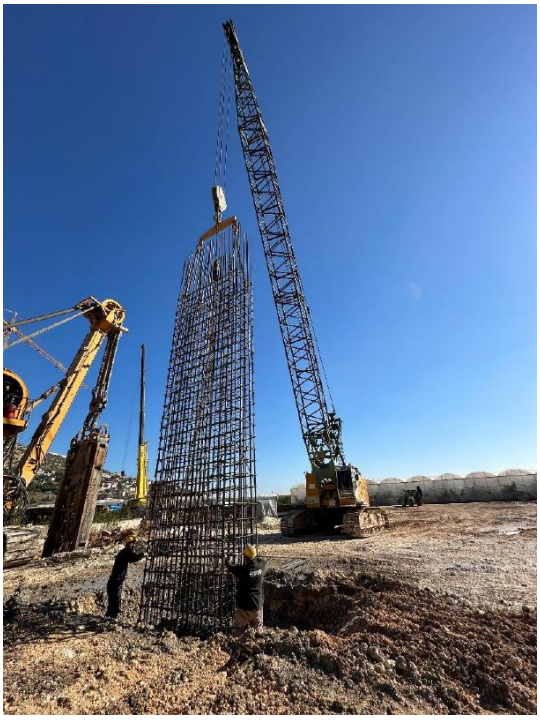


Figure 7. Placement of the reinforcement cage into the borehole

2.1.4. Concrete Production

In the concreting stage of barrette piles, the tremie method is most commonly preferred. This method allows fresh concrete to be placed starting directly from the bottom of the borehole, making it particularly suitable for achieving a homogeneous concrete mass in deep and large-section piles [10]. Pouring concrete continuously through tremie pipes is extremely important as it maintains the flow continuity of the concrete, thereby preventing segregation and void formation. In this way, free fall of the concrete is avoided and the risk of aggregate separation is eliminated. Continuous concrete placement ensures that the concrete fully contacts the walls and forms a void-free, durable structure. In practice, the tremie pipe is gradually lifted so that it always remains embedded in the concrete level, allowing the placement process to proceed in a controlled manner. This facilitates the natural settlement of the concrete under its own weight while also ensuring the integrity of the reinforced concrete pile. In the example project, concreting operations were meticulously carried out using the tremie method, and detailed site observations were recorded throughout the quality control processes (Figure 8).



Figure 8. Application of concrete pouring by the tremie method

2.1.5. Construction of Barrette Pile Foundation

Within the scope of the example project, since the blinding concrete (grobeton) sublevel was determined to be at depths of 5.30 meters and 5.80 meters relative to the existing ground level, the excavation works were planned to reach these levels. The implementation process was shaped according to the positions of the guide (guide trench) walls constructed prior to excavation. In the first stage, a controlled excavation was performed between these guide walls, fully exposing the wall surfaces. Subsequently, the guide walls were broken at the levels specified in the project, and excavation continued (Figure 9). Excavation proceeded until reaching the grobeton level designated as the foundation bottom elevation, where the surface flatness was checked and the blinding concrete was poured. After the grobeton had set, the foundation reinforcement was placed in accordance with the project, and detailed reinforcement arrangements were completed with the necessary concrete covers. In the final stage, the foundation concrete was poured in a way that integrated with the barrette piles (Figure 10), thus completing the construction of the barrette pile foundation system. This implementation process was carried out in harmony with local ground conditions and structural requirements to ensure that the foundation system demonstrated high performance both in terms of load-bearing capacity and on-site execution quality.



Figure 9. Preparation of completed barrette piles for foundation construction

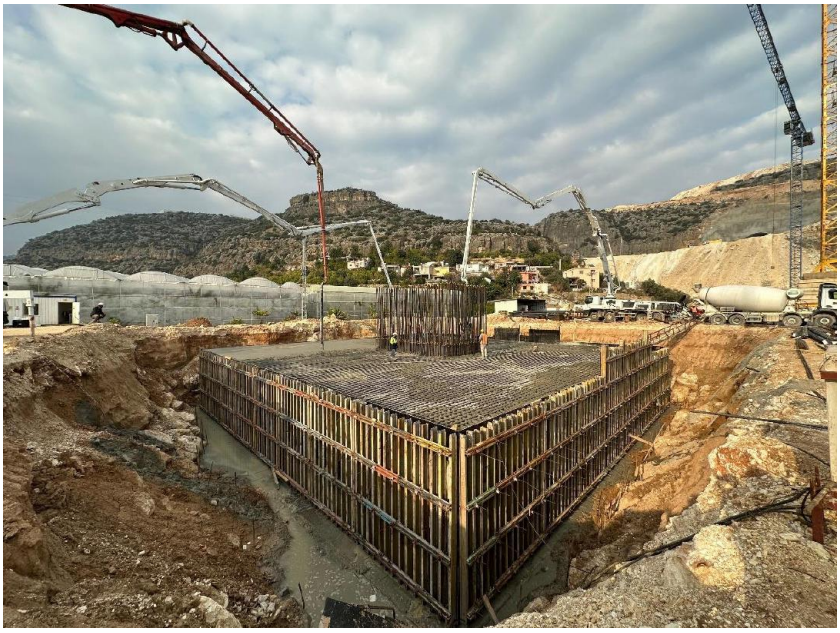


Figure 10. Reinforcement placement and concrete production for the barrette pile foundation

3. Results

The implementations and analyses carried out within the scope of this study have clearly demonstrated that barrette pile systems provide high load-bearing capacity across different soil types. In particular, barrette piles with a rectangular cross-section offer significant advantages by maximizing shaft friction through their larger peripheral surface area. Regular monitoring of properties such as the viscosity, density, and pH of the bentonite suspension used during excavation ensured the stability of borehole walls, thereby minimizing the risks of collapse and deformation. The circulation of bentonite balanced the groundwater pressure and created an impermeable filter cake along the borehole walls, ensuring safe excavation. Additionally, the precise positioning of the guide walls enabled the excavation equipment to operate along the predetermined alignment, thereby preserving the geometric accuracy that is critical for large-section barrette piles. The reinforcement cages were prepared with the required rigidity and dimensions according to the project and carefully lowered into the borehole, directly influencing the load-bearing capacity and stiffness of the reinforced concrete pile. The use of support devices ensured that the cages were placed without deformation. The tremie method ensured a continuous concrete placement, minimizing segregation and voids, and resulting in a uniform and durable reinforced concrete structure. Moreover, keeping the tremie pipes below the concrete level and gradually lifting them allowed full contact between the concrete and the borehole walls, preventing undesirable cavities. All these practices were carried out in light of both field observations and literature-backed theoretical foundations, certifying the engineering solution as a high-quality and reliable project according to scientific standards.

Future research can focus on the long-term performance and durability of barrette piles under varying groundwater conditions and loading scenarios. Comparative studies with other deep foundation systems may further highlight the distinct advantages of barrette piles. Advanced numerical modeling techniques can be utilized to simulate pile-soil interaction with greater accuracy. Additionally, the environmental impact and optimization of bentonite usage across different soil profiles represent valuable areas for further investigation.

Acknowledgements

We would like to extend our sincere thanks to Mr. Yasin Karslı, Project Manager of the Çeşmeli–Kızkalesi Highway Project, Mr. Sina Kiziroğlu, Geotechnical Manager, and Mr. Oktay Karaçıl, Tunnel Manager, for their valuable contributions.

4. References

[1] Yıldız K. Baret kazıklı temellerde düşey taşıma kapasitesi. MSc Thesis, İstanbul Teknik Üniversitesi, İstanbul, Türkiye, 2011.
[2] Tomlinson MJ, Woodward J. *Foundation Design and Construction*. 7th ed. Harlow, United Kingdom, Pearson Education Limited, 2015.
[3] Poulos HG, Chow HSW, Small JC. "The use of equivalent circular piles to model the behaviour of rectangular barrette foundations". *Geotechnical Engineering Journal of the SEAGS & AGSSEA*, 50(3), 106–109, 2019.
[4] Thakare SW, Dhatrak AI. "Performance of groups of barrettes subjected to vertical loading in different types of soils". *International Journal of Research in Civil Engineering and Technology*, 3(2), 37–41, 2022.
[5] Yeşiltepe Ö, Erginer M, Örnek M. "The behavior of axially compression loaded barrette and bored piles with model tests". *Journal of Civil Engineering and Urbanism*, 14(1), 1–10, 2024.
[6] Akl SA, Ismail AN, Abdelmoghni M, Hegazy YA. "Numerical analysis of laterally loaded barrettes performance in cohesionless soils". *Journal of Engineering and Applied Science*, 70(1), 51, 2023.
[7] Ramaswamy SD, Pertusier EM. "Construction of barrettes for high-rise foundations". *Journal of Construction Engineering and Management*, 112(4), 455–462, 1986.
[8] Köpüklü H. *Baret kazıklı temeller ile ilgili hesap yöntemleri, imalat ve kalite kontrol esasları*. MSc Thesis, İstanbul Teknik Üniversitesi, İstanbul, Türkiye, 2015.
[9] Shmeleva N. "Drilling fluids in deep foundation construction: Principles and quality control". *Geotechnical Journal*, 45(2), 123–136, 2019.
[10] Wilkes C, Kumar K, Biscontin G. "Investigating the thixotropic behaviour of Tremie Concrete using the Slump-flow test and the Material Point Method". *Cement and Concrete Composites*, 143, 105222, 2023.

Rheological Properties of Permeation Grout Using Sodium Silicate–Boric Acid Mixtures

Melih UYSAL[1] , Eyübhan AVCI [2], Emre DEVECİ [3]

[1,2,3]Department of Civil Engineering, Faculty of Engineering and Natural Sciences, Bursa Technical University, Bursa, Turkey.
[1]melih.uysal@btu.edu.tr, [2]eyubhan.avci@btu.edu.tr, [3]emre.deveci@btu.edu.tr
[1](ORCID: 0000-0001-6449-8854), [2](ORCID: 0000-0001-7206-0158),[3](ORCID: 0000-0003-3841-4391)

Abstract
In this experimental study, the rheological properties of sodium silicate–boric acid mixtures prepared at different ratios were investigated for potential use in permeation grouting applications. The sodium silicate (SS) used in the mixtures was supplied by Merck, a German manufacturer, while boric acid (BA) was obtained from the Eti Maden Company in Türkiye. The study focused on how the concentration of sodium silicate and the boric acid content (used as a reactant) would affect gelation time and gel formation. The gelation times, syneresis and viscosity values were measured at 20°C and 30°C for 35 mixtures differed in reactant content and SS/solution ratio. The results showed respectively that the ambient temperature immediately impacted on the minimum gelation time of each mixture. In syneresis percentage, as boric acid content increased the syneresis percentage decrease and higher sodium silicate caused increase in the syneresis. In relation to viscosity, an increase in the concentrations of boric acid and sodium silicate led to a significant increase in viscosity. These combined results provide useful information for the design and strength improvement of sodium silicate–based grout series in engineering practice.

Keywords: permeation grouting, sodium silicate, boric acid, gelation time, syneresis, viscosity

1. Introduction

Soil improvement remains a central focus within geotechnical engineering, with chemical grouting methods gaining significant attention in recent academic research for their practical application and strong performance in treating fine-grained soils with low permeability. As a common method for ground improvement, permeation grouting is mainly used to improve soil stability and decrease its permeability. Sodium silicate is an important material in this area, that has unique properties such as viscosity and gelation characteristics. This information is vital to understand and optimize the grout behavior in various soil formations especially when it comes to permeation grouting using sodium silicate. Boric acid has been researched as an effective reactant. It serves as a linker for silica chains to combine and generate the suitable silica gels. (Böschel and Roggendorf, 2002; Obut and Girgin, 2006). When mixing sodium silicate, it reacts with water and undergoes hydrolysis and ionization as the following reaction. When the basic solution is treated with an aqueous boric acid solution, the weakly acidic and inorganic boric acid inserts between the polymeric chain structures of sodium silicate, resulting in a shift to a practically neutral pH of the medium. The reaction pathway between sodium silicate chains and tetragonal borate anions is considered as an esterification reaction with H_2O removal.

The gelation time, or the time it takes for sodium silicate to go from a liquid state to a solid state (from a solution to a gel) is an important part of managing the effective use of grouts. The gelation takes place with the concentration of sodium silicate as well as external conditions like temperature and other additives. Higher sodium silicate concentrations can be used to decrease gelation time providing faster operational capabilities for example where rapid construction is required (Kermani et al., 2015). This rapid gelling allows engineers to alter their grouting strategies to specific soil conditions (Wang and Zhang 2013).

Syneresis is the expulsion of liquid from a gel; when this occurs with sodium silicate, it can negatively affect permeation grout effectiveness. This phenomenon might result in loss of grout volume during the time, and may also weaken structural reinforcement (Lucas et al. 2011). Managing syneresis is crucial because soil structure could be weakened once the grout has excessive water release. The formulation of sodium silicate has been investigated to reduce syneresis and provide a longer lasting grout (Yang et al. 2013).

[1]Corresponding author

The viscosity of a grout, which is based on sodium silicate solution, very much affects the penetration through voids in subsurface. The results of this research showed that viscosity can change along with the SS/solution ratio, which have different effects on the grout flowability and permeating ability during permeation progress (Hu et al., 2024). Additionally, the effect of grout flow behavior between sandy soils due to the viscosity differences results in various efficiency on grouting for each type of grout. (Wang et al., 2023; Wang et al., 2023). For example, higher viscosities may have inhibited the flow of grout and lower viscosities would encourage its penetration into small soil voids (Fraccica et al., 2022).

In this study, interactions between permeability, gelation time and syneresis suggest that a balance of both conditions should be considered for making optimal decisions on the grout properties during the sodium silicate based permeation grouting. Combinations of ratios and additives to be employed can improve performance, which makes soil stabilization feasible while mitigating the drawbacks related to grout applications in different geotechnical conditions.

2. Materials Used in the Experiments

Sodium Silicate

Due to its adjustable viscosity, chemical reactivity and low cost, sodium silicate (water glass) is one of the most widely material globally used in grouting (AASHTO, 2004). Sodium silicate ($Na_2(SiO_2)_nO$ is the chemical material used in this study, a product from Merck (Germany), with product number 105621 (Figure 1). The physical and chemical properties of the sodium silicate used in the experiments are presented in Table 1 (Supelco, 2025).

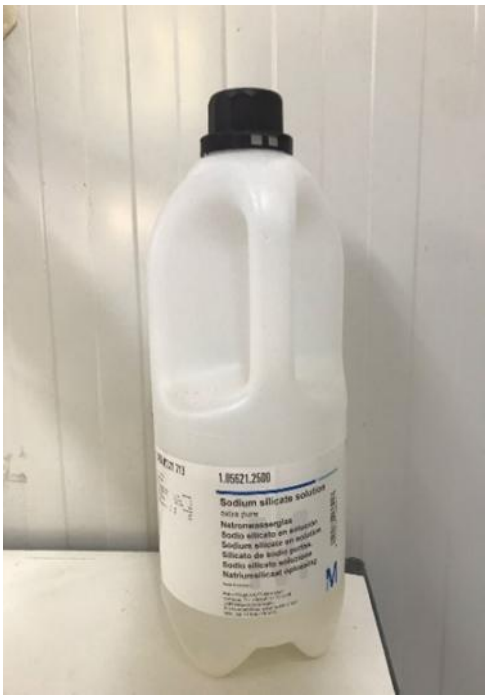


Figure 1. Sodium silicate used in the experiments

Table 1. Properties of sodium silicate.

Physical and chemical properties	Values
Coloring	Colorless, clear liquid
Formulation	Na_2SiO_3
pH (20 °C)	11.80
Weight modules ($SiO_2/NaO_)$	3.0-3.3
Molecular modules ($SiO_2/NaO_)$	3.0-3.4
Density (20 °C, gr/cm³)	1.35-1.40
Be' (20 °C)	38-41
Viscosity (20 °C, cP)	75-150
Na_2O (%)	8.5-9.5
SiO_2 (%)	25.5-30.5

Boric Acid

For the experiments boric acid with chemical formula H_3BO_3 were used. It appears in white crystalline powder and soluble in water. Boric acid, in the case of this study, was the main reactant in grout mixtures based in sodium silicate, acting as a gelation reaction starter and interacting with silicate ions. Boric acid (Figure 2) was procured from the Eti Maden Company, Türkiye. The physical and chemical properties of the boric acid are presented in Table 2 (Eti Maden, 2025).



Figure 2. Boric acid used in the experiments.

Table 2. Properties of boric acid.

Physical and chemical properties	Values
Specific Gravity (20 ºC , g/cm³)	1.51
Molar Weight (g/mol)	61.83
Melting Point (ºC)	450
Boiling Point (ºC)	1860
Purity (%)	Min 99.92
B_2O_3 (%)	Min 56.25
SO_4 (ppm)	Max. 300
Cl (ppm)	Max. 5
Fe (ppm)	Max. 4

3. Experimental Studies

This study was carried out to understand the rheological characteristics of sodium silicate–boric acid mixtures on a more detailed experimental program. An overall 35 different grout mixtures were prepared by changing the proportions of sodium silicate (SS) and boric acid (BA). The experiments were focused on assessing three critical parameters of grout behavior that have significant impacts on grouting performance, namely: gelation time, syneresis and viscosity.

Gelation time measurements were used to assess the setting behavior of the mixtures under different conditions. Syneresis tests on the formed gel were carried out by determining how much water was released over time. Finally, viscosity tests determined the flow properties of the grout mixtures are key factors to investigate the utility of a mixture as grout.

Gelation Time

In applications that use sodium-silicate-based grouting, there are various factors that influence the gelation time such as the nature and quantity of reactants used, ambient temperature and pH, SiO_2/Na_2O ratio, and sodium silicate concentration. In this study, the gelation behavior of sodium silicate-boric acid solutions were investigated concerning ambient temperatures, sodium silicate concentrations, and reactant (boric acid) ratios.

Before preparing the chemical grouts, boric acid was dissolved in water at 40 °C and 60 °C to ensure complete solubility. Each of these grout mixtures were prepared with increasing sodium silicate content, starting from lower concentrations. A total of 35 different mixtures were prepared using various sodium silicate and boric acid proportions and tested under two temperature conditions: 20 °C and 30 °C.

Initially as clear liquids, the mixtures slowly turned into white gel-like solids over time. The stages of gelation observed in the solutions are shown in Figure 3. Table 3 summarizes the gelation times for all mixtures made at both ambient temperatures. The effect of sodium silicate content on gelation time at 20 °C is illustrated in Figure 4, while Figure 5 presents the influence of boric acid content on gelation time at the same temperature. Finally, Figure 6 shows the overall impact of ambient temperature on gelation time across all mixture types.



Figure 3. Colour change of mixtures during gelation.

Table 3. Gelation times of solutions at 20 ºC and 30 ºC.

Mixtures	Sodium silicate (ml)	Water (ml)	Boric acid (g)	Ambient temperature (ºC)	Gelation time (min)	Ambient temperature (ºC)	Gelation time (min)
M1	28	72	1,8	20	550	30	478
M2	28	72	2,0	20	87	30	72
M3	28	72	2,1	20	44	30	36
M4	28	72	2,2	20	39	30	30
M5	30	70	2,0	20	305	30	206
M6	30	70	2,2	20	70	30	60
M7	30	70	2,4	20	35	30	22
M8	30	70	2,5	20	20	30	12
M9	30	70	2,6	20	15	30	8
M10	32	68	2,2	20	317	30	173
M11	32	68	2,4	20	52	30	32
M12	32	68	2,6	20	20	30	10
M13	32	68	2,8	20	6	30	3
M14	34	66	2,2	20	493	30	465
M15	34	66	2,4	20	97	30	48
M16	34	66	2,6	20	24	30	23
M17	34	66	2,8	20	8	30	3
M18	36	64	2,4	20	135	30	85
M19	36	64	2,6	20	32	30	26
M20	36	64	2,8	20	14	30	11
M21	36	64	3,0	20	7	30	3
M22	38	62	2,5	20	175	30	146
M23	38	62	2,6	20	157	30	95
M24	38	62	2,7	20	55	30	49
M25	38	62	2,8	20	30	30	15
M26	38	62	3,0	20	15	30	10
M27	40	60	2,6	20	322	30	284
M28	40	60	2,7	20	231	30	177
M29	40	60	2,8	20	141	30	47
M30	40	60	3,0	20	21	30	11
M31	40	60	3,3	20	10	30	4
M32	42	58	2,7	20	270	30	250
M33	42	58	2,8	20	151	30	142
M34	42	58	2,9	20	93	30	64
M35	42	58	3,0	20	35	30	25

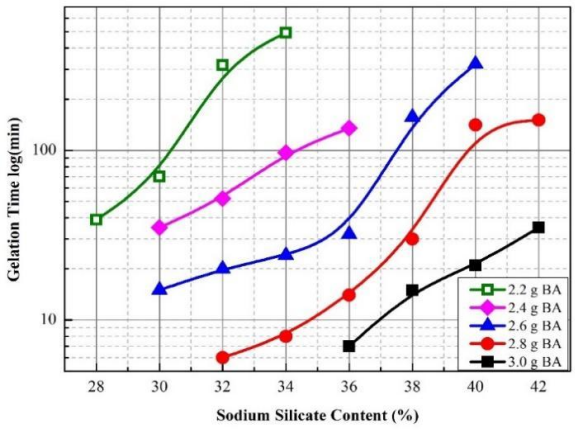


Figure 4. Effect of sodium silicate content on gelation time.

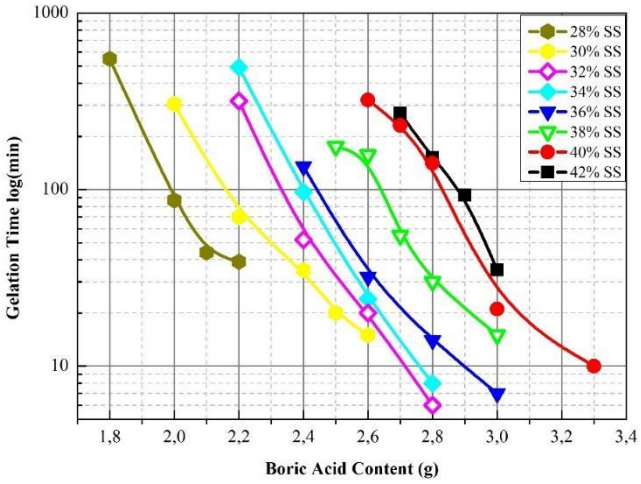


Figure 5. Effect of boric acid content on gelation time

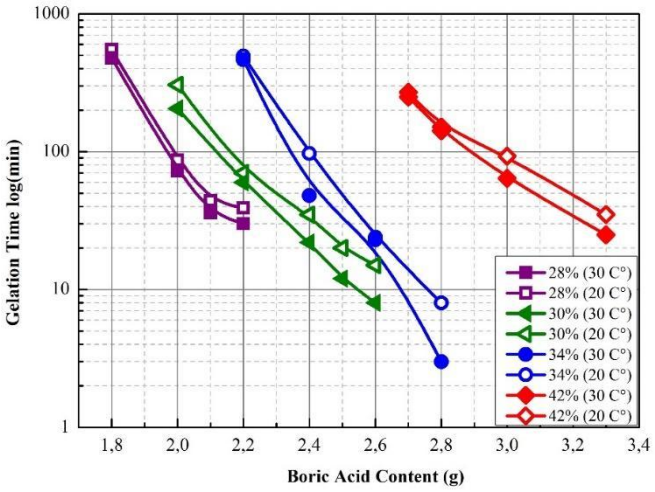


Figure 6. Effect of ambient temperature on gelation time

Syneresis

Syneresis is the process in which a mixture, after completing gelation, releases internal fluid and undergoes a reduction in volume. In this study, Syneresis measurements of the mixed preparations were determined on day 3, 7, 14, 28, 56 to 90 and a final time at day 150 following gelation. Syneresis is related to the quantity of reactant material and sodium silicate present in the mixture.

Syneresis at low levels is better for the stabilization of the gel and it does not mean that the mixture shows all necessary criteria to be used as a potential grout. Other important parameters—such as gelation time, viscosity, and the structural integrity of the gel—are also considered when evaluating overall grout performance. Mixtures that completed their syneresis process are shown in Figure 7, and the corresponding syneresis percentages are presented in Table 4. Effect of sodium silicate and boric acid on syneresis percentage is shown in Figure 8 and Figure 9, respectively.



Figure 7. Mixtures after completing the syneresis process.

Table 4. Syneresis percentages of solutions

Mixtures	Sodium silicate (ml)	Water (ml)	Boric acid (g)	Day 3 syneresis (%)	Day 7 syneresis (%)	Day 14 syneresis (%)	Day 28 syneresis (%)	Day 56 syneresis (%)	Day 90 syneresis (%)	Day 150 syneresis (%)
M1	28	72	1,8	7,27	42,75	63,38	73,23	80,50	85,10	86,89
M2	28	72	2,0	1,33	8,48	27,80	51,01	67,22	73,70	77,28
M3	28	72	2,1	5,70	13,71	29,01	45,88	60,80	68,28	71,15
M4	28	72	2,2	6,55	11,88	25,08	42,37	58,92	65,91	68,16
M5	30	70	2,0	16,24	29,40	45,38	60,69	70,70	79,54	84,60
M6	30	70	2,2	4,47	11,57	22,86	41,04	53,09	63,84	73,03
M7	30	70	2,4	6,12	13,03	22,66	36,88	48,66	59,73	70,49
M8	30	70	2,5	6,12	13,81	23,24	36,15	46,91	59,06	69,61
M9	30	70	2,6	5,58	9,89	16,07	27,53	37,95	48,10	57,01
M10	32	68	2,2	7,57	22,59	41,41	58,82	70,50	79,26	83,94
M11	32	68	2,4	1,94	15,61	36,03	52,16	64,43	74,57	80,04
M12	32	68	2,6	6,85	12,95	22,21	37,50	48,58	60,23	69,59
M13	32	68	2,8	3,61	9,18	18,31	31,48	40,51	51,41	62,27
M14	34	66	2,2	33,01	49,58	63,18	73,53	79,77	84,02	86,06
M15	34	66	2,4	7,42	20,83	40,59	58,56	69,26	77,50	82,14
M16	34	66	2,6	1,74	6,70	25,02	44,97	56,45	68,31	76,24
M17	34	66	2,8	1,34	3,30	14,23	30,65	43,00	56,70	69,04
M18	36	64	2,4	7,98	21,75	41,63	59,90	71,44	80,01	84,60
M19	36	64	2,6	4,26	15,86	36,32	54,61	65,30	73,94	79,07
M20	36	64	2,8	7,69	18,23	32,06	49,14	60,91	72,17	79,84
M21	36	64	3,0	9,27	17,55	28,98	43,48	53,63	65,04	73,64
M22	38	62	2,5	15,35	33,43	53,30	69,36	78,51	83,77	86,45
M23	38	62	2,6	14,09	25,19	40,97	57,02	68,10	77,83	83,91
M24	38	62	2,7	8,31	23,02	48,47	66,91	77,22	83,83	86,91
M25	38	62	2,8	5,68	22,69	42,58	60,37	70,87	78,52	82,44
M26	38	62	3,0	9,76	19,90	35,90	53,09	64,14	73,71	78,58
M27	40	60	2,6	9,29	26,08	50,85	67,74	77,58	84,32	87,46
M28	40	60	2,7	14,31	34,08	55,42	70,89	79,96	85,48	88,24
M29	40	60	2,8	4,53	28,44	47,95	64,14	73,76	80,55	83,97
M30	40	60	3,0	10,08	30,62	48,08	63,38	71,68	78,79	81,86
M31	40	60	3,3	2,17	11,58	27,72	45,43	57,12	68,12	74,67
M32	42	58	2,7	8,78	30,01	53,67	69,49	78,43	84,38	87,56
M33	42	58	2,8	13,09	32,89	53,55	69,40	78,37	83,26	85,51
M34	42	58	2,9	22,13	44,74	65,35	77,47	83,82	85,09	85,26
M35	42	58	3,0	8,18	37,91	62,54	75,85	82,50	83,44	83,53

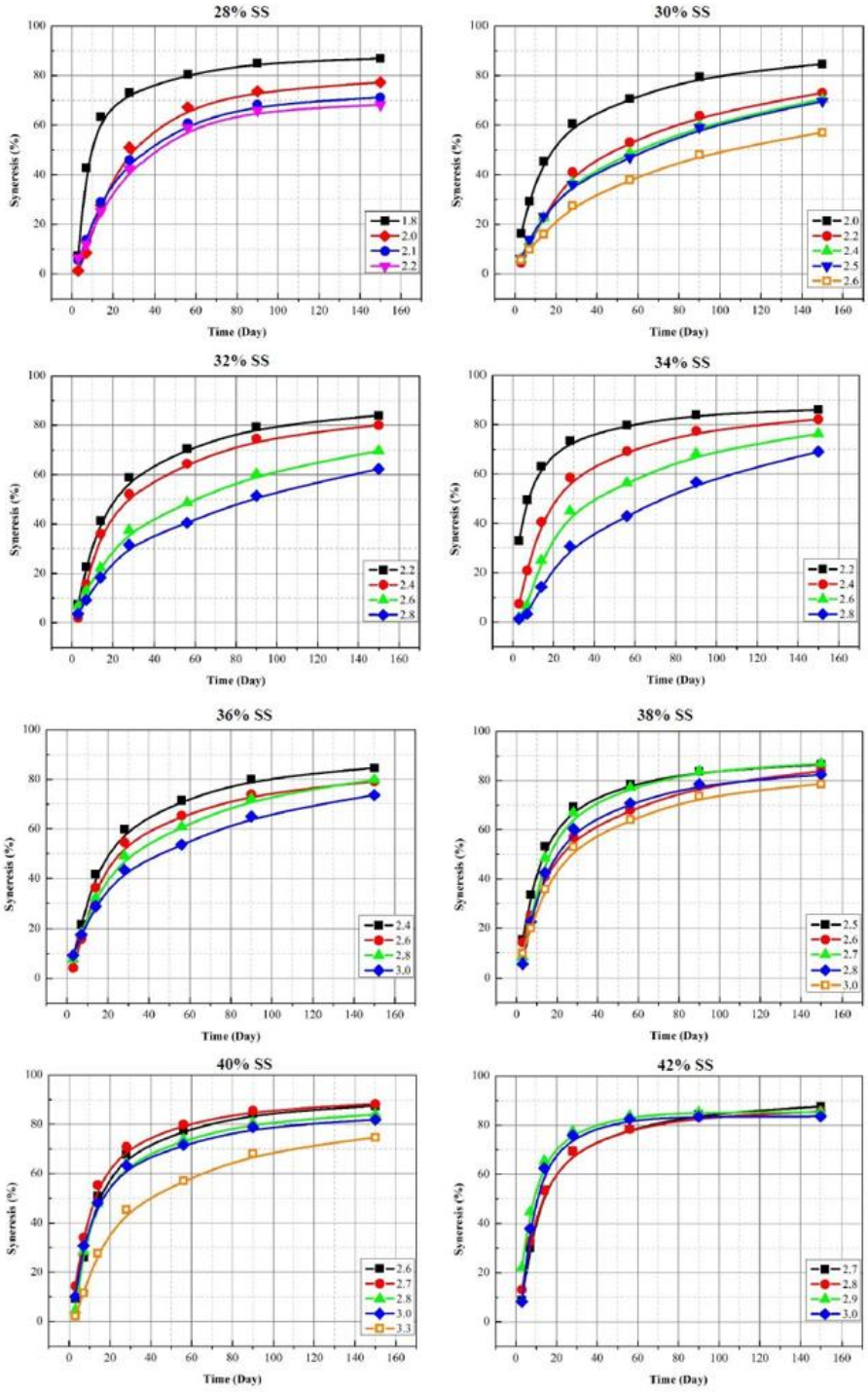


Figure 8. Effect of sodium silicate content on syneresis.

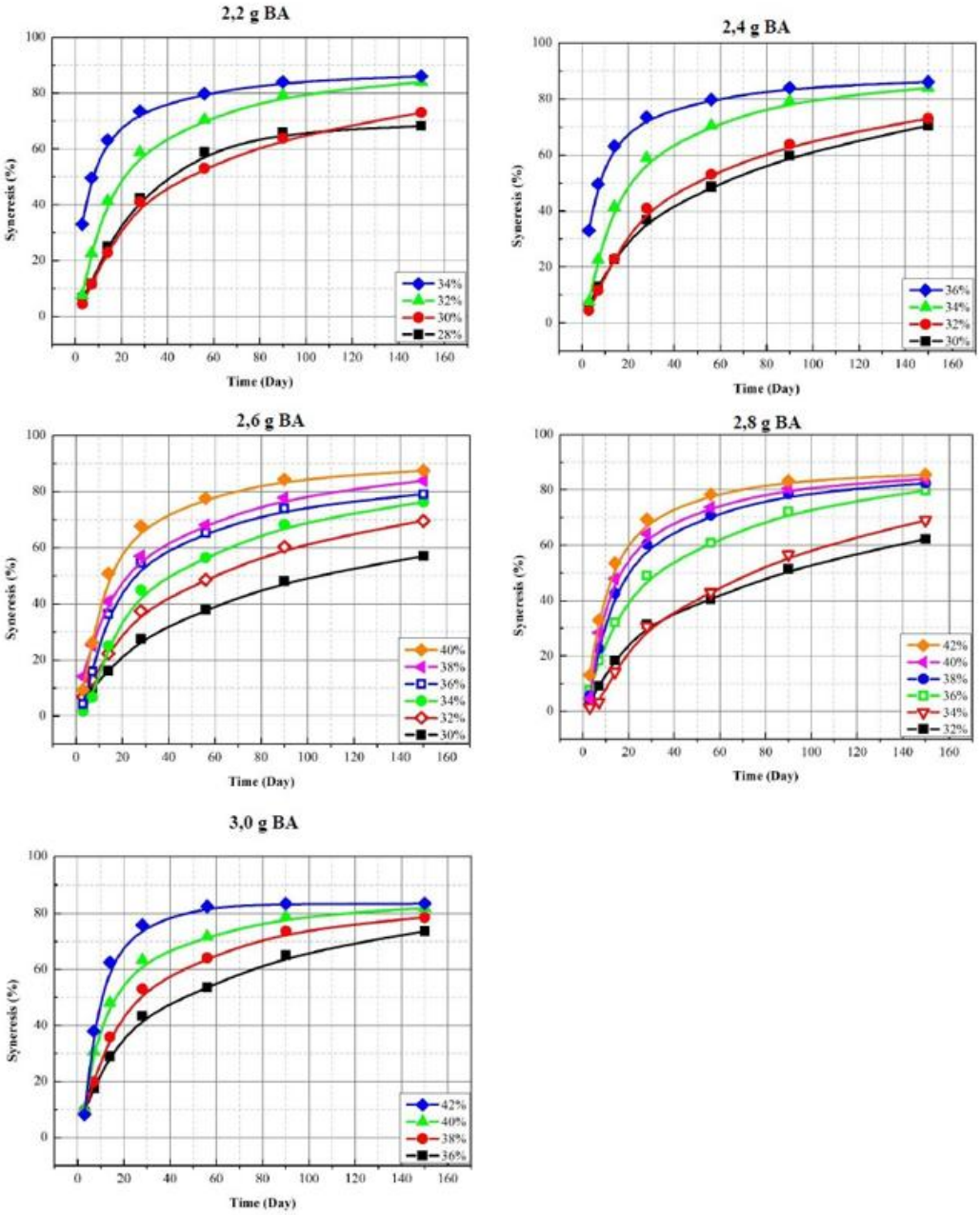


Figure 9. Effect of boric acid content on syneresis.

Viscosity

Viscosity the physical quality of a fluid that measures its resistance to flow, also considered as its internal friction. In chemical grouting applications, viscosity plays a critical role as it directly affects the penetrability of the grout material into the soil. Although viscosity alone is not sufficient to determine the ideal grout mix, it is an important test that should be considered alongside gelation time, syneresis, and gel structure for a comprehensive assessment. In this study, the viscosity of sodium silicate–boric acid solutions were measured using a rheometer. The results, presented in Table 5, show that the increase of the sodium silicate contents in each mixture increased its viscosity. Furthermore, boric acid had a higher viscosity when added with the same sodium silicate concentration. The results suggest that the rheological properties of grout are greatly governed by the silicate content and reactant content.

Table 5. Viscosity values of solutions

Mixtures	Sodium silicate (ml)	Water (ml)	Boric Acid (g)	Viscosit (cP)
M1	28	72	1,8	2,22
M2	28	72	2,0	2,29
M3	28	72	2,1	2,34
M4	28	72	2,2	2,39
M5	30	70	2,0	2,74
M6	30	70	2,2	2,79
M7	30	70	2,4	2,85
M8	30	70	2,5	2,88
M9	30	70	2,6	2,93
M10	32	68	2,2	2,97
M11	32	68	2,4	2,99
M12	32	68	2,6	3,02
M13	32	68	2,8	3,04
M14	34	66	2,2	3,06
M15	34	66	2,4	3,13
M16	34	66	2,6	3,18
M17	34	66	2,8	3,21
M18	36	64	2,4	3,25
M19	36	64	2,6	3,28
M20	36	64	2,8	3,30
M21	36	64	3,0	3,33
M22	38	62	2,5	3,35
M23	38	62	2,6	3,48
M24	38	62	2,7	3,56
M25	38	62	2,8	3,65
M26	38	62	3,0	3,69
M27	40	60	2,6	3,86
M28	40	60	2,7	3,92
M29	40	60	2,8	4,01
M30	40	60	3,0	4,05
M31	40	60	3,3	4,09
M32	42	58	2,7	4,28
M33	42	58	2,8	4,32
M34	42	58	2,9	4,38
M35	42	58	3,0	4,42

4. Results

This study investigated the rheological properties of sodium silicate–boric acid mixtures with different ratios was tested to apply as permeation grout material. This involved testing the gelation time, syneresis and viscosity of a total of 35 mixtures at two different ambient temperatures (20 °C and 30 °C). Based on the experimental results, the following conclusions can be drawn:

- It was observed that solutions containing up to 42% sodium silicate by volume were able to undergo gelation, whereas at higher concentrations gelation was either during mixing or prevented completely.
- The reactant content, SS/solution ratio and ambient temperature significantly influenced the gelation time. Increased sodium silicate or boric acid content also led to a faster gelation time.
- Mixtures prepared at an ambient temperature of 30 °C gelled more rapidly compared to those prepared at 20 °C. This indicates that higher ambient temperatures shorten gelation time.
- Syneresis behavior was influenced by both sodium silicate and boric acid concentrations. Higher boric acid content reduced syneresis, indicating improved gel stability, whereas higher sodium silicate content increased syneresis percentages.

- Examination of the graphs showed that the rate of increase in syneresis percentage decreased considerably after the 56th day.
- The solutions were examined, and it was observed that an increase in the sodium silicate percentage in the mixture led to higher viscosity values. For solutions with the same sodium silicate percentage, increasing the amount of reactant (boric acid) also resulted in higher viscosity.
- Assessing the gel time, syneresis and the viscosity together are necessary to avoid reaching a conclusion based on one parameter. Therefore, these parameters in engineering practice should be taken together instead of individually for the output optimum grout performance.

Hereafter, to evaluate the mechanical behavior and durability of sodium silicate–boric acid grouts in different field conditions would be an interesting aspect for further research. Additionally, examining alternative reactants, different silicate moduli and environmental factors such as groundwater chemistry could offer additional guidance to help create optimum grout formulations for a wider range of geotechnical applications.

5. References

[1] AASHTO (2004). Classification of Soils and Soil-Aggregate Mixtures for Highway Construction Purposes M 145-91: American Association of State Highway and Transportation Officials.

[2] Böschel, D., & Roggendorf, H. (2002). Water-based sodium borosilicate sols and their gelation-characterisation by viscosimetry, UV/VIS-and infrared spectroscopy. Journal of sol-gel science and technology, 25(3), 191-200. https://doi.org/10.1023/A:1020283206199

[3] Eti Maden (2025). Physical and Chemical Properties of Boric Acid. https://www.etimaden.gov.tr/storage/pages/March2019/1-borik-asit1.pdf Accessed: 17.03.2025.

[4] Fraccica, A., Spagnoli, G., Romero, E., Arroyo, M., & Gómez, R. (2022). Permeation grouting of silt-sand mixtures. Transportation Geotechnics, 35, 100800. https://doi.org/10.1016/j.trgeo.2022.100800

[5] Hu, H., Lu, Y., Deng, C., Gan, B., Xie, Z., Cai, Y., & Chu, A. (2024). Experimental study on grout–soil interaction effects in sandy soil under different water-to-cement ratios. Buildings, 14(4), 947. https://doi.org/10.3390/buildings14040947

[6] Kermani, M., Hassani, F., Aflaki, E., Benzaazoua, M., & Nokken, M. (2015). Evaluation of the effect of sodium silicate addition to mine backfill, gelfill − part 1. Journal of Rock Mechanics and Geotechnical Engineering, 7(3), 266-272. https://doi.org/10.1016/j.jrmge.2015.03.006

[7] Lucas, S., Tognonvi, M. T., Gelet, J., Soro, J., & Rossignol, S. (2011). Interactions between silica sand and sodium silicate solution during consolidation process. Journal of Non-Crystalline Solids, 357(4), 1310-1318. https://doi.org/10.1016/j.jnoncrysol.2010.12.016

[8] Obut, A., & Girgin, İ. (2006). Preparation of sodium borosilicates using sodium silicate and boric acid. Journal of sol-gel science and technology, 39(2), 79-83.

[9] Supelco (2025). Physical and Chemical Properties of Sodium Silicate. https://www.merckmillipore.com/TR/tr/product/msds/MDA_CHEM-105621?Origin=PDP Accessed: 17.03.2025.

[10] Wang, L. and Zhang, Y. (2013). Influence of additives on modification of sodium silicate and its molding sand properties. Advanced Materials Research, 634-638, 2702-2706. https://doi.org/10.4028/www.scientific.net/amr.634-638.2702

[11] Wang, G., Hong, B., Liu, X., Sun, D., Shao, Z., & Yao, Y. (2023). Experimental study on the shear properties of soil around piles with permeation grouting. Applied Sciences, 13(1), 621. https://doi.org/10.3390/app13010621

[12] Wang, H., Dong, L., Zhang, Q., Li, Z., & Zhang, P. (2023). Permeation grouting mechanism of viscous time-varying fluid considering diffusion path. Frontiers in Earth Science, 10. https://doi.org/10.3389/feart.2022.1100196

[13] Yang, H., Wang, H. P., Tang, B., & Qiao, J. (2013). Effect of different concentrations of sodium silicate impregnated on the performance of carbon/copper composites. Materials Science Forum, 749, 308-315. https://doi.org/10.4028/www.scientific.net/msf.749.308

High-Performance Spyware Classification with Wavelet Transform Features

Murat YAVUZ[1], Eren ÖNDER [2]

[1]Institute of Science, Firat University, Elazığ, Türkiye.
[2]Digital Forensics Engineering, Firat University, Elazığ, Türkiye.
[1]201137201@firat.edu.tr, [2]eerenonder0@gmail.com
[1](ORCID: 0000-0002-9896-9383)

Abstract
Spyware is a type of malicious software that infiltrates systems without user consent, capable of capturing screenshots, logging keystrokes, and monitoring network activity. Disguised as legitimate programs, they pose serious threats by stealing sensitive data and compromising system integrity, especially in widely used Windows-based environments. Timely and accurate detection of such threats is essential to mitigate cyber risks.
This study proposes a method for classifying four core behaviors Screenshot, Mouse Location, Keylogger, and Normal by applying the Daubechies (db1) wavelet transform to network packet features. The approximation coefficients were combined with the original attributes and classified using Decision Trees with 10-fold cross-validation.
The proposed method achieved superior performance compared to well-known approaches like Genetic Algorithm (GA), 1D-Gaussian filtering, and Recursive Feature Elimination (RFE). The model attained an accuracy of 97.06%, precision of 90.81%, recall of 89.94%, F1-score of 90.37%, and a geometric mean of 94.21%. Notably, improvements were observed in the classification of “Keylogger” and “Normal” categories. These findings demonstrate that wavelet-based feature enrichment is an effective alternative for robust spyware detection and behavior classification.

Keywords: spyware, network traffic analysis, wavelet transform, machine learning.

1. Introduction

With the widespread adoption of digital technologies, individuals and organizations are increasingly performing online activities, creating new opportunities for cybercriminals. Among the most insidious and effective types of malware are spyware programs, which infiltrate systems without the user's knowledge and secretly collect sensitive information. These malicious applications can operate in the background, capturing screenshots, logging keystrokes (keyloggers), monitoring network activity, and even altering system configurations. Consequently, spyware can lead to severe security breaches such as data leaks, financial losses, and reputational damage on both personal and corporate levels.

Spyware typically masquerades as legitimate software, making it difficult to detect using conventional antivirus tools. This evasion capability highlights the limitations of traditional defense mechanisms and the growing need for more advanced, behavior-based detection strategies. Microsoft Windows, being one of the most widely used operating systems, is a prime target for spyware attacks due to its prevalence across user groups. In this context, analyzing anomalous system behaviors through real-time network traffic classification is crucial for effective spyware detection.

In recent years, several approaches have been employed to address this problem, including Genetic Algorithms (GA), Recursive Feature Elimination (RFE), and 1D-Gaussian filtering for feature selection and extraction. However, many of these methods suffer from limited accuracy or underperformance in overlapping classes such as "Keylogger" and "Normal." Moreover, the potential of frequency-domain information processing remains underutilized in the literature.

This study proposes an alternative approach to spyware behavior classification by leveraging wavelet-based feature enrichment. A Daubechies (db1) wavelet transform was applied to a set of 26 features extracted from network packets, and the approximation coefficients were combined with the original attributes to form an expanded feature vector. Classification was performed using a Decision Tree algorithm under 10-fold cross-validation. The resulting performance 97.06% accuracy and high F1-score demonstrates the effectiveness of wavelet transform in improving classification performance over classical feature selection techniques. This research addresses a significant gap in the literature and provides a robust example of how frequency-time analysis can enhance malicious behavior detection.

[1]Corresponding author

In the second part of the study, existing approaches in the literature related to malware detection and particularly the classification of spyware behaviors are examined. The following section presents the proposed method in detail, including data preprocessing steps, feature enhancement process, and the classification model. In the fourth section, the experimental results of the developed model are evaluated, and the findings are analyzed based on various performance metrics. Finally, the conclusion section summarizes the overall outcomes of the study, emphasizes its contributions to the literature, and offers suggestions for future research.

2. Literature Review

Recent studies on spyware detection have increasingly focused on behavioral classification through network traffic analysis. Network-based approaches are particularly attractive because they allow for real-time monitoring without requiring direct access to system resources. As a result, numerous researchers have employed various machine learning algorithms and feature processing techniques to classify spyware behaviors effectively.

Pandey et al. used a data mining approach with Naive Bayes classification for spyware detection. By transforming binaries into byte sequences and applying frequency-based feature extraction, they effectively identified novel spyware [1].

Javaheri et al. (2018) used decision tree algorithms to detect spyware and ransomware, obtaining an accuracy rate of 92.32%. Their work shows decision trees are efficient for interpretable and fast classification [2].

McLaren et al. focused on detecting malware C&C traces using rules derived from decision trees, achieving a false positive rate of just 1.93%, which demonstrates the precision of their approach [3].

Tokmak and Küçüksille applied PCA for feature extraction from Windows executables and used deep learning to achieve 100% accuracy. The study highlighted PCA's role in improving model performance through dimensionality reduction [4].

Erginay (2019) compared various machine learning models (ANN, SVM, kNN, DT, NB) for anomaly detection in network traffic. Decision trees achieved the highest accuracy of 99.9%, confirming their strong performance in supervised tasks [5].

Mehtab et al. introduced AdDroid, a rule-based machine learning framework for Android malware analysis. Their method achieved over 99.11% accuracy, showcasing the power of properly tuned rule-based systems [6].

Pektaş and Acarman developed a malware detection system based on SDNE (Structural Deep Network Embedding) applied to API call graphs in Android apps. They reported a high accuracy of 98.86%, emphasizing the effectiveness of graph-based learning [7].

Bakour and Ünver proposed DipVisDroid, which combines image-based features with deep learning to detect Android malware. Their method achieved an accuracy of 98.96%, indicating strong potential in visual analysis techniques [8].

Utku proposed a deep learning-based LSTM model for Android malware detection using network traffic analysis. The study achieved an accuracy of 95%, demonstrating that effective real-time detection is possible through network-level data [9].

Yaman and Kılınç introduced a method based on 1D-Gaussian filtering to enhance the quality of input features. By applying statistical smoothing to network packet data, their model improved classification accuracy to 96.65%. Nonetheless, the accuracy for the "Normal" class remained limited at 74.21%, indicating persistent challenges in distinguishing benign activities from malicious behaviors [10].

Kılınç et al. proposed an approach combining Genetic Algorithms (GA) for feature selection with Decision Tree (DT) classification. Their model selected 18 features out of 26 using GA and achieved an overall accuracy of 89.35%. However, the performance for the "Keylogger" class was relatively low (71.85%), highlighting the limitations of heuristic-based feature selection methods in distinguishing certain overlapping behaviors [11].

Table 1 provides a summary of notable studies on spyware detection carried out between 2015 and 2024. These studies span various platforms and utilize a wide range of techniques, from traditional data mining and classification methods to advanced deep learning and network traffic analysis approaches. The progression over the years highlights a clear trend toward more sophisticated and accurate detection models.

Table 11. Summary of Recent Studies on Spyware and Malware Detection

Author(s)	Platform	Method	Accuracy (%)
Pandey et al. [1]	Binary Files	Naive Bayes + Data Mining	-
Javaheri et al. [2]	Windows	Decision Tree	92,32
McLaren et al. [3]	Windows	Decision Tree Ruleset	FPR: 1,93
Tokmak and Küçüksille [4]	Windows	PCA + Deep Learning	100
Erginay [5]	Network	Multiple ML models (ANN, SVM, kNN, DT, NB)	DT: 99,9
Mehtab et al. [6]	Android	Rule-based ML (AdDroid)	99,11
Pektaş and Acarman [7]	Android	SDNE on API graphs	98,86
Bakour and Ünver [8]	Android	Image + Deep Learning	98,96
Utku [9]	Android	LSTM on network traffic	95
Kilic and Yaman [10]	Network	1D-Gaussian Filter + Decision Tree	97,04
Kiliç et al. [11]	Windows	Genetic Algorithm (GA) + Decision Tree	89,35

Collectively, these studies demonstrate progress in spyware behavior classification but still face challenges such as low accuracy in overlapping classes, limited use of frequency-based analysis, and reliance on complex or unstable feature selection methods.

Addressing these gaps, our study proposes a wavelet-based feature enrichment approach for spyware classification. Using the Daubechies (db1) wavelet transform, approximation coefficients were extracted and combined with the original 26 features to create an expanded feature vector. Classification was performed using Decision Trees with 10-fold cross-validation, achieving 97.06% accuracy. The method demonstrated substantial improvements in previously weak classes such as “Keylogger” and “Normal,” offering a significant contribution to the field through robust frequency-time domain analysis.

3. Material Method

In this study, a methodology is proposed for the classification of spyware behaviors based on network traffic data. The approach integrates time-frequency feature extraction using wavelet transform with a supervised classification model. Specifically, a Daubechies (db1) wavelet transform is applied to the raw features to extract frequency-based components, which are then combined with the original attributes to form an expanded feature set. The resulting data is classified using a Decision Tree algorithm, and model performance is evaluated using 10-fold cross-validation.

This method diverges from conventional statistical feature selection techniques by capturing both time-domain and frequency-domain characteristics, thereby improving the ability to represent complex behavioral patterns. In particular, it addresses common challenges in spyware classification such as distinguishing between overlapping behavior classes like “Keylogger” and “Normal” by enriching the feature space with more informative representations.

3.1. Dataset

The dataset used in this study is derived from the work of Yaman and Kılınç [10], where a custom spyware behavior dataset was constructed in a controlled Windows environment. It comprises four primary behavior classes: Screenshot, Mouse Location, Keylogger, and Normal. Each class represents a distinct type of activity generated through simulated spyware operations and benign usage scenarios. The dataset includes a total of 12,226 labeled samples, each represented

by 26 numerical features extracted from captured network packets. The class names and the number of samples for each class in the dataset used in this study are provided in detail in Table 2, offering a clear summary of the dataset structure.

Table 12. Class names and sample counts in the dataset used in this study

Class No	Class	Number of data
1	Screenshot	6819
2	Mouse Location	899
3	Keylogger	508
4	Normal	4000

These features were derived from common IP and TCP header attributes, such as packet length, time-to-live (TTL), port numbers, flag settings, and byte distributions. Raw network traffic was recorded using Wireshark and processed with the Tshark command-line utility to generate the feature vectors. The resulting dataset provides a diverse and behavior-rich representation of real-time spyware and normal traffic, enabling robust behavioral classification. The reuse of this dataset allows for a direct and meaningful performance comparison with prior studies that employed traditional feature selection or statistical filtering techniques [10].

3.2. Method

In spyware behavior classification, incorporating features that carry both time-domain and frequency-domain information significantly improves the detection accuracy of malicious activities. In this context, the proposed method in this study consists of four main stages: wavelet transform for frequency-based feature extraction, feature combination with original attributes, classification using a Decision Tree algorithm, and evaluation through 10-fold cross-validation. Compared to traditional statistical filtering or heuristic feature selection techniques, this framework provides a more robust and interpretable approach to classification. Furthermore, the use of cross-validation enhances the generalizability and reliability of the model's performance metrics.

The flow diagram illustrating the operational process of the system is presented in detail in Figure 1.

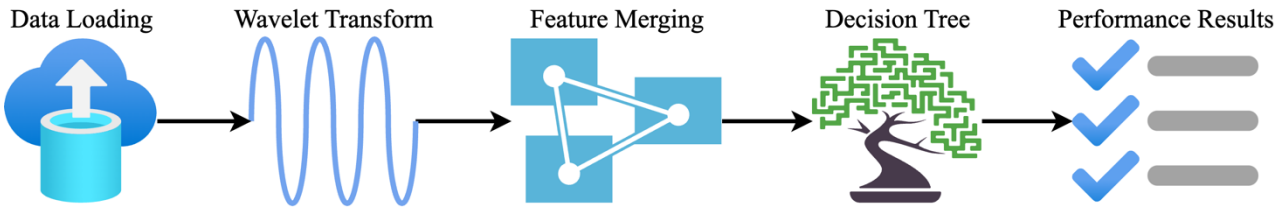


Figure 1. Flow diagram illustrating the operational process of the system

In this study, a frequency-based feature extraction approach was adopted to enhance the classification of spyware behaviors. As a first step, each of the original features was subjected to a level-1 Daubechies (db1) wavelet transform. Wavelet transforms are capable of analyzing signals simultaneously in both the time and frequency domains. The resulting approximation coefficients captured low-frequency behavioral patterns, which are essential for detecting subtle and short-duration malicious activities such as spyware operations. Among the Daubechies wavelet family, db1 was selected due to its low computational cost and stable decomposition performance.

The features obtained through wavelet transformation were then combined with the original 26 attributes to construct an expanded feature matrix of 52 dimensions. This enriched representation enabled the model to learn from both raw data characteristics and frequency-based signal patterns. In particular, this hybrid structure improved the model's ability to distinguish between overlapping behavioral classes, such as “Keylogger” and “Normal,” which are often

difficult to separate using traditional features alone. The names and descriptions of the 26 features used in the study are presented in Table 3.

Table 13. Names and descriptions of the 26 features used in the study

No	Feature name	Explanation
1	ip.version	IP address version
2	ip. hdr _len	IP address header length
3	ip.flags .rb	IP packet reserved bit
4	ip.flags .df	Unfragmented IP packet
5	ip.flags .mf	Multiple fragmented IP packets
6	ip.frag _offset	IP packet fragment offset value
7	ip.ttl	IP lifetime
8	ip.proto	IP protocol
9	ip.len	IP length
10	tcp.srcport	TCP source port
11	tcp.dstport	TCP destination port
12	tcp.seq	TCP sequence number
13	tcp.ack	TCP acknowledgment (ACK) number
14	tcp.len	TCP segment length
15	tcp.hdr _len	TCP header length
16	tcp.flags .fin	TCP Finnish flag
17	tcp.flags .syn	TCP Syn flag
18	tcp.flags .reset	TCP Reset flag
19	tcp.flags .push	TCP Push flag
20	tcp.flags .ack	TCP acknowledgment (ACK) flag
21	tcp.flags .urg	TCP Urgent flag
22	tcp.flags .cwr	TCP Reduced Congestion Window (CWR)
23	tcp.window _size	Calculated window size
24	tcp.urgent _pointer	TCP Urgent pointer
25	tcp.options .mss_val	TCP MSS value
26	frame it	Frame length

For classification, the Decision Tree algorithm was employed. Decision Trees iteratively split data based on threshold values to arrive at final class predictions. They were chosen in this study due to their inherent suitability for multi-class problems, low sensitivity to hyperparameters, fast computation, and ease of interpretability.

To validate the model's performance, a 10-fold cross-validation strategy was applied. The dataset was randomly partitioned into 10 subsets; in each iteration, one subset was used as the test set while the remaining served as training data. This technique mitigates overfitting and allows for a more robust estimation of the model’s generalization capability.

Finally, the classification performance was evaluated using widely accepted metrics: accuracy, precision, recall, F1-score, and geometric mean (G-mean). These metrics not only reflect the overall correctness of the model but also assess class-wise balance, offering a comprehensive view of practical effectiveness.

4. Results and Discussion

The performance of the proposed model was evaluated using 10-fold cross-validation. The model achieved an overall classification accuracy of 97.06%, indicating a robust generalization capability across different subsets of the dataset. The graph illustrating the model's accuracy performance is presented in Figure 2.

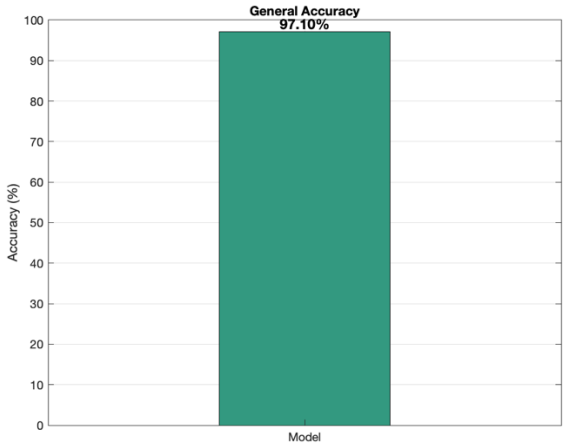


Figure 2. Model's accuracy performance graph

Class-wise performance revealed that the highest accuracy was obtained in the Mouse Location class with 99.13%, followed by Screenshot with 98.78%, Keylogger with 85.09%, and Normal with 76.77%. The relatively lower performance in the Normal class suggests a behavioral overlap with malicious classes, making accurate separation more challenging. The confusion matrix supports this observation, as most misclassifications were concentrated within the Normal category. Class-wise performance metrics are presented in Figure 3.

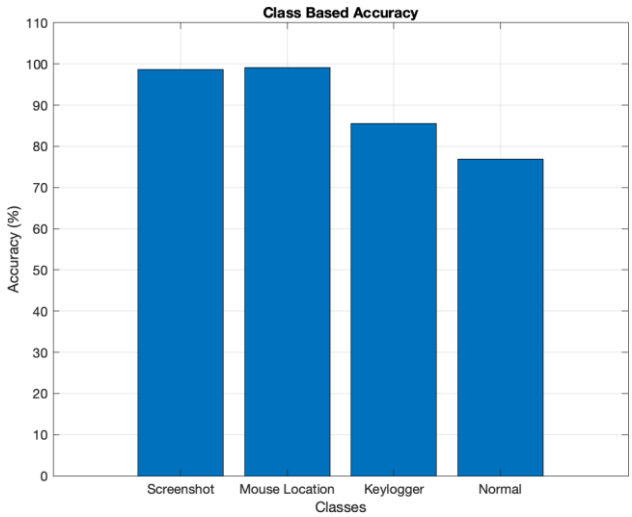


Figure 2. Model's class-based accuracy graph

In terms of evaluation metrics, the model demonstrated consistency across precision, recall, and F1-score. The average precision was 90.81%, recall was 89.94%, and F1-score reached 90.37%. These values indicate the model's balanced performance in minimizing both false positives and false negatives. The geometric mean (G-Mean) was also calculated to

assess performance across class imbalances and was found to be 94.21%, highlighting the model's stability across all classes. Figure 3 illustrates the average performance values of the model based on the Precision, Recall, F1-score, and G-Mean metrics.

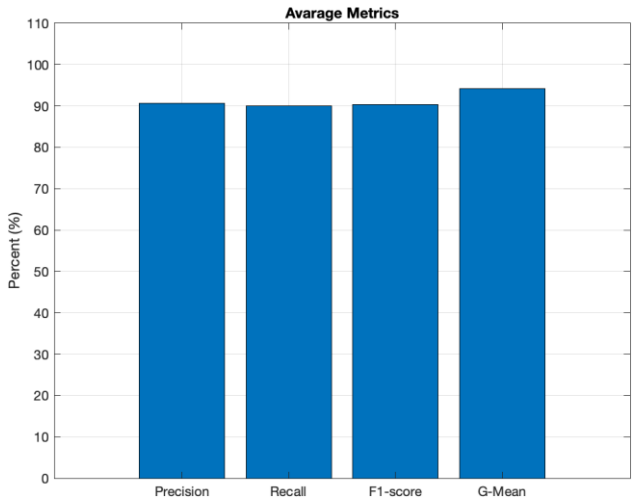


Figure 3. Average performance values of the model based on precision, recall, f1-score, and g-mean metrics

To further illustrate class-wise effectiveness, visualizations were generated for accuracy, precision, recall, and F1-score values. Additionally, the confusion matrix was visualized to depict misclassification distributions. These figures provide a comprehensive view of the classifier's strengths and inter-class discrimination performance. he confusion matrix is presented in Figure 4 to provide a clearer view of the model's class-wise prediction performance and misclassification patterns.

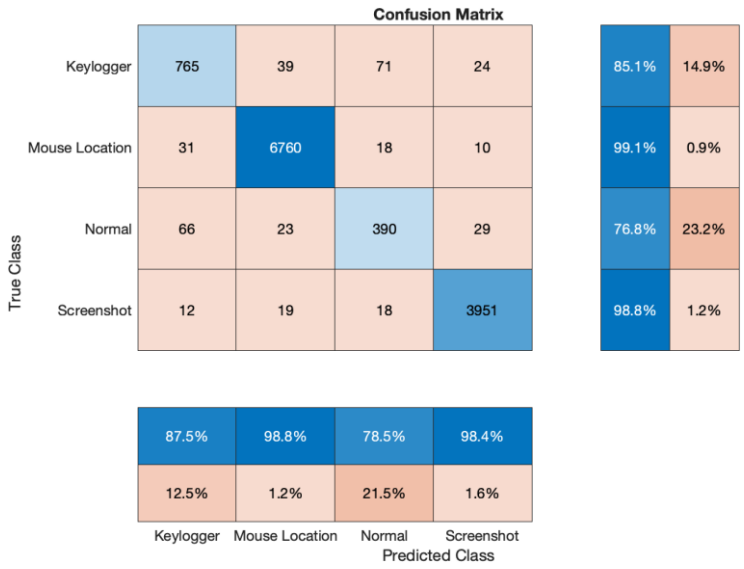


Figure 4. Confusion matrix showing the model's class-wise prediction performance

5. Conclusion

Spyware represents a significant cybersecurity threat, capable of infiltrating systems undetected and harvesting sensitive user data such as keystrokes, screenshots, and mouse activity. Its detection through network traffic analysis remains a challenge due to the behavioral resemblance between malicious and benign actions. Therefore, effective spyware detection

plays a critical role not only in protecting information security but also in advancing behavioral modeling and system monitoring.

In this study, a novel classification approach was proposed by enriching network-based features with time-frequency components using Daubechies (db1) wavelet transformation, followed by classification with a Decision Tree algorithm. The expanded feature space, constructed by combining the approximation coefficients with the original attributes, enabled a more discriminative representation of behavior. The model achieved an overall accuracy of 97.06%, alongside high performance in key metrics such as precision (90.81%), recall (89.94%), F1-score (90.37%), and G-Mean (94.21%). Notably, the proposed method performed better than previous studies using the same dataset, particularly improving accuracy in behaviorally ambiguous classes such as Keylogger and Normal.

The improvements observed in class-wise performance especially for complex and often misclassified categories highlight the benefit of wavelet-based representations in behavioral detection. Beyond overall accuracy, the model also offers stronger class balance and interpretability, making it a valuable contribution to the spyware detection literature.

Looking ahead, this approach may serve as a foundation for real-time systems, scalable to larger datasets and adaptable to hybrid machine learning models. Furthermore, the methodology can be extended to detect various types of malware or support more granular behavioral categorization in advanced cybersecurity infrastructures.

6. References

[1] K. Pandey, M. Patil, M. Naik, and J. Qamar, "Spyware Detection Using Data Mining," *International Journal of Engineering and Techniques*, vol. 1, 2015, [Online]. Available: http://www.ijetjournal.org

[2] D. Javaheri, M. Hosseinzadeh, and A. M. Rahmani, "Detection and elimination of spyware and ransomware by intercepting kernel-level system routines," *IEEE Access*, vol. 6, pp. 78321–78332, 2018, doi: 10.1109/ACCESS.2018.2884964.

[3] P. McLaren, G. Russell, and B. Buchanan, "Mining malware command and control traces," in *Proceedings of Computing Conference 2017*, Institute of Electrical and Electronics Engineers Inc., Jan. 2018, pp. 788–794. doi: 10.1109/SAI.2017.8252185.

[4] M. Tokmak and E. U. Küçüksille, "Kötü Amaçlı Windows Çalıştırılabilir Dosyalarının Derin Öğrenme İle Tespiti," *Bilge International Journal of Science and Technology Research*, vol. 3, no. 1, pp. 67–76, Mar. 2019, doi: 10.30516/bilgesci.531801.

[5] E. Erginay, "Ağ Trafiğinde Anormallik Tespiti İçin Veri Seti Oluşturma ve Test Yöntemlerinin Karşılaştırılması," Fen Bilimleri Enstitüsü, 2019.

[6] A. Mehtab *et al.*, "AdDroid: Rule-Based Machine Learning Framework for Android Malware Analysis," *Mobile Networks and Applications*, vol. 25, no. 1, pp. 180–192, Feb. 2020, doi: 10.1007/s11036-019-01248-0.

[7] A. Pektaş and T. Acarman, "Deep learning for effective Android malware detection using API call graph embeddings," *Soft comput*, vol. 24, no. 2, pp. 1027–1043, Jan. 2020, doi: 10.1007/s00500-019-03940-5.

[8] K. Bakour and H. M. Ünver, "DeepVisDroid: android malware detection by hybridizing image-based features with deep learning techniques," *Neural Comput Appl*, vol. 33, no. 18, pp. 11499–11516, Sep. 2021, doi: 10.1007/s00521-021-05816-y.

[9] A. Utku, "Using network traffic analysis deep learning based Android malware detection," *Journal of the Faculty of Engineering and Architecture of Gazi University*, vol. 37, no. 4, pp. 1823–1838, 2022, doi: 10.17341/gazimmfd.937374.

[10] I. Kilic and O. Yaman, "A New Approach to Spyware Classification Using a 1D-Gaussian Filter on Network Packets," in *8th International Artificial Intelligence and Data Processing Symposium, IDAP 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/IDAP64064.2024.10710346.

[11] İ. Kılıç, O. Yaman, E. Erdoğan, and M. İ. Aslan, "A Hybrid Method Based On A Genetic Algorithm That Uses Network Packets To Classify Spyware," *Journal of Physical Chemistry and Functional Materials*, vol. 7, no. 2, pp. 148–157, Dec. 2024, doi: 10.54565/jphcfum.1579687.

MRI-Based Brain Tumor Classification Using Deep Learning Models

Rawan ABOALQARAYA[1] , Fatih ÖZYURT[2], Canan Koç[3], Esra YÜZGEÇ ÖZDEMİR[4]

[1,2,3,4] Department of Software Engineering, Engineering Faculty, Fırat University, Elazığ, Turkey.
[1] rawan.qaraya@gmail.com , [2] fatihozyurt@firat.edu.tr, [3] canan.koc@firat.edu.tr , [4]eyuzgec@firat.edu.tr
[1](ORCID: 0009-0008-0995-2629), [2](ORCID: 0000-0002-8154-6691), [3](ORCID: 0000-0002-2651-9471)

Abstract
This study proposes an automatic classification approach for brain tumors using magnetic resonance imaging (MRI). The dataset comprises four classes: glioma, meningioma, pituitary tumor, and no tumor. We evaluated a custom-designed convolutional neural network (CNN) alongside transfer learning models such as VGG16, VGG19, MobileNetV2, and ResNet50. To enhance model generalization, data augmentation techniques were employed during training. Model performance was measured using accuracy, classification reports, and confusion matrices. The custom CNN model achieved the highest test accuracy of 90.08%, followed closely by MobileNetV2. These findings suggest that transfer learning models do not always outperform custom architectures, and dataset-specific, optimized models can yield superior results.

Keywords: brain tumor, deep learning, convolutional neural network, transfer learning, MRI classification, mobilenetv2

1. Introduction

Brain tumors are complex neoplasms that affect the central nervous system and are often associated with severe morbidity. These tumors are classified into various subtypes, including gliomas, meningiomas, and pituitary adenomas, each of which exhibits distinct biological behaviors and clinical outcomes. Early and accurate diagnosis is crucial for improving patient survival and quality of life, particularly in malignant tumors such as gliomas. However, conventional diagnostic methods are often based on invasive procedures and radiologic expert opinion and are prone to human error and interpreter dependency.

Magnetic Resonance Imaging (MRI) is the most widely used imaging modality for the non-invasive diagnosis of brain tumors. MRI enables high-resolution examination of tumor structure, localization, and its relationship with surrounding tissues, thanks to its soft tissue contrast. However, since manual analysis of MRI data is both time-consuming and subjective, there is an increasing need for fast and accurate automated diagnostic systems. Especially in resource-limited clinical settings, such decision support systems reduce the workload and standardize the diagnostic process. In recent years, the success of deep learning techniques in image processing has enabled the development of automated systems for brain tumor diagnosis. In this study, we propose a classification system that can distinguish between four different classes of brain tumors based on MRI images. A specially designed convolutional neural network (CNN) model is compared with common transfer learning architectures (VGG16, VGG19, MobileNetV2, and ResNet50), and the generalizability of the model is improved by data augmentation methods. The results show that architectures specifically optimized for specific datasets can be more successful than pre-trained deep networks.

Brain tumors are serious health problems affecting the central nervous system, and early diagnosis is crucial for effective treatment. Medical imaging techniques such as Magnetic Resonance Imaging (MRI) are commonly used for the detection and classification of tumors. However, manual analysis of these images is time-consuming and subject to interpretation. Therefore, the need for artificial intelligence-based automatic systems has increased [1]. The use of deep learning methods in medical imaging has gained significant momentum in recent years. In particular, convolutional neural networks (CNNs) and more advanced architectures have achieved great success in brain tumor classification [2].
In this study, we propose a deep learning-based method to classify brain tumors from MRI images. Our goal is to develop a model that can accurately analyze brain tumor images and distinguish among four classes: glioma, meningioma, pituitary tumor, and no tumor [3].

[1]Corresponding author

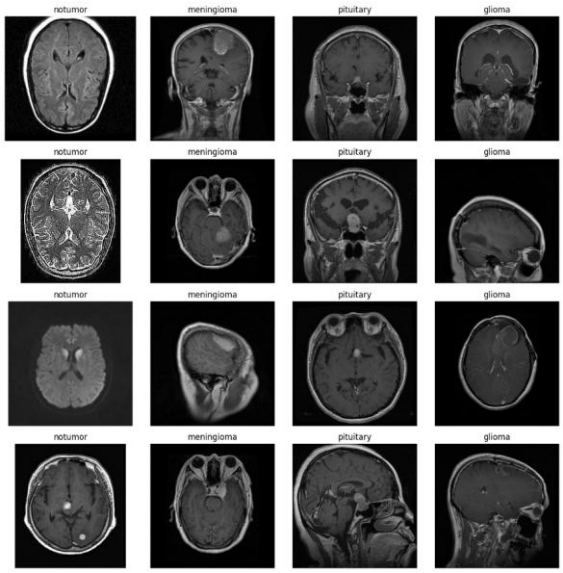


Figure 1. Sample MRI images from the dataset

Brain tumor classification has made significant progress in recent years with the widespread use of deep learning methods in the field of medical imaging. Below are some important studies conducted in this area:

Afshar et al. (2019) used Capsule Networks (CapsNet) for the classification of brain tumor types from MRI images. CapsNet aims to overcome the limitations of traditional convolutional neural networks (CNNs). In their study, the proposed CapsNet architecture was tested on the Figshare dataset and achieved an accuracy of 86.56%. This result was attributed to CapsNet's ability to work effectively with limited data and to model spatial relationships more efficiently. Moreover, this architecture is more robust to tumor orientation and location variations and can learn effectively even with less data [1].

Yang et al. (2023) compared deep learning and traditional machine learning algorithms in brain tumor classification. Using 2D Convolutional Neural Networks (CNN) and convolutional autoencoder models, they achieved an accuracy of 93.45%, demonstrating the effectiveness of deep learning methods. The study showed that the 2D CNN model achieved the best performance with an AUC score of 0.99, highlighting the significant advantage of deep learning methods over traditional approaches [4].

Kang et al. (2021) proposed a hybrid method combining deep learning-based feature extraction with traditional machine learning classifiers for brain tumor classification. The study adopted a transfer learning approach, using 13 different pre-trained CNN models to extract deep features from MRI images. These features were then evaluated using nine different machine learning classifiers, including Support Vector Machines (SVM), Random Forests (RF), k-Nearest Neighbors (k-NN), Gaussian Naive Bayes, and Extreme Learning Machines (ELM). For each dataset, features obtained from the top three performing CNN models were selected and combined to create a "feature-level ensemble." This combined feature set was again classified using various machine learning algorithms. Experimental results showed that, especially for large datasets, the SVM classifier with a radial basis function (RBF) kernel achieved the highest accuracy rates (98.67% in binary classification and 93.72% in multi-class classification). This study demonstrates that integrating deep learning and traditional machine learning techniques is an effective approach to achieve high accuracy in brain tumor classification tasks [5].

Haq et al. (2021) demonstrated how a deep learning-based paradigm achieved high success rates in classifying glioma, meningioma, and pituitary tumors using the Figshare and BraTS 2018 datasets. The model achieved 96.5% accuracy and a Dice Similarity Coefficient (DSC) of 94.3% on the BraTS 2018 dataset. Furthermore, performance

improvements were made using techniques such as data augmentation, intensity normalization, and conditional random fields (CRF) [6]. Nassar et al. (2023) developed a hybrid deep learning model and achieved an accuracy rate of 99.31% on 3,064 T1-weighted contrast-enhanced MRI images from 233 patients. Pre-trained models such as GoogleNet, AlexNet, and SqueezeNet were employed for brain tumor classification, and their outputs were combined using majority voting to obtain more reliable results [7].

Dal, Eliaçık, and Işık (2023) proposed a deep learning-based approach for brain tumor classification using convolutional neural networks (CNN). Their study aimed to accurately classify different tumor types (glioma, meningioma, and pituitary tumors) from brain MR images. By leveraging the automatic feature extraction capability offered by deep learning, the classification process was performed without the need for traditional image processing steps. A large dataset was used for training and testing the developed model; the reported accuracy at the end of training was 98%. Additionally, the proposed CNN architecture provided significant advantages over traditional methods in terms of training time and classification accuracy. This work highlights that deep learning-based approaches have the potential to provide highly accurate diagnoses in medical image analysis [8]. Aslan (2022) focused on the automatic detection of brain tumors by developing a deep learning-based model. Using convolutional neural networks (CNN), brain tumors were accurately classified from MRI images. The author reported high accuracy for this deep learning model and indicated that it outperformed traditional methods. The model's accuracy was reported as 96%, demonstrating how powerful deep learning can be in brain tumor diagnosis systems. The study also discussed the impact of pre-processing techniques used to improve the model's accuracy. The results show that deep learning techniques can play an important role in medical image analysis and potentially be useful in clinical applications [9].

Table 1. Approaches used for brain tumor detection in the related literature

Author(s)	Year	Method	Dataset	Performance / Results
Afshar Et Al.	2019	Capsule Networks (CapsNet)	Figshare dataset	86.56% accuracy
Yang Et Al.	2023	2D CNN, Convolutional Autoencoders	MRI images	93.45% accuracy, 0.99 AUC score
Kang Et Al.	2021	Deep features + Machine Learning classifiers	MRI images	98.67% accuracy (RBF-SVM, binary class), 93.72% accuracy (RBF-SVM, multi-class)
Haq Et Al.	2021	Deep learning-based model (BraTS 2018)	BraTS 2018, Figshare	96.5% accuracy, 94.3% DSC
Nassar Et Al.	2023	Hybrid deep learning, majority voting	T1-weighted MRI images	99.31% accuracy
Dal, Eliaçik, And Işik	2023	Deep learning, convolutional neural networks (CNN)	MRI images	98% accuracy
Aslan	2022	Deep learning	MRI images	96% accuracy

2. MATERIAL AND METHOD

2.1. Dataset

The brain MRI images used in this study were obtained from the Kaggle platform and consist of a combined dataset from Figshare, SARTAJ, and Br35H datasets. The dataset contains a total of 7,023 images, divided into four classes: glioma, meningioma, pituitary, and no tumor. Some mislabeled images in the dataset were corrected, and the images underwent preprocessing steps, including resizing and edge cleaning, before training. This dataset was used for training deep learning models for brain tumor classification.

Since the images in the dataset have varying dimensions, all images were resized to 240x240 pixels prior to model training. Additionally, pixel values were normalized from the range 0–255 to 0–1. These operations enable the model to learn more quickly and efficiently. Due to the limited number of images in medical imaging, data augmentation techniques were applied to reduce the risk of overfitting and improve the model's generalization ability. Random rotations (±20 degrees), zooming (10%), and horizontal flipping were applied to the training images, creating different variations of each image. This approach made the model more robust against the variability within the dataset.

2.2. Models

In this study, both a simple CNN model designed from scratch and pre-trained models, including VGG16, VGG19, MobileNetV2, and ResNet50, were used. For the pre-trained models, transfer learning was applied by freezing the base model weights and adding new classification layers at the end. The models were trained for 20 epochs using the Adam optimizer and categorical cross-entropy loss function. Additionally, EarlyStopping and ReduceLROnPlateau callbacks were employed to prevent overfitting. This model is a custom-built neural network created from scratch. It consists of three convolutional (Conv2D) layers, followed by max-pooling layers to increase depth. Finally, a flatten, fully connected (Dense), and dropout layer lead to the output. A dropout rate of 50% was applied to reduce the overfitting tendency. This architecture was chosen as a basic starting point due to its low computational cost and ability to deliver results quickly.

VGG16 is a 16-layer deep convolutional neural network pre-trained on the ImageNet dataset [10]. In this study, transfer learning was applied by freezing the pre-trained layers and retraining only the last layers. Due to its deep architecture and small filter sizes, VGG16 is effective in extracting detailed features. It is a widely preferred powerful architecture, especially for image classification problems. VGG19 is an improved version of the VGG16 model, consisting of 19 layers [10]. Although deeper, it uses a similar structure with small filter sizes and repeated blocks. This model was also adapted via transfer learning with only the last layers retrained. While its depth increases the capacity to learn more complex features, it requires longer training times and has a higher risk of overfitting.

MobileNetV2 is a lightweight CNN architecture optimized for mobile and embedded systems [11]. It has lower depth but high efficiency, enabling faster operation with fewer parameters. In this study, the base weights were frozen and only the classification layers were trained. MobileNetV2 provided balanced performance with low resource usage and high accuracy.

ResNet50 is a 50-layer deep convolutional network characterized by the use of "residual blocks" [12]. This structure was developed to mitigate the vanishing gradient problem commonly encountered in deep networks. In this study, the model was adapted with transfer learning; however, its deep architecture made training more challenging compared to other models. Especially when trained with limited data and augmented samples, its performance was lower.

The performances of the models were evaluated on the test dataset using various metrics. These included accuracy, confusion matrix, and classification reports comprising precision, recall, and F1-score. Accuracy graphs for each model's training process were plotted to analyze learning trends. Additionally, class-wise correct and incorrect classifications were visualized through confusion matrices. The CNN model achieved the highest test accuracy of 90.08%, followed by MobileNetV2 with 89.09%. The VGG16 and VGG19 models attained accuracies of 80.7% and 79.32%, respectively, while ResNet50 showed the lowest performance with 67.42%. These results reveal the discrimination power of the models across classes and their fit to the dataset.

To assess the success of the classification models, the relationship between the predicted class labels and the true class values was examined. This evaluation was conducted based on true positive (TP), false positive (FP), true negative (TN), and false negative (FN) values obtained from the confusion matrix. Several evaluation criteria were used to measure model performance, including accuracy, which indicates the overall success rate; recall (sensitivity), which measures the proportion of true positive samples correctly identified; and specificity, which reflects the

proportion of true negatives correctly classified. Additionally, precision, which measures the accuracy of positive predictions; F1-score, the harmonic mean of precision and recall; and Matthews Correlation Coefficient (MCC), which provides more reliable results in cases of class imbalance, were included. These metrics not only reflect overall success but also reveal strengths and weaknesses of the model at the class level, allowing for detailed performance analysis. The formulas for these evaluation criteria used in this study are presented in Equations (1) - (6) [8-13].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (1)$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100 \quad (2)$$

$$Specificity = \frac{TN}{TN + FP} \times 100 \quad (3)$$

$$Recall = \frac{TP}{TP + FP} \times 100 \quad (4)$$

$$F1 - Score = 2 \times \frac{Sensitivity \times Specificity}{Duy + Specificity} \times 100 \quad (5)$$

$$MCC = 2 \times \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (6)$$

3. Experimental Results

The experimental studies were carried out to evaluate the effectiveness of the proposed model and to present a comparative performance analysis with other common deep learning architectures. In this context, a specially developed convolutional neural network (CNN) model and common transfer learning-based models such as VGG16, VGG19, MobileNetV2, and ResNet50 were tested on a four-class dataset of MRI images. Model performances are analyzed in detail through accuracy, classification reports (precision, recall, F1-score), and complexity matrices. In addition, the effect of data augmentation techniques applied during training on generalization performance was observed. The findings presented below highlight the primary contributions of the study, encompassing both model performance and architectural choices.

Table 2 presents in detail the classification performance of the proposed CNN model on a four-class brain tumor dataset. The model achieved a high success rate of 90.08% in terms of the overall accuracy metric, demonstrating the reliability of the model. In particular, the highest performance was observed in the “No Tumor” class with 100% recall and an F1 score of 0.93%, indicating that the model has a high sensitivity in recognizing healthy samples. However, for the “Pituitary” tumor class, a very successful classification was also achieved with an F1 score of 0.96% and a recall of 0.99%.

On the other hand, the “Meningioma” class shows a relatively lower performance, with a recall of 0.76% and an F1 score of 0.80%, indicating that the model has a more cautious success in discriminating this tumor type. This is likely due to the fact that meningioma images are more visually similar to those of the other classes. In terms of macro and weighted averages (macro avg / weighted avg), 0.90% was achieved in all three metrics, suggesting that the model maintains its generalizability despite the imbalance between classes. The results show that a specially designed CNN architecture can offer a competitive alternative to transfer learning-based models, emphasizing the importance of dataset-specific, optimized architectures.

Table 2. CNN Classification Results

Class	Precision	Recall	F1-Score	Support
Glioma	0.95	0.83	0.88	300
Meningioma	0.86	0.76	0.80	306
No Tumor	0.87	1.00	0.93	405
Pituitary	0.94	0.99	0.96	300
Accuracy			**0.90**	1311
Macro Avg	0.90	0.89	0.90	1311
Weighted Avg	0.90	0.90	0.90	1311

Table 3 details the performance of the transfer learning-based VGG16 model on the four-class brain tumor classification task. The model achieves an overall accuracy of 79%, which is lower than that of the custom-designed CNN architecture. Especially for the "No Tumor" class, a very successful classification was achieved with an F1 score of 0.91% and a recall of 0.93%. This shows that the VGG16 model can distinguish healthy samples with high accuracy. In addition, the 0.97% recall and 0.84% F1 score for "Pituitary" tumor shows that the model can effectively recognize this class as well.

In contrast, the model performs poorly for the "Meningioma" class, with a recall of 0.50% and an F1 score of 0.58%, indicating that this tumor type has a high misclassification rate. For the "Glioma" class, the F1 score of 0.76% indicates that the model performs at an average level in this category. The macro and weighted average metrics are 0.77% and 0.78% respectively, indicating that the imbalance between classes affects the model performance. These results suggest that transfer learning-based architectures may not always achieve the performance of custom models optimized for the dataset.

Table 3. VGG16 Classification Results

Class	Precision	Recall	F1-Score	Support
Glioma	0.80	0.73	0.76	300
Meningioma	0.69	0.50	0.58	306
No Tumor	0.89	0.93	0.91	405
Pituitary	0.75	0.97	0.84	300
Accuracy			**0.79**	1311
Macro Avg	0.78	0.78	0.77	1311
Weighted Avg	0.79	0.79	0.78	1311

Table 4 shows the performance of the VGG19 transfer learning architecture on the four-class brain tumor classification task. The model produced an overall accuracy of 79%, which is almost on par with VGG16. The 0.91% F1 score and 0.93% recall value obtained for the "No Tumor" class show that the model can recognize healthy individuals with high accuracy. Similarly, the "Pituitary" tumor class performed well with 0.97% recall and 0.84% F1 score. These results show that VGG19 offers stable prediction capability in some classes.

On the other hand, a very poor performance was observed for the "Meningioma" class with 0.50% recall and 0.58% F1 score. This indicates that the VGG19 model has difficulty in distinguishing meningioma tumors from other classes. With an F1 score of 0.76% for the "Glioma" class, it can be said that the model recognizes this class at a moderate level. The macro and weighted averages are 0.77%-0.78% respectively, indicating that the overall success of the model is limited due to the imbalance between classes. As a result, although the VGG19 architecture shows acceptable performance in certain classes, it lags behind the custom CNN architecture in terms of overall success.

Table 4. VGG19 classification report

Class	Precision	Recall	F1-Score	Support
Glioma	0.80	0.73	0.76	300
Meningioma	0.69	0.50	0.58	306
No Tumor	0.89	0.93	0.91	405
Pituitary	0.75	0.97	0.84	300
Accuracy			**0.79**	1311
Macro Avg	0.78	0.78	0.77	1311
Weighted Avg	0.79	0.79	0.78	1311

Table 5 presents the performance of MobileNetV2, a transfer learning architecture, in a four-class brain tumor classification task. The model achieved an overall accuracy of 89%, which is very close to that of the CNN architecture developed specifically for this task. Especially in the "No Tumor" class, the model's ability to recognize healthy individuals with high accuracy is remarkable, with an F1 score of 0.96 and a recall value of 0.97. Similarly, the "Pituitary" tumor class showed a very strong classification performance with 1.00% recall and 0.91% F1 score. These results demonstrate that MobileNetV2 exhibits strong generalization performance despite its parametric efficiency.

The model produced an F1 score of 0.88% for the "Glioma" class and an F1 score of 0.78% for the "Meningioma" class, indicating a very good performance in these two tumor types. In particular, the performance in the "Meningioma" class reveals that MobileNetV2 offers a more balanced approach in this class, where larger models such as VGG16 and VGG19 lag behind. The macro and weighted averages were each in the range of 0.88% to 0.89%, indicating a relatively balanced distribution across classes. In this context, the MobileNetV2 model can be considered a viable and competitive solution in terms of both performance and efficiency.

Table 5. MobileNetV2 Classification Results

Class	Precision	Recall	F1-Score	Support
Glioma	0.94	0.83	0.88	300
Meningioma	0.83	0.74	0.78	306
No Tumor	0.94	0.97	0.96	405
Pituitary	0.84	1.00	0.91	300
Accuracy			**0.89**	1311
Macro Avg	0.89	0.88	0.88	1311
Weighted Avg	0.89	0.89	0.89	1311

Table 6 shows the performance of ResNet50, a deep transfer learning model, on the four-class brain tumor classification problem. The overall accuracy of the model is 67%, which is quite low compared to both the specially designed CNN architecture and other transfer learning based models. For the "No Tumor" class, an F1 score of 0.80% and a recall value of 0.86% were obtained, and this class was the most successfully recognized category by the model. The "Pituitary" tumor class also performed relatively well with an F1 score of 0.72%, but these levels were not enough to improve the overall success.
The model had serious classification problems, especially in the "Meningioma" class. The 0.32% recall and 0.41% F1 score obtained for this class indicate that the model frequently misclassifies meningioma samples. Similarly, only a 0.65% F1 score was obtained for the "Glioma" class. The macro and weighted averages are in the range of 0.65%-0.66% for all metrics, indicating that the model fails to provide a balanced and reliable classification across classes. These results suggest that the ResNet50 model exhibits low generalization capability on the current dataset and that more optimized or lightweight architectures may be needed for such medical imaging tasks.

Table 6. ResNet50 Classification Results

Class	Precision	Recall	F1-Score	Support
Glioma	0.64	0.67	0.65	300
Meningioma	0.57	0.32	0.41	306
No Tumor	0.74	0.86	0.80	405
Pituitary	0.66	0.80	0.72	300
Accuracy			**0.67**	1311
Macro Avg	0.65	0.66	0.65	1311
Weighted Avg	0.66	0.67	0.66	1311

Table 7 compares the test accuracies of the proposed custom CNN model and four different transfer learning architectures (MobileNetV2, VGG16, VGG19, and ResNet50). The results show that the custom CNN architecture shows the highest performance with an accuracy of 90.08%. This was followed by MobileNetV2, which achieved an accuracy of 89.09%, demonstrating a highly competitive performance despite its parametric efficiency and low computational requirements. VGG16 and VGG19 produced very close results with 80.70% and 79.32% accuracy, respectively, but lagged behind the specialized CNN and MobileNetV2 architectures. The lowest accuracy was 67.42% for ResNet-50, indicating that the model is not well-adapted to the current dataset. This comparison strongly suggests that custom architectures optimized for specific datasets can outperform transfer learning-based approaches.

Table 7. Comparison of The Models

Model	Test Acc (%)
CNN	90.08%
MobileNetV2	89.09%
VGG16	80.70%
VGG19	79.32%
ResNet50	67.42%

4. Conclusion

In this study, a deep learning approach for automatic classification of brain tumors from MRI images is presented. The specially developed convolutional neural network (CNN) model was able to discriminate four different tumor classes with high accuracy and outperformed all transfer learning models with a test accuracy of 90.08%. The high F1 scores, especially in the “No Tumor” and “Pituitary” classes, showed that the model can effectively classify both healthy and pathological samples. MobileNetV2 architecture was the closest competitor with an accuracy of 89.09%, showing a strong performance despite the low number of parameters.

The results of the comparative analysis show that transfer learning-based models do not always guarantee the best performance; on the contrary, custom models optimized for the data set may produce better results. The findings demonstrate the value of customized artificial intelligence solutions in the development of clinical decision support systems and pave the way for future studies with different imaging types, tumor subclasses, and multiple modalities. Furthermore, it is suggested that data augmentation and weighted learning strategies be integrated more effectively to overcome important problems, such as class imbalance.

6. References

[1] Afshar P, Mohammadi A, Plataniotis KN. “Brain tumor type classification via capsule networks”. 2019 IEEE International Conference on Image Processing, Taipei, Taiwan, 22–25 September 2019.

[2] Chakrabarty A, Bhunia AK. “A survey on brain tumor detection using image processing techniques”. International Journal of Scientific & Technology Research, 9(1), 1727–1731, 2020.

[3] Nickparvar M. “Brain Tumor MRI Dataset”. https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset.

[4] Yang W, Al-Khassaweneh M, Alzghool M, Fraiwan L, Aljarrah I, Al-Fahoum A. “MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques”. BMC Medical Informatics and Decision Making, 23(1), 1–19, 2023.

[5] Kang J, Ullah Z, Gwak J. “MRI-based brain tumor classification using ensemble of deep features and machine learning classifiers”. Sensors, 21(6), 2222, 2021.

[6] Haq EU, Azam MA, Khan MA, Saba T, Rehman A, Tariq U, Miceli M. “A hybrid deep learning-based paradigm to detect brain tumor from MRIs”. Journal of Ambient Intelligence and Humanized Computing, 13, 225–241, 2021.

[7] Nassar M, Almotairi S, Alshmrany S, Alsheikhy A. “A hybrid deep learning model for brain tumor classification based on majority voting technique”. The Journal of Supercomputing, 2023. https://doi.org/10.1007/s11227-023-05549-w

[8] Dal A, Eliaçık B, Işık AH. “Beyin tümörü MR görüntülerinin evrişimsel sinir ağları ile sınıflandırılması”. Uluborlu Mesleki Bilimler Dergisi, 6(1), 71–91, 2023.

[9] Aslan M. “Derin öğrenme tabanlı otomatik beyin tümör tespiti”. Fırat Üniversitesi Mühendislik Bilimleri Dergisi, 34(1), 399–407, 2022.

[10] Simonyan K, Zisserman A. “Very deep convolutional networks for large-scale image recognition”. International Conference on Learning Representations (ICLR), San Diego, USA, 7–9 May 2015.

[11] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. “MobileNetV2: Inverted residuals and linear bottlenecks”. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 18–23 June 2018.

[12] He K, Zhang X, Ren S, Sun J. “Deep residual learning for image recognition”. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 27–30 June 2016.

[13] Japkowicz N, Shah M. Evaluating Learning Algorithms: A Classification Perspective. Cambridge, UK, Cambridge University Press, 2011.

Disease Prediction Based on Medical Symptoms: A Comparative Study of Machine Learning Models

Fatma Zehra ÇELİK[1], Mehmet ARZU[2], Mahmut KAYA[3], Muhammed Emre ÇOLAK[4]

[1,2,3,4]Department of Artificial Intelligence and Data Engineering, Fırat University, Elâzığ, Türkiye.
[1]230291007@firat.edu.tr, [2] marzu@firat.edu.tr, [3]mahmutkaya@firat.edu.tr, [4]memrecolak@firat.edu.tr
[1](ORCID: 0009-0008-1069-4946), [2](ORCID: 0000-0001-6610-2788),[3](ORCID: 0000-0002-7846-1769), [4](ORCID: 0000-0003-2317-2385)

Abstract
Early diagnosis is an important area of research that contributes to explainable and clinical decision-making processes in disease diagnosis. In this study, machine learning methods were applied to predict possible diseases based on patients' symptoms. Using the "Diseases and Symptoms Dataset" available on Kaggle, data preprocessing was performed, and rare disease classes with fewer than 0.49% of the sample were excluded from the analysis. During the feature engineering phase, the most discriminative 49 features were selected using the ANOVA F-test and mutual_info_classif methods. During the modeling process, Logistic Regression, Random Forest, Support Vector Machines, XGBoost, LightGBM, and K-Nearest Neighbors algorithms were optimized and compared. Experimental findings show that the Logistic Regression model demonstrated the most successful performance with an accuracy rate of 83.4% in the testing phase. In addition to accuracy, the models were also evaluated using comprehensive metrics such as F1 score, recall, specificity, and AUC-ROC. The results revealed that the models were able to make both consistent and reliable predictions. The performance of the models was also improved through hyperparameter optimization using methods such as grid search. Furthermore, a variable importance analysis was conducted in the study, and it was determined that certain symptoms play a significant role in the diagnostic process.

Keywords*:* symptom-based prediction, feature selection, class imbalance, machine learning

1. Introduction

In recent years, there has been a significant transformation in the diagnosis and treatment processes in the healthcare sector. One of the main reasons for this transformation is the integration of artificial intelligence (AI) and automation systems into medical applications. In the past, disease diagnoses were based on doctor experience and laboratory results, but today, with AI-supported systems, the diagnostic process is faster and more systematic. For example, Chaurasia and Pal [1] achieved high accuracy in breast cancer diagnosis using machine learning algorithms, while Shen et al. [2] reported that artificial intelligence systems achieved accuracy levels similar to those of clinical experts. Miotto et al. [3] also demonstrated that unstructured data obtained from clinical notes could be successfully analyzed using deep learning.

Thanks to advancing technologies, various types of data, such as patients' previous examination information, genetic characteristics, and reported symptoms, can be analyzed holistically to make more accurate assessments. The identification of previously unnoticed important findings contributes to the decision-making process of physicians. However, relying solely on test results or imaging techniques to reach a diagnosis may not be sufficient to accurately assess the situation. Symptoms, which are often precursors to diseases, provide the first clues in the diagnostic process and should be carefully evaluated. Pain felt in the same area of the head may sometimes indicate migraine and sometimes sinusitis; what distinguishes these two diagnoses are the small but meaningful differences in the symptoms accompanying the pain. At this stage, AI-supported systems play a critical role in distinguishing between diseases that present clinically similar symptoms.
Smart prediction mechanisms that work based on symptoms not only reduce the burden on physicians, but also make meaningful contributions to the diagnostic process in areas with limited access to healthcare services. In a 2022 study by Ciecierski-Holmes et al. [4], it was found that these systems play a critical role in accurate referral and planning processes in resource-constrained areas. Recent research has shown that such systems accelerate early intervention and lead to more effective outcomes.

[1]Corresponding author

Having prior knowledge about potential diseases enables individuals to make more informed decisions. Esteva et al. [5] developed an artificial intelligence system that can classify skin cancer with the same accuracy as dermatologists using deep convolutional neural networks. Additionally, Chavoshi et al. [6] demonstrated that deep learning models provide higher sensitivity (93%) and similar specificity (91%) compared to radiologists in detecting COVID-19 pneumonia. These results highlight that AI-supported systems can support physician performance in clinical applications and play an important role in early diagnosis.

Disease prediction based on symptoms has been an important area of research in recent years. However, the scarcity of real-life data on rare disease classes makes accurate prediction difficult. In this context, machine learning approaches for the early detection of rare diseases have been developed using labeled clinical records obtained from Uruguay. In this study, models such as SVM, LSTM, and BERT were compared, and it was reported that the SVM model achieved a high accuracy rate [7].

For example, Yi et al. showed in 2023 that XGBoost achieved high accuracy even in imbalanced classes in the classification of Alzheimer's disease. [8]. Ambrish and Ganesh, in their 2022 study, showed that logistic regression was preferred in clinical datasets due to its simplicity and explainability, and yielded successful results with 86.5% accuracy [9].
In their 2019 study, Jamian et al. demonstrated that rule-based methods were more successful than machine learning models in the diagnosis of systemic sclerosis, with an F-score of 91% [10]. In 2024, Tokede et al. improved the accuracy of periodontal diagnosis from 71.8% to 79.6% through modifications to rule-based systems [11]. In 2017, Esteva and his team conducted a deep learning-based study and demonstrated that artificial intelligence provides similar accuracy to dermatologists in the diagnosis of skin cancer [12].

In this study, a machine learning-based approach is presented that predicts possible diseases based on patient symptoms and aims to support clinical decision-making processes. To develop a high-performance diagnostic model suitable for clinical applications, various algorithms such as Logistic Regression, Random Forest, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and XGBoost were evaluated together, and the model with the best performance was carefully analyzed. This study aims to establish a robust support system that will guide physicians in making faster and more accurate decisions during the preliminary diagnosis process. Through analyses conducted on an accessible symptom-disease dataset, the models that yielded the most meaningful and effective results were examined in depth.

2. Materials and Methods

2.1. Dataset

In this study, the “Diseases and Symptoms Dataset” available on the Kaggle platform was used [13]. The dataset contains more than 40 disease classes corresponding to symptoms reported by patients. This dataset consists of a total of 246,945 records and 378 columns, with the first column being the disease label (target variable) and the remaining 377 columns consisting of features where patients' symptoms are coded as 0 or 1.

The dataset is suitable for multi-class and multi-variable classification problems due to its large number of classes and high-dimensional feature structure. However, there are significant imbalances in the distribution of sample numbers between classes. He and Garcia (2009) emphasized that such class imbalances negatively affect classification success and cause bias in the learning process [14]. Therefore, the original dataset was processed with various preprocessing steps prior to modeling.

Figures 1 and 2 show the sample sizes and proportional distribution of disease types in the dataset.

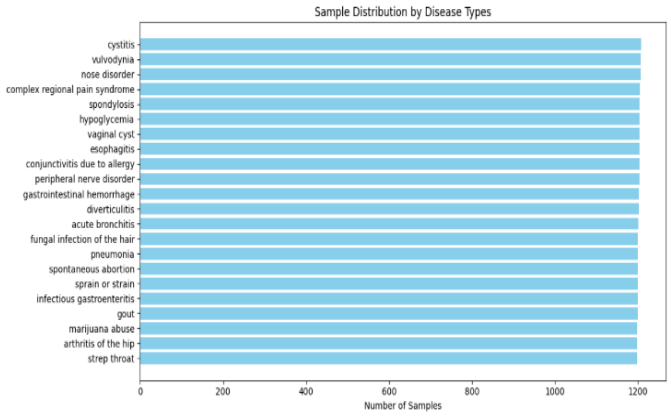


Figure 1. Distribution of sample numbers in the dataset according to disease type

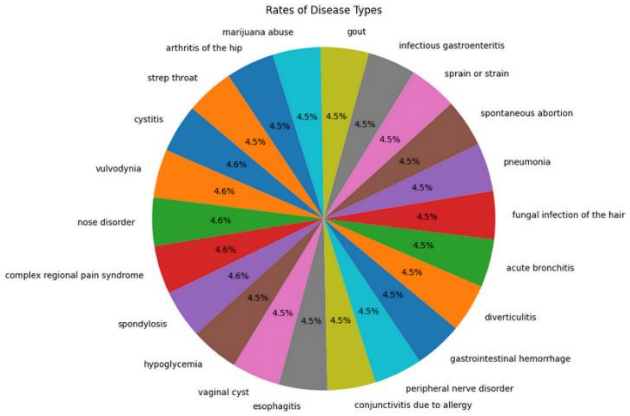


Figure 2. Proportional distribution of disease types in the dataset

2.2. Data Preprocessing

The following preprocessing steps were applied to improve data quality and classification performance prior to model training:

- **Removal of rare classes:** Rare disease classes with a sample size of less than 0.49% of the total data in the dataset were removed. This threshold value represents the minimum sample size required for the model to learn meaningfully. Additionally, since classes with very few examples negatively affect classification performance and cause excessive bias in the model, removing these classes has created a more balanced data structure.
- **Feature selection:** In order to reduce the complexity of the model and prevent noise caused by unnecessary features, ANOVA F-test and mutual_info_classif methods were applied, excluding the target variable, and 49 features with the highest discriminative power were determined. This number was selected as the optimal value that best balances the overall performance of the model and minimizes the risk of overfitting, based on experimental evaluations.

As a result of these preprocessing steps, a clean, balanced, and meaningful data structure has been obtained for model training.

Table 14. Preprocessing steps applied to the dataset

Stage	Number of Lines	Number of Columns	Description
Raw Data	246.945	378	Original dataset
Class Balancing	26.441	378	Rare classes removed
Feature Selection	26.441	49	The most meaningful features were selected

Each record consists of features containing one or more symptoms and the corresponding disease label. Based on the selected features, the dataset can be effectively applied to classification problems involving multiple classifications and high dimensionality.

2.3. Methods

In this study, various machine learning algorithms were applied for symptom-based disease prediction, and their performances were compared. The models used are as follows:

- **Logistic Regression:** Due to its high interpretability, it is a linear machine learning model that is frequently preferred, especially in multi-class classification problems [15].
- **Random Forest:** It is a powerful ensemble learning method in which a large number of decision trees make decisions by consensus, and it provides high performance in complex datasets [16].
- **Support Vector Machines (SVM):** An algorithm that can create effective decision boundaries in high-dimensional datasets and generalize well [17].
- **K-Nearest Neighbors (KNN):** A method that classifies test examples based on the closest training data, which has high computational costs despite its simple structure [18].
- **XGBoost:** An advanced boosting algorithm that offers improved accuracy and fast training advantages, performing particularly well on imbalanced datasets [19].
- **LightGBM:** A fast and efficient histogram-based gradient boosting algorithm that delivers high performance on large datasets [20].

The Grid Search method was used for model parameter optimization. For the Random Forest model, hyperparameters such as n_estimators, max_depth, and min_samples_split were optimized; for Logistic Regression, parameters such as C, penalty, solver, and l1_ratio were optimized. As a result of the optimization, the Random Forest model achieved the best result with settings such as n_estimators=100 and max_depth=20, while the Logistic Regression model achieved the best result with C=1 and penalty='l2'.
The n_estimators, max_depth, learning_rate, subsample, and colsample_bytree parameters for the XGBoost model; the C, kernel, and gamma parameters for SVM; and the n_neighbors, weights, and metric parameters for KNN were optimized using Grid Search.
Parameter searches for all models were tested using the 3-fold cross-validation method. After the best parameters were determined, the models were retrained with these settings and their performance was measured on the test data.
The LightGBM model was trained and monitored on the validation set using the early stopping method with fixed parameters such as n_estimators=1000, learning_rate=0.05, num_leaves=31, max_depth=-1, min_child_samples=20, subsample=0.8, and colsample_bytree=0.8.
The dataset was split using stratified sampling, which preserves class distributions, with 80% for training and 20% for testing. Additionally, the target variable was converted to numerical form using LabelEncoder to make it suitable for a multi-class classification structure.

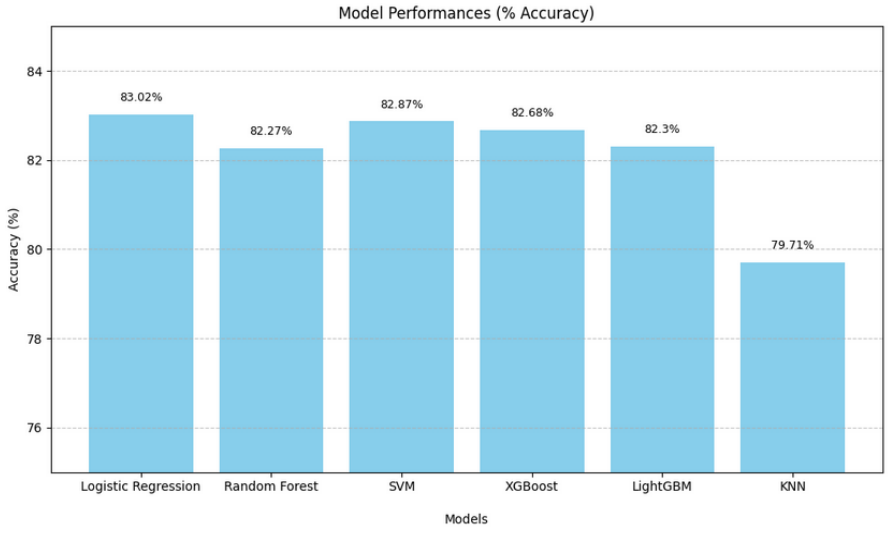


Figure 3. Accuracy of different models.

When the accuracy values obtained were examined, it was observed that the Logistic Regression model achieved the highest accuracy with 83.02%. This finding supports the reasons for preferring the Logistic Regression model in health data-based applications, especially with its high accuracy and balanced performance. Additionally, the model's high interpretability facilitates its use in medical decision support systems.

3. Performance Metrics

The following basic statistical metrics were used to evaluate model performance:

Accuracy: The ratio of correctly classified examples in all examples. It is generally preferred to measure model success in datasets where classes are balanced [21]. Accuracy is calculated as shown in Equation 1.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Precision: Indicates the proportion of positively predicted examples that are actually positive. It is a preferred metric when false positives are of high importance [22]. Precision is expressed by the formula in Equation 2.

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

Recall: Measures how many true positives are correctly identified. It is of critical importance in situations where false negatives need to be minimized [23]. Sensitivity is shown in Equation 3.

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

F1 Score: Obtained by taking the harmonic mean of precision and sensitivity. It is used to evaluate model performance more balanced, especially in imbalanced datasets [24]. The F1 score is calculated using the formula given in Equation 4.

$$F1\ Score = 2 * \frac{Precision*Recall}{Precision+Recall} \quad (4)$$

These metrics play an important role in understanding the strengths and weaknesses of the model. They are preferred for providing a balanced evaluation, especially in multi-class classification problems.

4. Experimental Results

In this study, the performance of Logistic Regression, Random Forest, SVM, XGBoost, LightGBM, and KNN models developed and optimized for symptom-based disease prediction was compared on training and test sets. The accuracy and F1 scores of the models were evaluated as basic performance indicators.

Table 2. Comparison of model accuracy and F1 score

Model	Accuracy	F1 Score
KNN	%79.7	%79.7
Random Forest	%82.2	%82.2
LightGBM	%82.3	%82.2
XGBoost	%82.7	%82.6
SVM	%82.9	%82.8
Logistic Regression	**%83**	**%82.9**

Table 2 compares the accuracy and F1 scores of all models. The Logistic Regression model showed the highest performance on the test set, with the SVM and XGBoost models achieving similar results. The LightGBM and Random Forest models showed competitive and balanced performance, while the KNN model showed relatively lower performance compared to other models.

These results show that all models performed the classification task in a generally balanced manner and that the Logistic Regression and SVM models in particular achieved high accuracy in disease prediction. Since accuracy alone can be misleading, it is important to use balanced metrics such as the F1 score. Saito and Rehmsmeier have emphasized that the F1 score provides a more reliable measure of success, especially in datasets with imbalanced class distributions [25].

The feature importance rankings obtained with the logistic regression model are presented in Figure 4. Some symptoms stand out significantly in disease prediction, which makes the model's decision-making process more understandable. Feature importance analyses are valuable in revealing which symptoms play a critical role, especially in clinical decision support systems. Similarly, Wang et al.'s 2025 study emphasized that the application of multi-stage feature selection and importance analysis approaches in high-dimensional electronic health records (EMR) increases the accuracy of models and provides higher acceptability in clinical settings [26].

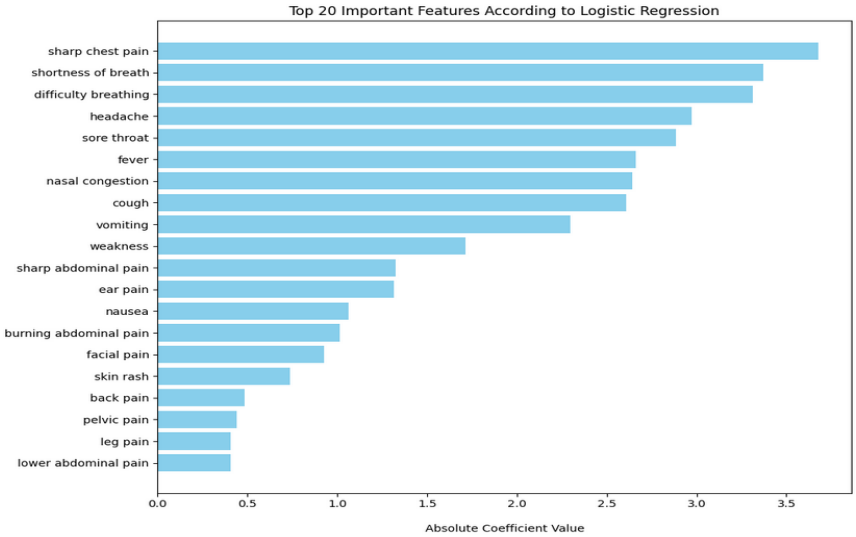


Figure 4. The most important features according to the logistic regression model

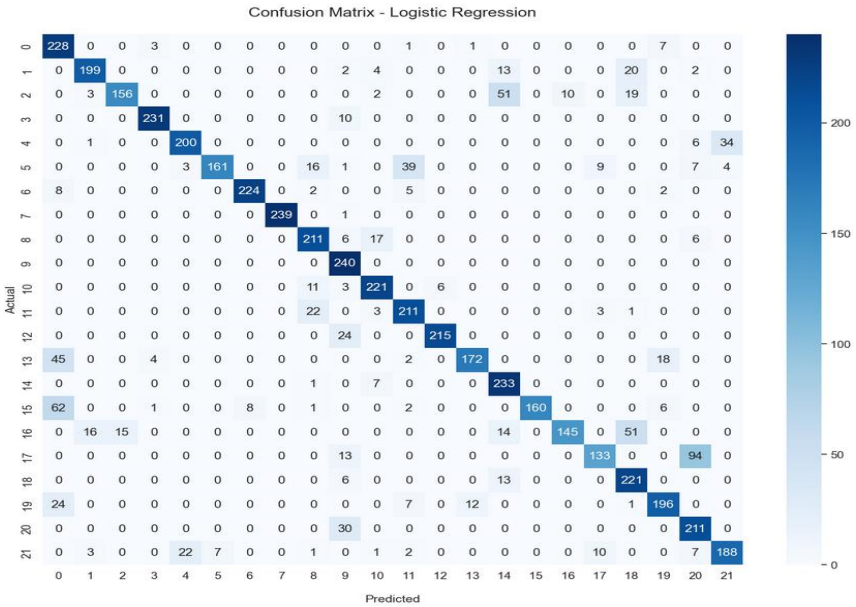


Figure 5. Confusion matrix for the logistic regression model

The confusion matrix created to examine the classification performance of the logistic regression model in detail is shown in Figure 5. Although the model generally exhibits high accuracy, some confusion was observed between certain disease classes.

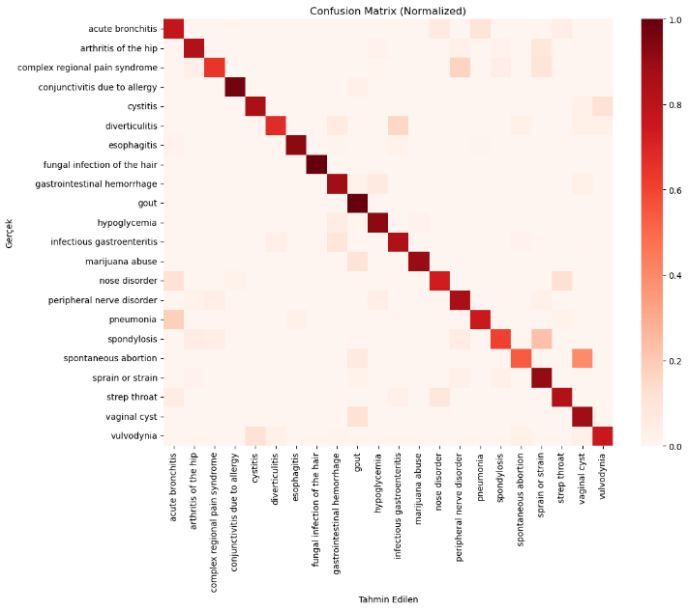


Figure 6. Visualization of the class pairs that the model confuses the most

The confusion observed between certain disease classes stems from the fact that these classes share similar symptoms. The most commonly confused classes are illustrated in Figure 6. This analysis provides important information for the more detailed modeling of symptoms in frequently confused classes and the development of targeted strategies for future model improvements.

ROC curves were plotted and AUC values were calculated to examine the classification performance of the models in more detail. Below are the ROC curves for Logistic Regression, Random Forest, and other models used. These graphs allow for a visual comparison of the overall discrimination capabilities of the models. In studies on ROC curves and performance evaluation metrics, Sokolova and Lapalme (2009) have noted that such measurements play an important role in comparative analysis [22].

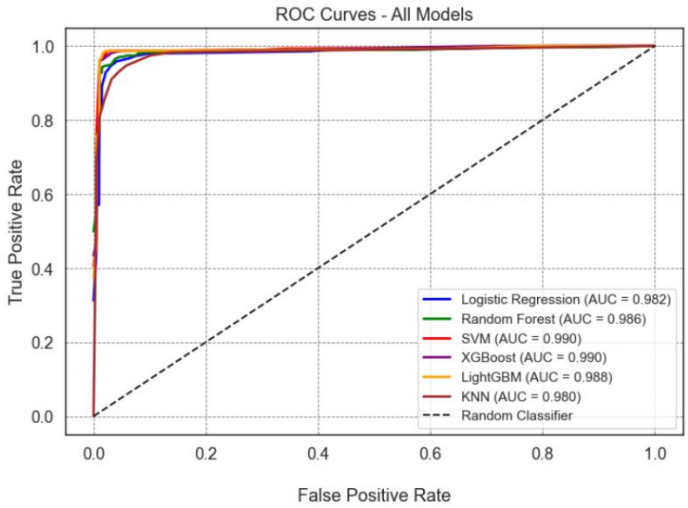


Figure 7. ROC curves and macro-average AUC values for different classification models

Figure 7 shows the ROC curves for the Logistic Regression, Random Forest, SVM, XGBoost, LightGBM, and KNN models. The overall performance of each model was evaluated using the macro-average AUC value for multi-class classification problems. The black dashed line represents the performance of the random classifier as a reference.

5. Discussion

The consistency between the training and testing performance of the developed models shows that the system can deliver effective results across different data subsets. Feature importance analyses can help identify specific symptoms that should be prioritized by physicians during the preliminary diagnosis process. Such AI-based decision support systems can provide rapid and critical decision support, particularly in regions with limited resources, while awaiting laboratory and imaging results.

The classification models developed in this study have demonstrated successful performance in disease prediction based on symptom data. In particular, the logistic regression model stands out as a preferred option in practical clinical applications with an accuracy rate of 83% and high interpretability. In addition, SVM and XGBoost models have also shown competitive performance with accuracy and F1 scores close to logistic regression. LightGBM and Random Forest models offer balanced results thanks to their ability to capture complex data patterns, while the KNN model has shown lower accuracy compared to the others.

When confusion matrices and model analyses were examined, high discriminative power was achieved between certain disease classes, but confusion persisted in some classes with high symptom similarity. This situation demonstrates that the discriminatory power of models is limited, especially for disease pairs with naturally overlapping symptom profiles. To reduce the impact of class imbalance in the dataset, a threshold of 0.49% was set to exclude rare disease classes from the analysis, thereby preventing the model from focusing excessively on imbalanced classes. As a result, the models learned all disease classes more evenly and produced reliable results.
ROC curves and AUC analyses confirmed that the models had high overall interclass discriminative power, with the logistic regression model in particular demonstrating balanced and robust performance in terms of AUC value. Additionally, SHAP analyses have detailed the decision-making mechanisms of the models and identified which symptoms are more decisive in the prediction process. This could enhance clinical experts' confidence in model predictions and facilitate the adoption of decision support systems.

As a result, the developed system's ability to provide effective disease predictions based on patient symptoms can make significant contributions to clinical decision-making processes, especially in situations where laboratory results are delayed or unavailable. In the future, the integration of data balancing techniques such as SMOTE and deep learning-based natural language processing methods such as Word2Vec and BERT may improve the prediction performance of rare classes. Additionally, testing the system with real-time patient data and under different clinical settings is critical for evaluating the model's effectiveness and applicability in daily practice.

6. Conclusion

In this study, a machine learning-based approach is proposed for disease prediction based on symptoms. The “Diseases and Symptoms Dataset” made publicly available on Kaggle was used as the dataset; rare classes were removed to reduce class imbalance, and the 49 most predictive features were selected using the ANOVA F-test and mutual_info_classif methods.

During the modeling process, Logistic Regression, SVM, XGBoost, LightGBM, Random Forest, and KNN algorithms were optimized and evaluated using Grid Search. The obtained test accuracies are as follows: Logistic Regression 83.0%; SVM 82.9%; XGBoost 82.7%; LightGBM 82.3%; Random Forest 82.2%; KNN 79.7%. These results show that logistic regression offers both high accuracy and interpretability advantages; SVM and XGBoost perform similarly to logistic regression; LightGBM and Random Forest provide competitive results with their ability to capture complex patterns; and KNN performs lower than other methods. Variable importance analyses showed that some symptoms significantly affect predictive power; confusion matrices revealed that challenges persist in distinguishing classes with similar symptom profiles.

ROC–AUC analyses demonstrate that the models have strong discriminatory power in multi-class classification problems. This high performance supports the potential for integration into clinical decision support systems in settings with limited access to healthcare services.

7. References

[1] Chaurasia V, Pal S. "A novel approach for breast cancer detection using data mining techniques". *International Journal of Innovative Research in Computer and Communication Engineering*, 2(1), 2456–2465, 2014.

[2] Shen J, Zhang CJP, Jiang B, Chen J, Song J, Liu Z, He Z, Wong SY, Fang P, Ming W. "Artificial intelligence versus clinicians in disease diagnosis: systematic review". *JMIR Medical Informatics*, 7(3), e10010, 2019. https://medinform.jmir.org/2019/3/e10010/.

[3] Miotto R, Li L, Kidd BA, Dudley JT. "Deep patient: An unsupervised representation to predict the future of patients from the electronic health records". *Scientific Reports*, 6, 26094, 2016. https://doi.org/10.1038/srep26094.

[4] Ciecierski-Holmes T, Singh R, Axt M, Brenner S, Barteit S. "*Artificial intelligence for strengthening healthcare systems in low- and middle-income countries: a systematic scoping review*." *npj Digital Medicine*, 5(1), 162, 2022. https://doi.org/10.1038/s41746-022-00700-y.

[5] Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. "Dermatologist-level classification of skin cancer with deep neural networks." *Nature*, 542(7639), 115–118, 2017. https://doi.org/10.1038/nature21056.

[6] Chavoshi M, Zamani S, Mirshahvalad SA. "Diagnostic performance of deep learning models versus radiologists in COVID-19 pneumonia: A systematic review and meta-analysis." *Clinical Imaging*, 107, 110092, 2024. https://doi.org/10.1016/j.clinimag.2024.110092.

[7] Rolando M, Raggio V, Naya H, Spangenberg L, Cagnina L. "A labeled medical records corpus for the timely detection of rare diseases using machine learning approaches." *Scientific Reports*, 15, 6932, 2025. https://doi.org/10.1038/s41598-025-90450-0.

[8] Yi F, Yang H, Chen D, Qin Y, Han H, Cui J, Bai W, Ma Y, Zhang R, Yu H. "XGBoost-SHAP-based interpretable diagnostic framework for Alzheimer's disease." *BMC Medical Informatics and Decision Making*, 23, 137, 2023. https://doi.org/10.1186/s12911-023-02238-9.

[9] Ambrish G, Ganesh B, Ganesh A, Srinivas C, Dhanraj, Mensinkal K. "Logistic regression technique for prediction of cardiovascular disease." *Global Transitions Proceedings*, 3(1), 127–130, 2022. https://doi.org/10.1016/j.gltp.2022.04.008.

[10] Jamian L, Wheless L, Crofford LJ, Barnado A. "Rule-based and machine learning algorithms identify patients with systemic sclerosis accurately in the electronic health record." *Arthritis Research & Therapy*, 21, 305, 2019. https://doi.org/10.1186/s13075-019-2092-7.

[11] Tokede B, Brandon R, Lee C, Lin G, White J, Yansane A, Jiang X, Kalenderian E, Walji M. "Development and validation of a rule-based algorithm to suggest periodontal diagnosis using electronic health record data." *Journal of Clinical Periodontology*, 51(5), 547-557, 2024. https://doi.org/10.1111/jcpe.13938.

[12] Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. "Dermatologist-level classification of skin cancer with deep neural networks". *Nature*, 542(7639), 115–118, 2017. https://doi.org/10.1038/nature21056.

[13] Stark B. "Disease and symptoms dataset 2023". Mendeley Data, V1, 2025. https://doi.org/10.17632/2cxccsxydc.1.

[14] He H, Garcia EA. "Learning from imbalanced data". *IEEE Transactions on Knowledge and Data Engineering*, 21(9), 1263–1284, 2009.

[15] Hosmer DW, Lemeshow S, Sturdivant RX. *Applied Logistic Regression* (3rd ed.). Hoboken, New Jersey, USA, Wiley, 2013. https://doi.org/10.1002/9781118548387.

[16] Breiman L. "Random forests". *Machine Learning*, 45(1), 5–32, 2001. https://doi.org/10.1023/A:1010933404324.

[17] Cortes C, Vapnik V. "Support-vector networks". *Machine Learning*, 20(3), 273–297, 1995. https://doi.org/10.1007/BF00994018.

[18]Halder RK, Uddin MN, Uddin MA, Aryal S, Khraisat A. "Enhancing K-nearest neighbor algorithm: a comprehensive review and performance analysis of modifications". *Journal of Big Data*, 11, 113, 2024. https://doi.org/10.1186/s40537-024-00973-y.

[19]Chen T, Guestrin C. "XGBoost: A scalable tree boosting system". *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 785–794, 13–17 August 2016. https://doi.org/10.1145/2939672.2939785.

[20]Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Ye Q, Liu T-Y. "LightGBM: A highly efficient gradient boosting decision tree". *Advances in Neural Information Processing Systems*, 30, 3149-3157, 2017.

[21]Powers DMW. "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation". *Journal of Machine Learning Technologies*, 2(1), 37–63, 2011. https://arxiv.org/abs/2010.16061.

[22]Sokolova M, Lapalme G. "A systematic analysis of performance measures for classification tasks". *Information Processing and Management*, 45(4), 427–437, 2009. https://doi.org/10.1016/j.ipm.2009.03.002.

[23]Davis J, Goadrich M. "The relationship between precision-recall and ROC curves". *Proceedings of the 23rd International Conference on Machine Learning*, 233–240, 2006. https://doi.org/10.1145/1143844.1143874.

[24]Chicco D, Jurman G. "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation". *BMC Genomics*, 21(1), 6, 2020. https://doi.org/10.1186/s12864-019-6413-7.

[25]Saito T, Rehmsmeier M. "The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets". *PLOS ONE*, 10(3), e0118432, 2015. https://doi.org/10.1371/journal.pone.0118432.

[26]Wang H, Zhang M, Mai L, Li X, Bellou A, Wu L. "An effective multi-step feature selection framework for clinical outcome prediction using electronic medical records". *BMC Medical Informatics and Decision Making*, 25(84), 2025. https://doi.org/10.1186/s12911-025-02922-y

Fuel Optimization in Flight Routes Using Machine Learning and Graph-Based Methods

Burak BEYAZIT[1]

[1] Turkish Technology, Küçük Çamlıca, Üçpınarlar Street No:104/1, 34696 Üsküdar, Istanbul, Turkey
[1]bbeyazit@thy.com

Abstract
This study presents a modular framework for fuel-efficient flight routing that combines tree-ensemble prediction with graph-based optimization. Segment- level costs—covering fuel, distance, and time—are learned from historical trajectories and ERA5 meteorology using XGBoost, LightGBM, Random Forest, and Extra Trees, then embedded as edge weights in Dijkstra and A* searches to yield wind-aware paths. Python-based simulations against EUROCONTROL data show consistent improvements in route length and theoretical fuel burn relative to flown tracks. The architecture supports rapid re-weighting under updated weather; future work will integrate live airspace constraints and field validation.

Keywords: fuel optimization; flight trajectory prediction; tree ensembles; graph-based routing; A*; wind-aware planning; EUROCONTROL; ERA5.

1. Introduction

Commercial aviation today operates under intensifying economic and environmental pressures. Fuel remains one of the largest operating expenses for airlines, commonly accounting for between one quarter and one third of direct costs, while the corresponding emissions position aviation as a notable contributor to global radiative forcing [1], [2]. Policy instruments such as the EU Emissions Trading System and ICAO's CORSIA further internalize the cost of emissions, creating a direct financial incentive to cut fuel burn through operational improvements in addition to fleet renewal and sustainable fuel adoption [3], [4]. At the same time, airspace programs like Free Route Airspace (FRA) and performance-based arrival procedures illustrate that carefully designed routing and procedure changes can unlock measurable savings without waiting for disruptive propulsion technologies [3], [4].

Conventional pre-tactical planning methods typically minimize geometric distance or follow fixed airway structures. These deterministic approaches are robust and interpretable, yet they are intrinsically limited in their ability to react to dynamic meteorological structure, sector configurations, and day-of-operations constraints [6]. Recent work has therefore moved toward hybrid frameworks that couple data-driven prediction with graph- theoretical optimization. In such systems, machine-learning models trained on historical trajectories and co-located weather learn to forecast deviations or fuel costs, and graph search methods compute low-cost paths under edge weights informed by those predictions [5]–[8]. Tree-ensemble algorithms—Random Forest, Extra Trees, XGBoost, and LightGBM—have proved particularly effective for aviation prediction tasks because they capture nonlinear interactions while remaining comparatively fast to train and deploy [8]–[10]. On the optimization side, Dijkstra and A* remainthe canonical choices for shortest-path search on fixed graphs, combining optimality guarantees with practical runtime when equipped with admissible heuristics [6], [7], [17].
This study proposes a two-stage framework designed for production-minded fuel optimization. A deviation/fuel prediction module first learns segment-level relationships between plan–actual kinematics and meteorological context, yielding predicted fuel surrogates for candidate graph edges. A graph-based optimization module then constructs a local lattice around the filed route and computes minimum-cost paths using Dijkstra and A*, where edge weights blend geometric length with predicted fuel and soft penalties for turning and constraint proximity. The focus is deliberately restricted to tree-based ensembles and classical graph search in order to maintain reproducibility, predictable latency, and clear optimality properties, avoiding bio-inspired heuristics that would complicate certification and line operations. The empirical scope covers EUROCONTROL-supervised European flights from 2021 to 2023 and ERA5 reanalysis meteorology sampled to each trajectory point [5], [13]–[15]. The main result is

[1]Corresponding author

that the predictive module attains high accuracy across aircraft types and weather regimes, and the optimization module delivers theoretical savings in the ≈3–18% range under relaxed airspace constraints, with the largest gains on long- haul legs exposed to strong jet streams.

2. Data and Methodology

The empirical pipeline integrates multi- source data acquisition, rigorous preprocessing, supervised learning, and graph-based search in a modular architecture designed for reproducibility and low latency. EUROCONTROL trajectory/operations records supply filed plans and realized tracks with latitude– longitude, flight level, timestamps, aircraft type and operator attributes, and route sequencing across the pan-European network, forming the operational backbone of the study [5]. To capture atmospheric forcing, ERA5 reanalysis from the Copernicus Climate Data Store contributes hourly u/v wind components, ambient temperature, sea-level pressure, and cloud fraction; these variables are matched to trajectory points by spatiotemporal interpolation so that each segment is paired with a meteorological state reflective of its time and location [13]. In order to keep the optimization stage interpretable and type- aware without invoking full performance simulators, we adopt BADA-style conventions to calibrate surrogate fuel parameters per aircraft class, thus preserving consistent scaling across narrowbody and widebody fleets [14], [15]. To avoid sample-selection bias, flights are stratified by market segment, ensuring that short/medium-haul and long-haul operations contribute balanced coverage to both training and evaluation splits.

Preprocessing aims to homogenize spatial– temporal resolution and remove artifacts that could leak into the predictor. Realized tracks are resampled along fixed arc-length intervals to standardize point density, which stabilizes derivative-like features such as segment length and heading change. Conservative filters remove corrupted observations (e.g., nonphysical coordinate jumps, duplicated or backward timestamps, altitude spikes), and missing values—rare after synchronization—are imputed only when physically justified. Feature engineering then assembles four families of predictors. Geometric variables quantify differences between filed and realized states (Δ latitude, Δ longitude, Δ flight level) together with great-circle spacing to the next waypoint; these capture the inherent spatial structure of the planned route. Temporal indicators (hour-of-day, day-of-week, month) model recurrent traffic waves and seasonality. Meteorological projections convert ERA5 winds into headwind and crosswind relative to the segment heading, while gradients in pressure and temperature summarize synoptic variability known to correlate with rerouting [13]. Finally, path-normalized coordinates and sequence indices encode longitudinal context so that models can learn position-dependent behaviors along the route. This combination yields a compact but expressive design matrix that balances interpretability and predictive power.

To represent fuel at segment resolution without sacrificing computational speed, we introduce a linear surrogate of the form

$$Cij = \theta 0 + \theta 1 \Delta sij + \theta 2 \Delta hij + \theta 3 HWij + \theta 4 \Delta tij,$$

where Δsij is the great-circle distance, Δhij the altitude change, $HWij$ the headwind projection (positive into wind), and Δtij the elapsed time. Parameters θ are estimated per aircraft type via ridge- regularized regression on historical segments, ensuring variance control and consistency with BADA conventions [14], [15]. Although not a full performance model, this surrogate captures differential effects across candidate paths so that cost rankings remain physically sensible as weather and geometry change. The predictive module overlays the surrogate with data-driven residual learning using four tree-ensemble regressors. Random Forest provides a robust baseline with straightforward interpretability; Extra Trees reduces variance via randomized split thresholds and, in our experiments, attains the lowest MAE; XGBoost, with explicit regularization and shrinkage, offers strong control of extreme errors; and LightGBM delivers competitive accuracy with fast histogram-based training. To prevent leakage, flight-wise splits are enforced so that all segments from a given flight reside exclusively in either train or test. Five-fold cross-validation yields reliable generalization estimates. We report MAE, RMSE, and R^2, and assess feature importance both via model-specific gains and permutation tests for a model-agnostic view [8]–[12]. Hyperparameters are tuned with coarse grid search and early stopping to favor stability over aggressive overfit, reflecting the deployment objective where predictable behavior is preferable to marginal gains.

Optimization proceeds on a directed, weighted graph constructed locally around the filed path. At each filed waypoint we generate a lattice of candidate offsets subject to bounded heading changes and altitude step limits, thereby producing a sparse graph whose edges represent feasible micro-deviations consistent with controller workload and passenger comfort. Each admissible edge from node *i* to node *j* receives the composite weight

$$wij = \lambda 1 \Delta sij + \lambda 2 \mathcal{C}ij + \lambda 3 \phi(\text{turn}) + \lambda_4 \psi(\text{constraints}),$$

Where $\mathcal{C}ij$represents predicted segment fuel obtained by combining the linear surrogate with machine learning residuals, ϕ(turn)applies penalties for excessive heading changes,and ψ(constraints)penalizes proximity to regulatory or capacity boundaries.Dijkstra's algorithm is employed to establish optimal baselines with full guarantee of solution accuracy, while the A* algorithm accelerates computation using an admissible heuristic that integrates great-circle lower bounds with non- negative wind adjustments to avoid overestimation of residual costs [6],[7],[17].
The modularity of this architecture allows updated meteorological inputs or recalibrated aircraft coefficients to propagate seamlessly from the predictive stage to the optimization stage without refactoring. Consequently, the system supports dynamic in-flight re-planning by simply re-weighting graph edges, maintaining both efficiency and adaptability for operational use.

3. Findings and Discussion

We evaluate along two axes: predictive accuracy with and without meteorology, and optimization benefits under distance- only versus wind-aware cost structures. On the predictive axis, geometry-only features already capture much of the spatial organization of planned routes; however, the models under-estimate large deviations, particularly during synoptically active periods. Adding ERA5 covariates— headwind and crosswind projections together with pressure and temperature— materially improves realism and compresses tail errors, in line with prior reports on weather-mediated trajectory prediction [8]–[12]. Using our full corpus of ≈185,000 flights (with over five million resampled segments), Extra Trees yields MAE ≈ 7,234 m, RMSE ≈ 31,008 m, and R^2≈ 0.77, while XGBoost achieves a nearly identical RMSE (≈ 31,005 m) and R^2 (≈ 0.77) with slightly higher medians but tighter control of extreme errors; Random Forest is stable yet exhibits mean-reversion under disturbed conditions; Gradient Boosting underperforms on this task. These magnitudes, summarized in Table 1, set realistic expectations for downstream routing.

Table 1. Comparative performance of machine learning models for deviation prediction

Model	MAE (m)	RMSE (m)	R^2
RandomForest	7,417.78	35,140.43	0.71
ExtraTrees	7,234.49	31,008.04	0.77
GradientBoosting	14,201.47	63,667.64	0.03
XGBoost	12,384.40	31,004.51	0.77
LightGBM	13,014.80	36,700.19	0.68

A detailed feature-importance analysis clarifies mechanisms. Under benign weather, filed geometry—waypoint longitude/latitude, requested flight level, and the city-pair—dominates the signal. As synoptic forcing strengthens, headwind projections and pressure fields rise in rank, and their marginal contribution becomes strongly state-dependent: precisely when variability is high, meteorology adds the most predictive power. Spatial error maps concentrate near FIR boundaries and convective regions where sector-capacity interventions and tactical reroutes are common; temporal error peaks align with morning/evening traffic waves. Two systematic biases persist: (i) under-prediction on long vectors, and (ii) reversion toward mean patterns during heavy regulation. Both effects are mitigated by rolling re-training and explicit time-of- day indicators that proxy traffic pulses [12]. In additional stratified checks, narrowbody short/medium-haul segments show slightly lower dispersion than widebody long-haul, while winter months exhibit broader error tails consistent with stronger baroclinic activity—again supporting the value of meteorological covariates.

To gauge calibration, residuals are inspected as a function of predicted magnitude. Geometry-only models are over-confident for large deviations; adding weather reduces slope and narrows prediction-interval miss rates. While we do not compute formal conformal intervals here, empirical coverage on held-out folds improves monotonically once ERA5 variables are included, suggesting that the ensembles are not only more accurate but also better calibrated for operational risk assessment [8]–[12].

On the optimization axis, we contrast flown tracks with two hypothetical planners. The distance-only configuration sets $wij = \lambda 1 \Delta sij$, providing an upper bound on geometric straightening. The wind-aware configuration activates the full composite weight so the search may accept slightly longer distances to exploit tailwinds or avoid headwinds. Across a stratified evaluation of thousands of flights drawn from the 185,000-flight corpus—covering 600–900 NM short-haul and 3,000–5,000 NM long-haul legs—distance-only paths reduce length by ≈ 1–10% (largest where plans were already circuitous). Wind-aware routes consistently outperform on days with strong jet structure, achieving ≈ 3–18% theoretical savings, with the highest gains on long-haul segments where latitude shifts align with persistent tailwinds. Importantly, runtime remains within dispatch budgets: short-haul problems complete well under a second on commodity hardware, long-haul in a few seconds—reflecting the sparse lattice and A* pruning.

To make these improvements concrete, consider two stylized cases. On a 700 NM sector under moderate headwind, distance- only straightening trims ~2–3% by removing filed dog-legs; wind-aware adds ≤ +1 pp unless the wind field is highly anisotropic. By contrast, on a 4,200 NM leg with strong winter jets, the wind-aware planner migrates one to two latitude bands to harvest tailwinds and avoid adverse shear, yielding double-digit savings with only a modest increase in geometric distance—behavior consistent with dispatcher heuristics and recent literature on energy-aware routing [6], [17].

Ablation studies quantify design choices. Replacing the ML residual with the linear surrogate alone reduces savings by ~1–3 pp on average, indicating that residuals capture actionable nuance in wind–turn interactions and local vertical profiles. Removing ϕ(turn)produces visually implausible zig- zagging and increases node expansions without material benefit; re-introducing a mild penalty restores smoothness with negligible impact on optimum value. Incorporating a wind-adjusted admissible heuristic reduces A* node expansions by ~33–50% in adverse weather while preserving optimality [6], [17]. Sensitivity to λshows a broad plateau: distance and fuel terms dominate objective value, while the turn term primarily governs path regularity; within a reasonable band, savings are stable, which is desirable for operations.

We also examine scalability and latency. Node and edge counts grow linearly with the number of candidate offsets per waypoint; because offsets are locally constrained and the lattice is sparse, memory usage remains modest. With meteorology cached along the lattice and vectorized weight updates, a full re-weight and re-solve can be executed each time a new ERA5 cycle arrives. This supports near real-time re-planning, and the clean separation between prediction and optimization allows the system to ingest updated type coefficients or revised weather without code changes—only edge weights change.From an operational perspective, single- digit savings are economically meaningful given fuel price volatility and thin margins. With an emissions factor of ≈ 3.16 kg CO_2 per kg fuel, a 10-ton reduction on a transoceanic flight avoids > 30 t CO_2; scaled to a network, the effect is material. Accurate deviation forecasting also reduces buffer fuel and schedule padding, improving aircraft and crew utilization. Per-flight fuel/emissions accounting supports ETS/CORSIA compliance and dovetails with corporate sustainability reporting [3], [4]. Strategically, the method complements FRA: greater lateral freedom at upper levels enlarges the candidate lattice and amplifies the value of fuel-aware edge weights [3], [4].

Nevertheless, limitations remain. The prototype uses soft penalties rather than hard operational constraints; sector capacities, NOTAM activation, and TMIs are not explicitly enforced, so optimized routes should be viewed as upper bounds rather than ready-to-clear flight plans. The focus on Europe prompts questions about transferability to NAT tracks or monsoon- affected corridors; domain adaptation and rolling re-training will be required. Deterministic ERA5 fields omit forecast uncertainty; robust routing with ensemble NWP (e.g., quantile costs) would better represent risk [11], [13]. Finally, adoption hinges on human-in-the-loop tooling— transparent explanations for route choice, confidence overlays, and simple controls for editing constraints—so that dispatchers and ATC coordination staff can interrogate and adjust recommendations.

Overall, the results indicate that combining tree-ensemble prediction with graph-based optimization yields consistent, interpretable gains that are operationally plausible in both runtime and behavior. The approach not only shortens or re-phases paths but also aligns trajectories with favorable wind structure, realizing savings that purely geometric methods cannot access. These findings reinforce the view that data- science-driven, graph-optimized flight planning is a practical near-term lever for cost efficiency and decarbonization while longer-horizon technologies mature.

4. Conclusion

Coupling tree-ensemble prediction with classical graph search provides a practical and reproducible path to fuel-aware routing. With co-located ERA5 meteorology, the predictive module attains stable accuracy on held-out flights ($R^2 \approx 0.77–0.78$), and embedding the resulting fuel surrogates into a waypoint lattice enables the optimizer to trade geometric efficiency against wind structure and turn smoothness in a transparent way. In controlled experimentsagainst EUROCONTROL trajectories, purely geometric straightening yields ~1–10% path reductions, while wind-aware routing achieves ~3–18% under strong jet patterns; both regimes execute with latencies compatible with near real-time dispatch, underscoring operational feasibility rather than only theoretical promise. Beyond headline savings, the architecture contributes a clean division of concerns between prediction and optimization that simplifies maintenance and certification: model recalibration or weather updates alter only edge weights, not solver logic, which in turn preserves optimality guarantees on fixed graphs (Dijkstra) while exploiting admissible heuristics for speed (A*). The approach also enhances explainability—a key adoption barrier—because costs decompose into interpretable terms (distance, predicted fuel, turn penalties, constraint proximity), allowing dispatchers to audit why a recommended path deviates from the filed plan and to gauge the sensitivity of that recommendation to wind shifts or λ-weight adjustments.

Economically and environmentally, even single-digit improvements are material at network scale. Reductions in planned distance and predicted burn translate directly into lower operating costs and avoided CO_2, and the same predictive signals that power routing can support ancillary decisions—buffer-fuel sizing, schedule padding, and crew/aircraft rotations—thereby compounding value. Strategically, the method complements Free Route Airspace and performance- based navigation initiatives: as lateral degrees of freedom expand, the benefit of fuel-aware edge weights grows, positioning the framework as a near-term lever for decarbonization while longer-horizon technologies (new airframes, SAF, propulsion) mature.

Limitations highlight the roadmap. The present prototype enforces soft, not hard, airspace constraints; integrating live sector capacities, NOTAM/TMI activations, STAM/UDPP measures, and procedure- level requirements is essential to yield flyable outputs. Deterministic reanalysis cannot express forecast risk; coupling the planner to ensemble NWP and quantile- aware objectives would enable robust choices under uncertainty and support dispatcher-tunable risk profiles. External validity beyond Europe will require domain adaptation across oceanic tracks and monsoon-affected corridors, together with rolling re-training to track seasonal regimes. Finally, human-centered interfaces should surface confidence overlays, counterfactual explanations (“what if $\lambda_2\uparrow$?”), and lightweight constraint editing to keep operators in the loop.

With these extensions—hard constraint integration for operational compliance, probabilistic weather and risk-sensitive objectives, cross-regional adaptation, and explainable operator tooling—a data- science-driven, graph-optimized planner can progress from promising prototype to deployable decision support that systematically advances cost efficiency and sustainability in day-to-day airline operations.

Acknowledgement

The author would like to express sincere gratitude to his academic advisor and faculty at Marmara University for their continuous guidance and support throughout this research. Special thanks are extended to Turkish Technology for providing valuable insights and an inspiring research environment that greatly contributed to the development of this study. The author also acknowledges EUROCONTROL for granting access to flight trajectory and operational

datasets, which formed the empirical basis of the analysis. Their contributions and support have been indispensable in completing this work.

5. References

[1] V. Mgbachi, "AI in Business Aviation Route Optimization: Reducing Fuel Consumption and Environmental Impact," *Journal of Business and Strategic Management*, vol. 9, no. 5, pp. 48–82, 2024.
[2] Airspace Intelligence, "AI-assisted flight planning deployments," 2021.
[3] FAA, "FAA Adds Fuel-Saving Arrival Routes for 11 Airports," 2023. [Online]. Available: https://www.faa.gov/newsroom/faa-adds- fuel-saving-arrival-routes-11-airports
[4] EUROCONTROL, "Free Route Airspace (FRA) Benefits," 2024. [Online]. Available: https://www.eurocontrol.int/concept/free- route-airspace
[5] EUROCONTROL, "Pan-European flight datasets" and STATFOR market segments, 2023–2024.
[6] D. Sun, et al., "Review of shortest-path methods for dynamic air-traffic networks," 2017.
[7] P. E. Hart, N. J. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *IEEE Trans. Systems Science and Cybernetics*, 1968.
[8] Y. Zhang, et al., "Tree-ensemble prediction of flight fuel and delay with weather and operations data," 2020.
[9] J. M. Klötergens, et al., "Predicting deviation of flight entry into air sector using machine learning techniques," in *Proc. IEEE/AIAA 42nd DASC*, 2023.
[10] R. V.-A. Gracia-Berná, et al., "Data- driven models for accurate estimation of fuel consumption using deep learning techniques," in *Proc. IEEE/AIAA 42nd DASC*,2023.
[11] F. Li, R. Wang, M. Ma, Z. Du, and X. Xu, "Flight trajectory prediction … with Interacting Multiple Model and Informer framework," *Sensors*, vol. 25, no. 8, 2025.C. Liu, Z. Chen, et al., "Airport delay prediction with Temporal Fusion Transformers," in *Proc. 17th ACM SIGSPATIAL IWCTS (GenAI & Smart Mobility)*,2024.
[12] H. Hersbach, et al., "ERA5: Global reanalysis—variables and access via Copernicus CDS," 2023.
[13] S. Chati, et al., "Linearized fuel-burn surrogates for trajectory optimization with BADAparameters," 2022.
[14] EUROCONTROL,BADA documentation, 2023–2024.
[15] "Machine Learning-Based Flight Route Optimization with Real-Time Weather Insights," Int. Res. J. of Modernization in Engineering, Technology and Science, vol. 7, no. 4, pp. 2828–2834,2025.
[16] J. Xie and M. Zhong, "A*-based energy-aware route planning under adverse weather," 2018.