

ISSN: 2757-9255



ÇUKUROVA ÜNİVERSİTESİ MÜHENDİSLİK FAKÜLTESİ DERGİSİ

CUKUROVA UNIVERSITY
JOURNAL OF THE FACULTY OF
ENGINEERING

EYLÜL
SEPTEMBER
2025

Cilt/Volume 40

Sayı/Issue 3



ÇUKUROVA ÜNİVERSİTESİ
MÜHENDİSLİK FAKÜLTESİ DERGİSİ
Cukurova University,
Journal of the Faculty of Engineering



Sahibi (Owner)

Ali KESKİN
Ç.Ü. Mühendislik Fakültesi Dekanı

Editör (Editor)

Serin MEZARCIÖZ

EDİTÖR KURULU (Editorial Board)

Ali KESKİN (Çukurova Üniversitesi)
Umut ORHAN (Çukurova Üniversitesi)
Mutlu AVCI (Çukurova Üniversitesi)
Mustafa GÖK (Çukurova Üniversitesi)
Hüseyin VAPUR (Çukurova Üniversitesi)
Hüseyin AKILLI (Çukurova Üniversitesi)
Ulus ÇEVİK (Çukurova Üniversitesi)
Ali KOKANGÜL (Çukurova Üniversitesi)
Mesut BAŞIBÜYÜK (Çukurova Üniversitesi)
Hacı Murat ARSLAN (Çukurova Üniversitesi)
Oğuz DEMİRYÜREK (Erciyes Üniversitesi)
Hüsamettin BULUT (Harran Üniversitesi)
Metin GÜRÜ (Gazi Üniversitesi)
Mehmet Emin AYDIN (University of the
West of England, İngiltere)
Azra BİLGİN (University of Colorado Boulder,
ABD)

Serin MEZARCIÖZ (Çukurova Üniversitesi)
Hasan ÇETİN (Çukurova Üniversitesi)
Özen KILIÇ (Çukurova Üniversitesi)
Hüseyin ERTEN (Çukurova Üniversitesi)
Belgin BAYAT (Çukurova Üniversitesi)
Mustafa ÖZCANLI (Çukurova Üniversitesi)
R. Tuğrul OĞULATA (Çukurova Üniversitesi)
Onur ERMAN (Çukurova Üniversitesi)
Yusuf KUVVETLİ (Çukurova Üniversitesi)
Hüseyin R. YERLİ (Çukurova Üniversitesi)
Ergun ERÇELEBİ (Gaziantep Üniversitesi)
Nihan ÇETİN DEMİREL (Yıldız Teknik Üniversitesi)
Serter ATABAY (American University of Sharjah,
United Arab Emirates)
Muammer KOÇ (Hamad Bin Khalifa University,
Qatar)
Tamio ENDO (Mie University-Japonya)

Hakemli Bir Dergidir. Yılda Dört Sayı Basılır. / Refereed Journal. Published Quarterly.

Dergimiz **TÜBİTAK ULAKBİM** Mühendislik ve Temel Bilimler Veri Tabanı Listesi'nde taranmaktadır.
Indexed by **TÜBİTAK ULAKBİM** Mühendislik ve Temel Bilimler Veri Tabanı Listesi

Derginin eski adı: Çukurova Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi
(ISSN: 1019-1011)

YAZIŞMA ADRESİ/CONTACT

Ç.Ü. Mühendislik Fakültesi Dergi ve Yayın Kurulu Başkanlığı Balcalı 01330 Sarıçam/ADANA
Tel : (322) 338 63 57 – (322) 338 70 50 E-posta: cu.mmfdergi@gmail.com



AMAÇ ve KAPSAM

Çukurova Üniversitesi Mühendislik Fakültesi Dergisi, yılda dört sayı olarak yayımlanan hakemli bilimsel bir dergidir. Mühendislik alanındaki gelişmeleri takip etmek, meslek kuruluşları ve bireylerin ulusal ve uluslararası gelişimlerine katkıda bulunmak ve bu alanlarda Türkçe bir kaynak oluşturmak amacıyla yayımlanmaktadır. Derginin yazım dili Türkçe/İngilizcedir. Daha önce başka bir yerde yayımlanmamış tüm özgün ve güncel bilimsel araştırma ve uygulamaları kapsayan makaleler kabul edilmektedir. Tüm mühendislik alanları ile ilgili teorik/deneysel ve uygulamaya yönelik çalışmalar Dergi'nin ilgi alanındadır. Çeviri ve derleme makaleler kabul edilmemektedir. “**Araştırma ve uygulama içeren makaleleri**” yayınlamaktadır.

AIM and SCOPE

Journal of the Engineering Faculty of Cukurova University is a peer-reviewed scientific journal which is currently published quarterly. The journal is published to follow the developments in the field of engineering, to contribute to the national and international development of professional organizations and individuals, and to create a Turkish resource in these areas. The language of the journal is Turkish/ English. All original and up to date scientific research and applications articles that have not been published elsewhere are accepted. Theoretical/experimental and practical studies related to all engineering fields are in the scope of the Journal. Translation and compilation articles are not accepted. The journal publishes “**articles of research and practice**”.



İÇİNDEKİLER

Investigation of Konya Meteorological Snow Cover Depth Trend Using MK and ITA	501
<i>Konya Meteorolojik Kar Örtüsü Derinlik Yöneliminin MK ve ITA Yöntemleri Kullanılarak Araştırılması</i>	
Mehmet Fatih ŞAHAN, Büşra YEŞİLGÜL, İsmail ÜNSAL	
The Effects of Steaming Duration and Temperature on Pilling and Fastness Properties in Vigoureux Printing Technique.....	513
<i>Vigore Baskı Tekniğinde Buharlama Süresi ve Sıcaklığının Boncuklanma ve Haslık Özelliklerine Etkileri</i>	
Muhammed İdris AKTAŞ, Behzat YILDIRIM, Halil İbrahim İÇOĞLU, Mehmet TOPALBEKİROĞLU	
1x2 Dizilimli Grup Helisel Kazıklarda Mesafe Etkisinin Laboratuvar Model Deneyleriyle İncelenmesi	521
<i>Analysis of the Spacing Effect in 1x2 Configured Group Helical Piles using Laboratory Model Tests</i>	
Burak TİNTAŞ, Murat ÖRNEK	
Diagnosis of Hepatocellular Carcinoma - HCC Liver Cancer Using Federated Learning on MR Images	531
<i>MR Görüntülerinde Federasyonlu Öğrenme Kullanılarak Hepatosit Karsinomu - HCC Karaciğer Kanseri Tanısı</i>	
Burak UZDUR, Erkut TEKELİ, Turgay İBRİKÇİ, Harun Ur RASHID, Geetha RAMACHANDRAN	
Beam-Limited k-Step Lookahead for Computationally Efficient HMM Decoding	545
<i>Hesaplama Açısından Verimli HMM Kod Çözümü İçin Demet Sınırlamalı k-Adımlı İleri Bakış</i>	
Mehmet KURUCAN	
Portakal Kabuğu ve Portakal Ağacı Yaprığından Elde Edilen Ekstraktlarla Standart Çok Lifli Kumaşın Boyanması.....	559
<i>Dyeing of Standard Multifiber Fabric with Extracts Obtained from Orange Peel and Orange Tree Leaves</i>	
Sabiha SEZGİN BOZOK, R. Tuğrul OĞULATA	
Modeling and Performance Analysis of Battery Energy Storage Incorporated Grid-Connected Photovoltaic Plants.....	569
<i>Batarya Enerji Depolama Sistemine Sahip Şebeke Bağlantılı Fotovoltaik Santrallerin Modellenmesi ve Performans Analizi</i>	
Türker Aşkın KABADAYI, Özgür ÇELİK	

Detection of Imbalance Faults in Industrial Machines by Means of Frequency-Based Feature Extraction Using Machine Learning and Deep Learning Approaches.....	581
<i>Makine Öğrenimi ve Derin Öğrenme Yaklaşımlarını Kullanarak Frekans Tabanlı Öznitelik Çıkarımı ile Endüstriyel Makinelerde Dengesizlik Hatalarının Tespiti</i>	
Feyyaz ALPSALAZ	
Tıbbi Akıl Yürütme için Gemma 3 Modellerinin CoT Destekli SFT ve GRPO ile İncelenmesi.....	593
<i>Investigation of Low-Parameter Gemma 3 Models for Medical Reasoning Using CoT-Supported SFT and GRPO</i>	
İsmail İŞERİ, Alper YILDIRIM, Alihan ÖZTORUN, Tuğba TUNA, Arda TURAN	
Development of Clarke and Park Transforms Visualization Software Using Python.....	607
<i>Python Kullanılarak Clarke ve Park Dönüşümlerini Görselleştirme Yazılımının Geliştirilmesi</i>	
Oğuzhan TİMUR, Kadir Sami POYRAZ	
Çözelti Üfleme Tekniği ile Farklı Bitim İşlemi Kimyasallarının Tekstil Yüzeylerine İşletme Şartlarında Aktarılması ve Performanslarının Değerlendirilmesi.....	617
<i>Transfer of Different Finishing Chemicals to Textile Surfaces by Solution Blowing Technique Under Operating Conditions and Evaluation of Their Performance</i>	
Koray PEKTAŞ, Onur BALCI	
Akıllı Şebekelerde Elektrik Hırsızlığı Tespiti İçin Gelişmiş Makine Öğrenmesi Yöntemlerinin Uygulanması.....	627
<i>Application of Advanced Machine Learning Methods for Electricity Theft Detection in Smart Grids</i>	
Ömer Can TOLUN, Kasım ZOR, Önder TUTSOY	
Entegre Sabit Filmlili Aktif Çamur Sistemi ve Membran Biyoreaktör ile Farklı Yapıdaki Biyofilm Performanslarının Değerlendirilmesi	643
<i>Evaluation of Biofilm Performance of Different Structures with Integrated Fixed Activated Sludge and Membrane Bioreactor</i>	
Türker TÜRKEN, Özgün YILMAZ	
A Novel Multi-Head Attention Framework for COVID-19 Detection: Hybrid Integration of MobileNet and VGG19 with Enhanced Feature Learning	655
<i>COVID-19 Tespiti için Yeni Bir Çok Başlı Dikkat Çerçevesi: MobileNet ve VGG19 Tabanlı Geliştirilmiş Özellik Öğrenimi ile Hibrit Entegrasyon</i>	
Şafak KILIÇ	
Kullanıcı Eğilimlerinin Göreceli Modellenmesinin Popülerlik Yanlılığına Etkisi	671
<i>The Effect of Relative Modeling of User Trends on Popularity Bias</i>	
Nilüfer BALLI, Tuğba TÜRKÖĞLU KAYA	
Bağlayıcı Olarak Bazaltik Pomza Tozu Kullanımının Geopolimer Harçların Mekanik Dayanımına Etkisi.....	687
<i>The Influence of Using Basaltic Pumice Powder as a Binder on the Mechanical Strength of Geopolymer Mortars</i>	
Taha Salah Wahhab AL-ANTAKİ, Anıl NİŞ	

Structural Design of the Electric Vehicle Components Using Runge-Kutta Optimization Algorithm	699
<i>Runge-Kutta Optimizasyon Algoritması Kullanılarak Elektrikli Araç Bileşenlerinin Yapısal Tasarımı</i>	
Dildar GÜRSES	
Sağlık Alanında Açıklanabilir Yapay Zekâ Kullanımı Üzerine Bir Uygulama	707
<i>An Application on the Use of Explainable Artificial Intelligence in the Healthcare</i>	
Gülhan TOĞA	
Basamak Patlatma Tasarımı ve Analizi Yapabilen Bir Mobil Uygulama Yazılımının Geliştirilmesi	725
<i>Development of a Mobile Application Software Capable of Bench Blast Design and Analysis</i>	
Barış KADEM, Abdulkadir KARADOĞAN	
Thermophysiological Insights Into Madaline®-Based Multilayer Textiles For Medical Protection	743
<i>Tıbbi Koruyucu Giysilerde Kullanılan Madaline® Tabanlı Çok Katmanlı Tekstillerin Termofizyolojik Analizi</i>	
Pelin ALTAY	

Investigation of Konya Meteorological Snow Cover Depth Trend Using MK and ITA

Mehmet Fatih ŞAHAN^{1,a}, Büşra YEŞİLGÜL^{2,b}, İsmail ÜNSAL^{1,c}

¹Adıyaman University, Faculty of Engineering, Department of Civil Engineering, Adıyaman, Türkiye

²Adıyaman University, Institute of Graduate Science, Department of Civil Engineering, Adıyaman, Türkiye

^aORCID: 0000-0002-3334-3945; ^bORCID: 0000-0002-6708-7150; ^cORCID: 0000-0002-6324-7209

Article Info

Received : 11.04.2025

Accepted : 08.07.2025

DOI: 10.21605/cukurovaumfd.1673602

Corresponding Author

Mehmet Fatih ŞAHAN

mfs@adiyaman.edu.tr

Keywords

Climate change

Snow cover depths

Trend analysis

Mann-kendall

Innovative trend analysis

How to cite: ŞAHAN, M.F., YEŞİLGÜL, B., ÜNSAL, İ., (2025). Investigation of Konya Meteorological Snow Cover Depth Trend Using MK and ITA. Çukurova University, Journal of the Faculty of Engineering, 40(3), 501-511.

ABSTRACT

This study investigates the snow cover depths trend in Konya province according to different approaches. For this purpose, trend analysis was carried out using the annual maximum snow cover depths data of 5 different stations in Konya. Trend data analysis at the stations was performed according to the Mann-Kendall and Innovative Sen Methods. In addition, the slope of the trend was found with Sen's Slope Estimator Test. According to the Mann-Kendall method, an increasing trend was observed in the annual maximum snow cover depth recorded in Beyşehir, but no trend was observed in other stations. According to the innovative Sen method, an increasing trend was observed in Konya Airport, Beyşehir, and Cihanbeyli stations, and a decreasing trend was observed in Akşehir and Ereğli. Trends obtained with different approaches reveal that the available TS EN 1991-1-3 standard, which determines snow loads, needs to be re-evaluated for Konya Airport, Beyşehir, and Cihanbeyli.

Konya Meteorolojik Kar Örtüsü Derinlik Yöneliminin MK ve ITA Yöntemleri Kullanılarak Araştırılması

Makale Bilgileri

Geliş : 11.04.2025

Kabul : 08.07.2025

DOI: 10.21605/cukurovaumfd.1673602

Sorumlu Yazar

Mehmet Fatih ŞAHAN

mfs@adiyaman.edu.tr

Anahtar Kelimeler

İklim değişikliği

Kar örtüsü derinliği

Trend analizi

Mann-kendall

Innovative trend analizi

Atf şekli: ŞAHAN, M.F., YEŞİLGÜL, B., ÜNSAL, İ., (2025). Konya Meteorolojik Kar Örtüsü Derinlik Yöneliminin MK ve ITA Yöntemleri Kullanılarak Araştırılması. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 501-511.

ÖZ

Bu çalışmada Konya ilindeki kar örtüsü derinliği yönelimi farklı yaklaşımlara göre araştırılmıştır. Bu amaçla Konya'daki 5 farklı istasyonun yıllık maksimum kar örtüsü derinliği verileri kullanılarak yönelim analizi yapılmıştır. İstasyonlardaki eğilim verisi analizi Mann-Kendall ve Inovative Trend Analizi'ne göre yapılmıştır. Ayrıca eğilimin eğimi Sen'in Eğim Tahmin Testi ile bulunmuştur. Mann-Kendall yöntemine göre Beyşehir'de kaydedilen yıllık maksimum kar örtüsü derinliğinde artış eğilimi gözlenirken, diğer istasyonlarda eğilim gözlenmemiştir. Inovative Trend Analizi'ne göre Konya Havalimanı, Beyşehir ve Cihanbeyli istasyonlarında artış eğilimi, Akşehir ve Ereğli'de azalış eğilimi gözlenmiştir. Farklı yaklaşımlarla elde edilen eğilimler, kar yüklerini belirleyen mevcut TS EN 1991-1-3 standardının Konya Havalimanı, Beyşehir ve Cihanbeyli için yeniden değerlendirilmesi gerektiğini ortaya koymaktadır.

1. INTRODUCTION

In addition to dynamic and other static loads, snow loads also significantly impact the design of steel structures. Building collapses due to unexpected snow loads cause financial losses and threaten human life. As far as followed by the media, such situations are often encountered in our country [1]. However, in Turkey, insufficient statistical information on roof collapses and damages caused by snow accumulation prevents a precise evaluation of the issue. One of the main reasons for roof collapses in buildings is the excessive snow load accumulated on the roof, which is higher than the expectation of the design code. The characteristics of meteorological events in recent years have become a matter of concern for the design of steel structures.

Climate change is experienced in the world due to many factors. Climate change is defined as changes in the average state of the climate over a long period caused by different variables [2-4]. It is vital to know the change in meteorological data over time to reveal the possible consequences of the effects of snowfall on steel structures due to climate change. Changes in meteorological data over the years are determined using methods such as trend analysis with the help of statistical approaches. The trend means an increase or decrease in the values of a time-dependent random series over time [5-6]. Determining the existence of a trend in a time-dependent random series allows us to obtain information about how that series behaves from past to present. Based on the knowledge of the behavior of the series, predictions can be made about the possible future behavior of the variable in the series [7-9]. As a result of trend studies, future increases or decreases in snowfall can be predicted, and various precautions can be taken, or plans can be made to minimize the adverse effects of the possible situation.

One of the main reasons for collapses in roofs is that the snow load values taken into account in the design of steel buildings are smaller than the actual snow loads that occur during snowfall and thus accumulate on the building roof [10]. Especially considering the collapses that occurred in Türkiye in the past years due to very heavy snowfall, it is very important to reconsider and update snow loads to ensure that buildings can resist heavy snowfall [11]. Turkey has different climate types in its geography. For this reason, the snow loads to be used in building designs must be considered according to the current climate conditions of the location where the building will be built. If we want to ensure the safety of buildings in Türkiye, we must update the snow loads in the standards according to the current climate conditions.

Some researchers have presented studies revealing the deficiencies of the standards valid in Türkiye for snow load [1,12,13]. Durmaz and Daloğlu [1] made statistical analyses of snow cover depths obtained from nearly 100 stations in Türkiye. They converted the snow cover depth to snow-water equivalent. Using appropriate distributions, they obtained the ground snow load values for the last 50 years and compared them with those recommended in Turkish codes. Aladağ [12] revealed the deficiencies of the available codes for snow load by creating current snow load calculation values and maps with the help of appropriate probability distributions using the data obtained from the General Directorate of Meteorological Affairs. Türkeş et al. [4] investigated the observed changes in rainfall intensity and anomalies with the data obtained from 111 stations in their study. They used the Mann-Kendall (MK) and Innovative Trend Analysis (ITA) to examine the rainfall trend at the designated observation stations. Baltacı et al. [13] conducted a climatological analysis using daily new snow cover data from 93 meteorological stations over Turkey. Baltacı et al. [14] made a climatic analysis of snowfall events for winter periods in Istanbul. For this purpose, they first obtained the daily snow cover depths for 1971-2006 and then investigated the role of snowfall events under the influence of the Black Sea in the formation of snow bands.

Although snow cover depth has many adverse effects on our social and economic life, there are a limited number of snowfall studies for our country. In contrast, most studies focus on satellite snow cover data [13]. Despite the increasing literature on precipitation trends, drought, and satellite snow cover analyses due to climate change, the number of studies representing the snow cover depth trends of certain parts of our country is insufficient.

This study uses different approaches to determine the annual maximum snow cover depth trends in Konya city center and its four districts.

2. MATERIAL AND METHOD

2.1. Study Area and Data

Our study area, the Konya Closed Basin, has limited water resources and large agricultural lands. Konya has a continental climate. Summers are dry and hot; winters are cold and snowy. The temperature difference between day and night is between 16-22 degrees in summer. In spring and winter, due to humidity, this difference drops to 9-12 °C. Snow stays on the ground for an average of 3 months. Konya is greatly affected by the hot-cold air centers around it. Although it is located in the southernmost part of Central Anatolia, it is colder than other Central Anatolian cities. This is because the central Taurus Mountains completely block the effect of the sea [14].

As the scope of this study, five stations in Konya, established by the General Directorate of Meteorology Affairs, were considered. The stations are in Konya Airport, Akşehir, Beyşehir, Cihanbeyli, and Ereğli districts. They are numbered 17244, 17239, 17242, 17191, and 17248, respectively. Numerical information of the stations is given in Table 1, and their geographical locations are given in Figure 1. The data shows annual maximum snow cover depths.

Table 1. Location of the stations

Location	Station no	Altitude	Coordinate information
Konya Airport	17244	1016	37°59'01.3"N 32°34'26.4"E
Akşehir	17239	1025	38°22'07.7"N 31°25'46.9"E
Beyşehir	17242	1131	37°40'39.7"N 31°44'46.7"E
Cihanbeyli	17191	955	38°39'02.1"N 32°55'18.7"E
Ereğli	17248	1055	37°31'31.8"N 34°02'54.6"E

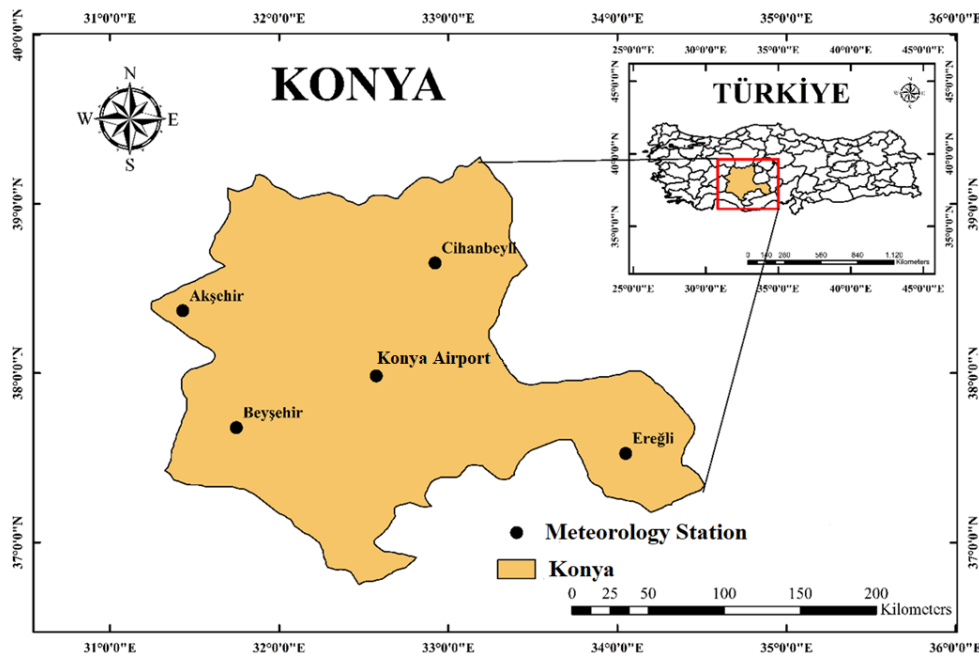


Figure 1. Map locations of stations

Figure 2-6 shows the distribution of meteorological data of Konya Airport, Akşehir, Beyşehir, Cihanbeyli, and Ereğli stations over time and their 5-year average values, respectively. Annual maximum snow cover depth data for the stations examined within the scope of the study were obtained from the General Directorate of Meteorology Affairs. The beginning of the time series for Konya airport is 1922. It is 1941 for Akşehir, 1931 for Beyşehir, and 1964 for Cihanbeyli and Ereğli. The data in the time series continues until 2023.

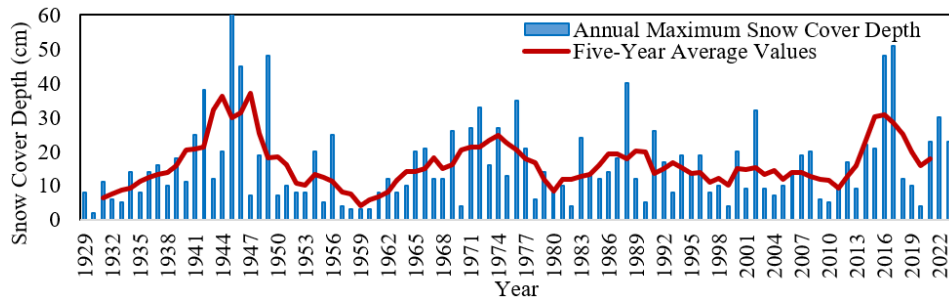


Figure 2. Time series of meteorological data of Konya Airport station

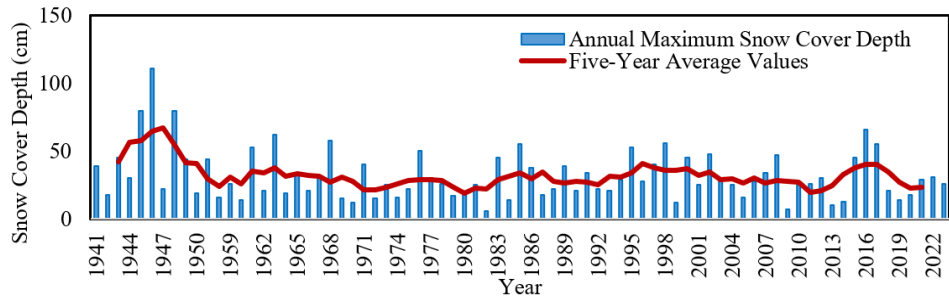


Figure 3. Time series of meteorological data of Akşehir station

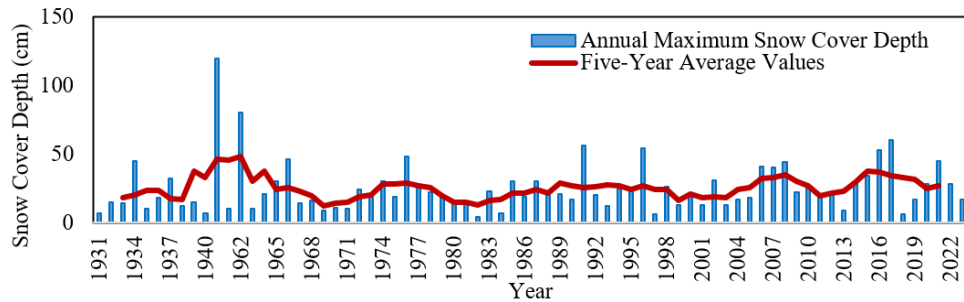


Figure 4. Time series of meteorological data of Beyşehir station

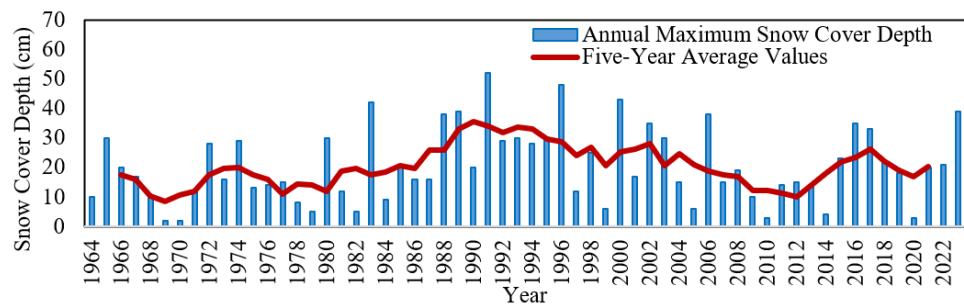


Figure 5. Time series of meteorological data of Cihanbeyli station

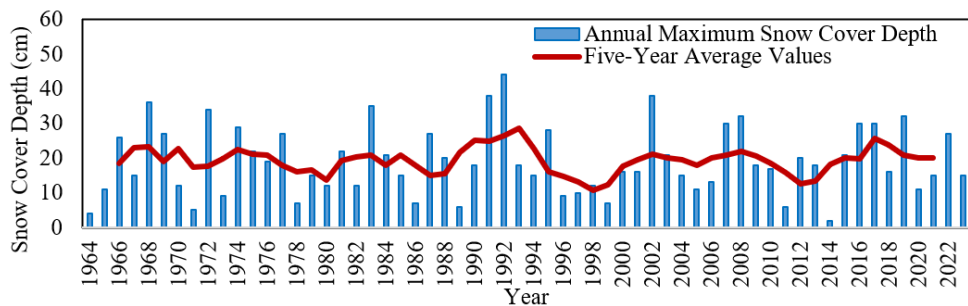


Figure 6. Time series of meteorological data of Ereğli station

It is observed from Figure 2-6 that the highest snow cover depth at Konya airport reached 65 cm in 1945. The highest snow cover depth reached 111 cm in Akşehir in 1946, 120 cm in Beyşehir in 1960, 52 cm in Cihanbeyli in 1991, and 44 cm in 1992 in Ereğli.

2.2. Mann-Kendall Trend Test

The Mann-Kendall Trend Test (MK) is one of the methods of testing the trend when considering a time series. This non-parametric method was developed by Mann [16] and Kendall [17]. The MK test statistic is calculated as follows:

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^n \text{sgn}(x_j - x_k) \quad (1)$$

$$\text{sgn}(x_j - x_k) = \begin{cases} 1 & (x_j - x_k) > 0 \\ 0 & (x_j - x_k) = 0 \\ -1 & (x_j - x_k) < 0 \end{cases} \quad (2)$$

Positive S values indicate an expanding (increasing) trend and a negative S value reveals a decreasing trend in the data time series. Here, n is the length of the dataset, x_j and x_k are the successive information esteems on occasion j and i , and $\text{sgn}(x_j - x_k)$ means the sign capacity that takes on the values 1, 0, or -1. Variance of S with normal distribution and zero mean is calculated by:

$$\text{Var}(S) = \frac{n(n-1)(2n+5)}{18} \quad (3)$$

The equation below is obtained for tests $n > 10$:

$$\text{Var}(S) = \frac{n(n-1)(2n+5) - \sum_i^r t_i(t_i-1)(2t_i+5)}{18} \quad (4)$$

The standardized test statistic Z value, which determines whether the Mann-Kendall test is significant or not, is calculated by utilizing the following equations:

$$Z = \begin{cases} \frac{S-1}{\sqrt{\text{Var}(S)}} & S > 0 \\ 0 & S = 0 \\ \frac{S+1}{\sqrt{\text{Var}(S)}} & S < 0 \end{cases} \quad (5)$$

The Mann-Kendall test determines the significance by comparing Z with the critical Z value. Depending on which α significance level the study is examined, if the $|Z| \leq Z_\alpha$ condition is valid, it is accepted that there is a trend. If there is a trend, it is concluded that there is an increasing trend if the Z sign is (+) and a decreasing trend if it is (-). Since this method does not have to comply with any distribution, it appears useful among researchers [18]. Many researchers have used this approach for their trend studies [19-22] because of this.

2.3. Innovative Trend Analysis

The Innovative Trend Analysis (ITA), developed by Sen [23], is a method that presents the data obtained as a result of the analysis in linear graphic form. The approach differs from other non-parametric trend analysis methods. It gives results even if the data do not follow a normal distribution, the data is small, and there is internal dependence on the data [23].

In this approach, the data time series is divided into two equivalent parts, and each sub-series is organized in ascending order independently. The sub-arrangement's first half (x_i) is put on the x axis, and the second half (x_j) is set on the y axis of a Cartesian coordinate system.

1:1 line is drawn on the graph. There is no trend if the data points in a scatter plot are collected in a straight line 1:1. If the data points accumulate in the triangular area below the 1:1 straight line, the time series shows decreasing trends. If above, it means that the time series shows increasing trend [23]. Another advantage that distinguishes this method from other non-parametric methods is that it allows time series to be divided into three categories: “low”, “medium” and “high” (Figure 7).

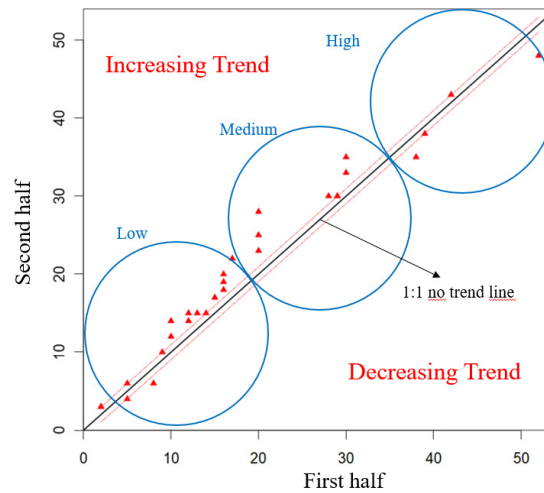


Figure 7. Graphical representation of the ITA

Sen [24] proposed Equation 6. to determine ITA's slope (S). The slope determined by the equation also shows the magnitude of the trend.

$$S = \frac{2(\bar{x} - \bar{y})}{n} \quad (6)$$

The standard deviation of the trend slope and confidence limit (CL) of the trend are calculated in equations 7 and 8, respectively.

$$\sigma_S = \frac{2\sqrt{2}}{n\sqrt{n}} \sigma \sqrt{1 - \rho_{xy}} \quad (7)$$

$$CL_{(n-1)} = 0 \pm S_{kr} \sigma_S \quad (8)$$

Here, n is the length of the dataset, $\bar{x}$ is the average of the first half-time series, and $\bar{y}$ is the average for the second half. ρ_{xy} in Equation (7) refers to the cross-correlation coefficient between the first and second half of the series, and σ_s is the standard deviation of the slope. Equation (8) was created to determine the upper confidence and lower confidence limit values, which are the confidence limits. The term S_{kr} in this equation refers to the critical Z value at the relevant confidence level. If the slope value S of each station is outside the lower and upper confidence limits, it indicates a trend in the time series. A positive S value indicates an increasing trend, while a negative value indicates a decreasing trend [24].

2.4. Sen's Slope Test

Sen [25] presented a non-parametric method to estimate the magnitude of trends in the data time series. In this method, the slope of “ n ” pairs of data ($(T_i) i = 1, 2, \dots, n$) can be estimated by using the following equation:

$$T = \frac{x_j - x_k}{j - k} \quad (9)$$

$$\beta = \begin{cases} T_{\frac{N+1}{2}} & n \text{ (odd)} \\ \frac{1}{2} \left(T_{\frac{x}{2}} + T_{\frac{N+2}{2}} \right) & n \text{ (even)} \end{cases} \quad (10)$$

3. RESULTS AND DISCUSSION

Within the scope of the study, annual maximum snow cover depths taken from 5 different stations in Konya province were considered, and trend analyses were performed using the MK and ITA. The slope was estimated with the help of Sen's Slope Estimator Test.

The MK test was used to evaluate trend analysis at a 0.05 significance level (95% confidence limit). If the absolute value of the standardized Z statistic obtained from the trend analysis was greater than 1.96, the trend was considered statistically significant. If the absolute value of Z was less than 1.96, it was concluded that there was no statistically significant trend.

MK trend analysis results of the annual maximum snow cover depth series at the stations are given in Table 2.

Table 2. Mann-Kendall test results

Stations	S	σ_s	Z	Trend state ($\alpha=5\%$)
Konya Airport	403	93425.67	1.315	None
Akşehir	-158	49620.00	-0.705	None
Beyşehir	471	44015.67	2.240	Statistically significant increasing trend
Cihanbeyli	134	134.00	0.871	None
Ereğli	79	79.00	0.511	None

According to the MK test, a statistically significant increasing trend was detected at the Beyşehir station at the 5% significance level ($Z = 2.240$, $p < 0.05$). Although the trend is upward in the data of Konya Airport, Beyşehir, Cihanbeyli and Ereğli, there is no trend in statistical significance. In the Akşehir data, the trend is decreasing and again no statistically significant trend is observed.

Within the scope of the study, Sen's innovative method was also used to investigate trends. In ITA, the slope value S of each station is compared with the lower and upper confidence limits. If the S value is outside the confidence limits, it is concluded that there is a trend in the time series. Trend analysis results obtained with ITA are given in Table 3. As a result of the ITA test conducted for the data of Konya Airport, Beyşehir, and Cihanbeyli stations, s values were calculated as 0.0095, 0.0753, and 0.0608, respectively. Comparing these slope values with the lower and upper confidence limits resulted in the conclusion that there is an increasing trend, considering that the s values are outside the limits and the s values are positive. s values for Akşehir and Ereğli were calculated as -0.0609, -0.0561, and it was concluded that there was a decreasing trend. Graphical results of the ITA are shown in Figure 8.

Table 3. Innovative trend analysis results based on confidence limit at 95%

Stations	Slope(s)	Indicator	Standart deviation	Correlation	Lower limit	Upper limit	Trend state
Konya Airport	0.0095	0.2807	0.0043	0.9869	-0.0083	0.0083	Increase
Akşehir	-0.0609	-0.6962	0.0178	0.9502	-0.0348	0.0348	Decrease
Beyşehir	0.0753	1.1620	0.0302	0.8701	-0.0592	0.0593	Increase
Cihanbeyli	0.0608	0.9302	0.0108	0.9801	-0.0211	0.0211	Increase
Ereğli	-0.0561	-0.8348	0.0099	0.9741	-0.0194	0.0194	Decrease

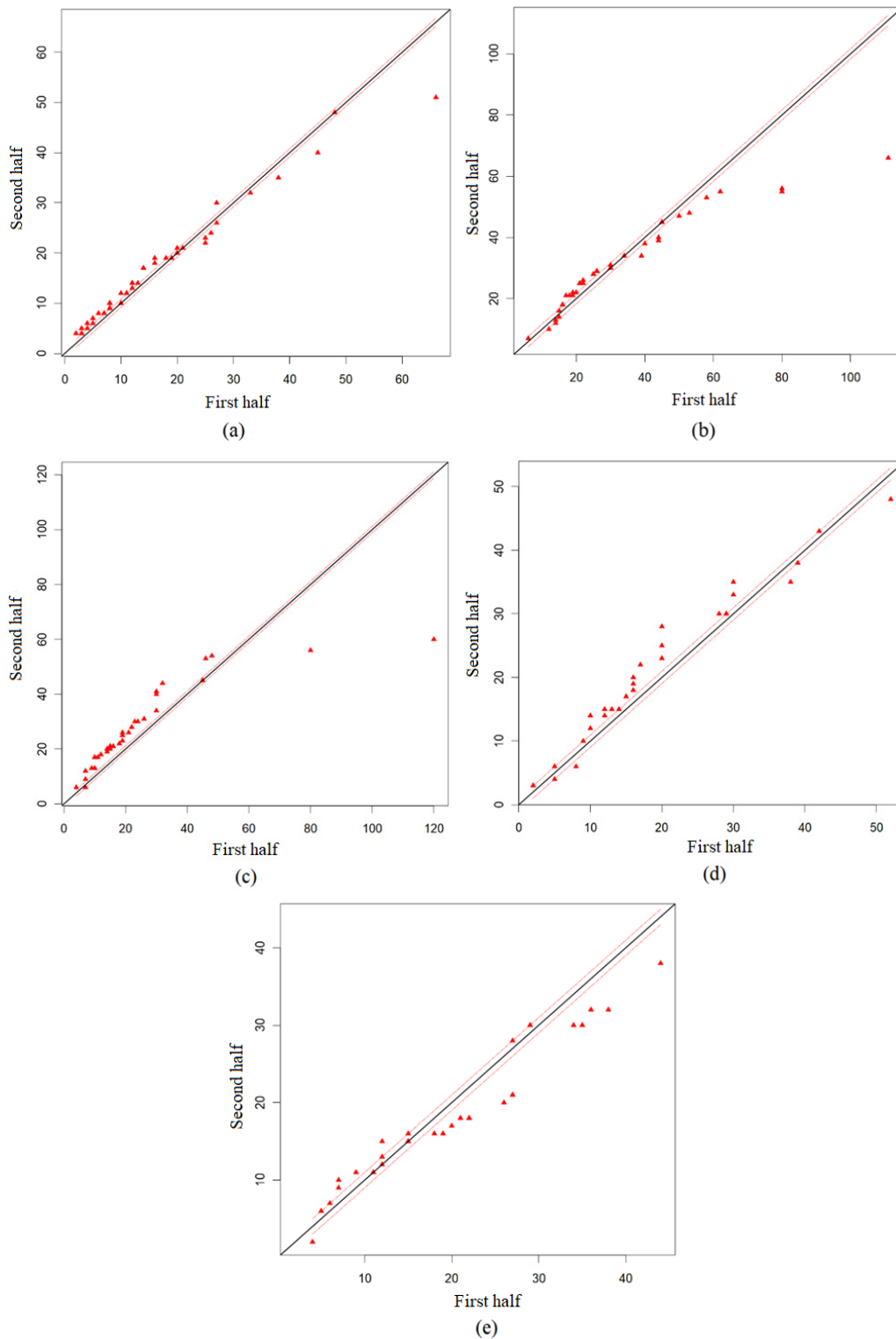


Figure 8. Trend graphs for snow cover depths obtained by the ITA (a) Konya Airport, (b) Akşehir, (c) Beyşehir, (d) Cihanbeyli, (e) Ereğli

Detailed information about the trends can be obtained by examining the graphs in Figure 8. In the ITA chart of Konya Airport station (Figure 8.a), it is understood that the data is mainly located in the upper triangle but in the lower area. This shows that the trend is increasing but at a low level. According to the chart given for Akşehir station (Figure 8.b), however, some of the data is located in the lower area of the upper triangle; most of the data is in the middle and the higher area of the lower triangle. Therefore, it is concluded that

the trend is decreasing. When the graph obtained for Beyşehir (Figure 8.c) is examined, most of the data is located in the lower area of the upper triangle. Thus, it is concluded that the trend is increasing. At Cihanbeyli station (Figure 8.d), almost all of the data is in the upper triangle, scattered in the lower and middle areas. This shows that the trend is increasing. At Ereğli station (Figure 8.e), while the majority of the data is in the lower triangle, it is seen that the data in the lower triangle is in the middle and higher area. According to ITA, a decreasing trend was determined at this station.

A comparison of the results of trend analysis using MK and ITA for the stations within the scope of the study is given in Table 4. The trend analysis conducted for Beyşehir station revealed that the trend was increasing in both methods. While no trend was observed for the other four stations according to MK, an increasing trend is seen in some stations and a decreasing trend in others, according to ITA.

Table 4. Comparison of trend tests

Stations	Man-Kendall test	Inovative Sen test
Konya Airport	None	Increase
Akşehir	None	Decrease
Beyşehir	Statistically significant increasing trend	Increase
Cihanbeyli	None	Increase
Ereğli	None	Decrease

Sen's Slope Estimator test was also performed on the stations' annual snow cover data evaluated within the study's scope (Table 5.). The test results were obtained as 0.0395, -0.0556, 0.1508, 0.0667, and 0.0303 at Konya airport, Akşehir, Beyşehir, Cihanbeyli, and Ereğli stations. As a result of Sen's slope estimator test, it shows that the trend is increasing as the slope values are positive in Konya airport, Beyşehir, Cihanbeyli, and Ereğli. In Akşehir, the slope value is negative, with a decreasing trend.

Table 5. Sen's slope estimator test results

Stations	Sen's slope estimator test
Konya Airport	0.0395
Akşehir	-0.0556
Beyşehir	0.1508
Cihanbeyli	0.0667
Ereğli	0.0303

4. CONCLUSION

The study discusses the annual maximum snow cover depths of Konya Airport, Akşehir, Beyşehir, Cihanbeyli, and Ereğli stations in Konya Province. Mann-Kendall, Innovative Trend Analysis and Sen's Slope estimator test were conducted on these data, and the obtained trend analysis results were compared.

- According to ITA, an increasing trend was determined at Konya Airport station, while no trend was found as a result of MK.
- According to ITA, a decreasing trend was determined at Akşehir station, while no trend was found due to MK.
- An increasing trend was determined at Beyşehir station according to both MK and ITA methods.
- According to ITA, an increasing trend was determined at Cihanbeyli station, while no trend was found due to MK.
- According to ITA, a decreasing trend was determined at Ereğli station, while no trend was found due to MK.
- According to Sen's slope test, it is seen that the slope values in Konya airport, Beyşehir, Cihanbeyli and Ereğli are positive and the trend is increasing, while in Akşehir the slope value is negative and the trend is decreasing.

As a result of all these studies, an increasing trend emerged in both methods at Beyşehir station, while dissimilar results emerged from the methods used for other stations. Although it was concluded that there

was no trend according to the MK approach at Konya Airport, Beyşehir, and Cihanbeyli stations, it was observed that there was an increasing trend according to the ITA approach. Considering this situation, there is a need to re-evaluate the snow loads for Konya Airport, Beyşehir, and Cihanbeyli in the current TS EN 1991-1-3 standard [26], which determines the snow loads in our country.

5. REFERENCES

1. Durmaz, M. & Daloğlu, A. (2005). Kar verilerinin istatistiksel analizi ve Dogu Karadeniz Bölgesinin zemin kar yükü haritasının oluşturulması. *Teknik Dergi*, 2(5), 3619-3642.
2. Aksoy, H., Onoz, B., Cetin, M., Yuce, M.I., Eris, E., Selek, B., Aksu, H., Burgan H.I., Esit, M., Orta, S. & Cavus, Y. (2018). SPI-based drought severity-duration-frequency analysis. *13th International Congress on Advances in Civil Engineering*, Izmir/Turkey.
3. Esit, M., Kumar, S., Pandey, A., Lawrence, D.M., Rangwala, I. & Yeager, S. (2021). Seasonal to multi-year soil moisture drought forecasting. *Npj Climate and Atmospheric Science*, 4(1), 1-8.
4. Türkeş, M. (1996). Spatial and temporal analysis of annual rainfall variations in Turkey. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 16(9), 1057-1076.
5. Yuce, M.I. & Esit, M. (2021). Drought monitoring in Ceyhan Basin, Turkey. *Journal of Applied Water Engineering and Research*, 9(4), 293-314.
6. Topçu, E. (2022). Appraisal of seasonal drought characteristics in Turkey during 1925-2016 with the standardized precipitation index and copula approach. *Natural Hazards*, 112, 697-723.
7. Çetin, M., Aksoy, H., Önöz, B., Eriş, E., Yüce, M.İ., Selek, B., Aksu, H., Burgan, H.İ., Eşit, M., Çavuş, Y. & Orta, S. (2018). Deriving accumulated precipitation deficits from drought severity-duration-frequency curves: A case study in Adana province, Turkey. *1st International Congress on Agricultural Structures and Irrigation*, Antalya/Turkey.
8. Eris, E., Aksoy, H., Onoz, B., Cetin, M., Yuce, M.I., Selek, B., Aksu, H., Burgan, H.I., Esit, M., Yildirim, I. & Karakus, E.U. (2019). Frequency analysis of low flows in intermittent and non-intermittent rivers from hydrological basins in Turkey. *Water Supply*, 19(1), 30-39.
9. Zamani, R., Mirabbasi, R., Nazeri, M., Meshram, S.G. & Ahmadi, F. (2018). Spatio-temporal analysis of daily, seasonal and annual precipitation concentration in Jharkhand state, India. *Stochastic Environmental Research and Risk Assessment*, 32(4), 1085-1097.
10. Durmaz, M. & Daloğlu, A. (2006) Frequency analysis of ground snow data and production of the snow load map using geographic information system for the Eastern Black Sea region of Turkey. *J. Struct. Eng. ASCE*, 132(7), 1166-1177.
11. Korkmaz, H.H., Dere Y., Özkılıç, Y.O., Bozkurt, M.B., Ecemiş, A.S. & Özdoner, N. (2022). Excessive snow induced steel roof failures in Turkey. *Engineering Failure Analysis*, 141, 106661.
12. Aladağ, E. (2019). Meteoroloji haritaları kullanılarak çelik çatıların kar yükü analizleri. *Yüksek Lisans Tezi*. Bilecik Şeyh Edebali Üniversitesi.
13. Baltacı, H., Arslan, H., Akkoyunlu, B.O. ve Gömes, H.B. (2020). Uzun vadeli değişkenlik ve Türkiye'de uzun süreli kış kar yağışı eğilimleri ve telebağlantı modellerinin rolü. *Meteoroloji Uygulamaları*, 27, e1891.
14. Kültür ve Turizm Bakanlığı, (2025). <https://konya.ktb.gov.tr/TR-370534/iklim.html>. Erişim tarihi: 08.05.2025.
15. Baltacı, H., da Silva M.C.L. & Gomes, H.B. (2021). Climatological conditions of the Black Sea-effect snowfall events in Istanbul, Turkey. *International Journal of Climatology*, 41, 2017-2028.
16. Mann, H.B. (1945). Non-parametric tests against trend. *Econometrica*, 13, 163-171.
17. Kendall, M.G. (1975). *Rank correlation methods*. Oxford University Press, New York.
18. Yu, Y., Zou, S. & Whittemore, D. (1993). Nonparametric trend analysis of water quality data of rivers in Kansas. *Journal Hydrology*, 150, 61-80.
19. Gümüş, V., Yenigün, K. ve Kırkgöz, M.S. (2016). Göksu Nehri-Himmetli İstasyonu aylık ortalama akım gidişlerinin belirlenmesi. *Çukurova Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 26(1), 27-36.
20. Ay, M. ve Özyıldırım, S. (2017). Yozgat'ın aylık toplam yağış ve aylık ortalama hava sıcaklıklarının trend analizi. *Çukurova Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 32(2), 65-76.
21. Yüce, M.I., Eşit, M. & Kalaycıoğlu, V. (2023) Investigation of trends in extreme events: a case study of Ceyhan Basin Turkey. *Journal of Applied Water Engineering and Research*, 11(3), 317-332.
22. Akbaş, E., Çelik, R. & Eşit, M. (2023) Assessment of Mann Kendall, Spearman Rho and Innovative Şen methods using meteorological data: a case study of Batman region. *Yüzüncü Yıl University Journal of the Institute of Natural and Applied Sciences*, 28(1), 239-251.

23. Şen, Z. (2012). Innovative trend analysis methodology. *Journal of Hydrological Engineering*, 17(9), 1042-1046.
24. Şen, Z. (2017). Innovative trend significance test and applications. *Theoretical and Applied Climatology*, 127(3-4), 939-947.
25. Sen, P.K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association*, 63, 1379-1389.
26. Türk Standartları Enstitüsü, (2007). *TS EN 1991-1-3 yapılar üzerindeki etkiler -Bölüm 1-3: Genel etkiler – Kar yükleri*, Ankara.

The Effects of Steaming Duration and Temperature on Pilling and Fastness Properties in Vigoureux Printing Technique

Muhammed İdris AKTAŞ^{1,a}, Behzat YILDIRIM^{2,b}, Halil İbrahim İÇOĞLU^{2,c},
Mehmet TOPALBEKİROĞLU^{2,d}

¹Boyar Kimya Inc., Gaziantep, Türkiye

²Gaziantep University, Faculty of Engineering, Department of Textile Engineering, Gaziantep, Türkiye

^aORCID: 0000-0002-7141-7984; ^bORCID: 0000-0002-7787-0595; ^cORCID: 0000-0003-0687-4721;

^dORCID: 0000-0003-4345-8815

Article Info

Received : 06.04.2025

Accepted : 08.08.2025

DOI: 10.21605/cukurovaumfd.1670936

Corresponding Author

Behzat YILDIRIM

behzatyildirim@gantep.edu.tr

Keywords

Vigoureux printing

Steaming

Temperature

Pilling

Fastness

How to cite: AKTAŞ, M.İ., YILDIRIM, B., İÇOĞLU, H.İ., TOPALBEKİROĞLU, M., (2025). The Effects of Steaming Duration and Temperature on Pilling and Fastness Properties in Vigoureux Printing Technique. Cukurova University, Journal of the Faculty of Engineering, 40(3), 513-520.

ABSTRACT

This study examines the impact of steaming duration and temperature on pilling resistance and fastness properties in fabrics printed using the Vigoureux technique. The Vigoureux printing method, which dates back to the 1800s, involves printing patterns on fiber tops, blending these with unprinted fibers, and producing mélange-like yarns used in textile applications. In this study, polyamide and wool fibers printed with acid dyes were subjected to varying steaming durations (50–70 minutes) and temperatures (85–105°C). Subsequently, the yarns were processed into knitted fabrics and subjected to a series of tests to assess their pilling, rubbing, washing, and water fastness properties. Results indicated that increased steaming duration and temperature led to reduced pilling resistance but enhanced fastness properties, likely due to improved dye penetration. The study emphasizes the critical role of steaming parameters in optimizing textile performance, providing insights for enhancing the durability of Vigoureux-printed fabrics.

Vigore Baskı Tekniğinde Buharlama Süresi ve Sıcaklığının Boncuklanma ve Haslık Özelliklerine Etkileri

Makale Bilgileri

Geliş : 06.04.2025

Kabul : 08.08.2025

DOI: 10.21605/cukurovaumfd.1670936

Sorumlu Yazar

Behzat YILDIRIM

behzatyildirim@gantep.edu.tr

Anahtar Kelimeler

Vigore baskı

Buharlama

Sıcaklık

Tüylene

Haslık.

Atf şekli: AKTAŞ, M.İ., YILDIRIM, B., İÇOĞLU, H.İ., TOPALBEKİROĞLU, M., (2025). Vigore Baskı Tekniğinde Buharlama Süresi ve Sıcaklığının Boncuklanma ve Haslık Özelliklerine Etkileri. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 513-520.

ÖZ

Bu çalışmada, vigore tekniği kullanılarak baskı yapılan kumaşlarda buharlama süresi ve sıcaklığının boncuklanma direnci ve haslık özellikleri üzerindeki etkisi incelenmektedir. 1800'lü yıllara dayanan Vigore baskı tekniği, elyafın üzerine desenler basmayı, bunları baskısız elyaflarla karıştırmayı ve tekstil uygulamalarında kullanılan melanj benzeri iplikler üretmeyi amaçlamaktadır. Bu çalışmada, asit boya ile baskı yapılan poliamid ve yün elyafları değişen buharlama sürelerine (50-70 dakika) ve sıcaklıklara (85-105 °C) tabi tutulmuştur. Baskılı elyaflar iplik haline getirilip örüldükten sonra boncuklanma, sürtünme, yıkama ve su haslığı özellikleri değerlendirilmiştir. Buharlama süresi ve sıcaklığının artması boncuklanma direncini azalttığını ancak muhtemelen iyileştirilmiş boya penetrasyonu nedeniyle haslık özelliklerini artırdığını göstermiştir. Bu çalışma vigore baskı tekniğinde buharlama parametresinin kritik rolünü önemini ortaya çıkararak vigore baskılı kumaşların kalitesini ve dayanıklılığını artırmak için öneriler sunmaktadır.

1. INTRODUCTION

One of the methods used to increase the value added in textile products is to create the visual effect techniques such as the surface appearance, colors, and patterns of the fabrics in clothing and home textiles. These techniques are of great interest in the industry [1]. Many methods are used to provide surface appearance effect on fabrics on fabrics, such as printing, dyeing, and mechanical finishing techniques like brushing or sanding, which help to alter the texture, color, and overall visual appeal of the fabric. In addition, different surface appearance effects are given to the fabrics by using many different printing methods in printing technologies. One of the methods that contains limited information in the literature and dates to 1800 years is the Vigoureux printing method. The Industrielle Vigoureux of St. Denis was granted patent rights in 1863, but the potential for extending the work was minimal since the patent reserved the exclusive right to utilize his privileges for the full length of the Grant [2,3]. In this method, the fiber tops are first printed with the desired patterns and colors, then the printed fiber tops are mixed with the white unprinted fiber tops and the yarns produced in ring spinning process have a similar effect mélange yarn and then evaluated in fabric production. Mélange yarns are the fancy yarns, which have fibers or filaments of different colors. They are usually made by mixing slivers of different colours or printed in stripes. The properties of the fibers that make up yarn are crucial for its functionality and appearance, and altering these fibers can enhance the overall quality of the yarn [4-9]. Vigoureux printing technique is used to obtain well blended colors in worsted spinning system. The traditional Vigoureux printing and continuous Vigoureux printing (Fleissner's system and Siro-print system) are the two prominent techniques in Vigoureux printing. [10,11]. This printing technique enhances the aesthetic value of yarn or fabric by generating a natural, vibrant, and multi-tonal pattern. It is widely utilized in the fashion, sportswear, and home textiles industries due to its ability to provide visual complexity while also concealing surface pilling, thus improving the overall durability and appearance of the fabric.

There are few studies on Vigoureux printing system in the literature. Kaimori et al. [12,13] studied on development of a new continuous dyeing and multi-color printing system for slivers to resolve some defects (color bleeding at blotch edge, and multi-color printing etc.) in the Vigoureux printing system. It was stated that the disadvantages in Vigoureux printing system were eliminated with the new system they developed, and multi-color printing, printing without gilling, and clear printed patterns were obtained. Additionally, a crucial assessment for the absorption of dye liquids by slivers was conducted using specially designed apparatus.

Studies were conducted on mélange or Vigoureux printing of wool by Delmenico in 1957. The study covered developing a method for increasing the rate of dye fixation in Vigoureux printing system. The duration for fixing dyes can be shortened from 1 or 2 hours under standard atmospheric conditions to mere minutes by employing the right solvents in the printing mixture and increasing the steam pressure. In experimental studies, the depth of color increased using suitable solvents were achieved with a given amount of dye [14].

Buxtorf and Wiazmitinow described the equipment, and the methods used in the Vigoureux printing process for wool, rayon staple, and polyamide fiber slubbing in their work published in 1953 [15]. Potter's work [2] was on the printing of wool including the mélange or Vigoureux printing of other staple fibers. The mélange process is giving a more intimate blend of colored and uncolored fiber. There are claims that the product has a more pleasant handle, but the primary appeal of this printing style stems from the special effects it generates, which cannot be replicated by blending dyed and undyed wool fibers. Bell et al. [16] studied on the effect of solvents in wool printing. The advantages of using benzyl alcohol in print pastes were highlighted in their work on Vigoureux printing with wool slubbing and standard printing on wool textiles.

There are many steps involved in Vigoureux printings. Single or multi fiber tops printing in this process can be done as a rope shown in Figure 1. The printing process is done with pressure rollers on tops. Usually diagonal and diagonal striped patterns are printed. Printing can be done in the desired number of colors. However, printing in more than two colors is not common. After the printing process, the printed tops first go through the pre-drying unit, which is located on the Vigoureux printing machine, and then they are placed inside the steaming machine. Steaming is done in the machine for 90 or 105 min at 100, 105 °C. After steaming, fiber tops are dried in a dryer, the fibers are transferred into tops again, yarn is obtained in the ring spinning system and fabric is produced by a flat knitting machine. Vigoureux printing process is

mostly applied to wool, polyester and polyamide tops. Especially, it is more preferred in worsted wool tops [17].

Considering the critical role of the steaming stage in the Vigoureux printing process and the limited number of studies conducted on Vigoureux printing, this study aimed to investigate the impact of varying durations and temperature levels during the steaming stage on the pilling performance and some important fastness values of wool and polyamide fibers printed by acid dyes using the vigorous printing method.

2. MATERIAL AND METHOD

2.1. Material

In this study, 3.3 dtex 100% polyamide fiber and 26.5 micron 100% wool fiber were used in the tops form. The staple length of wool and polyamide are 54.3 ± 7 mm and 56.8 ± 4 mm, respectively. The C.I. Acid Red 414, C.I. Acid Yellow 116, and C.I. Acid Blue 193 were used in Vigoureux printing process for both fiber types.

2.2. Printing Process

The Vigoureux printing process was given in Figure 1.

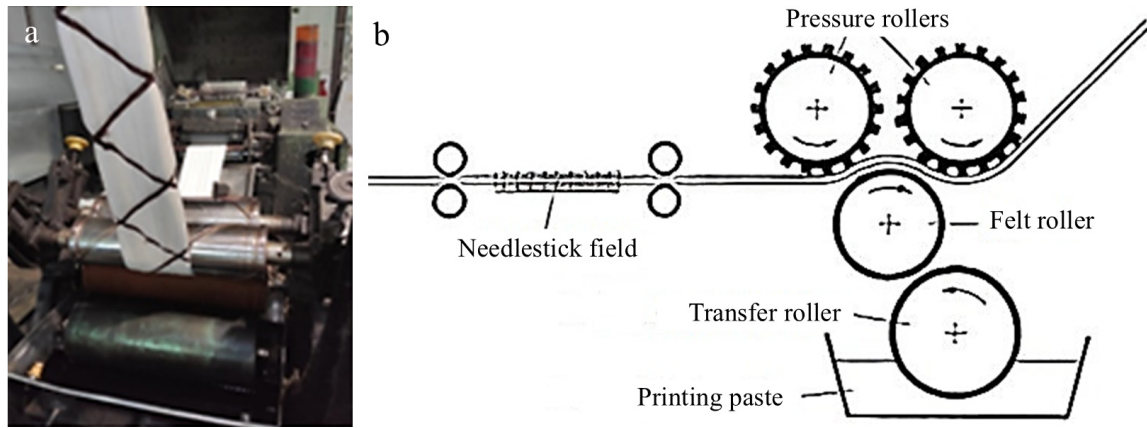


Figure 1. a) Vigoureux printing machine, and b) its components

After printing, the steaming process was applied to fiber samples via a HT batch steamer (Figure 2a). The fiber samples entered the steaming machine in bags (tulles) (Figure 2b). These bags were made of fine tulle and the total weight of the printed fiber was 4 kg. The reason for using tulles in steamers was to achieve complete dye penetration into the fiber by aiding the steaming process.



Figure 2. a) HT steamer, and b) tulles (bags) for printed fiber tops

In a steaming process, following the initial 3 minutes during which water is taken into the chamber, the temperature rises to 98 °C over the subsequent 10 minutes. The following 2 min are dedicated to the process of bluffing, which involves the removal of excess air from the chamber. Over the subsequent 5 min, the temperature is elevated to 105 °C for a period of 60 min. In this study, the duration and temperature at steaming step were varied. Applied temperature and duration values, and sample coding system were given in Table 1.

Table 1. Applied temperature and duration values, and sample coding system

Sample code	Fiber type	Temperature (°C)	Duration (min)
P85-50	Polyamide	85	50
P85-60			60
P85-70			70
P95-50		95	50
P95-60			60
P95-70			70
P105-50		105	50
P105-60			60
P105-70			70
W85-50	Wool	85	50
W85-60			60
W85-70			70
W95-50		95	50
W95-60			60
W95-70			70
W105-50		105	50
W105-60			60
W105-70			70

2.3. Knitted Fabric Production Process

Sample yarns with a yarn count of Nm 35/1 and a twist of 400 T/m were spun from printed polyamide and wool fiber tops using ring spinning machines with a feed speed of 24 m/min. The yarn parameters were selected in accordance with the fiber fineness and length to ensure optimal compatibility and performance. Yarn production parameters for polyamide and wool are total draft 28-31, twist 500-600 TPM, spindle speed 10,000 rpm and twist coefficient 134 μ m. Then, the plain knit fabrics were produced on Brother KH-864 manual flat weft-knitting machine with 5E fineness for pilling and fastness tests.

2.4. Characterization

The printed samples in fabric form were subjected to pilling resistance, rubbing fastness, washing fastness, and water fastness tests according to TS EN ISO 12945-1, TS EN ISO 105-X12, TS EN ISO 105-C06 (A2S), TS EN ISO 105-E01, respectively. The multifiber strip fabric was utilized as an adjacent material during washing and water fastness tests. The gray scale was employed for staining.

The results were subjected to statistical analysis of variance (ANOVA) to evaluate the effects of steaming temperature and duration on pilling and selected fastness properties using Design Expert 13.0 package program. The results were statistically significant at $p \leq 0.05$.

3. RESULTS AND DISCUSSION

3.1. Characterization

The pilling test results of the wool and the polyamide samples according to steaming temperature and duration are given in Figure 3.

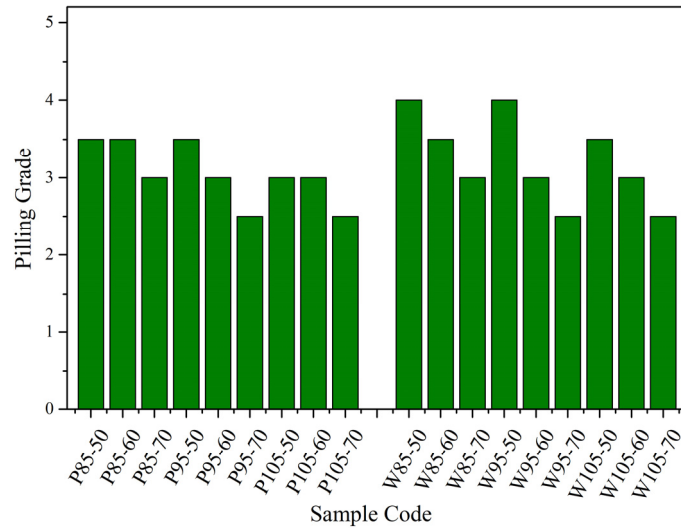


Figure 3. Pilling box test results

In pilling performance evaluation, a rating of 1 denotes a high degree of pilling, whereas a rating of 5 indicates the absence of any observable pilling on the fabric surface. The pilling performance was decreased with increasing of steaming duration for both wool and polyamide samples from 50 min to 70 min. The pilling performance was also decreased with increasing of steaming temperature for both wool and polyamide samples from 85 °C to 105 °C. This phenomenon can be attributed to the fact that elevated temperatures and extended periods of time result in greater damage to the fibers that comprise yarn structures. Because polyamide and wool fibers demonstrate distinct thermal and moisture sensitivities. Wool is prone to structural swelling and weakening under heat and humidity, while polyamide tends to undergo surface deformation due to softening at elevated temperatures. The effect of temperature ($p=0.0035$) and duration ($p<0.0001$) of the steaming process on the pilling performance were statistically significant.

3.2. Rubbing Fastness Test Results

Dry and wet rubbing fastness test results of the wool and the polyamide samples according to steaming temperature and duration are given in Table 2.

Table 2. Rubbing fastness test results

Sample code	Dry rubbing fastness	Wet rubbing fastness
P85-50	3-4	2-3
P85-60	4	3
P85-70	4	3-4
P95-50	4	3
P95-60	4	3
P95-70	4-5	3-4
P105-50	4	3-4
P105-60	4-5	4
P105-70	4-5	4
W85-50	4	3
W85-60	4-5	3-4
W85-70	4-5	3-4
W95-50	4	3-4
W95-60	4-5	4
W95-70	4-5	4
W105-50	4-5	4
W105-60	4-5	4-5
W105-70	4-5	4-5

The dry and wet rubbing fastness properties were generally increased with increasing of steaming duration for both wool and polyamide samples from 50 min to 70 min. Those properties were also increased with increasing of steaming temperature for both wool and polyamide samples from 85 °C to 105 °C. This may be related to the fact that raising the temperature accelerates diffusion, especially into the fiber. Longer duration also enhanced this. The effect of temperature ($p=0.0427$) and duration ($p=0.0108$) of the steaming process on the dry rubbing were statistically significant. Also, the effect of temperature ($p<0.0001$) and duration ($p=0.0012$) of the steaming process on the wet rubbing were statistically significant.

3.3. Washing Fastness Test Results

Washing fastness staining test results of the wool and the polyamide samples according to steaming temperature and duration are given in Table 3.

Table 3. Washing fastness test results

Sample code	Staining					
	CA	CO	PA	PES	PAN	WO
P85-50	1-2	2	2	2	2	2
P85-60	1-2	3	2	3	3	2-3
P85-70	2	3-4	2-3	4	4	3
P95-50	4	4	4	4	4	4
P95-60	4-5	4-5	4	4-5	4-5	4
P95-70	4-5	4-5	4-5	4-5	4-5	4-5
P105-50	4	4	4	4	4	4
P105-60	4-5	4-5	4-5	4-5	4	4-5
P105-70	4-5	4-5	4-5	4-5	4-5	4-5
W85-50	2	2	2	2	2	2
W85-60	2-3	3-4	2-3	3-4	3-4	2
W85-70	2-3	3-4	2-3	4	4	3
W95-50	4-5	4	4	4	4	4-5
W95-60	4-5	4	4-5	4-5	4-5	4-5
W95-70	4-5	4-5	4-5	4	4-5	4-5
W105-50	4-5	4-5	4	4	4	4
W105-60	4-5	4-5	4-5	4-5	4-5	4-5
W105-70	4-5	4-5	4-5	4-5	4	4-5

CA: Cellulose Acetate, CO: Cotton, PA: Polyamide, PES: Polyester, PAN: Acrylic, and WO: Wool.

The washing fastness staining performance was generally increased with increasing of steaming duration for both wool and polyamide samples from 50 min to 70 min. It was also increased with increasing of steaming temperature for both wool and polyamide samples from 85 °C to 105 °C. Raising the temperature speeds up diffusion into fibers resulting in higher washing fastness values. The effect of temperature ($p<0.0001$) and duration ($p<0.0001$) of the steaming process on the dry rubbing were statistically significant.

3.4. Washing Fastness Test Results

Water fastness staining test results of the wool and the polyamide samples according to steaming temperature and duration are given in Table 4.

Table 4. Water fastness test results

Sample code	Staining					
	CA	CO	PA	PES	PAN	WO
P85-50	1-2	1-2	2-3	3	1-2	2
P85-60	2	2	3	3-4	2	2-3
P85-70	2-3	2-3	3-4	4	3	3-4
P95-50	4	4	4	4	4	4
P95-60	4	4	4-5	4-5	4-5	4
P95-70	4-5	4-5	4	4-5	4-5	4-5
P105-50	4	4	4	4-5	4	4
P105-60	4-5	4-5	4-5	4-5	4-5	4-5
P105-70	4-5	4-5	4-5	4-5	4-5	4-5
W85-50	2-3	2	3	3-4	1-2	2
W85-60	2-3	2-3	3	4	2	2-3
W85-70	2-3	3	3-4	4	2-3	2-3
W95-50	4	4-5	4-5	4-5	4	4
W95-60	4-5	4	4-5	4-5	4-5	4-5
W95-70	4-5	4-5	4-5	4-5	4-5	4-5
W105-50	4-5	4-5	4	4-5	4-5	4-5
W105-60	4-5	4-5	4-5	4-5	4-5	4-5
W105-70	4-5	4-5	4-5	4-5	4-5	4-5

CA: Cellulose Acetate, CO: Cotton, PA: Polyamide, PES: Polyester, PAN: Acrylic, and WO: Wool.

The water fastness staining performance were increased with increasing of steaming duration for polyamide samples from 50 min to 70 min. It was also increased with increasing of steaming temperature for both wool and polyamide samples from 85 °C to 105 °C. The increase in temperature accelerates the diffusion into the fibers and results in higher water fastness values. The effect of temperature ($p<0.0001$) and duration ($p=0.0023$) of the steaming process on the dry rubbing were statistically significant.

4. CONCLUSION

This study examined the impact of steaming duration and temperature on the pilling resistance and various fastness properties of polyamide and wool fibers printed using the Vigoureux method with acid dyes. The results demonstrated that increased steaming temperature and duration tend to reduce pilling resistance, attributed to the fiber damage induced by prolonged heat exposure. However, these conditions significantly improved rubbing, washing, and water fastness properties due to enhanced dye diffusion and fixation within the fibers. The findings reveal that optimizing steaming parameters is crucial for balancing durability and aesthetic qualities in Vigoureux-printed fabrics. While higher temperatures and longer steaming durations may be necessary to achieve desired fastness levels, these settings should be adjusted carefully to mitigate adverse effects on pilling resistance. Upon examination of the test results in general, it was observed that the test results obtained at 105 °C were similar to those obtained at 95 °C. Based on these findings and considering the energy consumption, it can be posited that steaming conditions could be selected as 95 °C and 60 minutes. This study contributes to a deeper understanding of the Vigoureux printing process, providing valuable insights for textile manufacturers aiming to improve the quality and longevity of printed fabrics through controlled steaming.

5. REFERENCES

1. Mezarcıöz, S. & Sırlıbaş, S. (2021). Denim kumaşlara çevre dostu ön terbiyesiz boyama metodunun uygulanması. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 36(4), 1127-1136.
2. Potter, J.A. (1955). The printing of wool, including the mélange or vigoureux printing of other staple fibers. *Journal of the Society of Dyers and Colourists*, 71, 645-652.
3. Ostersetzer, A. (1892). Währungswechsel und aufnahme der baarzählungen, uebersichtliche darstellung der valutafrage. *Manz'sche Buchhandlung*, 1-93.
4. Hanumanth Naik, H.S. (2008). Production of melange yans - A review. *Man-made Textiles in India*, 161-165.

5. Gharehaghaji, A.A., Tavanaie, H. & Karim, S.K. (2007). Study on the interactions between melange yarn properties and fiber damage. *Fibers&Textiles in East Europe*, 15(3), 63-67.
6. Behera B.K., Hari, P.K., Bansal, S. & Singh, R. (1997). Effect of different blending methods and blending stages on properties of melange yarns. *Indian journal of Fiber and Textile Research*, 22, 84-88.
7. Ray S., Ghosh A. & Banerjee D. (2018). Analyzing the effect of spinning process variables on blow room blended cotton melange yarn quality. *Research Journal of Textile and Apparel*, 22(1), 2-14.
8. Wickramasinghe, G. & Dolawatte, S.U. (2020). Effect of intermingling process parameters on polyester/nylon melange fabric appearance. *Journal of Engineered Fibers and Fabrics*, 15, 1-12.
9. Özgün, S., Baykal, P.D. & Özkan, İ. (2025). Farklı bölge pamuklarından üretilen mamul örme kumaşların konfor özelliklerinin araştırılması. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 40(2), 251-263.
10. Fleissner Inc., (2025). *Continuous finishing lines for wool*. <https://www.thefreelibrary.com/Fleissner.-a019376455>, Erişim tarihi: 15.04.2025.
11. IWS Textile Technology, (1990). *Siroprint continuous vigoureux printing*.
12. Kaimori, M., Nobukuni, K. & Kitamura, K. (1995). Development of continuous dyeing and multi-color printing system for slivers part 1: development of continuous multi-color printing equipment. *Journal of the Textile Machinery Society of Japan*, 41(3), 75-79.
13. Kaimori, M., Nobukuni, K. & Kitamura, K. (1995). Development of continuous dyeing and multi-color printing system for slivers part 2: Results of fundamental test of suitable condition for dye liquid sorption. *Journal of the Textile Machinery Society of Japan*, 41(4), 109-115.
14. Delmenica, J. (1957). Melange or vigoureux printing of wool: A method for increasing the rate of dye fixation. *Textile Research Journal*, 27(11), 899-912.
15. Buxtorf, F. & Wiazmitinow, A. (1953). Melange or vigoureux printing. *Journal of the Society of Dyers and Colourists*, 69, 550-554.
16. Bell, J.W., Smith, P.J. & Stevens, C.B. (1963). A study of the effect of solvents in wool printing. *Journal of the Society of Dyers and Colourists*, 79, 305-312.
17. Tops (Vigore) Baskı, (2011). 542TGD59, T.C. Milli Eğitim Bakanlığı, Ankara.



1x2 Dizilimli Grup Helisel Kazıklarda Mesafe Etkisinin Laboratuvar Model Deneyleriyle İncelenmesi

Burak TİNTAŞ^{1,a}, Murat ÖRNEK^{1,b}

¹İskenderun Teknik Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, İnşaat Mühendisliği Bölümü, Hatay, Türkiye

^aORCID: 0009-0009-6427-9665; ^bORCID: 0000-0002-0809-2531

Makale Bilgileri

Geliş : 06.05.2025

Kabul : 18.07.2025

DOI: 10.21605/cukurovaumfd.1692961

Sorumlu Yazar

Burak TİNTAŞ

buraktintas.lee22@iste.edu.tr

Anahtar Kelimeler

Grup helisel kazık

Taşıma gücü

Kazıklar arası mesafe

Laboratuvar deneyleri

Atf şekli: TİNTAŞ, B., ÖRNEK, M., (2025). 1x2 Dizilimli Grup Helisel Kazıklarda Mesafe Etkisinin Laboratuvar Model Deneyleriyle İncelenmesi. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 521-530.

ÖZ

Helisel kazıklar derin temel uygulamalarında popülerliği gittikçe artan sistemlerdir. Bu çalışmada 1x2 diziliminde düzenlenmiş olan helisel kazık grubunda, kazıklar arası mesafenin ve kazıktaki helis varlığının, gevşek kum zeminde, basınç etkisi altında taşıma gücüne etkisi laboratuvar model deneyleri kullanılarak incelenmiş ve yük-deplasman eğrileri yorumlanmıştır. Deneylerde helis çapları ve sayısı sabit tutulmuştur. Kazıklar arası mesafe 200 mm'den 700 mm'ye kadar arttırılmış ve her aralık için yük deplasman grafikleri elde edilmiştir. Helisli ve helissiz kazıklar, kazıklar arası mesafe sabit tutularak kendi aralarında yorumlanmıştır. Deney sonuçlarından elde edilen yük-deplasman eğrilerinin nihai göçme yükünü bulmak için %5D ve %10D limit deplasman kriterleri uygulanmıştır. Sonuçlar, kazıklar arası mesafenin artmasının taşıma gücünde azalmaya sebep olduğunu, helis varlığının ise taşıma gücünü arttırdığını göstermiştir.

Analysis of the Spacing Effect in 1x2 Configured Group Helical Piles using Laboratory Model Tests

Article Info

Received : 06.05.2025

Accepted : 18.07.2025

DOI: 10.21605/cukurovaumfd.1692961

Corresponding Author

Burak TİNTAŞ

buraktintas.lee22@iste.edu.tr

Keywords

Group helical pile

Bearing capacity

Pile spacing

Laboratory tests

How to cite: TİNTAŞ, B., ÖRNEK, M., (2025). Analysis of the Spacing Effect in 1x2 Configured Group Helical Piles using Laboratory Model Tests. Çukurova University, Journal of the Faculty of Engineering, 40(3), 521-530.

ABSTRACT

Helical piles are increasingly popular systems in deep foundation applications. In this study, the effect of the distance between piles and the presence of helices on the load-bearing capacity under pressure in loose sand, within a group of helical piles arranged in a 1x2 configuration, was investigated using laboratory model tests, and the load-displacement curves were interpreted. In the experiments, the helix diameters and numbers were kept constant. The distance between the piles was increased from 200 mm to 700 mm, and load-displacement graphs were obtained for each interval. Piles with helix and without helix were interpreted among themselves while keeping the distance between the piles constant. To find the ultimate failure load from the load-displacement curves obtained from the test results, the %5D and %10D limit displacement criteria were applied. The results showed that the increase in the distance between the piles caused a decrease in load-bearing capacity, while the presence of the helix increased the load-bearing capacity.

1. GİRİŞ

Kazık sistemleri genel olarak zemin taşıma kapasitelerinin artırılması ve oturma miktarlarının azaltılması amacıyla kullanılan ve de özellikle taşıma kapasitesi zayıf zeminlerde fazlaca tercih edilen sistemlerdir. Bir veya daha fazla helisel taşıma plakasına sahip çelik kazık türleri de helisel kazık olarak tarif edilmektedir. Helisel kazıklar, eksenel basınç altında yüksek performans sergileyen alternatif bir temel sistemidir [1]. Helisel kazıklar temellerde, yüksek basınç ve çekme kapasiteleri, hızlı kurulum hızı, zorlu zeminlerde kullanım kolaylığı ve maliyetlerin etkin kullanılması gibi avantajlarla öne çıkmıştır [2]. Helisel kazıklar genel olarak uç ve uzatma bölgelerinden oluşur. Uzatma bölgesi şaft kısmını da içerir. Uç bölgesinde ise helisel taşıma plakaları yer alır [3]. Birçok araştırmacı helisel kazıklar üzerine çalışmalarda bulunmuştur [4-15].

Alwalan ve Alnuaim [16] tarafından yapılan çalışmada, kum zemine yerleştirilmiş farklı geometri ve konfigürasyonlardaki helisel kazık gruplarının davranışını değerlendirmek için sayısal modelleme yoluyla bir dizi analiz yapılmıştır. Bu analizlere göre kazıklar arası mesafe S ve helis çapı D 'nin oranları; $S/D \leq 4$ olduğunda helisel kazık grubu, helis bölgesinde oluşan derin bir toprak bloğuyla birlikte bütünsel davranış sergilerken; $S/D > 4$ durumunda kazıklar bireysel davranmaya başlamış ve geçiş noktası $S/D = 4$ olarak belirlenmiştir. $S/D > 5$ olduğunda ise grup içindeki kazık gömülme oranının helis çapına oranı L/D_h etkisi ihmal edilebilir düzeye ulaşmıştır. Eksenel basınç yükleri altındaki helisel kazık gruplarının performansını değerlendirmek için Elsherbiny ve El Naggar [17] tarafından helisel kazıklar üzerinde kapsamlı bir sonlu elemanlar analizi (FEA) çalışması yapılmıştır. Parametrik analizler, performansın esas olarak kazıklar arası mesafeye bağlı olduğunu, helisler arası mesafenin ise sınırlı etkide kaldığını göstermiştir. Lanyi-Bennett ve Deng [18] tarafından yapılan çalışmada buzul kil zemine gömülü eksenel basınç yükü altındaki tek ve 2x2 grup helisel kazıklar üzerinde saha deneyleri gerçekleştirilmiştir. Grup kazık davranışında ayrı kazık göçmesinin baskın olduğu ve grup veriminin düşük aralıklarda azaldığı belirlenmiştir. Mittal ve Mukherjee [19] tarafından yapılan çalışma eksenel yüklerin etkisi altındaki farklı sayıda helisel plakalı ankrajların (1, 2, 3 ve 4) performansının deneysel araştırmasını içermektedir. Ankraj sayısı arttıkça taşıma kapasitesinde artış gözlenmiş, ancak artış oranı daha yüksek değerlerde azalmaya başlamıştır.

Nowkandeh ve Choobbasti [20] tarafından sonlu elemanlar (FE) modellemesi kullanılarak hem kumlu hem de killi zeminde tek ve grup helisel kazıkların eksenel basınç davranışının kapsamlı bir araştırması yapılmıştır. Çalışma, zemin tipine bağlı olarak helisel kazıklar arasında önerilen minimum aralığın farklılık gösterdiğini ortaya koymuştur. Shaheen ve Demars [21] tarafından yapılan çalışmada granüler zemine gömülü helisel ankraj grubunun davranışı incelenmiştir. Sıkı kumda grup davranışını etkileyen kritik aralık belirlenebilmişken, gevşek kumda bu aralık bölgesel kayma nedeniyle net tanımlanamamıştır. Vignesh ve Muthukumar [22] tarafından yapılan çalışmada, yumuşak kil zemine yerleştirilen grup helisel kazıkların çekme ve yanal davranışları sonlu elemanlar modeli kullanılarak kapsamlı bir şekilde analiz edilmiştir. Çalışma, grup verimliliğinin sadece geometriye değil, aynı zamanda yükleme kriterlerine ve kazık düzenine de bağlı olduğunu göstermiştir.

Bak ve arkadaşları [23] tarafından yapılan çalışmada eksenel basınç yükü altında model grup ve tekil helisel kazıkların kum zemin davranışları incelenmiştir. Grup helisel kazık deneyleri, değişken helis aralığı (d , $3d$ ve şaft çapı) ve kazıklar arası mesafe ($3D$, $5D$, $6D$, $9D$, $12D$ ve $15D$) şeklinde gerçekleştirilmiştir. Grup helisel kazık deneylerinde $1d$ ve $3d$ helis aralığındaki kazıkların taşıma gücü yüklerinin $5D$ ve $6D$ kazık aralığında maksimum noktaya ulaştığı gözlenmiştir. Elsherbiny [24] tarafından yapılan çalışmada helisel kazık gruplarının performans özellikleri eksenel basınç ve yanal yüklere maruz iken birbirinden bağımsız olarak değerlendirilmiştir. Kum zemin ve kil zeminden oluşan iki farklı zeminde gerçekleştirilen beş adet eksenel basınç ve üç adet yanal yük deneyleri tam ölçekli şekilde yapılmış ve sonuçlar modellenmiştir. Çalışmada kazık yerleşimleri, helis çapının 1, 2 ve 3 katı kadar helisler arası mesafeye sahip çift helisli kazıklar ile tekil ve helis çapının 2, 3, 4, 5 ve 10 katında olmak üzere beş farklı kazık aralığına sahip grup kazıklar olarak seçildiği görülmektedir. Helisel kazıklar için tipik grup aralıklarında grup etkisinin ihmal edilebilir düzeyde olduğu tespit edilmiştir. Mendoza ve arkadaşları [25] tarafından yapılan çalışmada Brezilya Federal Bölgesi'nin tipik zemininde kurulan helisel vidalı kazıklarla inşa edilen standart grupların ve kazıklı radyelerin davranışı incelenmiştir. Deney sonuçları sonlu elemanlar yöntemine dayanan Abaqus yazılımı ile modellenmiştir. Grup içi kazıklar arasında gerilme-soğanı çakışması sınırlı düzeyde olup, sistem genelinde bireysel kazık davranışı ön plana çıkmıştır. Mukherjee ve Mittal [26] tarafından kum zeminde aynı seviyede yerleştirilmiş grup helisel ankrajların davranışlarını incelemek için bir deney serisi araştırmalar sunulmuştur. Deneyler farklı gömülme derinliklerinde gerçekleştirilmiş ve sığ ve derin temel

davranış biçimleri gözlemlenmiştir. Nihai çekme kapasitesinin, ankrajın montaj derinliği ve ankraj sayısı ile önemli ölçüde değiştiği gözlenmiştir. Ayrıca Plaxis 3D sonlu elemanlar yöntemi ile model deney sonuçları doğrulanmış ve mükemmel bir uyum sağladığı belirtilmiştir.

Sakr ve arkadaşları [27] tarafından yapılan çalışmada çekme yüklemesi altında toplam 60 adet model deney, grup ve tekil helisel kazıklar üzerinde gerçekleştirilmiştir. Çalışmada grup etkinliği, farklı sayıda kazık ve kazık aralığında sabit helis aralığı, gömülme derinliği, helis sayısı ve helis çapı gibi farklı parametrelerle değerlendirilmiştir. Ayrıca kazık gruplarının verimliliği etkisinin, zemin yoğunluğuna önemli ölçüde bağlı olduğu gözlenmiştir. Sonuç olarak grupta kazık aralığı, kazık sayısı arttıkça grup etkinliğinin de arttığı ve kum bağlı yoğunluğu arttıkça grup etkinliğinin düştüğü belirlenmiştir.

Literatür verileri grup helisel kazıklarda kazık aralığının incelendiği çalışmaların oldukça sınırlı sayıda olduğunu göstermiştir. Bu çalışma kapsamında büyük bir veri setinde kazık aralığının incelenmiş olması ve kazıklarda taşıma gücünün farklı yöntemlerle bulunmuş olmasının, çalışmanın özgün değerini ortaya çıkardığı düşünülmektedir.

Literatürde, grup helisel kazıkların taşıma kapasitesine etki eden en önemli parametrelerinin kazıkta helis varlığı ve kazıklar arası mesafe olduğu görülmektedir. Bu çalışmada, sadece 1x2 dizilimindeki grup helisel kazıklar için helisli ve helissiz sistemlerin ve kazıklar arası mesafelerin, taşıma kapasitesi üzerindeki etkisi laboratuvar model deneyleri ile incelenmiştir. Bu doğrultuda toplam 9 adet model deney yapılmıştır. Deneylerde, kazıklar arası mesafe sabit 20cm 30cm ve 40cm iken helis varlığının etkisi ve kazıklar arası mesafeler 20 cm'den 70 cm'ye kadar arttırılmış ve taşıma kapasitesindeki değişim yorumlanmıştır.

2. MATERYAL VE YÖNTEM

Bu çalışmada deneyler İskenderun Teknik Üniversitesi Mühendislik ve Doğa Bilimleri Fakültesi Geoteknik Laboratuvarı'nda yapılmıştır. Deney için 1500mm x 1200mm x 1000 mm boyutlarında Şekil 1'de görüldüğü üzere metal kum deney havuzu kullanılmıştır.



Şekil 1. Kum havuzu ve helezonik sistem

Çalışma esnasında kullanılan kum taş ocağından temin edilmiş ve laboratuvar ortamında 0.18 mm ile 2.00 mm arasında ölçülen özel eleklerden geçirilerek işlenmiştir. Bu kum zeminin özellikleri ve parametreleri Çizelge 1'de gösterilmiştir. Burada D_{30} danelerin %30'unun, D_{60} ise danelerin %60'ının geçtiği elek çapını ifade etmektedir.

Çizelge 1. Kum zemin parametreleri

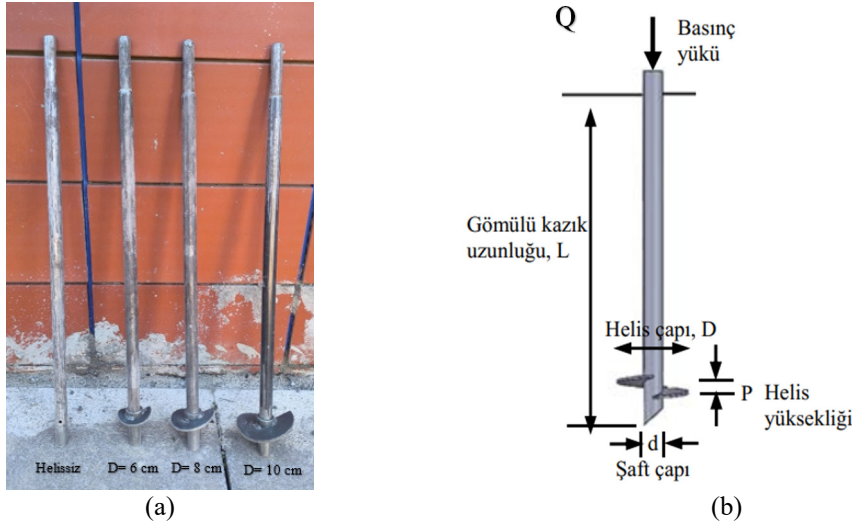
Parametreler	Değerler
Efektif dane çapı, D_{10} (mm)	0.260
D_{30} (mm)	0.441
D_{60} (mm)	0.775
Üniformluk katsayısı, C_u	2.972
Süreklilik katsayısı, C_r	0.962
Zemin sınıfı, (USCS)	SP

Rölatif sıklık, zeminin mevcut sıklığının en gevşek ve en sıkı durumları arasındaki konumunu gösteren, genellikle yüzde (%) cinsinden ifade edilen bir parametredir. Temel olarak, zeminin doğal hâlindeki sıklığı veya gevşekliliği hakkında bilgi verir. Deneylerde kullanılan kumun rölatif sıklığı %30–35 aralığında olup bu değer, ASTM D4253 ve D4254 standartlarına uygun olarak belirlenmiştir. Bu değer aralığı, zeminin gevşek sınıfta olduğunu göstermektedir. Kohezyonsuz zeminlerde rölatif sıklık değerlerine göre yapılan sınıflama Çizelge 2’de verilmiştir:

Çizelge 2. Kohezyonsuz zeminlerde sıklık [28]

Rölatif sıklık, D_r (%)	Sıklık
0-15	Çok gevşek
15-35	Gevşek
35-65	Orta sıkı
65-85	Sıkı
85-100	Çok sıkı

Kum havuzu sistemini doldurma ve boşaltma işlemleri helezonik makine ile yapılmıştır. Bu sistem, 50 rpm hızında dönebilmeye kapasitesine sahip bir elektrik motoru, 250 mm çapında ve 45° eğime sahip bir helisel boru ve 25° eğime sahip bir besleme haznesinden oluşmaktadır [29-30]. Zemin aktarma düzeneğinden zeminin deney kasasına yağmurlama şeklinde aktarılması için 1190 mm x 1190 mm ebatlarında göz açıklığı 5 mm x 5 mm olan bir ahşap çerçeve hazırlanmıştır. Bu sistem, deney kasasının üst kısmına yerleştirilen bir caraskal ile istenilen seviyede durabilmektedir. Kullanılan kazıklar çelikten imal edilmiş olup, şaft çapları sabit 22 mm olacak şekilde üretilmiştir. Bu çalışma kapsamında kullanılan kazıklar helissiz düz kazık ve 100 mm çapında helisli kazık şeklindedir. Laboratuvardaki kazık tipleri ve kazık geometrisi Şekil 2a ve 2b’de gösterilmiştir.



Şekil 2. (a) Model helisel kazıklar, (b) Helisel kazık geometrisi

Bu kazıklar özel üretilen bir tork motoru vasıtasıyla Şekil 3(a)’da görüldüğü üzere önceden kum ile doldurulmuş havuza yerleştirilmiştir. Yerleştirilen grup kazıkların üzerine 780 mm uzunluğunda delikli giriş sistemi eklenmiş ve onun üzerine hidrolik basınç sistemi monte edilmiştir. Yükleme işlemlerinde 5 ton kapasiteli S tipi bir yük hücresi kullanılmış ve taşıma gücü bu sayede ölçülmüştür. Sistemin kurulu hali Şekil 3(b)’de yer almaktadır.

Hidrolik basınç sistemi ile kazıkların oturma miktarı 20mm olana dek basınç kuvveti uygulanmış ve deplasman ölçerlerden oturma miktarı tespit edilmiştir. Yük hücresi ve deplasman ölçerden veri kaydetme ünitesi vasıtasıyla bilgisayara veriler çekilmiş ve alınan sonuçlar ile yük deplasman grafikleri oluşturulmuştur.



(a) Tork sistemi



(b) Kiriş ve hidrolik sistem

Şekil 3. Laboratuvar model deney sistemleri

Kazık yükleme deneylerinin sonuçlarından elde edilen yük-deplasman eğrileri; 20 mm'lik oturma altında, FHWA (%5D) ve ISSMFE (%10D) yöntemleri gibi farklı yöntemler kullanılarak değerlendirilmektedir. Bu yöntemler sayesinde kazıkların eksenel basınç kapasitelerini belirlenmekte ve yorumlanmaktadır.

FHWA (%5D) yöntemiyle belirlenen nihai eksenel basınç kapasitesi, kazık uç çapının %5'i kadar deplasmana karşılık gelen yük değeridir. Helisel kazıklar için uç çap helis çapına, çakma kazıklar için şaft çapına eşittir [31]. ISSMFE (%10D) yaklaşımı ise, kazık çapının %10'undaki deplasman değerini sağlayan yük miktarını nihai kazık kapasitesi olarak tanımlar. Bu yöntemin büyük çaplı kazıkların ve diğer kazık türlerinin kapasitesini yüksek tahmin ettiği söylenmektedir [1].

Bu çalışmada, deney sonuçlarından elde edilen yük-deplasman eğrileri hem kendi aralarında hem de nihai göçme yükünü belirlemek için 20mm'lik oturma altında, FHWA (%5D) ve ISSMFE (%10D) yöntemleri kullanılarak değerler elde edilmiştir. Bu değerler ışığında sonuçlar yorumlanmıştır.

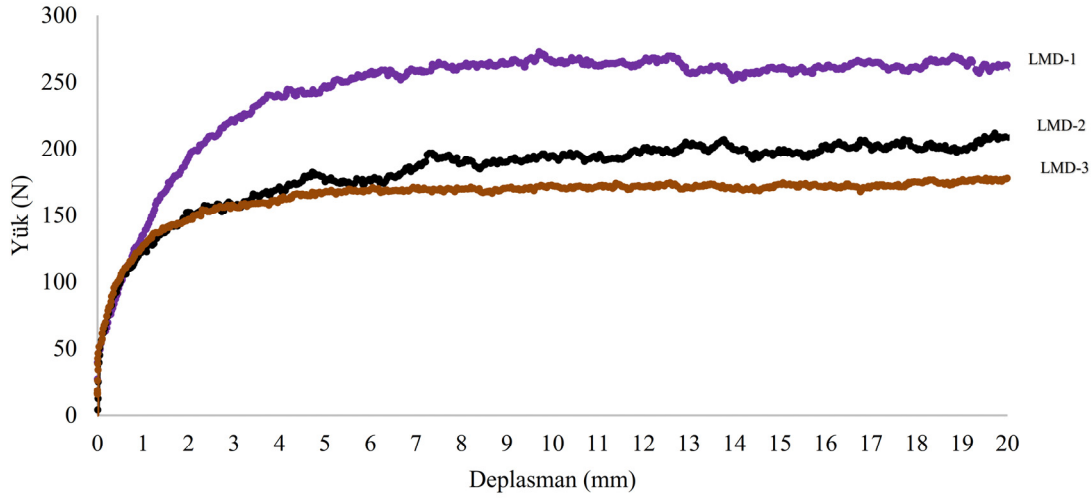
3. BULGULAR VE TARTIŞMA

Grup helisel kazık deneyleri 1x2 yerleşim düzeninde gerçekleştirilmiş ve Helisel Plaka Sayısı (N), Helisel Plaka Çapı (D) ve Kazıklar Arası Mesafe (S) değişken parametreleri kullanılmıştır. Bu parametreler ışığında 9 farklı laboratuvar model deneyi yapılmıştır. Yapılan deneyler ve deney isimleri Çizelge 3'te sunulmuştur.

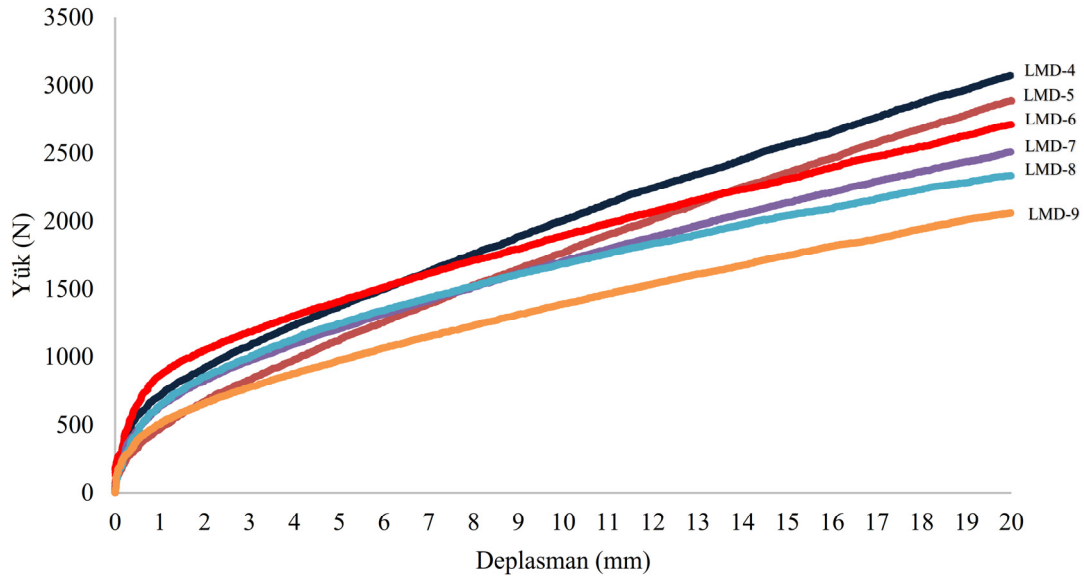
Çizelge 3. Laboratuvar model deneyleri

Deney adı	Helisel plaka sayısı, N (adet)	Helisel plaka çapı, D (mm)	Kazıklar arası mesafe, S (mm)
LMD-1	Helissiz	-	200
LMD-2	Helissiz	-	300
LMD-3	Helissiz	-	400
LMD-4	1	100	200 = 2D
LMD-5	1	100	300 = 3D
LMD-6	1	100	400 = 4D
LMD-7	1	100	500 = 5D
LMD-8	1	100	600 = 6D
LMD-9	1	100	700 = 7D

Deney sonuçları ile çeşitli yük deplasman eğrileri oluşturulmuştur. Model kazıklar ile yapılan deneylerde kazıklar arası mesafe, helis etkisi sabit tutularak Şekil 6'da sunulan helissiz durum için ve Şekil 4'te sunulan helisel kazıklı durum için grafiklerde incelenmiştir.



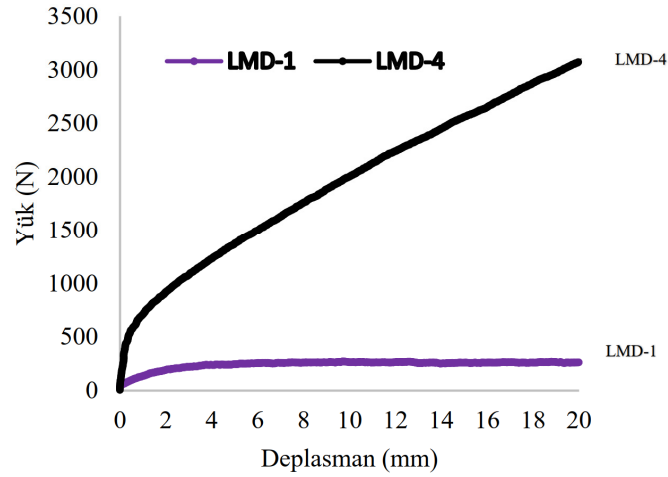
Şekil 4. Kazıklar arası mesafenin etkisi (Helissiz kazıklı durum)



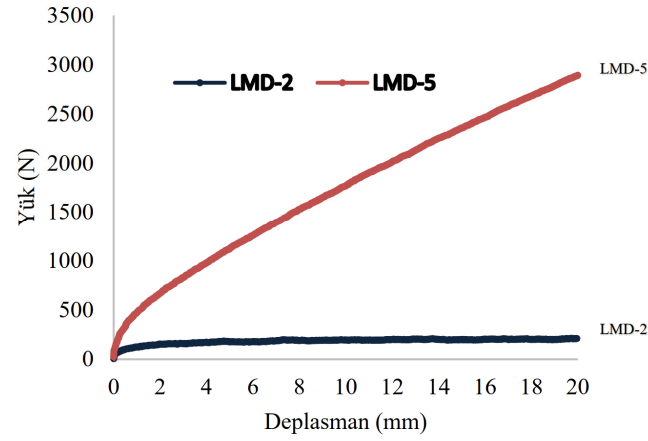
Şekil 5. Kazıklar arası mesafenin etkisi (Helisel kazıklı durum)

Şekil 4 ve Şekil 5 incelendiğinde laboratuvar model deney eğrileri için kazıklar arası mesafe azaldıkça taşıma gücünde artış meydana geldiği görülmüştür. Kazıklar arası mesafesi 200 mm olan helissiz kazık grubu; 300 mm olana göre 1.28 kat ve 400 mm olana göre 1.46 kat daha fazla taşıma kapasitesi sağlamıştır. Kazıklar arası mesafesi 200 mm (2D) olan helisel kazık grubu; 300 mm (3D) olana göre 1.06 kat, 400 mm (4D) olana göre 1.14 kat, 500 mm (5D) olana göre 1.22 kat, 600 mm (6D) olana göre 1.32 kat ve 700 mm (7D) olana göre 1.49 kat daha fazla taşıma kapasitesine sahiptir. Bu durumun sebebinin kazıklar arası mesafe arttıkça grup etkisinin azalması ve kazıkların daha bağımsız çalışmaya başlaması olduğu düşünülmektedir. Zemin içindeki gerilme dağılımı geniş bir alana yayıldığı için, aralarındaki zemin göçme mekanizmasından dolayı daha az sıkışmaktadır.

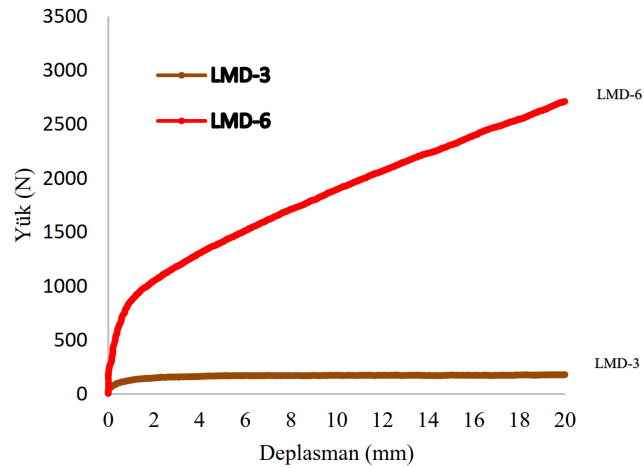
Farklı helisel kazık arası mesafeleri (S) için helissiz ve helisel durumları temsil eden yük-deplasman ilişkileri Şekil 6'da sunulmuştur. Şekil 6(a) kazıklar arası mesafenin 200 mm olduğu durumu, Şekil 6(b) kazıklar arası mesafenin 300 mm olduğu durumu ve Şekil 6(c) kazıklar arası mesafenin 400 mm olduğu durumu ifade etmektedir.



(a) S = 200mm



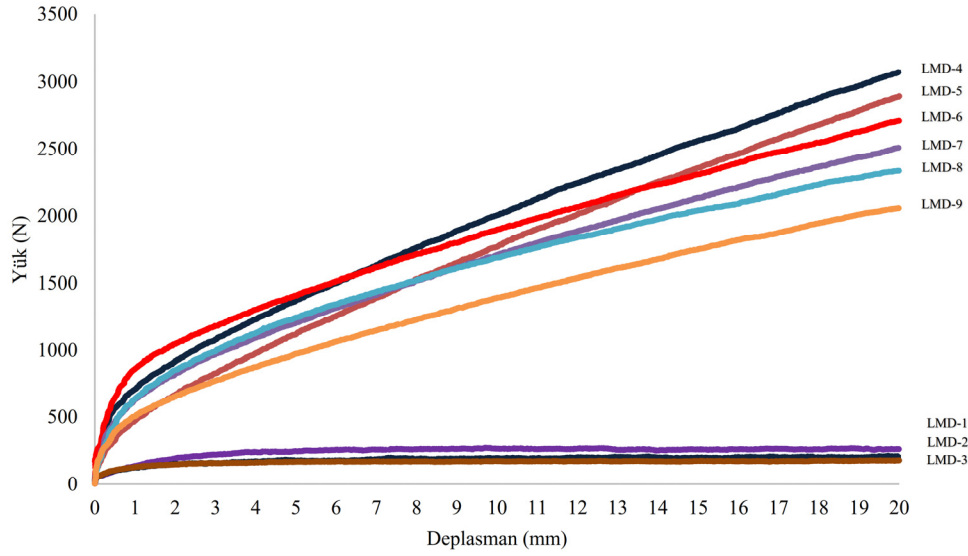
(b) S = 300mm



(c) S = 400mm

Şekil 6. Helisli ve helissiz durumlar için kazıklar arası mesafe etkisi

Şekil 6 incelendiğinde kazıklar arası mesafenin sabit olduğu durumlarda kazıkta helis bulunmasının taşıma gücünü önemli ölçüde arttırdığı tespit edilmiştir. Buna göre 200 mm aralıkta (Şekil 6(a)) helisli kazık, helissiz kazığa göre yaklaşık 12 kat, 300 mm aralıkta (Şekil 6(b)) yaklaşık 14.5 kat, 400 mm aralıkta (Şekil 6(c)) 15.4 kat daha fazla taşıma kapasitesi sağlamıştır. Bu durumun sebebi kazıkta helis varlığının, zeminde ilave yatay kesme yüzeyleri oluşturması ve göçme yüzeyinin daha geniş bir alana yayılması ile açıklanabilir.



Şekil 7. 1x2 dizilimli grup helisel kazıklarda kazıklar arası mesafe etkisi

Şekil 7’de yapılan tüm deneylerden elde edilen yük deplasman eğrileri toplu olarak tek bir grafikte verilmiştir. Deney sonuçları, helis varlığının taşıma gücünü artırma etkisini net bir şekilde ortaya koymuştur.

Çizelge 4. Deneysel nihai yük kapasiteleri

Deney adı	Helisel plaka çapı, D (mm)	Kazıklar arası mesafe, S (mm)	Nihai yük (% 5D) (N)	Nihai yük (%10D) (N)
LMD-1	-	200	141.921	200.086
LMD-2	-	300	126.216	152.391
LMD-3	-	400	131.103	150.646
LMD-4	100	200	738.685	952.438
LMD-5	100	300	497.436	704.341
LMD-6	100	400	886.771	1075.281
LMD-7	100	500	654.08	856.016
LMD-8	100	600	662.781	884.822
LMD-9	100	700	525.572	678.950

Literatürde grup helisel kazıkların basınç etkisi altındaki durumları incelendiğinde; kazıklar arası mesafe arttıkça yük taşıma kapasitesinin azaldığı ve kazığın helisli olmasının yük taşıma kapasitesini önemli ölçüde arttırdığı, %5D ve %10D yöntemleri ışığında optimum değerin 400 mm’de sağlandığı belirlenmiştir. Yapılan deneyler sonucunda elde edilen veriler literatürdeki önceki çalışmalarla paralellik göstermektedir [32-33].

4. SONUÇLAR

Bu çalışma verileri ile literatürde grup kazıkların deneysel çalışma eksikliğine katkı sağlaması amaçlanmıştır. Çalışmada gevşek kum zeminde 1x2 dizilimli grup helisel kazıklarda helis varlığı ve kazıklar arası mesafenin etkisi laboratuvar model deneyleri ile incelenmiştir. Bu amaçla helis sayısı sabit tutulup kazıklar arası mesafenin etkisine ve kazıklar arası mesafe sabit tutulup helis etkisine ait çeşitli değerlendirmeler yapılmıştır. Bu parametreler ışığında elde edilen verilerle yük-deplasman eğrileri elde edilmiş ve aşağıdaki sonuçlara ulaşılmıştır.

- Kazıklar arası mesafe etkisinde, helis sayısının 1 adet 10 cm çapında helis ile sabit tutulduğu durumda, 20 mm’lik kazık gömülme deplasman değeri sabit tutularak taşıma gücü değerlendirmesi yapılmış, kazıklar arası mesafe arttıkça taşıma gücünün azalma gösterdiği belirlenmiştir. Bu değerlendirmelere göre kazıklar arası mesafesi 20 cm olan kazık grubu; 30 cm olana göre % 5.6, 40 cm olana göre %

13.7, 50 cm olana göre % 22.4, 60 cm olana göre % 31.6 ve 70 cm olana göre % 49.4 daha fazla taşıma kapasitesine sahiptir.

- Kazıklar arası mesafe etkisinde, helis sayısının helissiz düz kazık ile sabit tutulduğu durumda, 20 mm'lik kazık gömülme deplasman değeri sabit tutularak taşıma gücü değerlendirmesi yapılmış, kazıklar arası mesafenin etkisinin helissiz kazıklarda da daha az farkla aynı sonuçları verdiği gözlenmiştir. Bu sonuçlara göre kazıklar arası mesafesi 20 cm olan kazık grubu; 30 cm olana göre % 28.1 ve 40 cm olana göre de % 46.3 daha fazla taşıma kapasitesi sağlamıştır.
- Helis varlığı etkisi incelendiğinde, kazıklar arası mesafe sabit 20 cm, 30 cm ve 40 cm olarak belirlenmiş, 10 cm çapında sabit helisli ve helissiz düz kazığın 20 mm göçmede ne kadar taşıma gücüne sahip olduğu değerlendirilmiştir. Bu değerlendirmeye göre 20cm aralıkta helisli kazık helissize göre yaklaşık %1200, 30 cm aralıkta yaklaşık % 1450 ve 40 cm aralıkta % 1540 daha fazla taşıma kapasitesi tespit edilmiştir.
- Kazıklar arası mesafe azaldığında ve helis sabit 10 cm çaplı tutulduğunda % 50'lere varan taşıma gücü artışı gözlenmiştir. Helis plakası olmayan durumda kazıklar arası mesafe azaldığında ise % 47'lere varan taşıma gücü artışı elde edilmiştir. Kazıkta helis bulunması da kazığın taşıma gücünü yaklaşık %1400 kadar artırmaktadır.
- %5D yöntemine göre yorumladığımızda, Helissiz kazıklarda taşıma gücü oldukça düşüktür ancak helisli kazıklarda, %5D kriterine göre nihai yük optimum değer için %600 - %700 kadar artmıştır. Optimum değere ulaşmasının ardından sabit kalma ve azalma eğilimine girmiştir. Bu durum, helisel plaka varlığının taşıma gücünü optimum değere kadar önemli ölçüde artırdığını, optimum değerden sonra bir miktar sabit kalma ve azalma eğiliminde olduğunu göstermektedir.

%10D yönteminde helisel plaka çapı 100 mm olan kazıklarda, taşıma gücü helissizlere göre optimum değer için %400 - %500 kadar artmıştır. Optimum noktadan sonra sabit kalma ve azalma eğilimine girmiştir. Bu veri, helisel plakaların taşıma gücünü optimum değere kadar önemli ölçüde artırdığını, optimum değerden sonra bir miktar sabit kalma ve azalma eğiliminde olduğunu göstermektedir.

5. KAYNAKLAR

1. Türedi, Y. (2021). Basınç yüküne maruz helisel kazık davranışının laboratuvar ve arazi deneyleri ile araştırılması. *Doktora Tezi*, İskenderun Teknik Üniversitesi, Lisansüstü Eğitim Enstitüsü, Hatay, 194.
2. Sakr, M. (2009). Performance of helical piles in oil sand. *Canadian Geotechnical Journal*, 46(9), 1046-1061.
3. Yılmaz, B. (2016). Helisel kazıklar. *Yüksek Lisans Tezi*, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul, 165.
4. Livneh, B. & El Naggar, M.H. (2008). Axial testing and numerical modeling of square shaft helical piles under compressive and tensile loading. *Canadian Geotechnical Journal*, 45(8), 1142-1155.
5. Sakr, M. (2011). Installation and performance characteristics of high-capacity helical piles in cohesionless soil. *DFI Journal - The Journal of the Deep Foundations Institute*, 5(1), 39-57.
6. Tsuha, C.H.C., Aoki, N., Rault, G., Thorel, L. & Garnier, J. (2012). Evaluation of the efficiencies of helical anchor plates in sand by centrifuge model tests. *Canadian Geotechnical Journal*, 49(9), 1102-1114.
7. Lutenege, A.J., Robson, D. & Mohan, P. (2013). Preservation of historic structures and monuments using minimally invasive screw piles and helical anchors – two cases. geotechnical engineering for the preservation of monuments and historic sites (Bilotta, Flora, Lirer & Viggiani, Eds.). *Taylor & Francis Group*, London, 515-522.
8. Harnish, J.L. (2015). Helical pile installation torque and capacity correlations. *MSc Thesis*, The University of Western Ontario London. 123.
9. Khan, M.U.S. (2016). Estimation of compressive load bearing capacity of helical piles using torque method and induced settlements. *MSc Thesis*, The University of British Columbia, 103.
10. Pérez, Z.A., Schiavon, J.A., Tsuha, C.H.C., Dias, D. & Thorel, L. (2018). Numerical and experimental study on influence of installation effects on behavior of helical anchors in very dense sand. *Canadian Geotechnical Journal*, 55(8), 1067-1080.

11. George, B.E., Banerjee, S. & Gandhi, S.R. (2019). Helical piles installed in cohesionless soil by displacement method. *International Journal of Geomechanics*, 19(7), 04019074.
12. Shahbazi, M., Cerato, A.B., El Naggar, M.H. & Elgamal, A. (2020). Evaluation of seismic soil–structure interaction of full-scale grouped helical piles in dense sand. *International Journal of Geomechanics*, 20(12), 04020228.
13. Ali, K., Ullah, S., Nawaz, H., Ahmad, M., Safdar, M. & Ullah, W. (2022). Experimental modelling of helical piles under axial loading. *Pakistan Journal of Engineering and Technology*, 5(2), 128-132.
14. Kaya, B. (2022). Kil zeminlerde helisel kazıkların davranışı. *Yüksek Lisans Tezi*, Çukurova Üniversitesi, Fen Bilimleri Enstitüsü, Adana, 101.
15. Çolak, A. (2023). Helisel kazıkların taşıma gücü kapasitelerinin analitik yöntemler ve sayısal analiz yöntemleri kullanılarak araştırılması. *Yüksek Lisans Tezi*, Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Kayseri, 61.
16. Alwalan, M. & Alnuaim A. (2022). Axial loading effect on the behavior of large helical pile groups in sandy soil. *Arabian Journal for Science and Engineering*, 47, 5017-5031.
17. Elsherbiny, Z.H. & El Naggar, M.H. (2013). Axial compressive capacity of helical piles from field tests and numerical study. *Canadian Geotechnical Journal*, 50(12), 1191-1203.
18. Lanyi-Bennett S.A. & Deng, L. (2019). Axial load testing of helical pile groups in a glaciolacustrine clay. *Canadian Geotechnical Journal*, 56(2), 187-197.
19. Mittal, S. & Mukherjee, S. (2013). Vertical uplift capacity of a group of helical screw anchors in sand. *Indian Geotechnical Journal*, 43(3), 238-250.
20. Nowkandeh, M.J. & Choobbasti, A.J. (2021). Numerical study of single helical piles and helical pile groups under compressive loading in cohesive and cohesionless soils. *Bulletin of Engineering Geology and the Environment*, 4001-4023.
21. Shaheen, W.A. & Demars, K.R. (1995). Interaction of multiple helical earth anchors embedded in granular soil. *Marine Georesources and Geotechnology*, 13, 357-374
22. Vignesh, V. & Muthukumar M. (2023). Experimental and numerical study of group effect on the behavior of helical piles in soft clays under uplift and lateral loading. *Ocean Engineering*, 268, 113500.
23. Bak, J., Choi, B.H., Lee, J., Bae, J., Lee, K. & Kim, D. (2019). Behaviour of single and group helical piles in sands from model experiments. *2nd International Conference on Building Materials and Materials Engineering (ICBMM 2018)*, 278.
24. Elsherbiny, Z. (2011). Axial and lateral performance of helical pile groups. *MSc Thesis*, The University of Western Ontario, London, Canada. 172.
25. Mendoza, C.C., Cunha, R. & Lizcano, A. (2015). Mechanical and numerical behavior of groups of screw (type) piles founded in a tropical soil of the Midwestern Brazil. *Computers and Geotechnics*, 67, 187-203.
26. Mukherjee, A. & Mittal, S. (2015). Vertical uplift capacity of a group of equally spaced helical screw anchors in sand. *6th International Conference on Structural Engineering and Construction Management*, 11-13 December, Kandy, Sri Lanka, 25-31.
27. Sakr, M., Nazir, A., Farouk, A. & Ads, A. (2015). Uplift capacity of helical piles in sand. *International Conference on Advances in Structural and Geotechnical Engineering (ICASGE'15)*, 6-9 April 2015, Hurgada, Egypt, 20.
28. Özyayın, K. (2016). *Zemin mekaniği*. Birsen Kitabevi, 326.
29. Türedi, Y., Emirler, B., Örnek, M. & Yıldız, A. (2023). Helisel kazıklarda helis çapının basınç yüküne etkisinin laboratuvar deneyleri ile araştırılması, *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 38 (4), 1013-1022.
30. Türedi, Y. ve Örnek, M. (2016). Kare temeller altında gerilme ve taşıma gücü analizi. *Çukurova Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 31(2), 59-66.
31. O'Neill, M.W. & Reese, L.C. (1999). Drilled shafts: construction procedures and design methods. Publication No. FHWA-IF-99-025, *Office of Infrastructure, Federal Highway Administration*, Washington D.C., 537.
32. Örnek, M., Yıldız, A. Emirler, B. & Türedi, Y. (2023) Arazide farklı yüklemelere maruz helisel kazıklar için optimum tasarım parametrelerinin önerilmesi. *TÜBİTAK Destekli Bilimsel Araştırma Projesi (TÜBİTAK 1001-No:218M571)*, 395.
33. Türedi, Y. & Örnek, M. (2020). Analysis of model helical piles subjected to axial compression, *Građevinar. Journal of the Croatian Association of Civil Engineers*, 9, 759-769.

Diagnosis of Hepatocellular Carcinoma - HCC Liver Cancer Using Federated Learning on MR Images

Burak UZDUR^{1,a}, Erkut TEKELİ^{1,b}, Turgay İBRİKÇİ^{1,c}, Harun Ur RASHID^{2,d},
Geetha RAMACHANDRAN^{3,e}

¹Adana Alparslan Türkeş Science and Technology University, Software Engineering Department, Adana, Türkiye

²Hankuk University of Foreign Studies, Department of Information and Communications Engineering, Seoul, South Korea

³S.A. Engineering College, Department of Computer Science and Engineering, Thiruverkadu, Chennai, India

^aORCID: 0009-0008-8093-1097; ^bORCID: 0000-0001-9468-5378; ^cORCID: 0000-0003-1321-2523;

^dORCID: 0000-0003-0874-7590; ^eORCID: 0000-0002-4541-3314

Article Info

Received : 05.07.2025

Accepted : 12.08.2025

DOI: 10.21605/cukurovaumfd.1735231

Corresponding Author

Ercut TEKELİ

etekeli@atu.edu.tr

Keywords

Federated learning

FedAvg

Liver tumor classification

Convolutional neural networks

How to cite: UZDUR, B., TEKELİ, E., İBRİKÇİ, T., RASHID, H.U., GEETHA, R. (2025). Diagnosis of Hepatocellular Carcinoma - HCC Liver Cancer Using Federated Learning on MR Images. Çukurova University, Journal of the Faculty of Engineering, 40(3), 531-544.

ABSTRACT

In recent years, Federated Learning (FL) has emerged as a powerful paradigm for training machine learning models across decentralized data sources while preserving data privacy. This study proposes an FL framework for the classification of liver tumors from the ATLAS dataset, which provides images of hepatocellular carcinoma cases. A comparative evaluation was performed utilizing CNN, EfficientNet, MobileNetV3, ResNet50, and VGG16 architectures within the federated environment. Among these models, the FL implementation based on EfficientNet achieved superior performance, reaching an accuracy of 93.75% and a ROC-AUC score of 99.19%. The results demonstrate that federated approaches can attain performance levels comparable to centralized learning while ensuring patient data confidentiality. The potential of developing a privacy-preserving collaborative model using the FL method has been demonstrated.

MR Görüntülerinde Federasyonlu Öğrenme Kullanılarak Hepatosit Karsinomu - HCC Karaciğer Kanseri Tanısı

Makale Bilgileri

Geliş : 05.07.2025

Kabul : 12.08.2025

DOI: 10.21605/cukurovaumfd.1735231

Sorumlu Yazar

Ercut TEKELİ

etekeli@atu.edu.tr

Anahtar Kelimeler

Federasyonlu öğrenme

FedAvg

Karaciğer tümörü sınıflandırması

Evrişimsel sinir ağları

Atf şekli: UZDUR, B., TEKELİ, E., İBRİKÇİ, T., RASHID, H.U., GEETHA, R. (2025). MR Görüntülerinde Federasyonlu Öğrenme Kullanılarak Hepatosit Karsinomu - HCC Karaciğer Kanseri Tanısı. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 531-544.

ÖZ

Son yıllarda, Federasyonlu Öğrenme (FÖ), veri gizliliğini korurken merkezi olmayan veri kaynakları arasında makine öğrenimi modellerini eğitmek için güçlü bir paradigma olarak ortaya çıkmıştır. Bu çalışma, hepatosellüler karsinom vakalarının görüntülerini sağlayan ATLAS veri setinden elde edilen Manyetik Rezonans Görüntülerini kullanarak karaciğer tümörlerinin sınıflandırılması için bir FÖ çerçevesi önermektedir. Federasyonlu ortamda Evrişimli Sinir Ağı, EfficientNet, MobileNetV3, ResNet50 ve VGG16 mimarileri kullanarak karşılaştırmalı bir değerlendirme yapılmıştır. Bu modeller arasında, EfficientNet tabanlı FÖ uygulaması, %93,75'lik bir doğruluk ve %99,19'luk bir ROC-AUC puanına ulaşarak üstün bir performans elde etmiştir. Sonuçlar, federasyonlu yaklaşımların hasta verilerinin gizliliğini sağlarken merkezi öğrenmeye benzer performans seviyelerine ulaşabileceğini göstermektedir. FÖ ile gizliliği koruyan işbirlikçi model geliştirme potansiyeli olduğu gösterilmiştir.

1. INTRODUCTION

The liver, as the largest internal organ in the human body, plays a vital role in numerous physiological functions, including detoxification, enzyme production, blood clotting, and metabolic regulation [1]. Given its critical importance, timely and accurate diagnosis of liver-related disorders is crucial for effective treatment and improved patient outcomes.

Magnetic Resonance Imaging (MRI) has become a cornerstone in liver disease diagnostics due to its non-invasive nature and its ability to generate high-resolution anatomical and functional images. The evolution of MRI techniques has significantly improved the early detection and characterization of hepatic abnormalities [2]. Alongside these advancements, Deep Learning (DL) techniques have increasingly gained attention for their ability to analyze complex medical imaging data with high precision.

DL, a subset of Machine Learning (ML), has shown transformative potential in healthcare, especially in radiology [3]. It allows for automated detection, classification, and segmentation of medical images while minimizing diagnostic variability and aiding clinicians in decision-making processes [4]. These advances are particularly important in diagnosing critical diseases like Hepatocellular Carcinoma (HCC).

HCC remains one of the most prevalent and fatal forms of liver cancer, with projections estimating over one million cases globally by 2025 [5]. It is commonly observed in regions with high hepatitis B and C virus prevalence, such as sub-Saharan Africa and Eastern Asia. Although dynamic Computed Tomography (CT) is a standard imaging technique for HCC diagnosis [6,7], its interpretation requires expert knowledge and is both time- and labor-intensive [8]. Recent research has demonstrated that automated methods, such as computer vision, can enhance diagnostic efficiency and reduce human error [9,10].

Pre-trained Convolutional Neural Networks (CNN), including MobileNetV3, EfficientNet, ResNet50, and VGG16, have proven effective in medical image classification tasks. However, their performance is highly dependent on large and diverse training datasets. In medical domains, data acquisition and sharing are constrained by strict privacy laws and ethical considerations.

As data privacy and security remain top concerns, Federated Learning (FL) has emerged as a viable alternative to centralized training approaches. FL allows multiple institutions to collaboratively train ML models locally, without sharing sensitive data. Only model parameters are exchanged and aggregated, preserving patient confidentiality while enabling large-scale collaborative research.

In this study, an FL-based classification framework is applied by using EfficientNetB5 for liver tumor identification in MRI images. To assess its effectiveness, we compared the FL model's performance with centralized implementations of widely used DL architectures, including CNN, EfficientNet, MobileNetV3, ResNet50, and VGG16. This comparison aims to evaluate FL's advantages not only in classification accuracy but also in addressing data privacy concerns.

The limitations of centralized data usage in medical imaging—particularly regarding privacy, ethics, and regulatory compliance—have led to increased interest in decentralized learning frameworks. FL provides a robust solution by enabling collaborative model training across institutions without transferring raw patient data. This privacy-preserving approach ensures compliance with legal standards while maintaining diagnostic performance. In the context of this study, FL was applied to liver tumor classification using MRI data. Figure 1 illustrates the architecture of the proposed FL framework and its operational flow in a healthcare setting, highlighting how the model aggregates local updates while preserving data confidentiality (adapted from [11]).

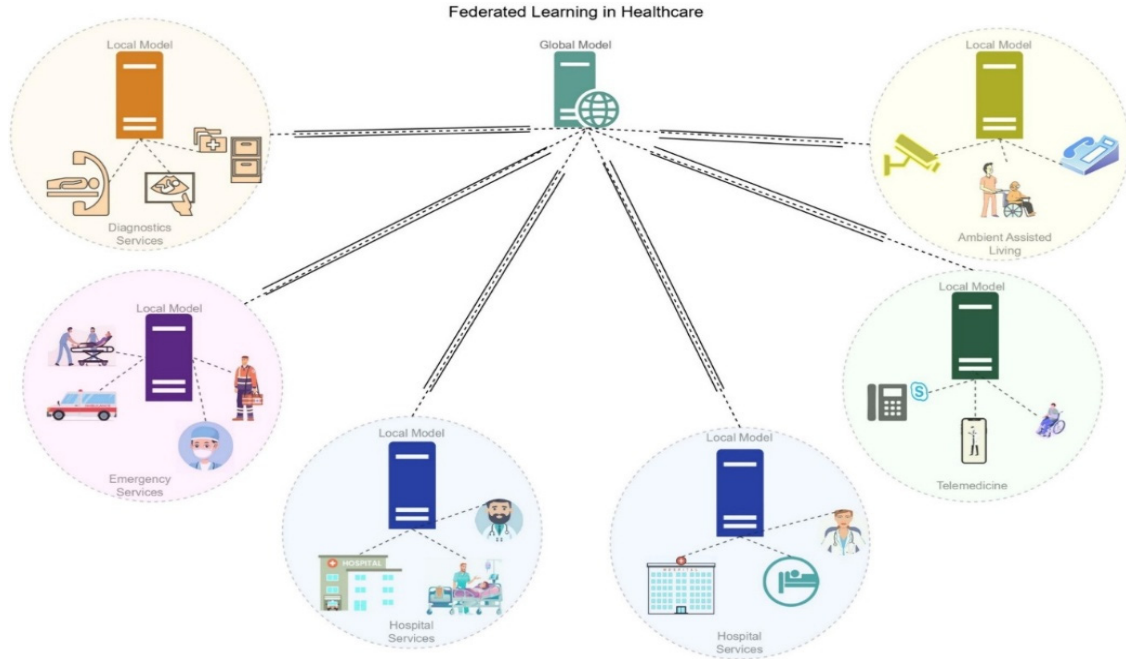


Figure 1. FL technology in the healthcare domain [11]

In conclusion, the study addresses the detection of HCC, a significant issue in medical imaging, by integrating FL to protect data privacy, which is of considerable importance in light of current concerns regarding patient data security. The study also observes that the combination of EfficientNetB5 with FL represents a novel approach and that the comparative analysis with traditional DL models adds significant value.

2. LITERATURE REVIEW

In the realm of medical imaging, the application of ML and DL techniques has led to significant progress, particularly in the classification and diagnosis of liver tumors using MRI images. This section examines relevant studies in the field, emphasizing key approaches, findings, and their connection to the present research (Table 1).

Roth et al. [12] demonstrated the effectiveness of FL in breast density classification by involving seven clinical institutions in a collaborative effort that ensured confidentiality. Their FL-trained model outperformed single-institution models with a 6.3% performance gain and showed a 45.8% improvement in generalizability across external datasets. Despite its success, the study lacks analysis of computational costs and overlooks challenges in large-scale FL deployment, such as infrastructure heterogeneity and data standardization.

Bernecker et al. [13] applied two FL methods, FedNorm and FedNorm+, to perform liver disease segmentation using CT and MRI data from 428 patients across six datasets. FedNorm+ demonstrated superior performance in comparison to local models and equaled the performance of centralized models by attaining a high Dice score of up to 0.961. However, the study is not accompanied by metrics that would allow for an assessment of its accuracy. These limitations, despite the study's initial success in demonstrating the efficacy of the segmentation approach, render it difficult to apply to actual clinical settings.

Mahloul et al. [14] proposed a novel classification model that integrates DL with the FL algorithm. The model was evaluated using the CT-small 2c and CT-large 3c datasets, yielding classification accuracies of 0.82 and 0.96, respectively. The researchers' findings indicate that classification systems developed in FL can offer high reliability and prove effective in clinical decision support systems.

Table 1. Summarization for related works

Authors	Year	Number of patients / images	Data set description	Accuracy / dice (%)	Methodology
Roth et al.	2020	715.000	Mammography (BIRADS, multi-institutional)	6.3 (avg.), 45.8 (gen.)	FL for breast density classification across 7 clinical sites
Bernecker et al.	2022	428	CT and MRI images from 6 different public datasets	Dice: 96.1	FedNorm and FedNorm+ algorithms for FL-based segmentation
Mahlool et al.	2022	253 (BT-small-2c), 3,264 (BT-large-3c)	BT-small-2c: MRI. BT-large-3c: MRI (3 tumor types, 500 healthy)	BT-small-2c Acc: 82, BT-large-3c Acc: 96	FL for brain tumor diagnosis
Trivedi et al.	2023	576	MRI	Acc: 99.59	AlexNet-based FL tested on IID Liver dataset with lightweight CNNs
Chai et al.	2024	733	Gene expression data from TCGA and GEO (cross-institutional, anonymized)	Acc: 54.2	FL with AdFed, DeepSurv-based survival prediction
Lusnig et al.	2024	41 patients / 4,400 images	41 whole-slide images (JPEG2000), divided into 1024×1024-pixel patches; balanced dataset with 1100 images per stage	Centralized Acc: 97%; Federated Acc: ~90%	FL applied to high-resolution histopathology images; images categorized into transplant-suitable vs. unsuitable; used HQNN for classification
Shankar et al.	2025		Multi-modal dataset (CT, MRI, ultrasound + lab values)	Acc: 79.05	Prediction of liver disease using FL from imaging and clinical data

Trivedi et al. [15] evaluated lightweight FL strategies for HCC classification using several pre-trained CNNs, with AlexNet achieving a peak accuracy of 99.59%. A distinguishing feature of their study is its emphasis on system efficiency and deployable architectures. However, the study lacks a detailed discussion on real-world application challenges such as dataset size and communication costs, or training latency.

Chai et al. [16] proposed AdFed, an FL framework for survival prediction in multiple cancers, including liver cancer. Using 733 genetic profiles, the model achieved an AUC of 0.605 for liver cancer, surpassing comparable FL approaches. A major strength is its biological interpretability, with half of the top genes already known to be associated with liver cancer. However, the limited sample size and lack of discussion on real-world FL challenges, such as scalability and communication overhead, weaken its practical applicability.

Lusnig et al. [17] introduced an FL framework using hybrid quantum neural networks (HQNNs) for classifying Non-Alcoholic Fatty Liver Disease (NAFLD) from histopathological biopsy images. The model achieved 97% accuracy under centralized training and around 90% under FL, demonstrating strong performance while preserving data privacy. Despite the promising results, the study is limited by a small dataset (41 patients) and potential scalability issues due to quantum infrastructure requirements.

Shankar et al [18] developed an FL framework that combines CT, MRI, and ultrasound medical imaging data with clinical test data (e.g., bilirubin, ALT, AST) for liver disease prediction. The approach is to achieve an acceptable accuracy of 79.05% while maintaining data confidentiality. However, this study lacks transparency regarding the dataset size and patient numbers. The objective of this study is to identify cases of general liver disease as opposed to tumor-specific classification.

Fofanah et al. [19] presented a comprehensive study that employed CNN and DL techniques for the purpose of detecting skin cancer. The proposed method in this study achieved an 84.3% accuracy rate in skin cancer detection.

Firat and Üzen [20] proposed a DL method for the classification of Alzheimer's disease on MRI. This method was based on Inception and CNNs. Utilising this approach, the researchers attained a classification accuracy of 98% on a four-class dataset.

3. METHOD

3.1. Dataset Description

This study utilizes a publicly available dataset from the ATLAS (A Tumour and Liver Automatic Segmentation) challenge, comprising contrast-enhanced T1-weighted MRI scans from 90 patients diagnosed with inoperable HCC [21]. The dataset includes segmentation masks for both liver and tumor regions, allowing for binary classification into "normal" and "tumor" categories. A total of 3.623 training and 1.553 test images were balanced across both classes, as illustrated in Figure 2. All images were resized to 224×224 pixels and preprocessed accordingly. The data was randomly split into 70% training and 30% testing sets to ensure generalizability. Additionally, Figure 3 presents representative examples that highlight variations in tumor morphology and anatomical structures.

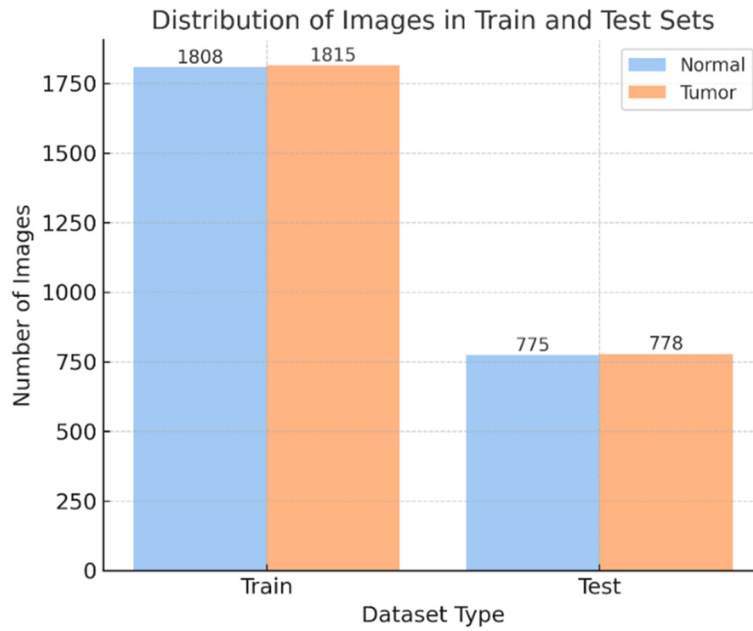


Figure 2. The numerical distribution of normal and tumor images in the training and test datasets

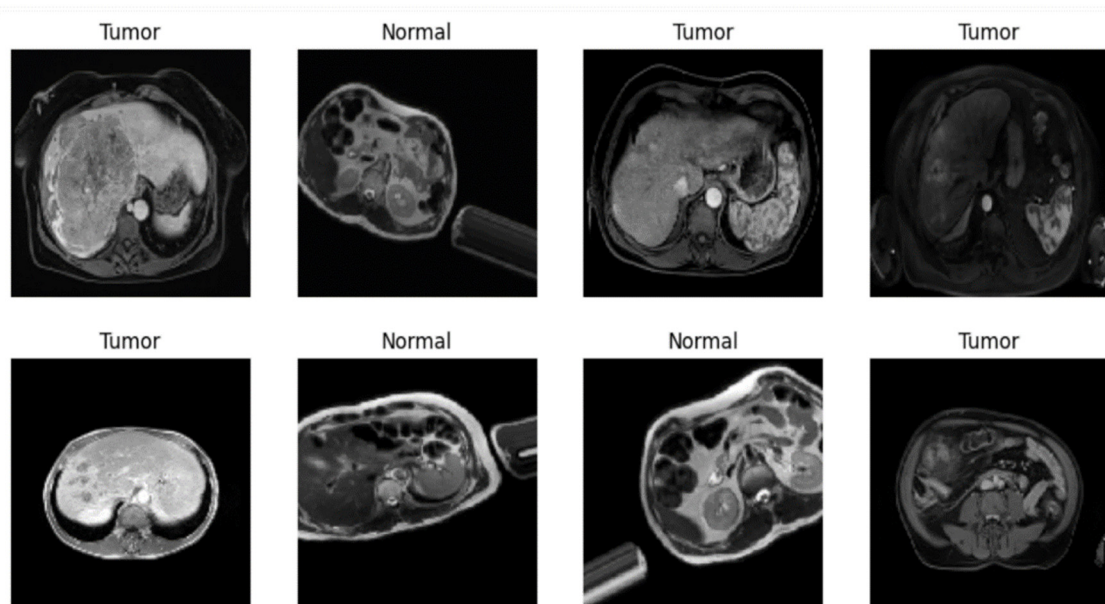


Figure 3. Sample images from the dataset

3.1.1. Data Preprocessing

In this study, a comprehensive data preprocessing pipeline was implemented to enhance the performance of the model and ensure consistency across the dataset. Initially, all MRI images were resized to 224×224 pixels, ensuring uniform input dimensions for all models. To improve the generalization capability of the model and mitigate overfitting, several data augmentation techniques were employed. These included rotation ($\pm 40^\circ$), horizontal flipping, zooming ($\pm 30\%$), brightness adjustments (range: 0.8–1.2), width and height shifting ($\pm 30\%$), and shear transformations ($\pm 30\%$). These augmentations not only expanded the diversity of the dataset but also contributed to improving the model's robustness. Additionally, all images were normalized by scaling pixel values to the $[0,1]$ range, which ensured numerical stability throughout the training process. To avoid class imbalance and reduce potential bias, a class-balancing strategy was applied, ensuring an equal representation of tumor and non-tumor images. Finally, the dataset was partitioned into training and test sets, with 70% of the images allocated for training and the remaining 30% used for evaluation. These preprocessing steps were essential for optimizing the performance of the FL-based classification model.

3.2. Federated Learning

AI has significantly transformed the field of medical imaging by enabling the automation of diagnostics and enhancing accuracy across diverse healthcare settings. However, developing robust AI models requires access to large and varied labeled datasets—a challenge in clinical environments due to stringent data privacy regulations, inter-institutional differences, and the limited availability of annotated medical data. To address these challenges, collaborative and privacy-preserving approaches have been introduced, with FL emerging as a leading decentralized technique. FL enables multiple medical institutions to train models locally on their private datasets and contribute to a shared global model by sending only model updates, thereby safeguarding sensitive patient information within the original institution while facilitating knowledge aggregation from varied data sources [22]. In this study, a strategy of full client participation was employed to maximize the collaborative potential of FL.

In this study, all three clients actively participated in the model training process during each federated round. Each client trained its local model for 3 epochs on its respective dataset, with this process being repeated across 10 federated rounds. The strategy of full participation guarantees that model updates from all clients are integrated into the global model in every round, thereby fostering more efficient and balanced learning.

The training dataset, comprising 90 ATLAS patients and 3,623 images, was partitioned equally and independently among three FL clients to simulate an Independent and Identically Distributed (IID) data setting. Each client received approximately one-third of the total data with balanced class distributions, ensuring statistical similarity across clients. This configuration represents an ideal FL scenario, where data is evenly and randomly distributed among clients, facilitating fair evaluation of model performance.

Federated Averaging (FedAvg) is one of the most widely used optimization algorithms in horizontal FL, where client data remains decentralized and is not shared with a central server. This method has been extensively studied for its convergence properties under various conditions, including data heterogeneity and variations in loss functions, demonstrating its robustness in collaborative learning environments [23]. In this study, FedAvg was employed as the aggregation mechanism to synchronize the local models of the three clients into a unified global model. After each client trained its local model over three epochs, the learned weights were sent to the central server, where layer-wise averaging was conducted. The resulting global weights were then redistributed to all clients for the subsequent training round. This iterative process was repeated across 10 federated rounds, allowing the global model to progressively improve by incorporating knowledge from all clients, all while maintaining data privacy.

While FedAvg facilitates collaborative model training without direct data sharing, it does not inherently guarantee complete privacy. FedAvg remains vulnerable to privacy attacks such as gradient inversion and membership inference, which can potentially expose sensitive client information. This study acknowledges these limitations and considers the integration of advanced privacy-preserving mechanisms as important directions for future work to enhance the robustness of the FL framework.

For the training of the FL, the binary cross-entropy loss function was employed, which is well-suited for the liver tumor classification task due to its binary nature. This loss function effectively measures the difference between predicted probabilities and actual class labels in binary classification tasks, making it a standard and reliable choice for such applications. The model was optimized using the Adam optimizer, renowned for its adaptive learning rate and strong performance in DL tasks. Alongside the minimization of the loss function, various performance metrics were monitored during both training and evaluation, including accuracy, precision, recall, F1-score, and ROC-AUC. These metrics provided a comprehensive evaluation of the model's classification capabilities, particularly in a clinical context where both sensitivity and specificity are crucial for decision-making.

Figure 4 illustrates the training and testing process implemented in this study, following the FL framework. In this setup, separate models are trained locally on distinct clients using their respective data. Following each training round, the local model weights are sent to a central server, where the FedAvg algorithm is applied to update the global model. This iterative process continues until the model achieves optimal accuracy. Once the final model is obtained, its performance is assessed using the central test dataset.

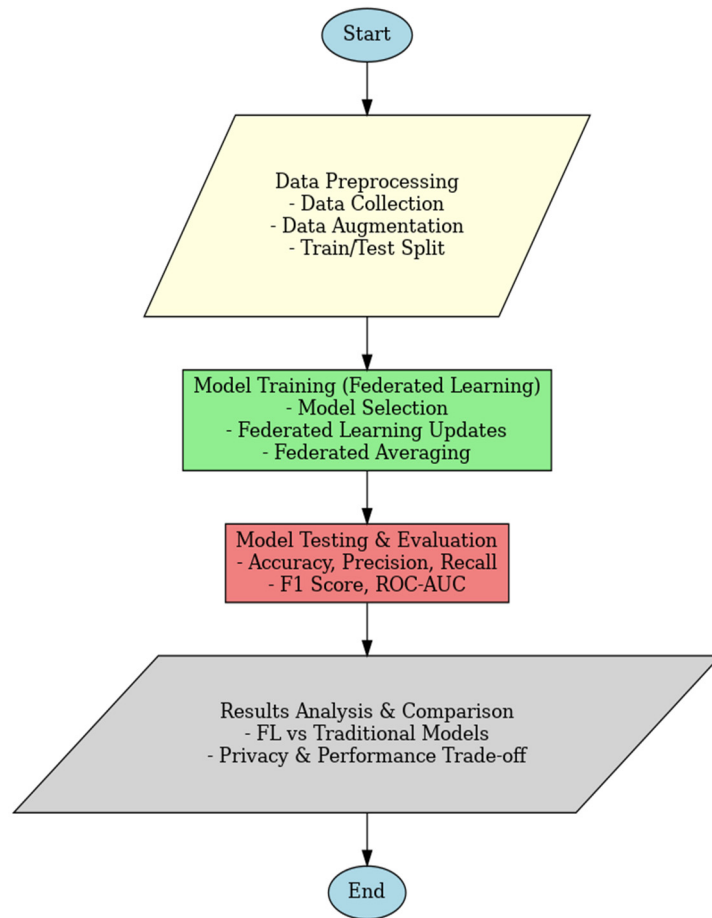


Figure 4. The flowchart of the proposed training and testing process for liver tumor classification

3.3. Comparison Models for Liver Tumor Classification

To assess the performance of the proposed FL model for liver tumor classification, five well-established DL architectures—CNN, EfficientNet, MobileNetV3, ResNet50, and VGG16—were implemented as baseline models. These models were selected for comparison and were trained under identical hyperparameter settings to ensure consistency and fairness in the evaluation process. Specifically, each model underwent training for 10 epochs with a batch size of 32, utilizing the Adaptive Moment Estimation-Adam optimizer and binary cross-entropy as the loss function. ReLU activation was applied to the hidden layers, while the output layer employed the sigmoid activation function, which is suitable for binary classification tasks. Table 2 provides an overview of the training configurations for these models.

Table 2. Training configurations for models

	Number of epochs	Batch size	Optimizer	Loss function	Activation functions	Output layer
CNN	10	32	Adam	Binary Crossentropy	Relu	Sigmoid
EfficientNet	10	32	Adam	Binary Crossentropy	Relu	Sigmoid
MobileNetV3	10	32	Adam	Binary Crossentropy	Relu	Sigmoid
ResNet50	10	32	Adam	Binary Crossentropy	Relu	Sigmoid
VGG16	10	32	Adam	Binary Crossentropy	Relu	Sigmoid

3.3.1. Convolutional Neural Networks

In this study, a custom CNN was developed as a baseline model for liver tumor classification. The architecture includes three convolutional layers with progressively larger filter sizes (32, 64, and 128), each followed by a MaxPooling2D layer to reduce spatial dimensions and mitigate overfitting. After the convolutional operations, the feature maps are flattened and passed through a fully connected dense layer containing 128 neurons, activated by the ReLU function. Finally, a sigmoid-activated output layer is employed to perform binary classification. This relatively straightforward yet efficient architecture serves as a foundational model for comparing the performance of more advanced, pre-trained models in liver tumor classification.

3.3.2. EfficientNet

In the comparative analysis, EfficientNet was employed to evaluate the performance of liver tumor classification. The model was fine-tuned by unfreezing the last 50 layers of the pre-trained base and retraining them on the target dataset. To tailor the model for binary classification, additional layers were incorporated on top of the base, including a GlobalAveragePooling2D layer, a dropout layer with a 0.3 rate to mitigate overfitting, and two dense layers—one with 128 neurons activated by ReLU, and a final output layer with a sigmoid activation function. This architecture enabled the model to capture domain-specific features while leveraging the benefits of transfer learning for improved performance.

3.3.3. ResNet50

The ResNet50 architecture with 50 layers deep was developed by Microsoft Research in 2015, considered to be among the most popular CNN architectures around, which employs skip connections to address the vanishing gradient problem, enabling effective training of deeper networks. This model has shown robust performance in medical imaging classification, particularly in identifying complex patterns within tumor images. However, its computationally intensive nature may pose challenges for deployment in real-world medical settings, where resources may be limited.

3.3.4. MobileNetV3

In this study, MobileNetV3 Small was adapted using a transfer learning approach specifically designed for medical image classification tasks. The model utilized pre-trained weights, with the initial layers frozen to retain low-level feature extraction capabilities. The MobileNetV3 small architecture was employed, followed by the addition of a Global Average Pooling layer, a fully connected dense layer with 128 neurons and ReLU activation, a Dropout layer with a 0.5 rate to mitigate overfitting, and a final output layer with a single neuron using sigmoid activation for binary classification. The training process utilized the Adam optimizer along with the binary cross-entropy loss function. Additionally, the learning rate was dynamically adjusted using the ReduceLROnPlateau technique, which monitored the validation loss and reduced the learning rate when further improvements plateaued.

3.3.5. VGG16

One of the comparison models employed in this study was the VGG16 architecture. The base of the model consisted of the VGG16 network, which was pre-trained on ImageNet, with its top layers removed. To minimize overfitting and allow the model to focus on training the newly added layers, the weights of the pre-trained layers were frozen. A Global Average Pooling layer was incorporated to flatten the feature

maps, followed by a fully connected dense layer comprising 128 neurons and ReLU activation, and a Dropout layer with a rate of 0.5 to reduce overfitting. The final output layer incorporated a sigmoid activation function to facilitate binary classification. The model was constructed with the Adam optimiser, with a learning rate of 0.001 and binary cross-entropy as the loss function.

4. RESULTS

To evaluate the effectiveness of the proposed FL model in liver tumor classification, five distinct DL architectures were implemented for comparison. The custom-designed CNN employed a simplified architecture consisting of three convolutional layers, each followed by max-pooling and fully connected layers. EfficientNet was fine-tuned by unfreezing its last 50 layers and enhanced with a Global Average Pooling layer, dropout regularization, and dense layers. Similarly, MobileNetV3 and ResNet50—both initialized with ImageNet pre-trained weights—were adapted with additional layers including Global Average Pooling, ReLU-activated dense layers, and dropout to improve generalization. The VGG16 model was also employed by freezing its convolutional base and appending dense layers for the binary classification task. All models were trained under consistent settings using the binary cross-entropy loss function and the Adam optimizer, which is an optimisation algorithm that is frequently employed in the training of DL models. This approach synthesises the merits of two other extensions of stochastic gradient descent (SGD): AdaGrad and RMSProp. The architectural and training configurations of these models are summarized in Table 3, ensuring a fair and systematic comparison of model performances.

Table 3. Architectural and training configurations of DL models used for image classification

Model	Pretrained weights	Trainable layers	Pooling type	Dropout rate	Dense layer (units)
CNN	No	All	MaxPooling2D	None	128
EfficientNet	ImageNet	Last 50 layers	GlobalAveragePooling2D	0.3	128
MobileNetV3	ImageNet	Frozen (feature extractor)	GlobalAveragePooling2D	0.5	128
ResNet50	ImageNet	Frozen (feature extractor)	GlobalAveragePooling2D	0.5	128
VGG16	ImageNet	Frozen (feature extractor)	GlobalAveragePooling2D	0.5	128

To comprehensively assess the performance of the proposed liver tumor classification model, five key evaluation metrics—accuracy, precision, recall, F1-score, and ROC-AUC score—were utilized. These metrics offer valuable insights into the model's ability to correctly classify MRI images, differentiate between normal and tumor cases, and perform effectively in a real-world diagnostic setting.

The evaluation of performance criteria constitutes a pivotal step in both ML and DL. The True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) outcomes are pivotal for the evaluation of the performance of classification models. These metrics facilitate comprehension of the extent to which their models can generate precise predictions and subsequently optimise them for discrete or particular decisions, thereby enhancing decision-making processes across diverse domains.

TP: The number of times the model correctly predicted the positive class.

TN: The number of cases where the model correctly predicted the negative class.

FP: The number of instances where the model incorrectly predicted the positive class.

FN: The number of cases where the model incorrectly predicted the negative class.

Accuracy is one of the most fundamental metrics for evaluating the overall performance of a classification model. It is defined as the ratio of correctly classified instances—both positive and negative—to the total number of predictions made, as shown in Equation 1. While accuracy can serve as a reliable indicator in balanced datasets, it may become misleading in imbalanced scenarios.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Precision is defined as the proportion of predicted positive cases that are actually positive and is formally expressed in equation 2. A high precision score indicates that the model effectively minimizes false positives, meaning its positive classifications are more reliable.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Recall assesses the model's ability to identify positive cases correctly and is mathematically defined in equation 3. This metric is particularly important in applications where failing to detect a positive case can have serious consequences, such as in medical diagnosis

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

F1-Score is the harmonic mean of precision and recall, as defined in Equation 4. It is particularly effective in evaluating model performance on imbalanced datasets, as it considers both false positives and false negatives. A high F1-score reflects the model's ability to strike a balance between precision and recall, effectively minimizing both types of errors. The F1-Score serves as a crucial metric, ensuring that the model not only avoids false alarms but also captures the majority of true positive cases.

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

ROC-AUC Score is a graphical representation that is used to evaluate the performance of a classification model at various threshold values. The Area Under the Curve (AUC) represents the area beneath the Receiver-Operating Characteristic Curve (ROC) and measures the model's overall discriminative ability. The ROC-AUC metric is particularly important for evaluating classification models on imbalanced datasets. As illustrated in Figure 5, the ROC curves of the models under investigation are displayed.

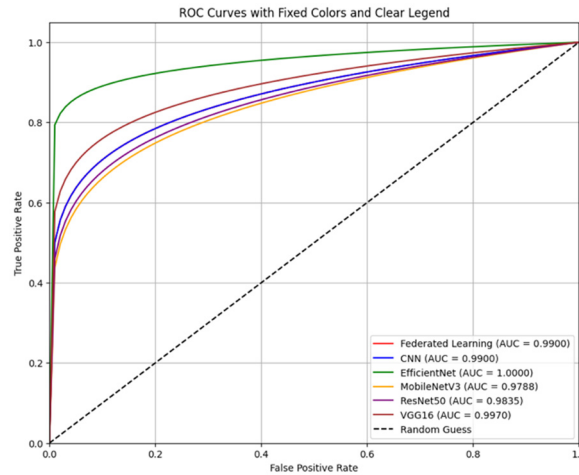


Figure 5. ROC curves

The classification performance of the models studied is shown in Table 4.

Table 4. Performance of models

	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC (%)
FL	93.75	99.71	87.79	93.37	99.19
CNN	98.58	99.74	97.43	98.57	98.59
EfficientNet	99.55	99.61	99.49	99.55	99.55
MobileNetV3	95.43	97.20	93.57	95.35	97.88
ResNet50	94.01	94.54	93.44	93.99	98.35
VGG16	98.71	99.74	97.69	98.70	99.70

Figure 6 presents a graphical representation that compares the classification performance of the models.

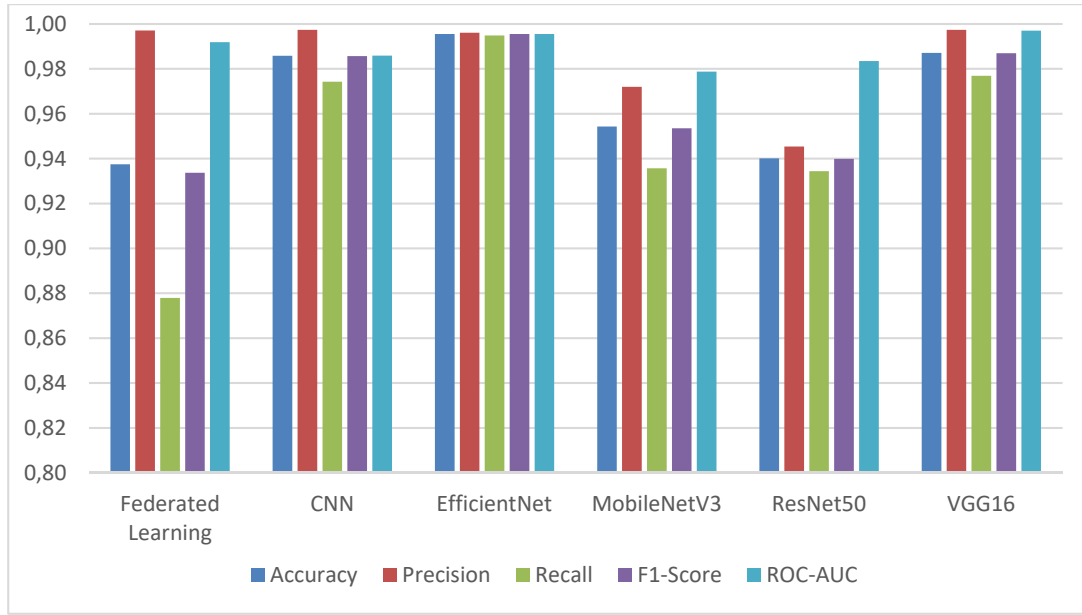


Figure 6. Comparison of model performances

To empirically justify our choice of hyper parameters in the FL setup, we conducted an ablation study by varying local epochs and global rounds. Table 5 summarizes the performance metrics across four different configurations.

Table 5. Ablation study of FL configuration

Local epochs	Rounds	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC (%)
3	10	93,75	99,71	87,79	93,37	99,19
1	5	92,14	90,80	93,83	92,29	97,40
1	10	92,98	96,65	89,07	92,71	98,49
1	20	82,68	75,02	98,07	85,01	97,62

As shown in Table 5, the configuration with 3 clients, 3 local epochs, and 10 rounds yielded superior results in terms of accuracy and ROC-AUC compared to the other evaluated configurations. Based on these findings, this setup was selected for use in the subsequent experiments.

As the final results, the FL-based model achieved an accuracy of 93.75% and a precision value of 99.71%. These results demonstrate that the model provides highly reliable and accurate positive results in the diagnosis of liver tumors, while also effectively differentiating between classes. The F1-Score was 93.37%, and the ROC-AUC value was measured at 99.19%, both of which indicate the high overall performance of the model and its potential for use in critical decision-support systems in healthcare.

It is observed that the centralized EfficientNet model achieves a higher accuracy (99.55%) compared to the FL model (93.75%). This performance gap is expected and can be attributed to the inherent challenges of FL, such as data heterogeneity across clients, communication constraints, and the absence of direct access to the full training dataset.

Despite this difference, the FL model still achieves a high level of accuracy and offers substantial privacy advantages by enabling model training without centralized data aggregation. In real-world medical environments, such as inter-hospital collaborations, data sharing is often limited by strict privacy regulations (e.g., HIPAA, GDPR). Consequently, the observed reduction in accuracy is justified by the critical need to maintain data privacy, particularly in healthcare settings where centralized data sharing is not feasible.

In the diagnosis of HCC using MRI, sensitivity (recall) rates are typically reported in the range of 85% to 95% [24, 25, 26]. Deep learning-based models have also achieved recall rates comparable to those of expert radiologists [24]. In our study, the achieved recall rate of 87.79% falls within the lower bound of this clinically accepted range. However, this corresponds to a 12.21% false-negative rate, which could lead to missed diagnoses in practice. According to the Liver Imaging Reporting and Data System (LI-RADS) developed by the American College of Radiology, sensitivity levels above 88% are considered desirable for MRI-based HCC diagnosis [27]. In our study, the recall rate of 87.79% falls within the lower limit of the range accepted as clinically acceptable in the literature.

In comparison, EfficientNet exhibited the highest performance with an accuracy of 99.55%. Its precision and recall values are very close, reflecting the model's success in accurately diagnosing tumors. Both VGG16 and CNN models also showed high accuracy rates (98.71% and 98.58%) and excellent precision values (99.74%). However, when compared to the FL model, these models exhibited higher recall values (97.69% and 97.43%), which is a crucial factor in healthcare applications, where false negatives can have significant consequences. This situation demonstrates that further development of FL is required in order to offer a more balanced and reliable solution.

MobileNetV3 and ResNet50 models, with accuracy rates of 95.43% and 94.01%, respectively, presented the lowest performance. Nevertheless, their precision and F1-score values are still at levels suitable for healthcare applications, providing useful results.

5. CONCLUSION AND DISCUSSIONS

Accurate diagnosis of HCC is essential for guiding life-saving treatments. In this study, it is shown that FL—a privacy-preserving ML approach—can reliably detect HCC from MRI medical imaging. The necessity to classify liver cancer while protecting patient privacy has become increasingly important. This study investigates the feasibility of using FL and DL to categorize liver cancer. The study also examines the comparisons of CNN, EfficientNet, MobileNetV3, ResNet50, and VGG16 with FL.

This study confirms that FL, particularly when integrated with EfficientNetB5, is a viable solution for privacy-preserving liver tumor classification in MRI imaging. The findings demonstrate the system's capacity to generate outcomes that are competitive with those of the baseline system, while utilizing fewer resources and ensuring superior data privacy. These results underscore the practical viability of FL for HCC classification in clinical environments.

Conversely, it is imperative to recognise that the Federated EfficientNet model exhibits a lower classification accuracy (93.75%) in comparison to its central counterpart (99.55%). This performance gap is expected, given the decentralized training constraints and the limited global information available to each client. However, this trade-off is justified in privacy-sensitive environments where direct data sharing is infeasible. Future work will aim to reduce this gap by incorporating more sophisticated aggregation techniques (e.g., FedProx, FedOpt), dynamic client selection, and hybrid learning approaches that strike a better balance between performance and privacy.

Although the proposed FL framework achieved promising accuracy and ROC-AUC scores, the recall value remained relatively lower compared to clinical expectations. This indicates that the model may still miss a portion of true positive cases, which is critical in hepatocellular carcinoma diagnosis. To address this limitation, future work will explore advanced strategies such as cost-sensitive learning, focal loss functions, and ensemble techniques tailored to enhance recall. Additionally, incorporating clinical metadata (e.g., patient history, lab results) alongside imaging data could improve diagnostic robustness. Finally, investigating personalization techniques in FL, such as federated fine-tuning or clustering-based model adaptation, may further reduce false negatives by tailoring models to client-specific distributions.

While FedAvg facilitates privacy-aware training by avoiding direct data sharing, recent studies have shown that it is still vulnerable to privacy leakage through indirect inference attacks. In particular, gradient inversion attacks [28] can reconstruct sensitive input data from shared model updates, while membership inference attacks [29] can expose whether a particular data point was part of a client's training set. These vulnerabilities highlight the need for complementary privacy-preserving mechanisms such as differential privacy, secure aggregation, or homomorphic encryption. Incorporating such techniques in future

implementations will be critical to ensuring stronger privacy guarantees without compromising model performance.

Future work may also involve the deployment of FL frameworks in multi-institutional hospital networks and the integration of multimodal clinical data.

6. REFERENCES

1. Singh, A. & Pandey, B. (2016). Diagnosis of liver disease by using least squares support vector machine approach. *International Journal of Healthcare Information Systems and Informatics*, 11(2), 62-75.
2. Çaviş, T. & Arda, K.N. (2024). Advanced magnetic resonance imaging techniques in the diagnosis of liver diseases. *The Turkish Journal of Current Gastroenterology*, 26(3), 130-141.
3. Chan, H.P., Samala, R.K., Hadjiiski, L.M. & Zhou, C. (2020). Deep learning in medical image analysis. In: Lee, G., Fujita, H. (eds) *Deep Learning in Medical Image Analysis. Advances in Experimental Medicine and Biology*, Springer, 181.
4. Kwak, L. & Bai, H. (2023). The role of federated learning models in medical imaging. *Radiology: Artificial Intelligence*, 5(3), 1-2.
5. Llovet, J.M., Kelley, R.K., Villanueva, A., Singal, A.G., Pikarsky, E., Roayaie, S., Lencioni, R., Koike, K., Rossi, J.Z. & Finn, R.S. (2021). Hepatocellular carcinoma. *Nature Rev. Dis. Primers*, 7(1), 6-34.
6. Heimbach, J.K., Kulik, L.M., Finn, R.S., Sirlin, C.B., Abecassis, M.M., Roberts, L.R., Zhu, A.X., Murad, M.H. & Marrero J.A. (2018). AASLD guidelines for the treatment of hepatocellular carcinoma. *Hepatology*, 67(1), 358-380.
7. Shao, Y.Y., Wang, S.Y. & Lin, S.M. (2021). Management consensus guideline for hepatocellular carcinoma: 2020 update on surveillance, diagnosis, and systemic treatment by the Taiwan liver cancer association and the gastroenterological society of Taiwan. *J Formos Med Assoc.*, 120(4), 1051-1060.
8. Chen, H., Gomez, C., Huang, C.M. & Unberath, M. (2022). Explainable medical imaging AI needs human-centered design: guidelines and evidence from a systematic review. *npj Digital Medicine*, 5(156), 1-15.
9. Song, L., Geoffrey, K. & Kaijian, H. (2020). Bottleneck feature supervised U-net for pixel-wise liver and tumor segmentation. *Expert Syst. Appl.*, 145(5), 1-11.
10. Song, D., Wang, Y., Wang, W. & Wang, Y. (2021). Using deep learning to predict microvascular invasion in hepatocellular carcinoma based on dynamic contrast-enhanced MRI combined with clinical parameters. *J. Cancer Res. Clin. Oncol.*, 147(12), 3757-3767.
11. Srinivasu, P.N., Lakshmi, G.J., Narahari, S.C., Shafi, J., Choi, J. & Ijaz, M.F. (2024). Enhancing medical image classification via federated learning and pre-trained model. *Egyptian Informatics Journal*, 27(1), 1-16.
12. Roth, H.R., Chang, K., Singh, P., Neumark, N., Li, W., Gupta, V., Gupta, S., Qu, L., Ihsani, A., Bizzo, B.C., Wen, Y., Buch, V., Shah, M., Kitamura, F., Mendonça, M., Lavor, V., Harouni, A., Compas, C., Tetreault, J., Dogra, P., Cheng, Y., Erdal, S., White, R., Hashemian, B., Schultz, T., Zhang, M., McCarthy, A., Yun, B.M., Sharaf, E., Hoebel, K.V., Patel, J.B., Chen, B., Ko, S., Leibovitz, E., Pisano, E.D., Coombs, L., Xu, D., Dreyer, K.J., Dayan, I., Naidu, R.C., Flores, M., Rubin, D. & Cramer, J.K. (2020). Federated learning for breast density classification: a real-world implementation. *Domain Adaptation and Representation Transfer, and Distributed and Collaborative Learning, DART 2020*, Peru.
13. Bernecker, T., Peters, A., Schlett, C.L., Bamberg, F., Theis, F., Rueckert, D., Weib, J. & Albarqouni, S. (2022). FedNorm: modality-based normalization in federated learning for multi-modal liver segmentation. *ArXiv preprint*, abs/2205.11096, 1-21.
14. Mahlool, D.H. & Abed, M.H. (2022). Distributed brain tumor diagnosis using a federated learning environment. *Bulletin of Electrical Engineering and Informatics*, 11(6), 3313-3321.
15. Triverdi, N.K., Shukla, S., Tiwari, R.G., Agarwal, A.K. & Gautam, V. (2023). Liver cancer diagnosis with lightweight federated learning using identically distributed images. *12th International Conference on System Modeling & Advancement in Research Trends (SMART-2023)*, Moradabad, India.
16. Chai, H., Huang, Y., Xu, L., Song, X., He, M. & Wang, Q. (2024). A decentralized federated learning-based cancer survival prediction method with privacy protection. *Heliyon*, 10(11), 1-11.
17. Lusnig, L., Sagingalieva, A., Surmach, M., Protasevich, T., Michiu, O., McLoughlin, J., Mansell, C., Petris, G.D., Bonazza, D., Zanconati, F., Melnikov, A. & Cavalli, F. (2024). Hybrid quantum image classification and federated learning for hepatic steatosis diagnosis. *Diagnostics*, 14(5), 1-16.

18. Shankar, P.U., Rahul, E.S., Rao, K.D., Satish, K., Kumar, U.D., Ravindra, D. & Subbarao, G. (2025). Liver disease prediction using federated learning. *International Journal of Innovative Science and Research Technology*, 10(4), 880-887.
19. Balla Fofanah, A., Özbilge, E. & Kırsal, Y. (2023). Skin cancer recognition using compact deep convolutional neural network. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 38(3), 787-797.
20. Fırat, H. & Üzen, H. (2024). MR görüntülerinden alzheimer hastalığının sınıflandırılması için inception ve sıkma-uyarma ağı tabanlı derin öğrenme modeli. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 39(2), 555-567.
21. Quinton, F., Popoff, R., Presles, B., Leclerc, S., Meriaudeau, F., Nodari, G., Lopez, O., Pellegrinelli, J., Chevallier, O., Gignac, D., Vrigneaud, J.-M. & Alberini, J.-L. (2023). A tumour and liver automatic segmentation (ATLAS) dataset on contrast-enhanced magnetic resonance imaging for hepatocellular carcinoma. *Data*, 8(5), 1-9.
22. Lotfinia, M., Tayebiarasteh, A., Samiei, S., Joodaki, M. & Arasteh, S.T. (2025). Boosting multi-demographic federated learning for chest x-ray analysis using general-purpose self-supervised representations. *European Journal of Radiology Artificial Intelligence*, 3(1), 1-13.
23. Overman, T. & Klabjan, D. (2025). Continuous-time analysis of federated averaging. *ArXiv preprint*, abs/2501.18870, 1-25.
24. Yasaka, K., Akai, H., Abe, O. & Kiryu, S. (2018). Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: A preliminary study. *Radiology*, 286(3), 887-896.
25. Ahn, J.C., Qureshi, T.A., Singal, A.G., Li, D. & Yang, J.D. (2021). Deep learning in hepatocellular carcinoma: current status and future perspectives. *World J Hepatol.*, 13(12), 2039-2051.
26. American College of Radiology (ACR), (2021). *Liver imaging reporting and data system (LI-RADS) v2018 manual*. Retrieved from <https://www.acr.org/Clinical-Resources/Clinical-Tools-and-Reference/Reporting-and-Data-Systems/LI-RADS>, Access date: 16/07/2025.
27. Choi, J.Y., Lee, J.M. & Sirlin, C.B. (2014). CT and MR imaging diagnosis and staging of hepatocellular carcinoma: Part II. Extracellular agents, hepatobiliary agents, and ancillary imaging features. *Radiology*, 273(1), 30-50.
28. Zhu, L., Liu, Z. & Han, S. (2019). Deep leakage from gradients. *33rd Conference on Neural Information Processing Systems (NeurIPS 2019)*, Vancouver, Canada.
29. Shokri, R., Stronati, M., Song, C. & Shmatikov, V. (2017). Membership inference attacks against machine learning models. *38th IEEE Symposium on Security and Privacy (SP)*, San Jose, CA.

Beam-Limited k-Step Lookahead for Computationally Efficient HMM Decoding

Mehmet KURUCAN^{1,a}

¹Adana Alparslan Türkeş Science and Technology University, Faculty of Computer and Informatics, Department of Computer Engineering, Adana, Türkiye

^aORCID: 0000-0003-4359-3726

Article Info

Received : 28.05.2025

Accepted : 13.08.2025

DOI: 10.21605/cukurovaumfd.1708178

Corresponding Author

Mehmet KURUCAN

mkurucan@atu.edu.tr

Keywords

Hidden markov models

Decoding problems

Beam limited search

Sequential estimation

How to cite: KURUCAN, M., (2025).

Beam-Limited k-Step Lookahead for Computationally Efficient HMM Decoding. Cukurova University, Journal of the Faculty of Engineering, 40(3), 545-558.

ABSTRACT

Hidden Markov Models (HMMs) are widely used in many sequential decision-making problems due to their ability to model time-related dependencies. The standard decoding methods in these models, such as the Viterbi algorithm, are limited by their dependence on past observations only. Thus, this leads to unpredictability when future information is available. In this work, we propose a decoding strategy called Beam-Limited k-Step Lookahead that looks k-step ahead, drawing parallels to k-step discrete control synthesis, to make use of future information. The proposed method achieves a balance between decoding accuracy and computational complexity by constraining the search space to the top M most promising paths. Experimental results on synthetic HMM data show that our new decoding strategy significantly improves decoding accuracy over classical Viterbi decoding. The findings highlight the potential of this new strategy to improve the performance of sequential decoding systems.

Hesaplama Açısından Verimli HMM Kod Çözümü İçin Demet Sınırlamalı k-Adımlı İleri Bakış

Makale Bilgileri

Geliş : 28.05.2025

Kabul : 13.08.2025

DOI: 10.21605/cukurovaumfd.1708178

Sorumlu Yazar

Mehmet KURUCAN

mkurucan@atu.edu.tr

Anahtar Kelimeler

Saklı markov modelleri

Kod çözme problemleri

Sınırlı ışın araması

Ardışık tahminleme

Atf şekli: KURUCAN, M., (2025).

Hesaplama Açısından Verimli HMM Kod Çözümü İçin Demet Sınırlamalı k-Adımlı İleri Bakış. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 545-558.

ÖZ

Gizli Markov Modelleri (HMM'ler), zamanla ilgili bağımlılıkları modelleme yetenekleri nedeniyle birçok ardışık karar verme probleminde yaygın olarak kullanılır. Bu modellerdeki standart kod çözme yöntemleri, Viterbi algoritması gibi, yalnızca geçmiş gözlemlere olan bağımlılıklarıyla sınırlıdır. Bu nedenle, gelecekteki bilgiler mevcut olduğunda öngörülemezliğe yol açar. Bu çalışmada, gelecekteki bilgileri kullanmak için (yani kontrol teorisindeki k-adımlı ayrık kontrol sentezine benzer bir yaklaşımla) k-adım ileriye gören Işın Sınırlı k-Adım İleriye Bakış adı verilen bir kod çözme stratejisi öneriyoruz. Önerilen yöntem, arama alanını en umut verici M yolla sınırlayarak kod çözme doğruluğu ve hesaplama karmaşıklığı arasında bir denge sağlar. Sentetik HMM verileri üzerindeki deneysel sonuçlar, yeni kod çözme stratejimizin kod çözme doğruluğunu klasik Viterbi kod çözmeye kıyasla önemli ölçüde iyileştirdiğini göstermektedir. Bulgular, bu yeni stratejinin ardışık kod çözme sistemlerinin performansını iyileştirme potansiyelini vurgulamaktadır.

1. INTRODUCTION

The utilization of machine learning-based models in scientific research has become firmly settled. Up-to-date works are now directed towards enhancing the effectiveness and efficiency of existing Machine Learning (ML) models through the combination of different models or tools. As is well-known, the Hidden Markov Model (HMM) is considered a powerful ML tool employed for modeling sequential data in several fields such as speech recognition[1], bioinformatics[2], and handwriting recognition[3].

Three distinct algorithms are employed for three fundamental problems in HMM. The decoding problem, which is the central focus of this study, is typically solved by the Viterbi algorithm. The algorithm gives the most probable sequence of underlying hidden states as its output given a sequence of observations [4]. However, the standard Viterbi algorithm, exclusively uses the past observations. Due to its inherent structure rooted in the Markov assumption, it potentially ignores the future observations. This causal dependency means that Viterbi cannot leverage potentially rich information contained in subsequent observations by its design. This aspect leads to suboptimal state estimations in scenarios where future context is available and relevant. This limitation naturally leads us to the following question: In the context of the decoding problem, could knowledge of observations a few steps into the future help us to determine more optimal and probable sequence of hidden states?

To find a proper answer to this question, we introduce a k-step lookahead decoding strategy that incorporates future knowledges when performing state estimations with the decoder algorithm. We suppose that utilizing information from the future in this manner will yield a more accurate estimation of the most likely hidden state sequence. However, a significant challenge arises in the computations: the computational complexity of the k-step lookahead calculation increases exponentially with the depth of k. Adapting a model that performs such expensive calculations to real-world applications will be difficult.

Effectively overcoming the computational complexity induced by the depth of the k value requires strategic method. Thus, we integrate the beam search pruning technique into the lookahead process. This method allows us to keep only the M most promising paths at each step, thereby aligning with the fundamental objective of decoding, which is to retain the path with the highest probability among the current possibilities. The proposed method offers a scalable framework that adeptly balances accuracy and computational efficiency.

Our work is determined by combining k-step lookahead with beam pruning in a lightweight and efficient manner. It is specifically applied for HMM decoding without the need for retraining or model restructuring.

1.1. Related Work

There are several works that have been proposed to improve decoding performance in Hidden Markov Models. The Viterbi algorithm [5] is the classical approach, finding the most likely sequence of hidden states based on given observations. However, this method does not exploit the information of future observations.

Lookahead decoding techniques have been investigated to address the use of future information. For instance, k-step lookahead strategies, as studied in [6] where the proposed decoder method uses future observations to improve prediction accuracy. The difference between this work and ours is positioning based on the lookahead module. Their approach depends on depth-first search and our study's lookahead is based on k-step discrete controller synthesis. We perform the decoding operation with a discrete controller for the decoding problem.

Beam search is a heuristic method widely used in natural language processing. [7] provides a practical solution by limiting the number of paths considered during decoding. Some works have applied beam search to sequence models [8,9]. However, in their works, integration with k-step lookahead has been limited.

Other approaches, such as particle filtering [10] and deep learning-based sequence models [11], have been studied for sequential decoding. Yet, these methods require large-scale computational resources.

1.2. Related Work

The remainder of this paper is organised as follows: Section 2 formally defines the problem, Section 3 describes the proposed method, Section 4 presents experimental evaluations, Section 5 discusses the results, and Sections 6 conclude with insights and future directions.

2. PROBLEM DEFINITION

An HMM is a well-known ML tool that provides a probabilistic framework for modelling sequential data, where transitions of the system between hidden states while emitting observable outputs [12]. The goal of HMM-based decoding is to determine the most probable sequence of hidden states that could have generated a given sequence of observations [5]. This is called *Decoding Problem* which is one of the classical problems of HMM. This is commonly solved using the Viterbi algorithm, which applies dynamic programming to identify the proper state path. However, the structure of the standard Viterbi is inherently causal [2]. It means that it only depends on the previous condition. Thus, its predictive capability is limited when future information is known and available. Yet, directly apply lookahead idea into the decoding process increases computational complexity dramatically. In this point, we need an efficient decoding strategy to utilize the future observations without suffering from exponential cost. Thus, we propose a novel decoding approach to addresses this problem.

2.1. Definition of Hidden Markov Model

Hidden Markov Model (HMM) is defined by the following components [13]:

- A finite set of hidden states : $S = \{s_1, s_2, \dots, s_N\}$,
A finite set of observations: $O = \{o_1, o_2, \dots, o_M\}$,
- A transition probability matrix: $A = [a_{ij}]$, where $a_{ij} = P(q_{t+1} = s_j \mid q_t = s_i)$
- An emission probability matrix: $B = [\beta_j(k)]$, where $\beta_j(k) = P(o_t = o_k \mid q_t = s_j)$,
- An initial state distribution: $\pi = [\pi_i]$, where $\pi_i = P(q_0 = s_i)$.

The goal is to decode the most probable hidden state sequence $Q = (q_0, q_1, \dots, q_T)$ depending on given observation sequence $O = (o_1, o_2, \dots, o_T)$.

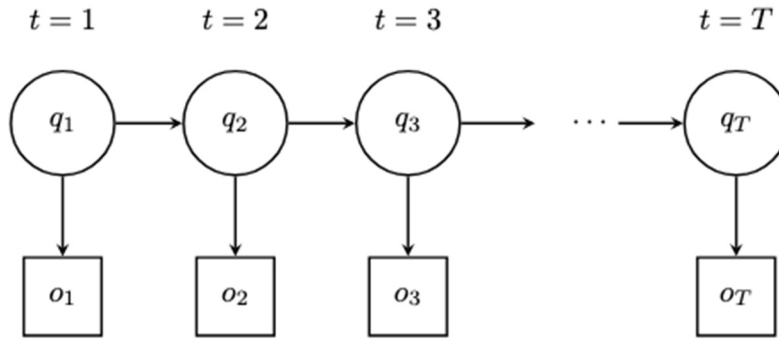


Figure 1. HMM structure showing transitions between hidden states q_t and corresponding emissions o_t

2.2. Standard Viterbi Algorithm

The standard decoding algorithm, Viterbi, is a dynamic programming approach to find the most probable hidden state sequence depending on given observation sequence. The formal computation is [14];

$$\hat{Q} = \underset{Q}{\operatorname{argmax}} P(Q|O) \quad (1)$$

It defines;

$$\delta_t(j) = \max_{q_1, \dots, q_{t-1}} P(q_1, \dots, q_{t-1}, q_t = s_j, o_1, \dots, o_t) \quad (2)$$

The steps of this algorithm are;

- Initial step: $\delta_1(j) = \pi_j b_j(o_1)$
- Recursion step: $\delta_t(j) = \max_i [\delta_{t-1}(i) a_{ij}] b_j(o_t)$

The optimal path is reconstructed using backtracking when $t=T$.

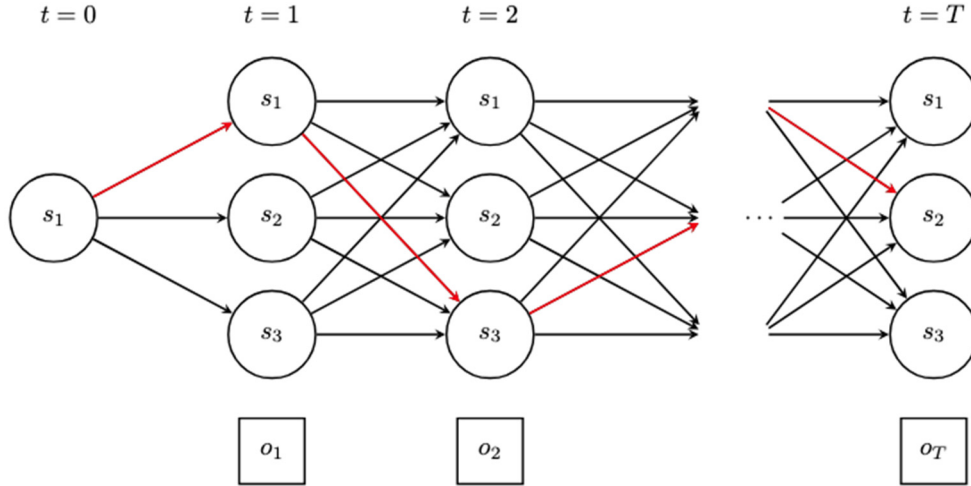


Figure 2. Viterbi decoding as dynamic programming over possible state paths that are illustrated with red arrows

2.3. Limitations and Definitions of the Proposed Model

The standard decoding algorithm on HMM is optimal under the assumption that only past informations (i.e. observations) are known [15]. Thus, it does not exploit the future observations during the process. On the other hand, while our proposed model facilitates the consider of future steps into the calculation, thus offering a novel strategy for decoding, extending the lookahead by k steps leads to an exponential increase in computational complexity.

The problem statement:

- **Given :**
 - A HMM (A, B, π)
 - An observation sequence $O = (o_1, \dots, o_T)$
 - A desired lookahead depth k
 - A beam width constraint M
- **Design a decoding**
 - Use k -step lookahead to consider future observations
 - Avoids exponential computation
 - Use near-optimal M value within practical runtime

This work proposes a novel decoding approach that directly addresses this problem via a beam-limited lookahead mechanism. The core challenge we tackle is utilizing k -step future observations efficiently while bounding computational complexity through beam pruning. The k -step lookahead strategy employed in this work can be designed as a form of k -step discrete control synthesis. This approach arises from the decision-making process at each time step. The system synthesizes a sequence of future hidden states where maximizes the overall likelihood of the given future observations. This process analogous to a controller optimizing future actions based on predicted states. Unlike traditional depth-first search (e.g. as [6]) methods that might explore all branches exhaustively, our approach implicitly guides this "synthesis" through probabilistic maximization. With this way, the proposed method prepares the beam pruning to efficiently manage the search space.

3. METHOD

In this section, we detail our proposed model, including definitions of the tools employed and their operational principles. As illustrated by the model's name, we will begin by defining the k-step lookahead. Next, we will discuss the application of beam search pruning to reduce the computational cost associated with the k-step. To clarify the operational principle, we will present a pseudocode representation and theoretically support the proposed strategy by providing relevant theorems and proofs.

3.1. k-Step Lookahead Formulation

HMM is traditionally designed to make sequential state estimations based only on past observations and the current hidden state. However, in this work, access to future observations can improve the accuracy of decoding. To utilize future information, we introduce a k-step lookahead strategy where at each step, future $\square$ observations are added into the state estimation process.

As mentioned the definition of HMM in Section 2.1, the objective is to maximize the joint probability of $P(q_t | o_1, \dots, o_T)$ at each time t . However, the proposed method instead maximizes the joint probability of $P(q_t, q_{t+1}, \dots, q_{t+k} | o_1, \dots, o_T)$. This joint probability can be expanded by using Bayes' rule and Markov assumptions :

$$P(q_t, q_{t+1}, \dots, q_{t+k}, o_t, o_{t+1}, \dots, o_{t+k}) \propto P(q_t) \prod_{i=0}^k \alpha_{q_{t+i}, q_{t+i+1}} \beta_{q_{t+i}}(o_{t+i}) \quad (3)$$

Here, multiple future states and observation symbols are considered during decoding at each time step.

3.2. Beam Search Pruning

The main challenge with k-step lookahead decoding is the exponential growth of the possible state paths. The computational complexity of the classical decoding algorithm is quadratic. However, the proposed method has N^k computation cost, and this causes the model to become unmanageable for practical applications. We integrate a Beam Search strategy to address this issue by limiting the number of paths explored at each step. It is introduced in two parts as;

Strategy:

- **Beam Width M** : Keep only M top paths with highest probabilities.
- **Pruning**: Discard the outside of the top M paths to reduce the computational complexity.

Processs:

- Expand the path by considering all possible next states for active paths at $t-1$.
- Calculate the scor for each extended path:

$$Score(path_j) = \prod_{i=0}^k \alpha_{q_{t+i}, q_{t+i+1}} \beta_{q_{t+i}}(o_{t+i}), \forall j = 1, \dots, N \quad (4)$$

- Sort the all scores.
- Keep only the top M paths for next expansion.

This strategy ensures that the calculation cost remains manageable while maintaining paths with high probabilities. The most significant value provided by the new strategy in this process is reducing the computational complexity. The exponential computation resulting from the classical k-step lookahead implementation is reduced to a manageable $N \times M^{k-1}$ with the new strategy where M is the Beam width. It is a small and feasible integer value such that $M \ll N^k$. The whole process is illustrated in Algorithm 1.

Algorithm 1: Beam-Limited k-Step Lookahead HMM Decoding	
Input:	
	• Transition matrix A • Emission matrix B • Initial probabilities π • Observation sequence $O = (o_1, o_2, \dots, o_T)$ • Lookahead depth k • Beam width M
Output:	
	Most likely state sequence $\hat{Q} = (\hat{q}_1, \hat{q}_2, \dots, \hat{q}_T)$
Initialization:	
	1. Create initial paths for each state : $P(s_i) = \pi_i b_i(o_i)$ 2. Keep top M paths depending on their probabilities.
Iteration for $t = 2$ to T:	
	1. For each active path at time $t - 1$: 1.a. Expand the path by considering all possible next states. 1.b. For each expansion, compute the cumulative path score by considering: $Score(path_i) = \prod_{j=0}^k A_{q_{t+j-1}, q_{t+j}} \times B_{q_{t+j}}(o_{t+j})$,where A represents the transition probability $P(q_{t+j} q_{t+j-1})$ and B represents the emission probability $P(o_{t+j} q_{t+j})$. Future emissions and transitions up to k steps are multiplied. 2. Collect all extended paths. 3. Sort the paths depending on their score in descending order. 4. Keep only the top M paths
Termination:	
	1. At $t = T$, select the path with max score 2. Output the corresponding state sequence $\hat{Q}$

3.3. Beam Width Selection

Selecting an appropriate beam width $\square$ is another important issue to handle for a good balance between decoding accuracy and computational efficiency in this work. Based on our experimental observations, a proper M value can be considered in the following aspects:

Trade-off Between Accuracy and Efficiency

- Small M values (e.g., $M=3$) yield faster runtime and less memory usage. However, it may discard promising paths and returns lower decoding accuracy.
- Large M values (e.g. $M=50$) preserve more promising paths and it improves the accuracy. But computational cost increases.
- The results suggest that moderate beam widths (e.g., $M=20$) typically provide a good trade-off, achieving high decoding accuracy with manageable runtime and memory usage.

Application-Specific Adaptation

- In real-time systems where runtime constraints are critical. A smaller beam width may be preferred for this kind of system. However, keep in mind, slight sacrifices are accepted for faster decoding.
- A larger beam width can be used to achieve maximum accuracy in the offline process without concern for runtime.

Adaptive Beam Width Adjustment

- Future implementations may benefit from the dynamic adjustment of $\square$ during decoding based on the following scenarios:
 - For instance, if score differences between paths are large, a smaller beam width might suffice. Thus, unnecessary calculations are discarded.
 - If many paths have close scores then expanding the beam width temporarily could prevent discarding optimal paths. In this scenario, the computational cost increases but the protection of the paths that have the highest probability among the competing paths is secured.

The precise adjustment of the beam width may vary depending on the available computational resources and application requirements. For instance, the following optimal M values are provided as examples based on the number of states utilized in two distinct models. It is probable that these values will exhibit variability according to the specific task being addressed.

4. IMPLEMENTATION DETAILS

In this section, we describe practical reviews that were included in the implementation of the proposed Beam-Limited k -Step Lookahead HMM decoding algorithm. These details are important for establishing stability and computational efficiency.

The procedure of the proposed algorithm is shown as a flowchart in Figure 3. Here, the k -step value is reduced due to the potential event of a condition such as $T-t < k$. In other words, the value of k is dynamically decreased based on the remaining number of steps. Another adjustment made to enhance stability is the utilization of logarithmic probabilities. It will improve numerical stability by using summation instead of multiplication.

4.1. Dynamic Adjustment of Lookahead Depth

The classical structure of standard k -step lookahead decoding processes over k future observations at every time step. However, towards the end of the given observation sequence, specifically when the remaining number of time steps ($T-t$) is less than the desired lookahead depth (k), the algorithm cannot look ahead by the full k steps. To handle this issue and ensure the algorithm always considers only available future observations:

- The k -step depth is dynamically adjusted:

$$k_{effective} = \min(k, T - t) \quad (5)$$

at each time step t . This adjustment ensures that as the decoding process approaches the end of the observation sequence: the lookahead depth gracefully decreases, preventing out-of-bounds access and maintaining computational accuracy. At the final time step, the decoding process secures only the available number of future observations that are added into the calculations of the path score. The stability is preserved in this way.

- At the final time step, the decoding process secures only available number of the future observations that are added into the calculations of path score. The consistency is maintained with this way.

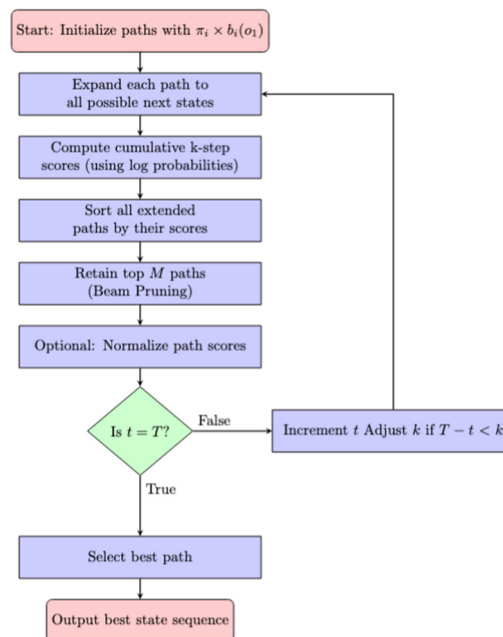


Figure 3. The flowchart of proposed method

4.2. Logarithmic Probability Computation

The result of direct multiplication may go underflow due the path score include the product of multiple transition and emission probabilities. Even, long sequences or small probabilities are also affect the result goes underflow. To address this:

- All probability products are calculated in logarithmic.
- Then, we sum logarithmic results:

$$\log(\Pi_i(p_i)) = \sum_i(\log(p_i)) \quad (6)$$

- The cumulative score calculation for a path:

$$\log Score(path) = \sum_{j=0}^k \left(\log a_{q_{t+j-1}, q_{t+j}} + \log b_{q_{t+j}}(o_{t+j}) \right) \quad (7)$$

- This process provides numerical stability and improves computational robustness on path scores without changing the order of path.

4.3. Pruning Strategy

After computing the score of all extended paths we need to update the pruning strategy depending on k-step adjustment and logarithmic computation. Thus,

- Path are sorted depending on their logarithmic calculated scores.
- Only save the top M paths.
- Any ties are solved arbitrarily or by chosen paths with fewer transitions. It is done depending on the application scenario.

This simple but effective pruning strategy makes the decoder remain computationally efficient while maintaining high-quality path candidates.

5. EXPERIMENTS

5.1. Experimental Setup

In this section, we evaluate the performance of the proposed decoding algorithm: Beam-Limited k-Step Lookahead HMM. We use synthetic datasets and varying algorithmic parameters to design a series of controlled experiments.

5.1.1. Data Generation

The synthetic data that is used in this work was created with an original HMM by the following parameters:

- The number of hidden states (N) is 5.
- The number of emission symbol (M) is 6.
- Transition matrix A is generated randomly and row-stochastic. Specifically, elements were initially sampled from a uniform distribution between 0 and 1, and each row was then normalized to sum to 1.
- Emission matrix B is generated randomly and row-stochastic. Similarly, elements were sampled from a uniform distribution between 0 and 1, and each row was then normalized to sum to 1.
- The initial state distribution π is randomly uniformly distributed over all hidden states.

All observation sequences were generated by simulating A and B matrices according to parameters of HMM. 1000 independent observation sequences were generated and the length of each is 100.

5.1.2. Algorithm Parameters

The algorithm parameters were varied systematically during evaluation as;

- k depth value: {1,2,3,4}
- M width: {5,10,20,50}

In testing part, each combination of (k,M) was tested to show the effects of different levels of lookahead and pruning depth on the decoding.

5.1.3. Evaluation Metrics

The metrics that are used to estimate the performance of the algorithm are:

- **Decoding Accuracy:** The quantity of correct decoded hidden states compared to the original hidden state sequence.
- **Runtime:** The average of time that is required to decode a sequence.
- **Memory Usage:** Calculate the memory consumption during decoding to measure using profiling tools.

We provide a comprehensive view of the trade-off between accuracy, computational time, and memory efficiency with these metrics.

5.1.4. Computational Environment

The following system specifications were used to conduct all experiments:

- Intel Core i7-12700H CPU
- 16 GB RAM
- Python 3.12, Numpy, SciPy
- Windows 11

We averaged the timing measurements over 10 independent runs for each experimental configuration to decrease variability.

5.2. Computational Environment

As mentioned in the previous section, the generated synthetic dataset was used to compare the performance of both the proposed and classical (i.e. Viterbi) decoding algorithms. We performed variety of tuple lookahead depth k and beam width M on decoding process to illustrate metric performance over on accuracy, runtime, and memory consumption.

5.2.1. Decoding Accuracy

Figure 4 illustrates the correlation between decoding accuracy and beam width (M) for varying lookahead depths (k). A stable trend across all tested k values reveals that increasing the beam width leads to a corresponding improvement in decoding accuracy. Especially, at limited beam widths, such as M=5, an observable drop in accuracy is evident when comparing the performance of full k-step lookahead decoding. However, even with quite expanded beam widths (e.g., M=20), we achieved accuracy levels closely approximating (within a 2-3% margin) those obtained by full lookahead decoding. Furthermore, the results show that greater lookahead depths (k=3,4) consistently yield superior decoding accuracy compared to casual lookahead depths (k=1,2). Thus, the advantage of containing future observations is underlined in the decoding process.

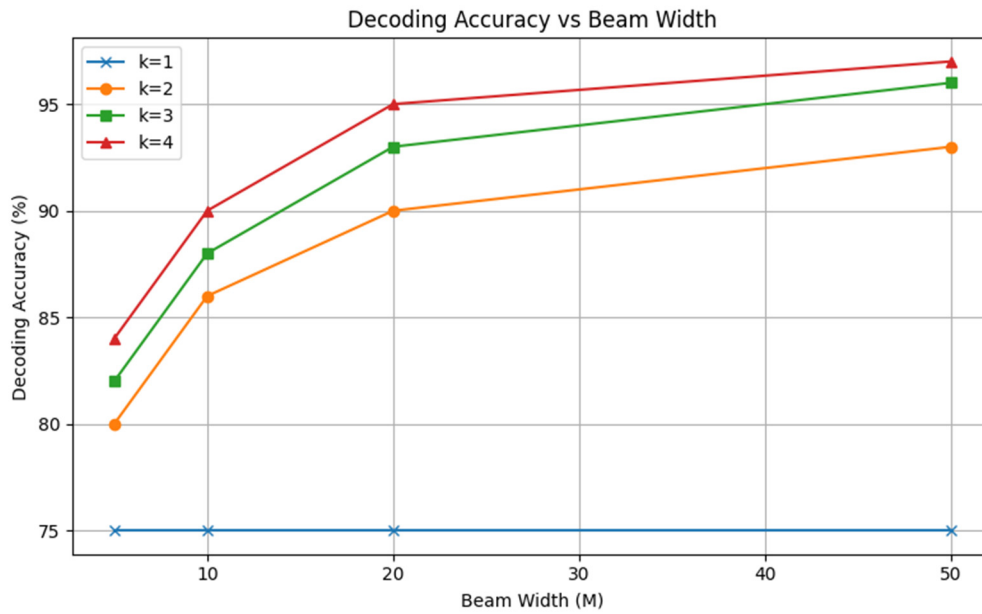


Figure 4. Decoding accuracy vs beam width

The corresponding graph visually represents this analysis by plotting beam width (M) on the x-axis and decoding accuracy (in percentage) on the y-axis with distinct lines defining the performance for each k value. This visualization effectively demonstrates the trade-off between computational efficiency (influenced by beam width) and decoding accuracy for different levels of future consideration.

5.2.2. Runtime Performance

Figure 5 presents the analysis of the average runtime per sequence. Here, again, we examine the relationship with varying beam widths (M) and lookahead depths (k). As expected, the data consistently exhibits that increasing the beam width leads to a corresponding rise in runtime. In particular, the full k -step lookahead decoding where executed without any pruning, shows the highest computational cost. On the contrary, the proposed method, Beam-Limited k -Step Lookahead, significantly shorten runtime when compared to the exhaustive search approach. This advantage becomes more notable at higher values of k .

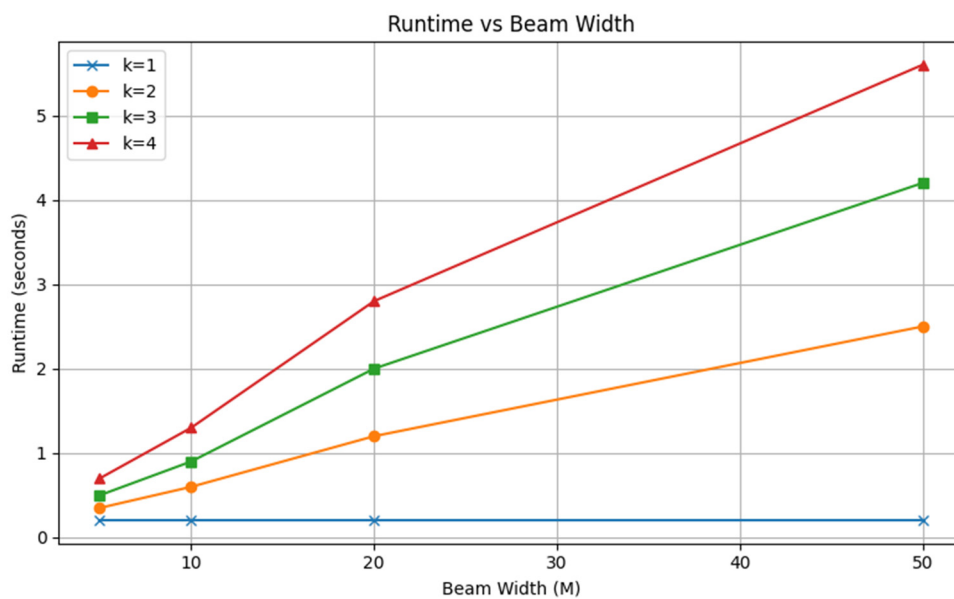


Figure 5. Runtime vs beam width

This valuable reduction in runtime is directly pointed to the effective avoiding of exponential growth in computation through the implementation of pruning techniques. The graph visually supports these observations, plotting the Beam Width (M) on the x-axis against the Average Runtime (in seconds) on the y-axis. Here, we see the efficiency obtained by the proposed method.

5.2.3. Memory Usage

Figure 6 illustrates the memory consumption depending on varying beam widths (M) for different lookahead depths (k). A key result is that the Beam-Limited decoding strategy significantly reduces memory usage when compared to a full lookahead decoding approach. While an increase in beam width (M) does cause higher memory consumption, the process remains well within feasible limits for practical values (e.g., $M \leq 20$).

This characteristic makes the Beam-Limited approach particularly important. Using a smaller M width for applications, such as real-time or embedded systems, which are operated under strict memory constraints presents a pretty trade-off due to balancing computational performance with efficient resource utilization.

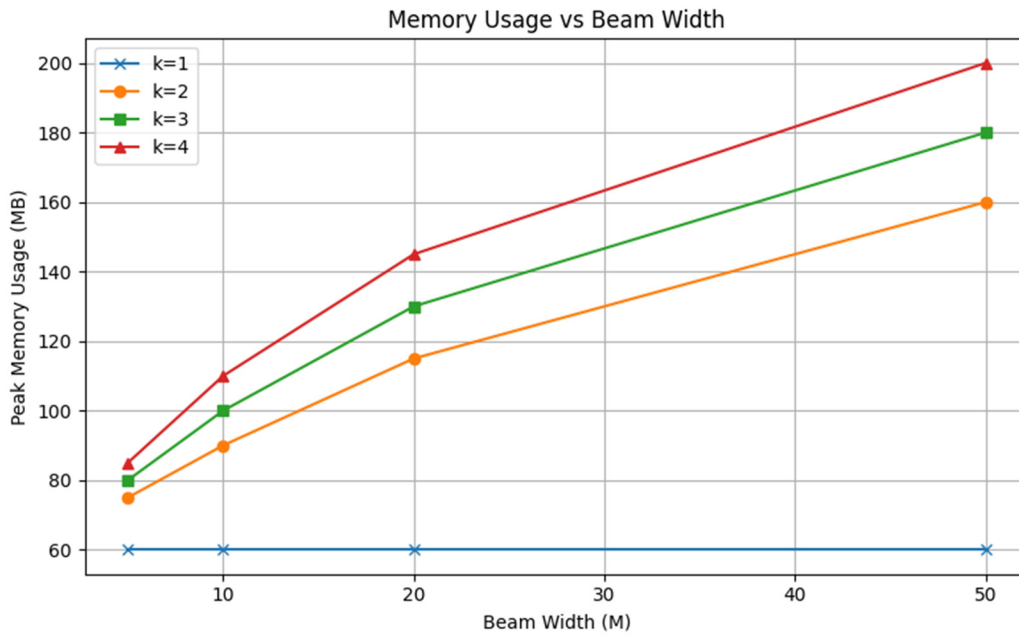


Figure 6. Memory usage vs beam width

The graph presents these relationships, illustrating peak memory usage (in MB) on the y-axis against beam width (M) on the x-axis. The separated lines show the memory profiles for each k value. This spots the practical benefit of beam pruning in optimizing memory traces without strongly compromising accuracy.

5.2.4. Trade-Off Analysis

Figure 7 effectively summarizes the complex trade-off between decoding accuracy and runtime across various beam widths (M). The plot readably reveals a smooth curve to illustrate where the beam width increases while decoding accuracy improves although at the cost of increased runtime. An obvious "sweet spot" comes out around $M=20$. This means an optimal balance where the algorithm achieves an exceptional level of accuracy without suffering high computational expense.

These findings have significant practical inferences as the beam width value can contribute as a crucial mechanism where it can be adapted as a decoder behaviour to specific application constraints. If the priority is speed or maximal accuracy then adjusting the beam width can allow for flexible optimization.

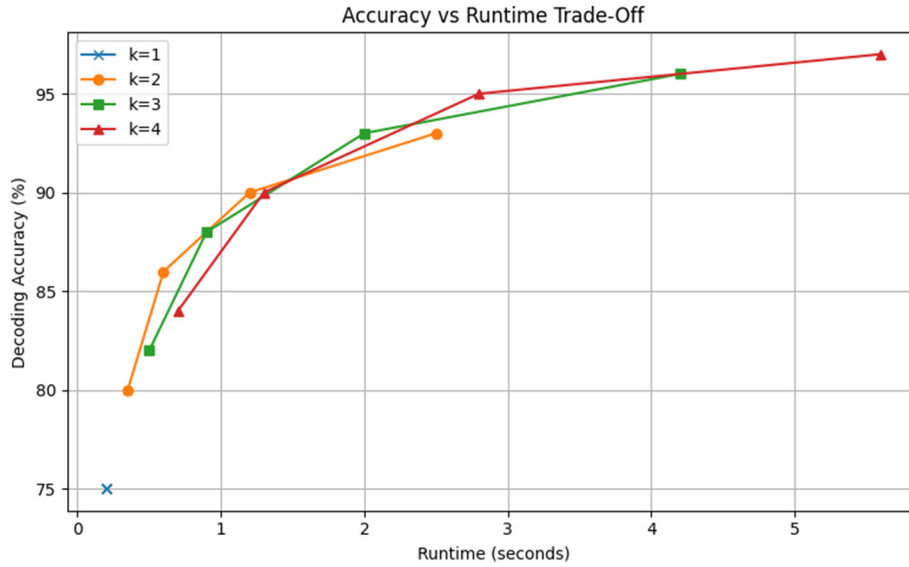


Figure 7. Accuracy vs runtime trade-off

The graph shows the terms which are runtime on the x-axis and decoding accuracy on the y-axis. Each point on the curve represents a different beam width value. This graph provides a clear roadmap for researchers to navigate the performance aspect and they can make proper decisions based on their system's requirements.

6. DISCUSSION

In this work, we proposed a new decoding strategy that is called Beam-Limited k -Step Lookahead for Hidden Markov Models (HMMs). This strategy uses the future information to find the most proper path in decoding. The calculation of the future information could increase the complexity, so we aim to balance decoding accuracy and decrease computational cost in this work. In this section, we discuss our experimental findings, limitations of the strategy, and directions for future work.

6.1. Insights from Experimental Results

We got several important trends to handle from the experiments, such as:

- **Improve accuracy with k -step lookahead:** As known that the classical decoding algorithm (i.e. lookahead value $k=1$) of HMM has no attribute to consider the future information in decoding processs, so we consider the beyond the $k=1$ step to anticipate future observations and increasing more informed decisions.
- **Control parameter - Beam Width (M):** This parameter plays a key role in the trade-off between accuracy and computational complexity. For instance, if faster decoding is required a small M value is enough, but, on the other hand, the result may be in suboptimal paths. If accuracy is important, then a larger beam width should be used; however, increased runtime must be accepted.
- **Classical Viterbi vs k -Step:** The proposed method with modest value k -Step and moderate beam width substantially outperforms classical Viterbi decoding (i.e. $k=1$) in terms of accuracy and validating the value of future observations.
- **Memory efficiency:** The proposed decoding algorithm has slightly increased memory usage according to standard decoding algorithm. However, the Beam-Limited approach remains practical and scalable for reasonable beam width values (i.e. $M \leq 20$).

6.2. Limitations

The proposed method has an exponential complexity of k -step lookahead decoding. But it still takes over the limitations such as:

- **Beam width parameter:** The value of beam width M is key parameter. Very small M may cause early pruning of optimal paths and this causes weak accuracy.
- **Dynamic environments:** The benefit of lookahead may be decreased while the environment is highly dynamic or non-stationary where the unpredictability of future information.
- **Scalability:** While the beam pruning effectively reduces the exponential computational cost of full k -step lookahead, the scalability of the proposed method may still present significant computational and memory challenges for extremely large-scale HMMs. In such scenarios, even with beam-limiting, the number of paths to manage and the computations per step can become restrictive. Addressing these limitations for ultra-large HMMs might require a combination of advanced optimization strategies beyond the current scope. These could include:
 - Distributed Computing: Using distributed systems to parallelize the path expansion and scoring across multiple processing units or nodes [16].
 - Hierarchical HMMs (HHMMs): For very complex systems, adopting hierarchical HMM structures could reduce the effective state space at each level, thereby simplifying the decoding problem.
 - Approximation Techniques: Exploring more aggressive approximation or early exit strategies within the beam search when confidence in a path becomes exceptionally high.
 - State Aggregation/Reduction: Pre-processing techniques to reduce the number of effective hidden states if the application allows for some loss of granularity.

While these approaches are beyond the focus of the current work, these strategies can represent crucial directions for extending the applicability of beam-limited lookahead decoding to highly complex real-world systems.

6.3. Future Work

Future research can be emerged from this work. There are several directions such as:

- **Adaptive beam width:** The beam width value M can be adjusted during decoding based on the confidence score.
- **Parallelization:** GPU based expansion could dramatically increase the fast of decoding.
- **Robustness in noise:** Research the robustness of the proposed decoding stragey under the noisy observations could be the next step of this work.

7. REFERENCES

1. Siddalingappa, R., Hanumanthappa, P. & Reddy, M. (2018). Hidden markov model for speech recognition system - a pilot study and a naive approach for speech-to-text model. *Advances in Intelligent Systems and Computing*, 77-90.
2. Li, J., Lee, J.Y. & Liao, L. (2021). A new algorithm to train hidden markov models for biological sequences with partial labels. *BMC Bioinformatics*, 22, 162.
3. Brakensiek, A. & Rigoll, G. (2004). Handwritten address recognition using hidden markov models. In: Dengel, A., Junker, M., Weisbecker, A. (eds) *Reading and Learning. Lecture Notes in Computer Science*, 2956. Springer, Berlin, Heidelberg.
4. Tataru, P., Sand, A., Hobolth, A., Mailund, T. & Pedersen, C.N. (2013). Algorithms for hidden markov models restricted to occurrences of regular expressions. *Biology (Basel)*, 2(4), 1282-1295.
5. Rabiner, L.R. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257-286.
6. Wang, Y.S., Kuo, Y.L. & Katz, B. (2020). Investigating the decoders of maximum likelihood sequence models: a look-ahead approach. *arXiv preprint*.
7. Graves, A. (2012). Sequence transduction with recurrent neural networks. *arXiv preprint*. arXiv: 1211.3711.
8. Hannun, A. (2017). *Sequence modeling with CTC*. <https://distill.pub/2017/ctc>.
9. Chorowski, J. & Jaitly, N. (2016). Towards better decoding and language model integration in sequence to sequence models. *arXiv preprint*. arXiv:1612.02695.
10. Doucet, A., Godsill, S. & Andrieu, C. (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10, 197-208.

11. Graves, A., Mohamed, A.-R. & Hinton, G. (2013). Speech recognition with deep recurrent neural networks. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6645-6649.
12. Xiaomin, L., Thomas, H. & Colin, C. (2023). Estimating hidden markov models (HMMs) of the cognitive process in strategic thinking using eye-tracking. *Frontiers in Behavioral Economics*, 2, 1225856.
13. Manouchehri, N. & Bouguila, N. (2023). Human activity recognition with an HMM-Based Generative Model. *Sensors*, 23(3), 1390.
14. Saize, S. & Yang, X. (2024). On the definitions of hidden markov models. *Applied Mathematical Modelling*, 125, 617-629.
15. Kurucan, M. & Wojtczak, D. (2024). The utilization of single-counter systems featuring final terminals with non-zero counter values. *Cukurova University, Journal of the Faculty of Engineering*, 39(4), 999-1014.
16. Tüfekçi, Z. & Dişken, G. (2022). Fast computation of parameters of the random variable that is logarithm of sum of two independent log-normally distributed random variables. *Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi*, 37(1), 261-270.

Portakal Kabuğu ve Portakal Ağacı Yaprağından Elde Edilen Ekstraktlarla Standart Çok Lifli Kumaşın Boyanması

Sabiha SEZGİN BOZOK^{1,a}, R. Tuğrul OĞULATA^{1,b}

¹Çukurova Üniversitesi, Mühendislik Fakültesi, Tekstil Mühendisliği Bölümü, Adana, Türkiye

^aORCID: 0000-0002-8464-0405; ^bORCID: 0000-0003-2783-5246

Makale Bilgileri

Geliş : 10.07.2025

Kabul : 18.08.2025

DOI: 10.21605/cukurovaumfd.1738580

Sorumlu Yazar

Sabiha SEZGİN BOZOK

sbozok@cu.edu.tr

Anahtar Kelimeler

Doğal boyarmadde

Portakal kabuğu

Portakal ağacı yaprağı

Renk analizi

Spektrofotometre

Atf şekli: SEZGİN BOZOK, S., OĞULATA, R.T., (2025). Portakal Kabuğu ve Portakal Ağacı Yaprağından Elde Edilen Ekstraktlarla Standart Çok Lifli Kumaşın Boyanması. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 559-568.

ÖZ

Çalışmada, portakal kabuğu ve portakal ağacı yaprağından ekstraktlar elde edilmiştir. Elde edilen doğal boyarmaddelerin lif yüzeyine bağlanabilmesi için farklı oranlarda alüminyum sülfat ($Al_2(SO_4)_3 \cdot 8H_2O$) ile yapılan mordanlama, boyama ile eşzamanlı olarak gerçekleştirilmiştir. Hazırlanan karışımlarla yün, akrilik, polyester, poliamid, pamuk ve asetat içeren çoklu lif kumaşı renklendirilmiştir. Boyama sonrası, spektrofotometre ile yapılan ölçümler yazılım aracılığıyla değerlendirilmiş; renk kuvveti (K/S), kırmızılık-yeşillik (a^*), sarılık-mavilik (b^*) ve açıklık-koyuluk (L^*) değerleri analiz edilmiştir. Tüm numunelere yıkama uygulanmış ve yıkama sonrası toplam renk farkı (ΔE) değerleri elde edilmiştir. Lif türleri arasında en yüksek renk verimi yün ve poliamid içeren numunelerde gözlemlenmiştir. Yünlü kumaşlarda en yüksek K/S değeri, %1 oranında alüminyum sülfat içeren portakal kabuğu ekstraktı ile elde edilirken; poliamid kumaşlarda ise %3 oranında alüminyum sülfat içeren portakal ağacı yaprağı ekstraktı ile en yüksek renk verimi sağlanmıştır. Bu numunelerde yıkama haslığı da diğer liflere göre daha yüksek bulunmuştur.

Dyeing of Standard Multifiber Fabric with Extracts Obtained from Orange Peel and Orange Tree Leaves

Article Info

Received : 10.07.2025

Accepted : 18.08.2025

DOI: 10.21605/cukurovaumfd.1738580

Corresponding Author

Sabiha SEZGİN BOZOK

sbozok@cu.edu.tr

Keywords

Natural dyestuff

Orange peel

Orange tree leaf

Color analysis

Spektrofotometer

How to cite: SEZGİN BOZOK, S., OĞULATA, R.T., (2025). Dyeing of Standard Multifiber Fabric with Extracts Obtained from Orange Peel and Orange Tree Leaves. Çukurova University, Journal of the Faculty of Engineering, 40(3), 559-568.

ABSTRACT

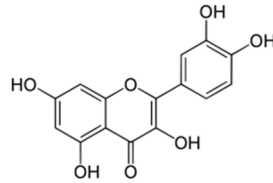
In the study, extracts were obtained from orange peels and orange tree leaves. Mordanting with different proportions of aluminum sulfate ($Al_2(SO_4)_3 \cdot 8H_2O$) was carried out simultaneously with dyeing to ensure the binding of the obtained natural dyes to the fibers. Multifiber fabrics containing wool, acrylic, polyester, polyamide, cotton, and acetate were dyed using the prepared dyeing solutions. After dyeing, measurements taken with a spectrophotometer were evaluated via software, and color strength (K/S), redness-greenness (a^*), yellowness-blueness (b^*), and apparent darkness (L^*) values were analyzed. All samples were washed, and total color difference (ΔE) values were obtained after washing. According to the results, the highest color yield among the different fiber types was observed in samples containing wool and polyamide. While the highest K/S value in wool fabrics was obtained with orange peel extract containing 1% aluminum sulfate, the highest color yield was achieved in polyamide fabrics with orange tree leaf extract containing 3% aluminum sulfate. Washing fastness in these samples is also higher than in other samples.

1. GİRİŞ

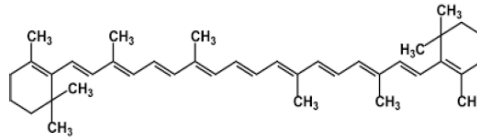
Günümüzde çevresel sürdürülebilirliğe olan ilgi arttıkça, doğal boyarmaddeler ve doğal boyama yöntemleri de tekrar popülerlik kazanmıştır. Bitkiler, hayvanlar ve minerallerden elde edilen doğal boyarmaddeler, sanayi devrimi öncesinde tekstil ürünlerinin renklendirilmesinde yaygın olarak kullanılmıştır. Ancak, sentetik boyarmaddelerin keşfi ve yaygınlaşmasıyla doğal boyarmaddelerin kullanımı büyük ölçüde azalmıştır. Son yıllarda, çevresel ve sağlıkla ilgili kaygılar, doğal boyarmaddelerin yeniden popülerlik kazanmasına neden olmuştur. Sentetik boyarmaddelerin üretimi ve kullanımı sırasında ortaya çıkan kimyasal atıklar ve bu boyarmaddelerin insan sağlığı üzerindeki olumsuz etkileri, tüketicileri ve üreticileri daha sürdürülebilir ve ekolojik alternatif olan doğal boyarmaddelere tekrar yönlendirmiştir. Meyve kabuğu ve yaprak gibi bitkisel atıkların tekstil sektöründe boyar madde olarak değerlendirilmesi, ekolojik açıdan önemli avantajlar sunmaktadır. Ayrıca bazı doğal kaynaklı boyarmaddeler, yüzeylere antibakteriyel özellik kazandırarak insan sağlığı açısından da ek faydalar sağlayabilmektedir [1].

Narenciye ürünleri de doğal boyarmadde kaynağı olarak farklı çalışmalarda kullanılmıştır. Portakal ağacı yaprakları ve portakal kabukları hem evsel ölçekte hem de meyve suyu üreticileri gibi gıda işleme endüstrisinden atık olarak büyük ölçekte toplanabilen ürünlerdir [2]. Bu da tekstil endüstrisinde kullanılabilmesi için büyük bir avantaj oluşturmaktadır. Özellikle bazı çalışmalarda portakal kabuğunu oluşturan bileşenlerden olan karotenoidler ile tekstil ürünlerine turuncumsu-sarı renkler kazandırılabilirdiği belirtilmiştir [3].

Benli ve Bahtiyari [4], viskon içerikli kumaşları renklendirmek için portakal ağacı yapraklarından (Şekil 1) iki farklı konsantrasyonda (25 ve 50 g/L) ekstraktlar elde etmişlerdir. Farklı mordanlar ile kumaşlara ön mordanlama işlemi uygulanmıştır. Sonuçlarda portakal ağacı yaprağının doğal boyarmadde elde etmek için uygun bir bitkisel kaynak olduğu, farklı mordanlar ile kumaşta farklı tonlarda renk elde edilebileceği tespit edilmiştir. Tayyab ve arkadaşları [5] yaptıkları çalışmada portakal kabuğundan ekstrakt elde ederek %100 lyosel içerikli kumaş numunelerini renklendirmişlerdir. Bu numuneler boyama öncesi demir sülfat ve bakır sülfat mordanları (%2 ve %4 konsantrasyonda) ile ön işlem görmüştür. Demir sülfat ile mordanlanan %4 konsantrasyonda ekstrakt ile renklendirilmiş numunelerin renk kuvveti (K/S) değerlerinin diğer numunelere göre daha yüksek olduğu görülmüştür. Edeen [6] portakal kabuğundan ekstrakt elde etmiş ve bu ekstraktı kullanarak emdirme yoluyla pamuklu kumaşları renklendirmiştir. Portakal kabuğunun yapısındaki renk verici bileşenlerden olan β -karoten (Şekil 2) propanon ile ayrıştırılıp boyarmadde olarak kullanılmıştır. Boyama öncesi ekstrakta sodyum alginat ve sodyum sülfat eklenmiştir. Sonuçlarda β -karotenin pamuklu kumaşa yeşilimsi sarı rengi kazandırdığı, boyanan numunelerin renk kuvvetinin yüksek, renk haslıklarının düşük olduğu görülmüştür.



Şekil 1. Portakal ağacı yaprağındaki flavonoid bileşenlerinden kuersetinin yapısı [7]



Şekil 2. Portakal kabuğunda bulunan β -karotenin yapısı [6]

Başka bir çalışmada portakal kabuğu ve limon kabuğundan elde edilen doğal boyarmaddeler ile pamuklu kumaşlar renklendirilmiştir [8]. Ekstraksiyon için organik çözücü olarak etanol kullanılmıştır. Kurutulmuş kabuklardan öğütülerek elde edilmiş parçacıklar farklı konsantrasyonda çözücü ile karıştırılmıştır. %30 mordanla işlem gören numunelerin boyarmadde alımının en fazla orana yaklaştığı, hem limon hem de portakal kabuğundan elde edilen boyarmadde ile yüksek renk veriminin elde edildiği tespit edilmiştir. Etanol ve kurutulmuş kabuk içeren 1:1 banyo oranıyla, kuru ve yağ sürtünmeye karşı renk haslığında en iyi sonuç elde edildiği ortaya konmuştur.

Taura ve arkadaşları [9], portakal kabuğundan doğal boyarmadde elde ederek pamuklu kumaşları renklendirmişlerdir. Konvansiyonel mordanlar yerine kumaşların boyanabilirliğini artırmak için sodyum hidroksit ile ortam pH seviyesini (3.34, 3.36, ve 3.34) değiştirmiştir. Boyamalarda sodyum hidroksit konsantrasyonu ve sıcaklık arttıkça numunelerin boya alımının arttığı sonucuna varmışlardır. Sodyum hidroksitin boyalı numunelerde dış etkiler sonucu solmasını azalttığını tespit etmişlerdir.

Portakal kabuğundan elde edilen doğal boyarmadde ile pamuklu kumaşların renklendirilmesi için mordan madde olarak tanik asitin kullanıldığı bir çalışmada mordanlama zamanı (boyama öncesi, sonrası ve boyama ile eş zamanlı), mordan madde konsantrasyonu, boyama sıcaklığı ve boyama süresi değişken parametre olarak belirlenmiştir [10]. Çalışma sonucunda tanik asitin doğal boyarmaddenin renk verimini artırdığı, boyama sırasında eklenen tanik asitin boyama süresinin ve sıcaklığın renge etkisinin düşük olduğu tespit edilmiştir. Diğer bir çalışmada, portakal kabuğu artıkları ile yine doğal boyarmadde elde edilmiş ve pamuklu kumaşlar renklendirilmiştir [11]. Ekstraksiyon için hekzan, aseton ve etanol olmak üzere üç farklı çözücü kullanılmıştır (50 °C'de 60 dk boyunca). Mordan madde olarak alüminyum ve bakır sülfat metal tuzları kullanılmıştır. Kabuktaki renklendirici bileşen olan karetonidin elde edilmesi bakımından çözücülerin uygunlukları azalan sırayla aseton, hekzan ve etanol şeklinde tespit edilmiştir. Bakır sülfat kullanımıyla ise en yüksek renk kuvveti (K/S) değeri elde edilmiştir. Ancak bakır sülfat üzerine yapılmış çalışmalarda hem ekolojik olarak hem de insan sağlığı açısından zararlı olduğu vurgulanmaktadır [12,13].

Doğal kaynaklardan elde edilen boyarmadde ile elyaf arasındaki afiniteyi sağlamak için metal tuzlar (metalik mordanlar), tanin ve tanik asitler veya yağ bazlı mordanlar kullanılabilir [14]. Metal tuzlarından olan alüminyum sülfat diğer bakır, krom bazlı mordanlara göre daha ekolojik ve güvenli bir üründür. Bir araştırmada alüminyum sülfat pamuklu kumaşları mordanlaması için kullanıldığında sadece %20-22 oranında kumaş tarafından emildiği geri kalanının ise atık su haline geldiği tespit edilmiştir [15,16]. Ancak alüminyum tuzların, toksik olmaması ve çevreye zararının neredeyse hiç olmaması bu kimyasal mordan madde olarak tercih edilebilir bir ürün haline getirmektedir. Bu sebeplerle bu çalışmada alüminyum sülfat ($Al_2(SO_4)_3 \cdot 18H_2O$) mordan madde olarak seçilmiştir.

Daha önce yapılmış çalışmalarda pamuk, rejenere selüloz ve yün haricinde başka hammadde içeren tekstil ürünlerine portakal ürünlerinden elde edilmiş ekstraktların renklendirme etkisinin karşılaştırılmadığı tespit edilmiştir. Bu çalışmada hem portakal kabuğu hem de portakal ağacı yaprağından doğal boyarmadde elde edilmiştir. Bu boyarmaddelerin yün, akrilik, polyester, poliamid, pamuk ve asetat gibi farklı hammadde içeren kumaşlardaki renklendirici etkisi karşılaştırılmıştır.

2. YÖNTEM

2.1. Materyal

Çalışma kapsamında renklendirme için kullanılan çoklu lif içeren standart kumaş (multifibre bezi) sırasıyla asetat (filament), pamuk (ağartılmış), poliamid (nylon 6.6), polyester, akrilik ve kamgarn yün içermektedir [17]. Bu kumaşın atkı ipliği tek bir cins elyaftan (pamuk) oluşup, çözgü ipliği farklı lif türlerinden meydana gelmektedir. Kumaş şeritler halinde farklı lif içeren bölgelerden oluşmaktadır.

Doğal boyarmaddenin kumaşa bağlanması için alüminyum sülfat ($Al_2(SO_4)_3 \cdot 18H_2O$, CAS No: 10043-01-3, 98%) metal tuzu kullanılmıştır. Portakal kabuğu ve yaprakları Adana ili bölgesinden Mayıs ayında toplanmıştır. Toplandıktan sonra etüv içerisinde 90 °C'yi geçmeyecek şekilde 3 saat boyunca kurutulmuştur. Kuruduktan sonra ekstraksiyon işlemine kadar gölgede bekletilmiştir.

2.2. Metot

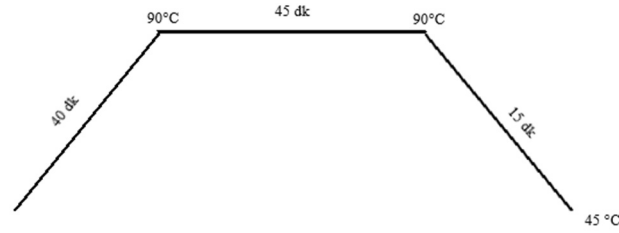
Ekstrakt hazırlamak için kurutulan portakal kabuğu ve yaprakları küçük parçalar (yaklaşık 0,5 cm kalınlıkta) haline getirilmiştir. Hem portakal kabuğu hem de portakal ağacı yaprağı 30 g/L konsantrasyonda olacak şekilde saf su ile karıştırılmıştır. Karışımlar manyetik karıştırıcı özelliği olan ısıtıcılarda tutulmuştur. Renklendirici bileşenlerin suya geçmesi için bu karışımlar 60 dk boyunca ısıtıcı üzerinde kaynatılmıştır. Bu işlem tamamlandıktan sonra filtre kağıdı yardımıyla parçacıklar süzülüş ve ekstrakt elde edilmiştir. Kumaşlar 5 cm (~1,5 g) ene sahip olacak şekilde kesilerek numuneler hazırlanmıştır. Boyama işlemi için kızılötesi ile çalışan laboratuvar tipi boyama makinesi (Termal) kullanılmıştır. Tüplerin içine 150 ml ekstrakt, mordan madde ve kumaş numunesi eklenmiştir. Banyo oranı yaklaşık 1:100 şeklindedir.

Alüminyum sülfat miktarının numunelerdeki renklendirici etkisini görmek için farklı oranlarda tüplere ilave edilmiştir. Numune kodları ve özellikleri Çizelge 1’de sunulmuştur.

Çizelge 1. Numune kodları ve özellikleri

Numune kodu	Ekstrakt içeriği	Alüminyum sülfat oranı
K-M1	Portakal kabuğu	%1
K-M2	Portakal kabuğu	%2
K-M3	Portakal kabuğu	%3
Y-M1	Portakal ağacı yaprağı	%1
Y-M2	Portakal ağacı yaprağı	%2
Y-M3	Portakal ağacı yaprağı	%3

Dolu boyama tüpleri cihaza yerleştirildikten sonra uygulanan boyama işleminin ayrıntılarını sunan boyama programı grafiği Şekil 3’te sunulmuştur.



Şekil 3. Boyama programı grafiği

Cihaz durdurulduktan sonra tüplerden çıkarılan numuneler su ile durulanmış ve kurumaya bırakılmıştır. Doğal boyalı numunelerin renk ölçümleri için Minolta CM 3600 model spektrofotometre ve renk analizleri için RealColor1.3® yazılımı kullanılmıştır. Ölçümler 400–700 nm dalga boyları arasında, D65 gün ışığında ve 10° gözlemci açısı ile yapılmıştır. Ölçümler sırasında parlaklık bileşeni (Speküler bileşen dahil, SCI) yansımaya dahil edilmiştir. Boyanan numunelerin renk verimini belirlemek için Kubelka-Munk eşitliği (1) ile 400–700 nm görünür spektrum dalga boyları arasında ölçüm yapılarak elde edilen K/S (renk kuvveti) değerleri hesaplandı. Denklemdaki 'R' örneklerin yansıma değerini temsil etmektedir. K/S değerleri numunelerin maksimum absorpsiyon dalga boyundaki R değeri esas alınarak hesaplanmıştır. Ayrıca renkli numunelerin CIELab renk sistemine göre 3 boyutlu renk koordinatları olan a*(kırmızılık-yeşillik), b*(sarılık-mavilik), C (kroma), L*(açıklık) değerleri elde edildi ve bu değerler değişken parametrelere göre kıyaslandı.

$$K/S = (1-R)^2 / 2R \quad (1)$$

Boyanan numunelerin yıkamaya karşı renk direncini tespit etmek için TS EN ISO 105-C06 standardında belirtilen şartlara göre tüm numuneler, yıkama haslığı cihazında standart ECE deterjanı (4 gr) ile 40 °C’de 30 dk (tüplerin içine çelik bilye ilave etmeden) yıkanmıştır. Çıkarıldıktan sonra durulanmış ve serbest halde kurumaya bırakılmıştır. Yıkama sonrası renk değişimini tespit etmek için numunelere tekrar spektrofotometre ile renk ölçümü yapılmıştır. Sonrasında yıkama öncesi ve yıkama sonrası toplam renk farkı (ΔE) değerleri elde edilmiştir.

3. ARAŞTIRMA BULGULAR

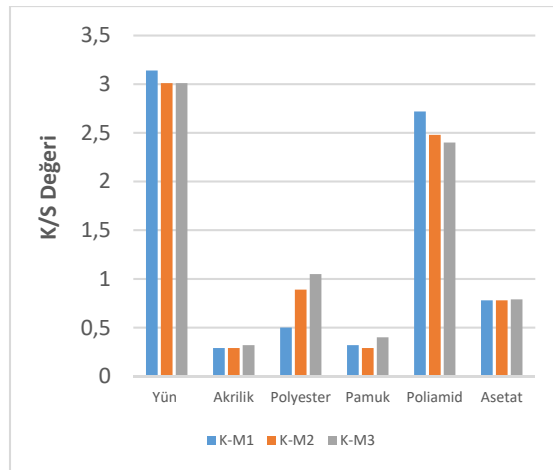
3.1. Numunelerin Renk Kuvveti

Portakal kabuğu ile boyanan numunelerin Kubelk-Munk formülüne göre hesaplanmış renk kuvveti değerlerinin bulunduğu grafik Şekil 4’te; portakal ağacı yaprağı ile boyanan numunelerin renk kuvveti değerleri ise Şekil 5’te sunulmuştur. K/S değerleri, minimum reflektans değerinin görüldüğü dalga boyu olan 400 nm’de tespit edilmiştir. Mordan madde olarak kullanılan alüminyum sülfat miktarındaki artışın portakal kabuğu ile hazırlanmış boyarmaddenin yün, poliamid ve asetat içerikli numunelere verdiği renk kuvvetine etkisinin pozitif yönde olmadığı dikkat çekmektedir. %3 konsantrasyonda kullanılan alüminyum sülfat polyester ve pamuk içerikli numunelerin renk kuvvetine az miktarda katkı sağlamıştır. Akrilik içeren numunede ise renk kuvveti değerinde 0,5’ten 1’e doğru bir artış sağlanmıştır. En yüksek K/S değeri yün ve onun ardından poliamid içerikli numunelere ait çıkmıştır. Yün lifinin yapısındaki amino (NH₂), karboksil

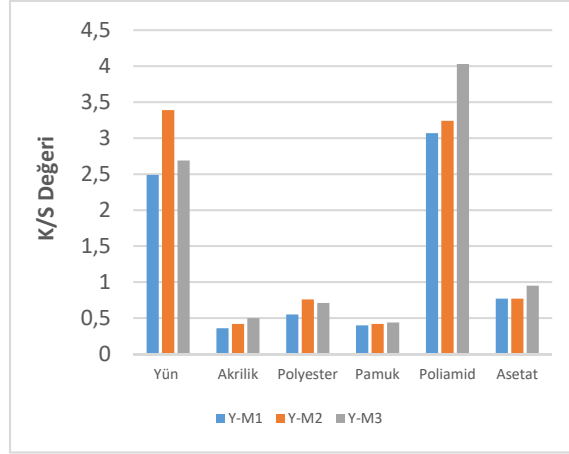
(COOH) gibi grupların portakal kabuğundaki fenolik bileşenlerden hidroksil (-OH) grubuyla hidrojen bağı kurarak etkileşim kurmuştur [18-19]. Poliamid, yapısında bulunan amid (-CONH-), aminil (-NH₂) ve karboksil grupları aracılığıyla doğal boyarmaddedeki hidroksil grupları ile hidrojen bağları oluşturmakta ve bu sayede kumaşa iyi renk tutuculuğu sağladığı ifade edilebilmektedir [20]. Portakal kabuğu ekstraktı içeren doğal boya içerisindeki alüminyum sülfat miktarındaki artışın, yün ve poliamid içeren numunelerde renk kuvvetini düşürdüğü dikkat çekmektedir. Özellikle doğal boyarmaddenin daha çok renk verdiği bu numunelerde mordan madde miktarındaki artışın kompleksleri sıklaştırarak boyarmaddenin difüzyonunu engellediği düşünülmektedir. Alüminyum sülfat miktarındaki artış adsorpsiyonu artırarak polyester kumaşın K/S değerini yükseltmiştir. Akrilik ve asetat içerikli liflerin K/S değerlerinin ise alüminyum sülfat miktarının artışından etkilenmediği söylenebilmektedir. Pamuk lifi ise yüzeyindeki hidroksil gruplara rağmen doğal boyarmadde ile bağ kuramamıştır. Doğal boyamada demir sülfat, bakır sülfat ve alüminyum sülfat mordan maddelerinin pamuklu numunelerdeki renk kuvvetinin karşılaştırıldığı bir çalışmada, alüminyum sülfat kullanılan numunelerin K/S değerlerinin diğerlerine kıyasla daha düşük olduğu belirlenmiştir [21]. Bu durum, doğal boyarmaddenin life tutunmasında, diğer mordanlara kıyasla daha zayıf bağların oluşmasına neden olmuştur. Bu çalışmada da benzer şekilde, bağlayıcı madde olarak kullanılan alüminyum sülfat, doğal boyarmaddenin pamuk lifine bağlanma mekanizması açısından yeterli uyum göstermemiştir.

Şekil 5'teki değerler incelendiğinde ise özellikle poliamid içeren numunelerde renk kuvveti değerinin yüksek olduğu dikkat çekmektedir. K/S değeri açısından poliamidin ardından ise yün içerikli liflerin geldiği görülmektedir. Portakal kabuğu ekstraktı ile olduğu gibi portakal ağacı yaprağında bulunan hidroksil (fenolik bileşenler), karboksil (fenolik asit) gibi gruplar metal tuzu olan alüminyum sülfatın da etkisiyle güçlü bağ kurmuş ve iyi bir renklendirme sağlanmıştır [22-23]. Alüminyum sülfat konsantrasyonunun %1'den %2'ye artışı yünlü kumaşın boyarmadde-lif arasındaki koordinat bağ oluşumunu güçlendirmiş, renk kuvvetini artırmıştır. Ancak konsantrasyon %3'e ulaştığında K/S değeri düşüş eğilimi göstermiştir. Bu durumun, portakal kabuğu ekstraktı ile yapılan boyamalarda olduğu gibi, artan mordan miktarının boyarmaddenin life nüfuzunu engellemesinden kaynaklandığı düşünülmektedir. Mordan konsantrasyonundaki artış, poliamid liflerde renk kuvvetinde doğrusal bir artış sağlarken; pamuk ve asetat liflerinde ise sınırlı bir etki göstermiştir. Alüminyum sülfat konsantrasyonunun %3 olduğu boyamalarda ise asetat liflerinde en yüksek renk kuvveti değeri elde edilmiştir.

Her iki ekstrakt ile polyester, akrilik, pamuk ve asetat liflerinde yeterli renk kuvveti sağlanamamıştır. Polyester lifinin hidrofobik yapıya sahip olması, doğal boyarmadde ile bağ kurabilecek aktif gruba sahip olmaması (alüminyum sülfat kullanılmasına rağmen) renk kuvvetinin az olmasına neden olduğu düşünülmektedir. Polyester kumaşın %10, 20, 30, 40 konsantrasyondaki portakal kabuğu tozu ile renklendirildiği bir çalışmada konsantrasyon artışının K/S değerine etkisinin dikkat çekici olmadığı tespit edilmiştir [24]. Akrilik ipliklerin doğal boyarmadde ile renklendirildiği bir çalışmada, pH seviyesinin 2-3 olduğu koşulların optimum boyama şartlarını sağladığı belirlenmiştir. Ancak akrilik liflerin, doğal boyarmaddede bulunan flavonoid ve fenolik bileşiklerle bağ kurabilecek aktif grup sayısının az olması ve apolar yapıda bulunması, renk kuvvetinin düşük olmasını açıklamaktadır [25]. Asetat lifinin asetillenme sebebiyle hidrojen bağı kurma potansiyeli düşüktür.



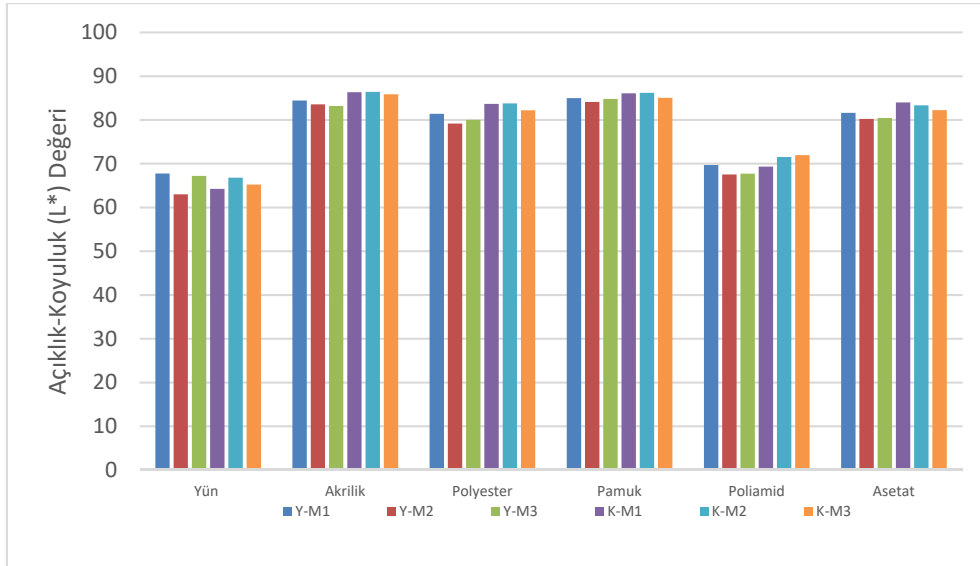
Şekil 4. Portakal kabuğu ile boyanan numunelerin K/S değerleri



Şekil 5. Portakal ağacı yaprağı ile boyanan numunelerin K/S değerleri

3.2. Numunelerin Farklı Renk Değerleri

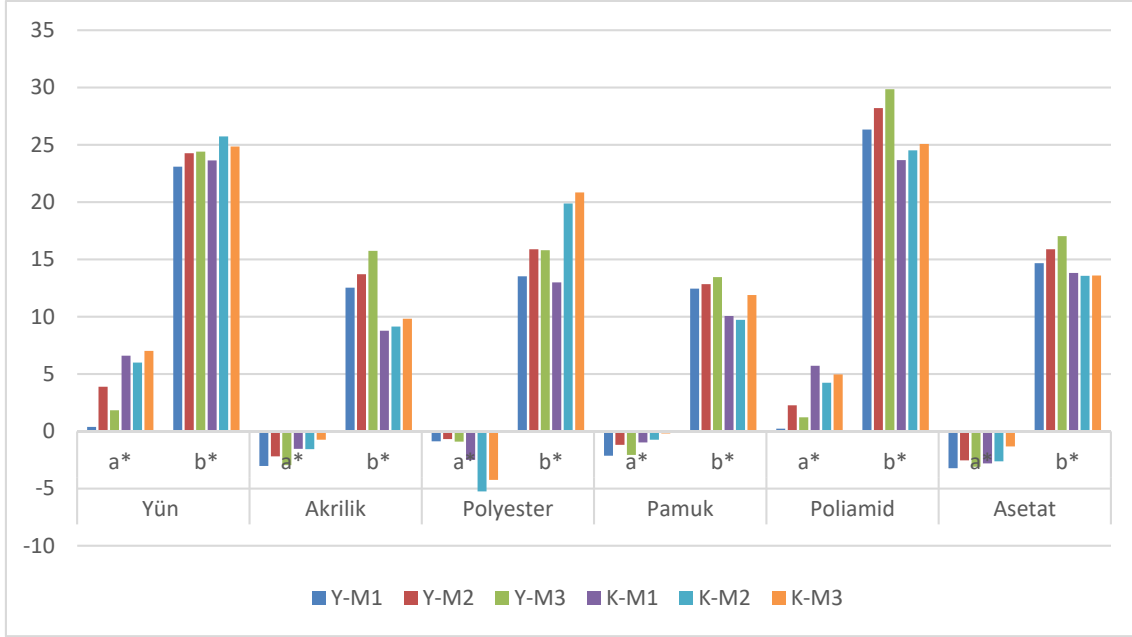
Portakal kabuğu ve portakal ağacı yaprağı ekstraktları ile boyanan numunelerin açıklık-koyuluk (L^*) değerleri Şekil 6'da sunulmuştur. Renk kuvvetleri (K/S) değerleri ile paralel olarak her iki farklı doğal kaynaktan elde edilmiş boyarmadde ile hem yünlü hem de poliamid içerikli numunelerde en koyu renk değerleri elde edilirken; polyester, akrilik, pamuk ve asetat liflerinin daha açık tonda olduğu görülmüştür. Doğal boya içerisindeki alüminyum sülfatın artan konsantrasyonunun renk koyuluğuna etkisi anlamlı değildir. Yünlü numuneler haricindeki tüm numunelerde genel olarak portakal kabuğu ile yapılan boyama işlemi ile portakal ağacı yaprağına göre daha koyu bir renk elde edilmiştir. Portakal kabuğunda flavonoid, β -karotenin gibi bileşenler mevcutken, portakal ağacı yaprağında sadece flavonoid ağırlıklı bileşenler bulunmaktadır. Koyu ve doygun renkli pigment olan karotenoidler sayesinde portakal kabuğu ekstraktı ile numunelerde daha yüksek renk yoğunluğu meydana geldiği söylenebilmektedir.



Şekil 6. Boyanan numunelerin L^* değerleri

Şekil 7'de boyanan numunelerin a^* ve b^* değerleri görülmektedir. Şekil 8'de ise boyanan numunelerin görselleri sunulmuştur. Her iki farklı doğal boyarmadde ile yapılan renklendirme işlemi ile polyester, akrilik, pamuk, asetat lifi içeren kumaşların a^* değerinin negatif yani yeşil renge yatkın oldukları görülmektedir. Portakal kabuğu ile pamuklu numunelerin renklendirildiği bir çalışmada renk ölçümü sonucunda kumaş renginin yeşile doğru eğilim gösterdiği tespit edilmiştir [6-10]. Yün ve poliamid lifi içeren numunelerde ise a^* değeri pozitif yani kırmızı renge eğilimli oldukları ortaya çıkmıştır. Doğal boyarmaddeler ile renklendirilen tüm numunelerin b^* değerleri pozitifdir. Yani tüm numunelerin rengi mavi değil sarı renge doğru eğilim göstermektedir. Yünlü kumaşların mordan kullanmadan portakal kabuğu ile

renklendirildiği başka bir çalışmada da a^* ve b^* değerleri pozitif olduğu ortaya koyulmuştur [26]. Ayrıca bu çalışmada alüminyum metal tuzunun kullanımıyla rengin daha sarı tonlara kaydığı tespit edilmiştir. Şekil 7'deki b^* değerleri incelendiğinde de tüm numunelerin sarılık değerinin alüminyum sülfat konsantrasyonu arttıkça yükseldiği görülmektedir. Alüminyum sülfat suda çözündüğünde Al^{3+} iyonları vermektedir. Bu iyonlar lifte bulunan hidroksil, amino gibi gruplarla ve doğal boyarmadde içerisindeki hidroksil ve karboksil grupları ile kompleks oluşturmaktadır [27]. Al^{3+} iyonları elektron yapısı sebebiyle rengi parlaklaştırır ve sarı rengin doygunluğunu artırır. Bu sebeple doğal boyarmaddenin lifte kazandırdığı renk yoğunluğunu ve doygunluğunu artırarak sarı rengi belirginleştirdiği söylenebilmektedir.



Şekil 7. Boyanan numunelerin a^* (kırmızılık-yeşillik) ve b^* (sarılık-mavilik) değerleri

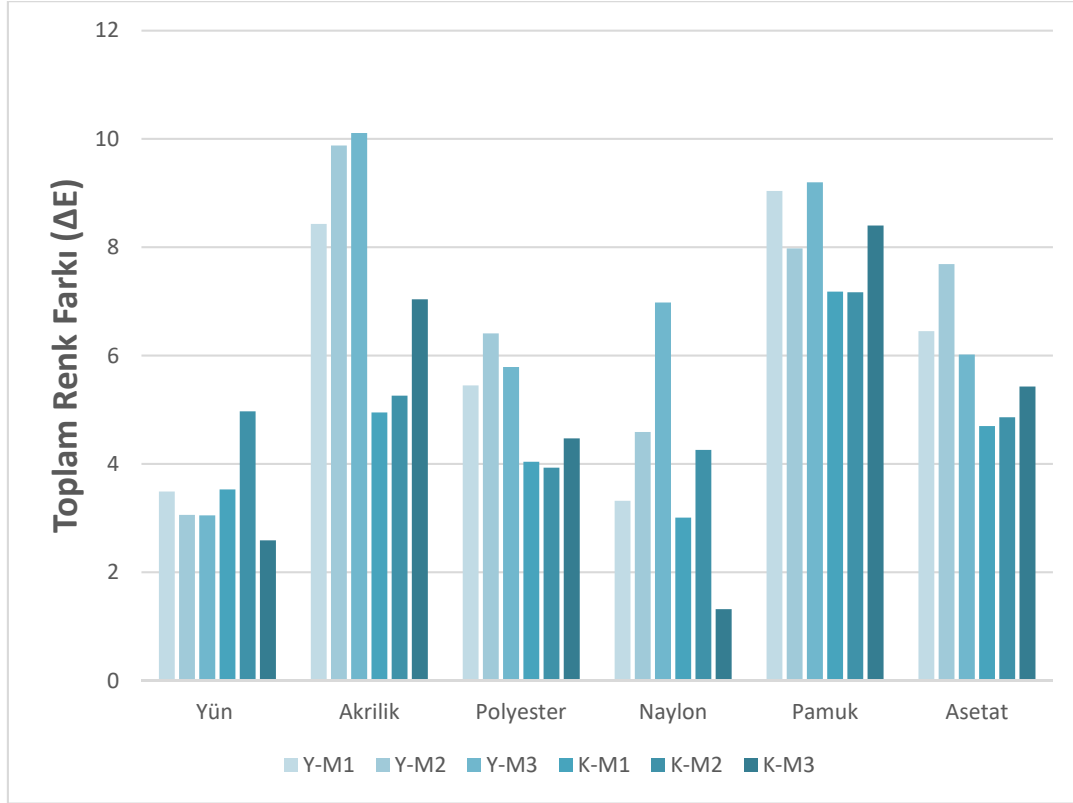
Lif türü	Boyanmamış	K-M1	K-M2	K-M3	Y-M1	Y-M2	Y-M3
Yün							
Akrilik							
Polyester							
Poliamid							
Pamuk							
Asetat							

Şekil 8. Renklendirilmemiş ve doğal boyanmış numunelerin görselleri

3.3. Boyanan Numunelerin Yıkama Sonucu Renk Değişimleri

Numunelerin yıkama öncesi renk değerleri ile yıkama sonrası renk değerlerinin karşılaştırılması sonucunda elde edilen toplam renk farkı değerleri Şekil 9'da sunulmuştur. Yıkama sebebiyle en az renk değişiminin

K/S değerleri en yüksek olan yün ve naylon içeren numunelerde olduğu görülmektedir. Portakal kabuğu ekstraktı, portakal ağacı yaprağına göre özellikle yün ve naylon liflerinde daha iyi sabitlenmiş ve yıkama sonrası diğer lif türlerine göre daha düşük renk farkı göstermiştir. Portakal kabuğu ekstraktı ile yün, poliamid ve selüloz asetat içerikli liflerin renklendirildiği bir çalışmada numunelerin K/S değerinin yıkama sonrası renk koordinatlarının bir miktar değiştiğini, K/S değerlerinin düştüğünü tespit etmişlerdir [28]. Bu çalışmada doğal boyanan numunelerin yıkama haslığının boya banyosunun pH'ı işlem sıcaklığı ve süresinin optimizasyonu ile iyileştirilebileceği sonucuna varılmıştır.



Şekil 9. Numunelerin yıkama öncesi ve sonrası toplam renk farkı değerleri

Akrilik ve Pamuk liflerinde yıkama sonrası (K-M1, K-M2, K-M3) ΔE değerleri oldukça yüksektir. Bu, bu liflerin doğal boyarmadde ile arasındaki bağın zayıf ya da boyarmaddenin stabilitesinin zayıf olduğunu göstermektedir. Özellikle akrilik için K-M2 ve K-M3 değerleri yaklaşık 10'un üzerindedir; bu, görsel olarak fark edilir derecede renk değişimi olduğu anlamına gelmektedir. Boyamada alüminyum sülfat konsantrasyonu artışı akrilik, polyester, naylon içerikli sentetik numunelerde yıkama sonrası renk değişiminin artmasına sebep olmuştur. Bu durumun metal tuzunun liflerle etkileşime girerek boyarmaddeyi sabitlemek yerine boyarmadde göçüne neden olabileceğini ve bu mordanın boyarmaddenin yıkama sırasında kaybını artırmasından kaynaklandığı düşünülmektedir. Alüminyum sülfatın doğal boyamada mordan olarak kullanıldığı bir çalışmada da numunelerin yıkama haslığının demir ve bakır sülfat kullanımına göre daha düşük olduğu tespit edilmiştir [21]. Alüminyum sülfatın yün ve naylon kumaşlarda yıkamaya karşı renk haslığını iyileştirebileceği, portakal kabuğu ekstraktı ile boyanan bu numuneler için uygun olan mordan konsantrasyonunun ise %3 olduğu sonucuna varılmıştır. Yaprak ekstraktı ile yapılan boyamalarda değişen alüminyum sülfat oranının yün ve naylon liflerinin yıkama haslığine etkisi anlamlı değildir. Bu grupta en az renk değişimi ise %2 ve %3 oranında mordan kullanılan yünlü numunelerde olduğu görülmüştür.

4. SONUÇLAR

Çalışmada farklı konsantrasyonda alüminyum sülfat mordan olarak kullanılarak portakal ağacı yaprağı ve portakal kabuğu ekstraktları ile yün, akrilik, polyester, poliamid, pamuk ve asetat içerikli farklı numuneler renklendirilmiştir. Yapılan renk analizleri doğrultusunda en yüksek renk verimi yün ve poliamid içerikli numunelerde elde edilmiştir. Yünlü kumaşlarda en yüksek K/S değeri %1 alüminyum sülfat içeren portakal

kabuğu ekstraktı ile elde edilirken, poliamid kumaşlarda %3 alüminyum sülfat içeren portakal ağacı yaprağı ekstraktı ile elde edilmiştir. Diğer numunelerde 400 nm dalga boyunda K/S değeri 1'den düşük çıkmıştır. Doğal boyarmadde ile renklendirilmiş yün ve poliamid numunelerinin renk tonu kırmızıya eğilimli iken; diğer numunelerin renginin kırmızı değil yeşil tona eğilimli çıkmıştır. Numunelerin yıkama sonrası renk değişimi davranışı incelendiğinde en az renk değişiminin (daha yüksek yıkama hasarının) yün ve poliamid içerikli numunelere ait olduğu görülmüştür. Diğer numunelerde ise boyarmadde stabilitesinin düşük olduğu görülmüştür. Daha sonra yapılacak çalışmalarda portakal ağacı ve portakal kabuğunun doğal boyama açısından çoklu liflere etkisini geliştirmek için boya banyosunun pH'ı, işlem sıcaklığı ve süresinde optimizasyon yapılması öngörülmektedir.

5. KAYNAKLAR

1. Sancar Beşen, B., Parlakyiğit, P. ve Yılmaz, E. (2024). Simultaneous coloring and antibacterial finishing to cotton through environmentally-friendly way. *Tekstil ve Konfeksiyon*, 34(1), 44-50.
2. Aishwariya, S. (2020). Textiles from orange peel waste. *Science & Technology Development Journal*, 23(2), 508-516.
3. Devi, O.R. & Saini, H. (2020). Utilization of orange peel waste in textile industry: A review. *International Journal of Chemical Studies*, 8(4) 5-8.
4. Benli, H. ve Bahtiyari, M.İ. (2011). Portakal ağacı yaprakları ile viskon kumaşların boyanması. *Akdeniz Sanat*, 4(8) 85-89.
5. Tayyab, N., Sayed, R.Y., Faisal, R. & Wang, W. (2020). Dyeing and colour fastness of natural dye from Citrus aurantium. *Industria Textila*, 71(1) 350-356.
6. Edeen, A.B. (2015). Dyeing of Egyptian cotton fabrics with orange peel using padding technique. *International Design Journal*, 5(2), 511-516.
7. Fernández-Cabal, J., Avilés-Betanzos, A., Cauich-Rodríguez, J.V., Ramírez-Sucre, M.O. & Rodríguez-Buenfil, I.M. (2025) Recent developments in Citrus aurantium L.: An overview of bioactive compounds, extraction techniques, and technological applications. *Processes*, 13(120), 1-34.
8. Senthil Kumar, C.S. & Dhinakaran, M. (2017). Extraction and application of natural dyes from orange peel and lemon peel on cotton fabrics. *International Research Journal of Engineering and Technology*, 4(5), 237-238.
9. Taura, U.H., Abubakar, M.A., Abubakar, A.M. & Kurgıya, M.U. (2024). Extraction and characterisation of natural dye from orange peel for textile applications. *Journal of Applied Life Sciences and Environment*, 57(1), 169-181.
10. Li, K., Ding, Q. & Zhang, H. (2022). Eco-friendly dyeing of cotton fabric using natural dye from orange peel. *The Journal of The Textile Institute*, 113(3), 360-366.
11. Baaka, N., Mahfoudhi, A. & Mhenni, M.F. (2017). Orange peels waste as a low cost cotton natural dye. *Moroccan Journal of Chemistry*, 5(2), 259-265.
12. Iqbal, S. & Ansari, T. (2021). Extraction and application of natural dyes. *Sustainable Practices in the Textile Industry* (s.13), Wiley, 368.
13. Chavan, R.B. (2011). Environmentally friendly dyes. *Handbook of Textile and Industrial Dyeing*. Woodhead Publishing, Cambridge.
14. Chakraborty, J.N. (2011). Metal-complex dyes. *Handbook of Textile and Industrial Dyes*. Woodhead Publishing, Philadelphia, 454.
15. Failisnur, F., Sofyan, S., Kasim, A. & Angraini, T. (2018). Study of cotton fabric dyeing process with some mordant methods by using gambier (Uncaria gambir roxb) extract. *International Journal on Advanced Science Engineering and Information Technology*, 8(4), 1098-1104.
16. Rosyida, A., Suranto, Masykuri, M. & Margono, M. (2022). Minimisation of pollution in the cotton fabric dyeing process with natural dyes by the selection of mordant type. *Research Journal of Textile and Apparel*, 26(1), 41-56.
17. SDL Atlas, (2025). www.sdlatlas.com/test-materials/multifiber/multifiber-adjacent-fabric-ioaisotic-10-meters, Erişim tarihi: 17.07.2025.
18. Özdemir, H. (2018). Bodur mürver (Sambucus ebulus L.) ile yün ipliklerin boyanması. *Çukurova Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 33(1), 219-228.
19. Koyuncu, M. (2022). Investigation of color, and fastness properties of wool yarn dyed with ternary plant combination: A new dyeing method for the production of eco-friendly textiles. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 37(4), 1077-1086.
20. Krishna, S., Sreedhar, I. & Patel, C.M. (2021). Molecular dynamics simulation of polyamide-based materials – A review. *Computational Materials Science*, 200, 1-14.

21. Sezgin Bozok, S., Ahmed, T., Bebekli, M. & Oğulata, R.T. (2023). Investigation of the effect of different mordant substances and nonionic surfactant on cotton fabrics dyed with eucalyptus leaf extract by using IR dyeing machine. *The Journal of The Textile Institute*, 115(12), 2421-2428.
22. Khettal, B., Kadri, N., Tighilet, K., Adjebli, A. & Dahmoune, F. (2016). Phenolic compounds from Citrus leaves: antioxidant activity and enzymatic browning inhibition. *Journal of Complementary and Integrative Medicine*, 20160030, 1-13.
23. Uysal, S., Cvetanović, A., Zengin, G., Đurović, S., Zeković, Z. & Aktumsek, A. (2018). Effects of orange leaves extraction conditions on antioxidant and phenolic content: Optimization using response surface methodology. *Analytical Letters*, 51(10), 1505-1519.
24. Hussain, T., Arain, M.F., Khan, I.A., Javed, K., Khan, H., Ahmed, A. & Khan, A. (2024). Recycling orange waste for the sustainable dyeing of polyester fabric. *Journal of the Indian Chemical Society*, 101, 1-9.
25. Dhahri, H., Guesmi, A. & Hamadi, N.B. (2019). Application of phenolic compounds as natural dye extracted from date-pits: dyeing studies of modified acrylic fibres. *Natural Product Research*, 33(9), 1329-1333.
26. Hou, X., Chen, X., Cheng, Y., Xu, H., Chen, L. & Yang, Y. (2013). Dyeing and UV-protection properties of water extracts from orange peel. *Journal of Cleaner Production*, 52, 410-419.
27. Bancroft, W.D. (2002). Alumina. *Mordants II*, ASC Publications, 36.
28. Ivanovska, A., Gajic', I.S., Ladarevic', J., Milošević', M., Savic', I., Mihajlovski, K. & Kostić, M. (2022). Sustainable dyeing and functionalization of different fibers using orange peel extract's antioxidants. *Antioxidants*, 11(2059), 1-15.

Modeling and Performance Analysis of Battery Energy Storage Incorporated Grid-Connected Photovoltaic Plants

Türker Aşkın KABADAYI^{1,a}, Özgür ÇELİK^{2,b}

¹Adana Alparslan Türkeş Science and Technology University, Graduate School, Department of Electrical & Electronics Engineering, Adana, Türkiye

²Adana Alparslan Türkeş Science and Technology University, Department of Energy Systems Engineering, Adana, Türkiye

^aORCID: 0009-0008-2544-5635; ^bORCID: 0000-0002-7683-2415

Article Info

Received : 18.07.2025

Accepted : 18.08.2025

DOI: 10.21605/cukurovaumfd.1745510

Corresponding Author

Özgür ÇELİK
ozgurcelik@atu.edu.tr

Keywords

Photovoltaic system

Peak shaving

Energy storage

Battery

How to cite: KABADAYI, T.A., ÇELİK, O., (2025). Modeling and Performance Analysis of Battery Energy Storage Incorporated Grid-Connected Photovoltaic Plants. Çukurova University, Journal of the Faculty of Engineering, 40(3), 569-579.

ABSTRACT

This study conducts a detailed performance comparison between two large-scale photovoltaic (PV) power plants, each with an installed capacity of 1 MW. One system follows a conventional PV configuration, while the other incorporates a battery energy storage system (BESS) for enhanced functionality. The BESS-integrated system employs a peak-shaving strategy, aiming to store excess energy generated beyond 1 MW during daytime hours. The primary goals are to extend the PV system's energy supply into peak demand periods following sunset and to improve responsiveness to energy demand fluctuations. Both configurations were independently modeled using the same simulation environment, with a focus solely on technical performance, excluding economic considerations. Simulation results reveal that the BESS-enhanced system delivers an additional 153.18 MWh annually to the grid and achieves a 4.8% improvement in performance ratio. These findings highlight the technical benefits of BESS integration in PV systems, particularly its contribution to improving energy continuity.

Batarya Enerji Depolama Sistemine Sahip Şebeke Bağlantılı Fotovoltaik Santrallerin Modellenmesi ve Performans Analizi

Makale Bilgileri

Geliş : 18.07.2025

Kabul : 18.08.2025

DOI: 10.21605/cukurovaumfd.1745510

Sorumlu Yazar

Özgür ÇELİK
ozgurcelik@atu.edu.tr

Anahtar Kelimeler

FV sistemler

Tepe-traşlama

Enerji depolama

Batarya

Atıf şekli: KABADAYI, T.A., ÇELİK, O., (2025). Batarya Enerji Depolama Sistemine Sahip Şebeke Bağlantılı Fotovoltaik Santrallerin Modellenmesi ve Performans Analizi. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 569-579.

ÖZ

Bu çalışma, her biri 1 MW kurulu güce sahip iki büyük ölçekli fotovoltaik (FV) enerji santrali arasında ayrıntılı bir performans karşılaştırması yürütmektedir. Sistemlerden biri geleneksel bir PV konfigürasyonunu takip ederken, diğeri gelişmiş işlevsellik için bir batarya enerji depolama sistemi (BEDS) içermektedir. BEDS entegreli sistem, gündüz saatlerinde 1 MW'ın üzerinde üretilen fazla enerjiyi depolamayı amaçlayan bir tepe traşlama stratejisi kullanmaktadır. Birincil hedefler, PV sisteminin enerji arzını gün batımından sonraki tepe talep dönemlerine kadar uzatmak ve enerji talebi dalgalanmalarına duyarlılığı iyileştirmektir. Her iki konfigürasyon da aynı simülasyon ortamı kullanılarak, ekonomik hususlar hariç tutularak yalnızca teknik performansa odaklanılarak bağımsız olarak modellenmiştir. Benzetim sonuçları, BEDS ile geliştirilmiş sistemin şebekeye yıllık ek 153,18 MWh sağladığını ve performans oranında %4,8'lik bir iyileştirme sağladığını ortaya koymaktadır. Bu bulgular, özellikle enerji sürekliliğini iyileştirmeye katkısı olmak üzere, FV sistemlerine BEDS entegrasyonunun teknik faydalarını vurgulamaktadır.

1. INTRODUCTION

With the increasing population and developing industry in the world, the energy demand is increasing rapidly. A large part of the energy demand is still provided by fossil fuels. Fossil fuels create serious concerns for the society in terms of the sustainability targets. For this reason, the tendency towards renewable energy in the world is increasing day by day and official steps are being taken and targets are being determined in this regard. Finally, it is the transition to completely renewable energy and zero carbon emission. Although renewable energy sources are accepted as an unlimited energy source in theory, in practice, the dependence on fossil fuels still continues today due to the fact that energy continuity and never and instantly respond [1,2]. By nature, photovoltaic power plants (PV) can only produce energy when sunlight directly reaches the PV modules. Therefore, solar power plants only have the potential to generate electricity during daylight hours. They are also directly affected by momentary weather changes. This directly affects the power plant's output, i.e. the amount of energy injected into the grid. On the other hand, energy consumption, i.e. energy demand, differs from this production cycle. Energy demand usually reaches its highest levels in the early morning and evening hours. In this context, solar power plants have difficulty responding to this demand. Because these plants can only reach their highest generation capacity at noon in clear and sunny weather [3,4].

Recent developments in battery technology have created a serious potential for the attainability of the ultimate zero carbon emission target. In this context, it is aimed to store the excess energy in batteries as soon as it is produced and inject it into the grid during periods of high demand. In addition, excess energy that exceeds the maximum output power of the plant within the scope of the contract is also transferred to the batteries. For these reasons, applications have begun to integrate battery systems into photovoltaic power plants. In the coming years, it is expected that this integration will accelerate further with the decrease in battery unit costs [5].

The three most commonly used methods with energy storage systems (ESS) in PV power plants are peak shaving, self-consumption optimization and load shifting applications. Peak shaving aims to store the excess production remaining above the peak of the production license in batteries and integrate it into the grid during periods when production decreases during the day or during periods when demand increases after sunset. Self-consumption optimization increases the use of this energy directly in the facility by storing the excess electricity generated from solar energy in batteries and increases the on-site usage rate instead of selling it to the grid. Load shifting is an effective strategy to reduce energy costs by charging the battery during low-cost hours and discharging it during high-tariff periods; this provides a great advantage especially in systems where time-of-use (TOU) tariffs are applied. These three methods are the most common and profitable applications of PV+BESS integration in terms of both energy efficiency and economic return [6].

In addition, battery energy storage systems (BESS) significantly facilitate the grid integration of PV systems. PV systems have an inherently variable and intermittent generation profile, which can lead to problems with grid frequency and voltage stability, especially during sudden generation changes such as cloud transitions. While this variability complicates direct integration into the grid, BESS solutions compensate for such fluctuations, making PV system outputs more stable and predictable. With services such as ramp rate control, frequency regulation and voltage support, BESS enables the safe and stable integration of PV systems into the grid. BESS integration to PV systems and grid integration of these systems has been a very popular research and development topic in recent years [7].

Rodriguez-Gallegos et al. conducted a study on the cost and benefit analysis of storage and non-storage systems in different geographical regions [8]. Okafor et al. presented a study on the techno-economic analysis of the integration of BESS into on-grid PV system [9]. Sharma et al. performed a comparative analysis of different battery technologies and discussed to determine the most suitable size for solar power plant applications [10]. Li et al. investigated performance of lithium-ion batteries in solar power plants in detail [11]. Rouholamini et al. examined the annual generation and grid interaction of grid-connected solar power plants with and without energy storage systems [12]. Hill et al. investigated the integration of solar energy into the system in order to overcome the problems that may arise from the intermittent nature of solar energy in the integration of these two systems as solar energy becomes more popular and the transition

to smart grids were examined [13]. Rahimi et al. conducted a study to discuss a simple method to achieve peak load reduction through energy storage systems for a utility [14]. Cervantes and Choobineh performed a study for optimal sizing of a nonutility-scale solar power system and battery energy storage system. Analysis shows that the size of the solar system is affected by labor cost and load size, while the battery storage size is sensitive to load size and battery cost [15]. Lange et al. performed a study for sizing battery energy storage systems for peak shaving based on a real-time control algorithm. A simulation environment to determine the behavior of the energy system, a real-time peak shaving control algorithm and a sizing process consisting of an optimization process to determine the battery algorithm parameters were proposed [16].

Consequently, this literature review focused on the performance analysis of on-grid PV systems with and without energy storage systems. The conducted literature survey show that energy storage systems have the potential to improve solar integration, increase grid stabilization and improve overall system efficiency. However, it is emphasized that local conditions, technological developments and economic factors should be carefully considered for optimal system design.

The primary objective of this study is to model and analyze the performance of grid-connected PV power plants integrated with BESS. The main focus is on enhancing energy dispatchability and grid interaction. By incorporating BESS into conventional PV systems, the research aims to evaluate how energy storage can mitigate the intermittent nature of solar energy, enable peak shaving, and improve system reliability during periods of high demand or low solar generation. Through comprehensive simulation-based modeling, the study investigates the impact of BESS integration on key performance metrics such as energy yield, performance ratio, and energy injected into the grid.

The rest of the paper is organized as follows: Section 2 presents material and methods. System description is performed in Section 3. In Section 4, the performance results of the PV power plant with and without battery energy storage system. Finally, conclusions are highlighted in Section 5.

2. MATERIALS AND METHODS

The simulation study for the Karacaören PV Power Plant project was carried out using a simulation software widely used in the sector. The system has a grid-connected structure and does not include a battery system, while the secondary project includes a BESS integrated into the same system.

The power plant is located in the Karacaören region of Karaman province in southern Türkiye. The geographic coordinates that used in the simulation are as follows [17]:

- Latitude: 37.50° North
- Longitude: 33.21° East
- Altitude: 1112 meters
- Time Zone: UTC+3

This location is quite suitable for photovoltaic energy production due to its high insolation potential. The meteorological data used in the simulation is based on the Meteonorm 8.1 dataset. These data are synthetic data based on multi-year averages between 2003-2010 and have 100% satellite coverage. Monthly average global horizontal irradiance (GlobHor) is taken as 1834 kWh/m² on an annual basis. The albedo value, which is the reflection coefficient of the ground surface, is defined as 0.20. This value is an average reflectivity ratio suitable for open terrain conditions.

The simulation software that used in this project is a professional simulation software widely used to perform performance analysis of photovoltaic systems. After the equipment used in the design stage were determined externally, the system configuration, energy production, system losses and annual performance rates in the Karacaören project are calculated by this software. The shadowing effect was not taken into account in the simulation and 3D scene definition was not made. Since the two variants are simulated in the same conditions, these ignored parameters have a minimal effect on the obtained results.

The photovoltaic panels used in the project are manufactured by HT-SAAE and the model name is HT66-18X-500. The fact that it is manufactured in a location close to the project area in Türkiye and after-sales support factors played an important role in the selection of this module. The basic technical features of the system are given in Table 1.

Table 1. Technical features of the system

Nominal power of modules at STC	500 W _p
Total number of modules	3060 units
Total installed PV power	1530 kW _p
Series connection configuration	170 strings x 18 modules

The output power of the system at 50°C operating temperature is calculated as 1406 kW_p. The operating voltage obtained at this temperature is 623 V and the current is 2256 A.

On the inverter side, Huawei Technologies' SUN2000-100KTL-M1-400Vac model was used. The reason for choosing this material is that it has proven itself with its previous versions as a string inverter and its after-sales support advantages. The basic technical features of the inverter configuration are presented in Table 2.

Table 2. The basic technical features of the inverter configuration

Unit power	100 kW _{ac}
Total number of inverters	11 units
Total AC power	1100 kW _{ac}
DC:AC power ratio	1.39
Operating voltage range	200 – 1000 V

The storage system is configured with BYD Battery Box Premium HVS 12.8 model batteries. The storage system features are given in Table 3.

Table 3. The storage system features

Total number of batteries	102 (2 series x 51 parallel)
Voltage	1024 V
Nominal capacity	1275 Ah
Total storage energy	1044.5 kWh
Minimum discharge level (SOC)	20%
Storage strategy	Peak shaving
Charge time	When grid power limit is exceeded
Discharge time	After sunset
Charger unit	380 kW _{dc} , 97% efficiency
Discharge inverter	280 kW _{ac} , 97% efficiency

The system without BESS integration is called Variant 0 (VC0), the system with BESS integration is called Variant 1 (VC1).

3. SYSTEM DESCRIPTION

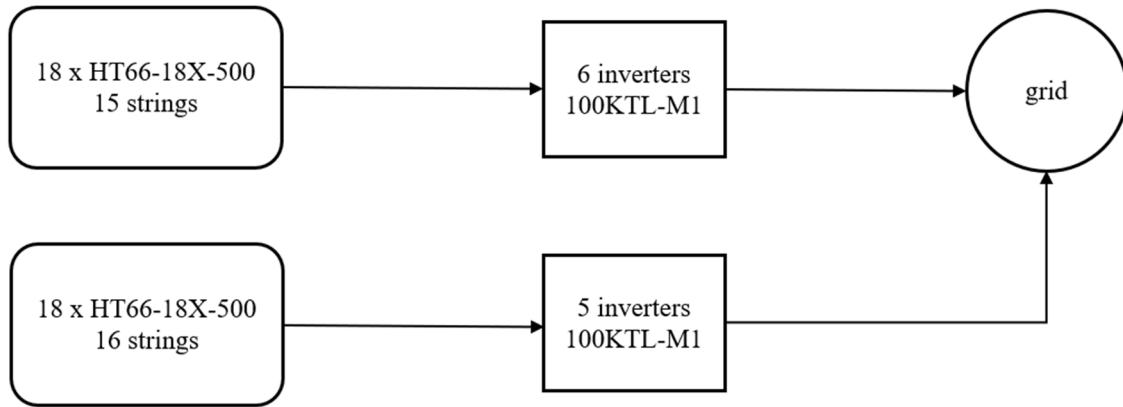
3.1. PV System without BESS and Peak Shaving

Karacaören PV Power Plant is configured as a grid-connected photovoltaic power plant with an installed capacity of 1530 kW_p. There is no BESS in this variant. The system is installed on a fixed inclined plane and shading factors are not taken into account. General system parameters are given in Table 4.

Table 4. General system parameters

Total PV power	1530 kW _p
PV module type	HT-SAAE HT66-18X-500 (500 W _p)
Number of modules	3060 pcs (170 strings × 18 modules)
Number of inverters	11 pcs
Total AC power	1100 kW _{ac}
System tilt / azimuth	32.5° / 0°
Load to grid	Unlimited
Albedo	0.20
Shading	No shadings

Single-line diagram of the constructed PV system without BESS is illustrated in Figure 1.

**Figure 1.** VC0 single line diagram

The limits of these inverters are depicted in Table 5.

Table 5. The limits of utilized inverters

Unit power	100 kW _{ac}
Total inverter power	1100 kW _{ac}
DC:AC power ratio	1.39
Operating voltage range	200 – 1000 V
Maximum output at rated power	110 kW _{ac} (at 33°C)

3.2. PV System with BESS and Peak Shaving

In this variant, the BESS system consisting of BYD Battery Box Premium HVS 12.8 model lithium-ion batteries are integrated into the PV system. The main purpose of the system is to ensure supply continuity, prevent instantaneous peaks and support grid loads.

Battery box general features are given in Table 6.

Table 6. Battery box general features.

Battery type	Lithium-ion
Model	BYD Battery Box Premium HVS 12.8
Total number of units	102 (2 series × 51 parallel)
Total storage capacity	1044.5 kWh
Nominal voltage	1024 V
Capacity	1275 Ah
Operating temperature	Fixed 20°C
Minimum discharge level (SOC)	20%

Single-line diagram of the constructed PV system with BESS is illustrated in Figure 2.

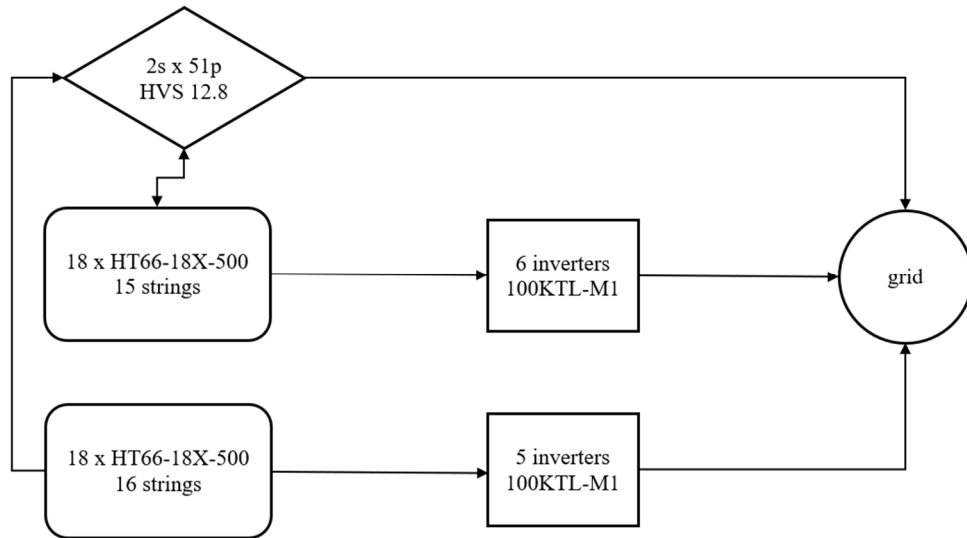


Figure 2. Variant 1 Single line diagram

Power conversion system is ignored in this project. The operating principle of BESS is determined as peak-shaving. With this strategy, excess energy beyond the grid limit is stored in batteries and injected into the grid when demand is high.

When the instantaneous output power of the PV system exceeds the 1000 kW threshold, the BESS is activated and stores the energy above this limit, while the discharge is active towards the grid after sunset. Charge/discharge parameters are given in Table 7.

Table 7. Charge/discharge parameters

Max. charge power	380 kWdc
Max. discharge power	280 kWdc
Efficiency	97% max

4. RESULTS AND DISCUSSIONS

Table 8 lists the monthly and annual total energy production data and performance ratios (PR) of the variants according to the simulation results:

Table 8. VC0/1 simulation results

Month	Production (kWh)-V0	Production (kWh)-V1	Production (%) - V0	Production (%) - V1
January	153,397	162,004	87.2	92.1
February	167,290	174,585	86.6	90.4
March	205,732	219,415	80.8	86.1
April	223,467	238,421	78.4	83.6
May	259,739	275,482	79.3	84.1
June	256,734	274,317	77.0	82.3
July	266,609	280,787	77.8	81.9
August	260,678	277,942	75.9	81.0
September	240,067	256,772	77.0	82.4
October	207,073	221,467	79.4	84.9
November	168,392	176,882	85.2	89.5
December	152,673	156,961	90.5	93.0
Total	2,561,851	2,715,034	80.2	85.0

It is calculated that 153.18 MWh is injected into the grid via BESS at VC1 after battery conversion losses, which cannot be injected into the grid at VC0 under normal conditions. It is also observed that VC1 offers 4.8% higher performance in terms of performance ratio.

When the daily power output curves are examined, it is seen that the power output at VC0 flattens out on clear days due to reaching its peak limit at noon as shown in Figure 3.

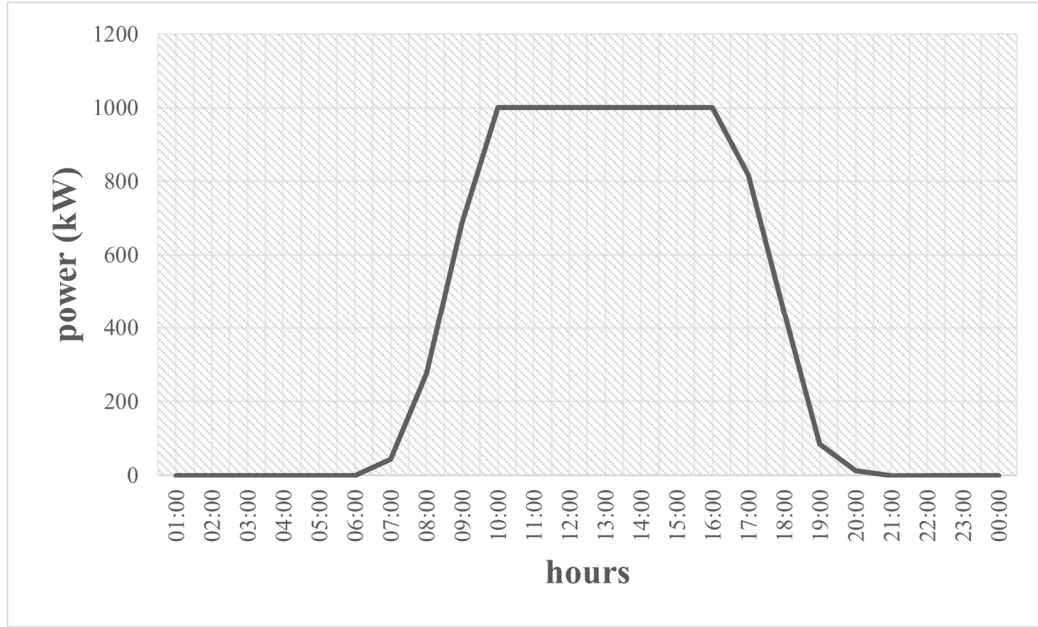


Figure 3. VC0 injected energy to grid in a clear sky day

The curve of the clear sky day presented for VC0 and the curve of the same day in the VC1 simulation are illustrated in Figure 4.

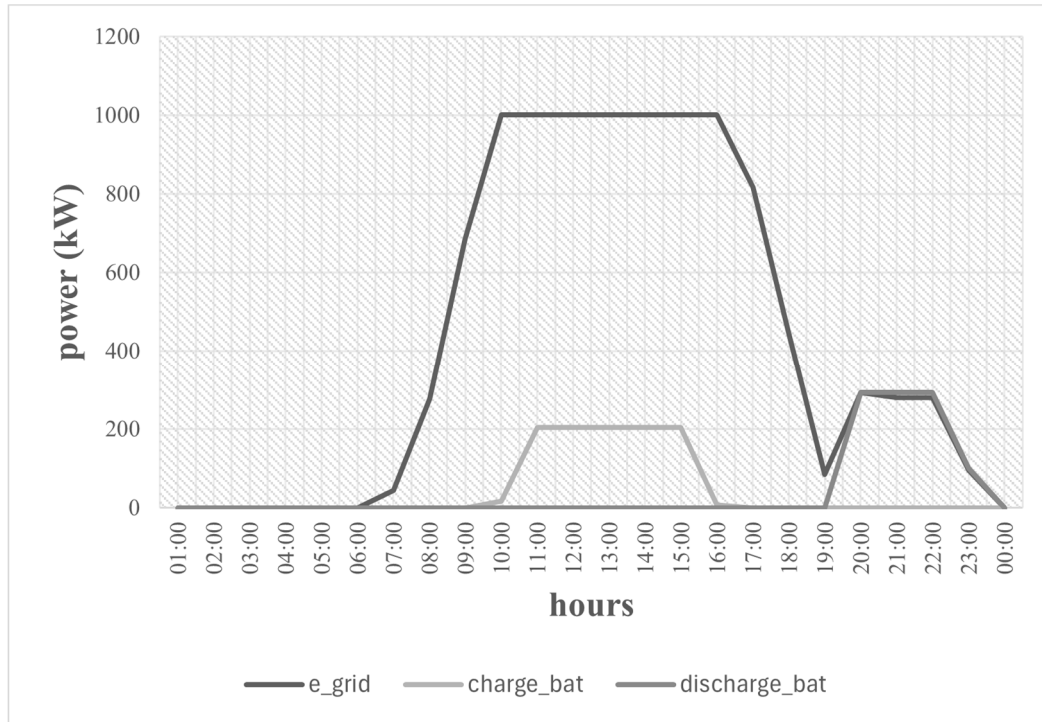


Figure 4. VC1 injected energy to grid in a clear sky day

When the daily power output curves are examined, it is seen that the power output in VC1 reaches its peak at noon and follows a horizontal course on clear days, while the energy remaining above the peak is charged to the batteries to be discharged in the evening.

When Figures 3 and 4 are compared, it is seen that VC0 cuts the power output increase when it reaches the grid injection point and is fixed at 1000 kW, that is, it limits the maximum output power to 1000 kW. However, since there is no limit criterion in VC1, the output power up to 1000 kW is injected into the grid, while the system keeps the output power fixed at 1000 kW without restricting itself for more energy output and stores the excess value in batteries to be injected into the grid after sunset when the energy supply increases. The simulation loss diagram for VC0 is presented in Figure 5.

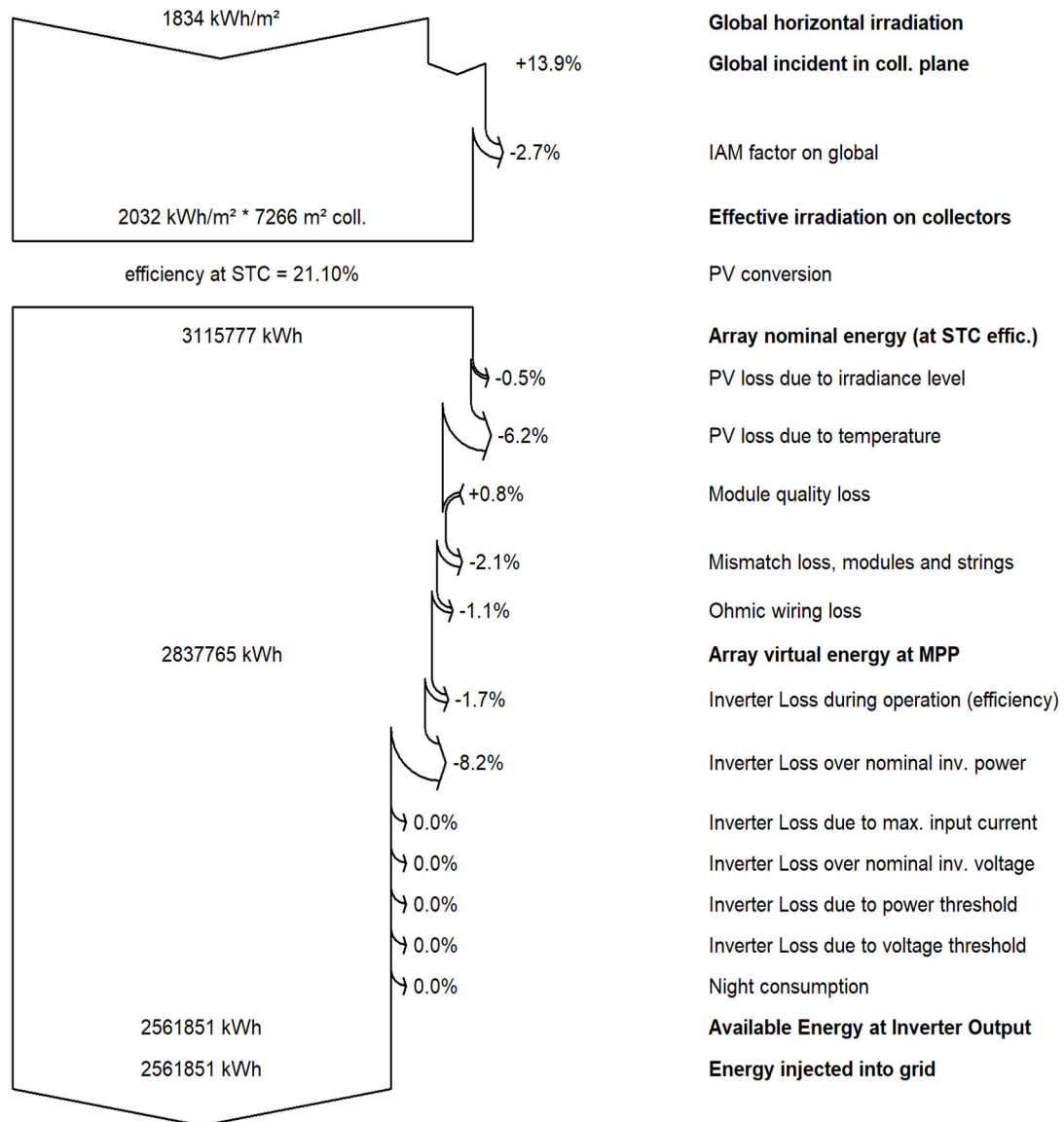


Figure 5. VC0 loss diagram

It has been determined that there is a serious loss at this point because the inverters cannot provide output power over 1000 kW and are limited according to the connection agreement. The simulation loss diagram for VC1 is presented in Figure 6.

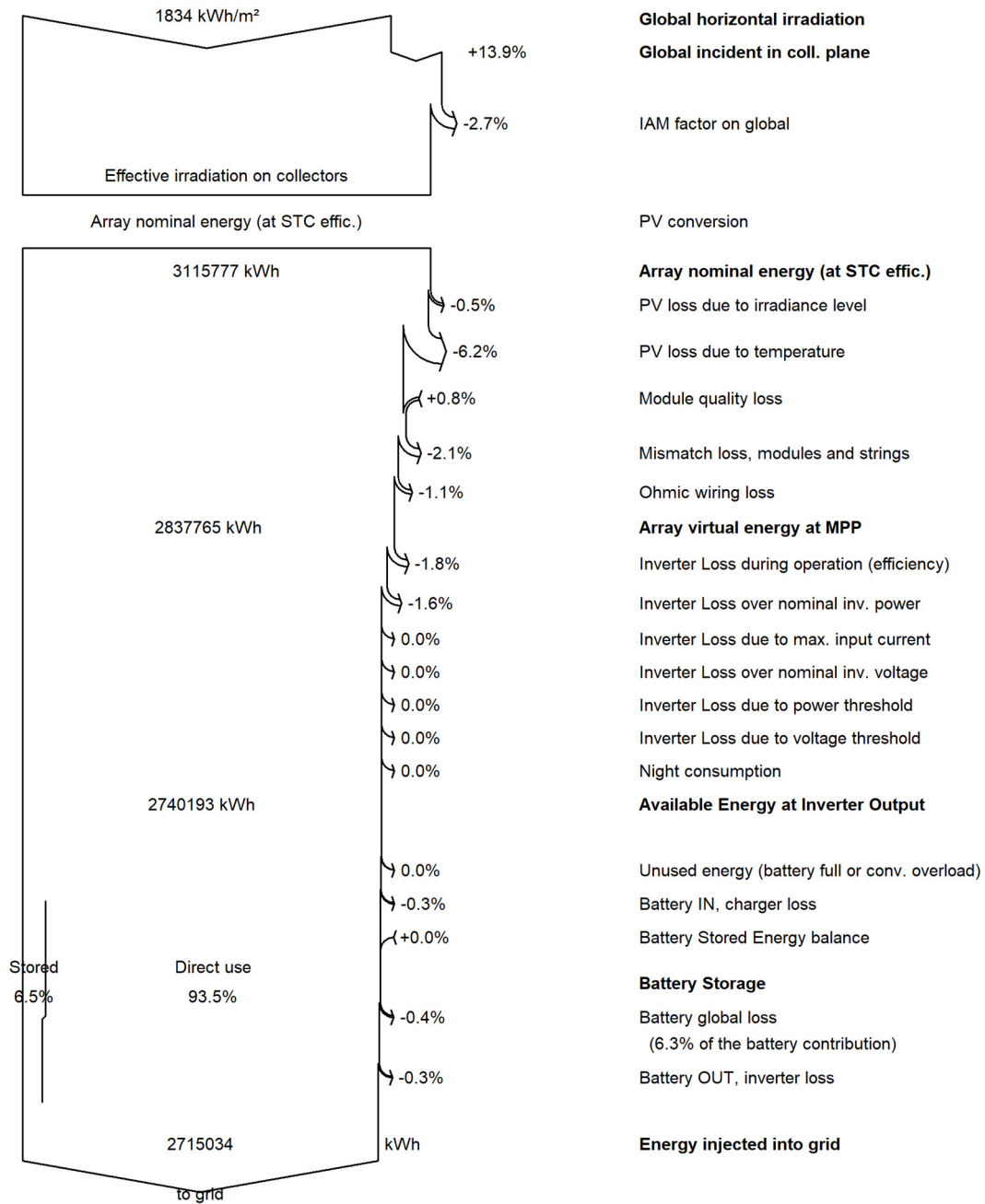


Figure 6. VC1 loss diagram

It is seen that the inverter nominal limit loss of 8.2% has decreased to 1.64% with the BESS integration and the system operates with less loss. In addition, it has been observed that the losses of the batteries during the charge/discharge cycle are in accordance with the battery label efficiency values and it is understood from the loss diagrams that the system operates stably.

As mentioned in the paragraph comparing Figures 3 and 4, on days with clear weather and long sunshine, the grid output power limitation in VC0 prevented the injection of more than 1000 kW into the grid, limiting the system and targeting a maximum output power of 1000 kW during this period. Therefore, VC0 remained below its current potential, meaning the energy it could generate. This is technically considered a loss. However, while 1000 kW was injected into the grid without any restrictions in VC1, no generation loss occurred as long as the remaining value did not exceed the maximum charging power, which is the threshold

value of 1380 kW for VC1 (1000 kW maximum grid injection power limit plus 380 kW maximum charging power). Based on this, the major loss item in Figures 5 and 6 can be explained.

It should be noted that there are many advantageous and disadvantageous of BESS assisted PV systems in technical, environmental, and economic perspectives. Providing energy time-shifting and peak shaving, reducing PV output fluctuations and curtailment, enhancing grid stability via frequency/voltage support, availability revenue from ancillary services, facilitating higher penetration of renewable energy for reducing greenhouse gas emissions make BESS integration an attractive option for modern power systems. However, limited cycle life and calendar degradation, requiring advanced energy management system and complex control structure, high initial investment cost and maintenance cost, safety issues, and recycling challenges makes them disadvantageous

5. CONCLUSION

This study focuses on the integration of BESS into PV systems in Karaman, Türkiye, and presents a comparison that focuses on technical efficiency without any economic concerns. The major goal is to provide a deeper technical understanding of the operational benefits and limitations of BESS-enhanced PV systems, offering valuable insights for optimizing system design, planning, and control strategies in modern renewable energy applications. Simulations were performed using the current software of a popular simulation program in the sector, without adding data manually, and with the correct data in the database of the software. First, VC0 was designed, then BESS was designed and integrated into the system with the peak-shaving method and the storage capacity that would work efficiently with VC0, and as a result, the VC1 variant was created. According to the simulation results of both variants, it is inferred that both systems work efficiently.

According to the simulation results obtained, BESS integration has provided significant benefits in terms of efficiency and grid stabilization to the PV system. In VC1, an increase of 153.2 MWh is observed compared to VC0, while a 4.8% increase is observed on the PR side. The main reason for these improvements is that the energy above the grid power limitation during peak of the inverters enters the useful energy class with VC1.

In different conditions and situations, for example; in a scenario where energy demand on the grid is high and supply is low, if hourly energy unit prices vary, it is possible in practice to pull the peak shaving limit below the grid power limitation so that BESS can store more and create supply during periods with high unit prices when demand is high.

In addition, regarding grid interaction, BESS strengthens the optimization of the PV system to the grid. Due to its nature, instant decreases and increases in the output power of PV systems can be seen depending on weather conditions. These situations can strain the grid frequency because keeping the grid frequency stable is the primary priority for energy continuity. In instant peak situations, BESS helps to keep the grid frequency constant by smoothing the power curve.

As a result, BESS integration has provided advantages to the PV system in many areas such as higher production and more efficient operation, and helping with grid frequency balancing. However, in order to be able to carry out such works, these advantages must be considered economically. Although large-scale battery systems are developing rapidly, they are relatively new technology and therefore they are quite costly materials, and operational expenses such as failure and maintenance should not be ignored because periodic maintenance and repairs are essential for the system to operate properly and stably, which creates additional labor and costs.

In addition, although there are now official regulations with almost complete boundaries for PV systems, it should not be ignored that the current regulations on BESS systems can be updated according to current conditions and situations. Since current regulations are in question, there is no financial energy purchase guarantee or additional support for current energy sales unit prices for BESS integrated PV systems in the location where the project is designed. Since these factors are techno-economic feasibility issues, it is recommended that the investment be made by the investor by examining the possible investments in detail in accordance with current regulations.

The environmental results of the BESS integration into PV systems are quite essential because, as mentioned at the beginning, the dependency on fossil fuels can only end when renewable energy sources become permanent energy sources. For this reason, the availability of batteries loaded with energy to respond to energy demand on-site can be interpreted as a promising step, although not completely at this stage, and it is possible to approach or reach the targeted zero carbon emission with the developing technology and official and social steps to be taken in this field in the coming years.

6. REFERENCES

1. Çelik, Ö. (2023). Analysis of current limiting algorithm with anti-windup control for transient stability of grid-forming converters. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 38(3), 671-681.
2. Erişti, H., Akdağlı, A., Baldan, E. & Sarı, A. (2025). Investigation of panel efficiency in photovoltaic systems under partial shading and different pollution conditions: An experimental study. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 40(2), 301-311.
3. Çelik, Ö., Zor, K., Tan, A. & Teke, A. (2022). A novel gene expression programming-based MPPT technique for PV micro-inverter applications under fast-changing atmospheric conditions. *Solar Energy*, 239, 268-282.
4. Çelik, Ö., Büyük, M. & Tan, A. (2022). Mitigation of power oscillations for energy harvesting capability improvement of grid-connected renewable energy systems. *Electric Power Systems Research*, 213, 108756.
5. Çelik, Ö. (2025). An efficient control scheme for operational performance enhancement of vehicular fuel cell integrated power system. *Journal of Power Sources*, 625, 235660.
6. Inaolaji, A., Wu, X., Roychowdhury, R. & Smith, R. (2024). Optimal allocation of battery energy storage systems for peak shaving and reliability enhancement in distribution systems. *Journal of Energy Storage*, 95, 112305.
7. Bozkurt, H., Çelik, Ö. & Teke, A. (2024). Power quality enhancement in hybrid PV-BES system based on ANN-MPPT. *Turkish Journal of Electrical Engineering and Computer Sciences*, 32(5), 662-681.
8. Gandhi, O., Kumar, D.S., Rodríguez-Gallegos, C.D. & Srinivasan, D. (2020). Review of stability and reliability implications on power systems with high PV penetration: Part II – Potential solutions and the way forward. *Renewable and Sustainable Energy Reviews*, 132, 110018.
9. Okafor, C. E., Gbadamosi, S. L., Krishnamurthy, S., Ratshitanga, M. & Moodley, P. (2025). Techno-economic analysis of battery storage technologies in distribution networks with integrated electric vehicles and solar PV systems. *Energy Reports*, 14, 579-599.
10. Sharma, D.D., Singh, S.N., Rajpurohit, B.S. & Longatt, F.G. (2015). Critical load profile estimation for sizing of battery storage system. In *Proceedings of the IEEE Power & Energy Society General Meeting*.
11. Li, H., Wang, L., Song, Y., Zhang, Z., Zhang, H., Du, A. & He, X. (2023). Significance of current collectors for high-performance conventional lithium-ion batteries: A review. *Advanced Functional Materials*, 33(49), 2305515.
12. Rouholamini, M., Wang, C., Nehrir, H., Hu, X., Hu, Z., Aki, H., Zhao, B., Miao, Z. & Strunz, K. (2022). A review of modeling, management, and applications of grid-connected Li-ion battery storage systems. *IEEE Transactions on Smart Grid*, 13(6), 4505-4524.
13. Hill, C.A., Such, M.C., Chen, D., Gonzalez, J. & Grady, W.M. (2012). Battery energy storage for enabling integration of distributed solar power generation. *IEEE Transactions on Smart Grid*, 3(2), 850-857.
14. Rahimi, A., Zarghami, M., Vaziri, M. & Vadhva, S. (2013). A simple and effective approach for peak load shaving using battery storage systems. *North American Power Symposium (NAPS)*, 1-5.
15. Cervantes, J. & Choobineh, F. (2018). Optimal sizing of a nonutility-scale solar power system and its battery storage. *Applied Energy*, 216, 105-115.
16. Lange, C., Rueß, A., Nuß, A., Öchsner, R. & März, M. (2020). Dimensioning battery energy storage systems for peak shaving based on a real-time control algorithm. *Applied Energy*, 280, 115993.
17. PVsyst SA, (c. 2022). *PVsyst Tutorials V7* [User's manual—Grid-connected systems].

Detection of Imbalance Faults in Industrial Machines by Means of Frequency-Based Feature Extraction Using Machine Learning and Deep Learning Approaches

Feyyaz ALPSALAZ^{1,a}

¹Yozgat Bozok University, Akdağmadeni Vocational Schools, Yozgat, Türkiye

^aORCID: 0000-0002-7695-6426

Article Info

Received : 18.03.2025

Accepted : 10.09.2025

DOI: 10.21605/cukurovaumfd.1660218

Corresponding Author

Feyyaz ALPSALAZ

feyyaz.alpsalaz@bozok.edu.tr

Keywords

Machine imbalance fault

MLP

Vibration analysis

Machine learning

How to cite: ALPSALAZ, F., (2025). Detection of Imbalance Faults in Industrial Machines by Means of Frequency-Based Feature Extraction Using Machine Learning and Deep Learning Approaches. Çukurova University, Journal of the Faculty of Engineering, 40(3), 581-592.

ABSTRACT

This study investigated the effectiveness of machine learning and deep learning models in diagnosing imbalance faults in industrial machines, using Fast Fourier Transform (FFT) for frequency-based feature extraction. As imbalance shortens equipment life and increases maintenance costs, vibration data was analysed and frequency components were extracted using FFT for classification. Support Vector Machine, Random Forest and Multi-Layer Perceptron models were then compared using the metrics of accuracy, precision, recall and F1 score. The Multi-Layer Perceptron model performed best with 99% accuracy, capturing the patterns extracted by FFT most effectively. Random Forests made successful predictions, but had a high error rate in some classes. Support Vector Machines, on the other hand, offered lower accuracy. Combining FFT with machine learning contributes to the diagnosis of faults in rotating machines. Model performance could be improved in future using larger data sets, hyperparameter optimisation and methods such as wavelet transformation.

Makine Öğrenimi ve Derin Öğrenme Yaklaşımlarını Kullanarak Frekans Tabanlı Öznitelik Çıkarımı ile Endüstriyel Makinelerde Dengesizlik Hatalarının Tespiti

Makale Bilgileri

Geliş : 18.03.2025

Kabul : 10.09.2025

DOI: 10.21605/cukurovaumfd.1660218

Sorumlu Yazar

Feyyaz ALPSALAZ

feyyaz.alpsalaz@bozok.edu.tr

Anahtar Kelimeler

Makine dengesizlik arızası

MLP

Titreşim analizi

Makine öğrenmesi

Atıf şekli: ALPSALAZ, F., (2025). Makine Öğrenimi ve Derin Öğrenme Yaklaşımlarını Kullanarak Frekans Tabanlı Öznitelik Çıkarımı ile Endüstriyel Makinelerde Dengesizlik Hatalarının Tespiti. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 581-592.

ÖZ

Bu çalışmada, endüstriyel makinelerde dengesizlik hatalarının teşhisinde makine öğrenimi ve derin öğrenme modellerinin etkinliği incelenmiş, frekans tabanlı özellik çıkarımı için Hızlı Fourier Dönüşümü (FFT) kullanılmıştır. Dengesizlik, ekipman ömrünü kısaltıp bakım maliyetlerini artırdığından, titreşim verileri analiz edilerek FFT ile frekans bileşenleri çıkarılmış ve sınıflandırma yapılmıştır. Destek Vektör Makinaları, Rastgele Ormanlar ve Çok Katmanlı Algılayıcı modelleri; doğruluk, hassasiyet, geri çağırma ve F1-skoru metrikleriyle karşılaştırılmıştır. Çok Katmanlı Algılayıcı, %99 doğrulukla en iyi performansı göstermiş, FFT ile çıkarılan örüntüleri en iyi yakalamıştır. Rastgele Ormanlar başarılı tahminler yapmış ancak bazı sınıflarda hata oranı yüksek bulunmuştur. Destek Vektör Makinaları ise daha düşük doğruluk sunmuştur. FFT ve makine öğrenimi kombinasyonu, döner makine arıza teşhisine katkı sağlamaktadır. Gelecekte, daha büyük veri setleri, hiperparametre optimizasyonu ve dalgacık dönüşümü gibi yöntemlerle model performansı artırılabilir.

1. INTRODUCTION

Imbalance faults, which are common in machine systems, reduce equipment performance, increase vibration levels and cause serious productivity losses in industrial processes due to increased maintenance costs. If such imbalance faults in machine elements are not detected and eliminated at an early stage, unexpected equipment failures may occur, resulting in long-term downtime in production processes. Therefore, the development of effective methods for diagnosing and preventing imbalance-related problems is of great importance for the reliability and sustainability of industrial applications. Improving the accuracy of fault diagnosis, particularly through the use of methods such as vibration analysis, signal processing techniques and artificial intelligence-based approaches, can make a significant contribution to extending the life of machinery and reducing operating costs.

In recent years, machine learning and artificial intelligence-based approaches have gained an important place in fault diagnosis and predictive maintenance processes [1]. Traditional methods are usually based on expert knowledge and cannot provide sufficient efficiency in automated decision-making processes with large data sets. In particular, the use of vibration analysis, signal processing techniques and deep learning-based models allows for more accurate detection of imbalance faults [2].

Accordingly, several studies have been conducted to evaluate the impact of different machine learning algorithms on fault diagnosis. Recent studies have shown that classification algorithms such as Support Vector Machines (SVM), Random Forest and Multilayer Perceptron (MLP) provide high accuracy in imbalance fault diagnosis [3]. However, a detailed comparison of model performance and their integration into industrial processes is of great importance for the development of an effective fault diagnosis approach [4].

In this study, a machine learning based model for imbalance fault diagnosis is developed, Fast Fourier Transform (FFT) is used for frequency-based feature extraction and the most appropriate method is determined by comparing different algorithms. The data set used consists of a wide range of vibration measurements including imbalance scenarios under different operating conditions. The time domain signals are decomposed into frequency components using FFT. Through this transformation, important frequency components in the vibration data are extracted as features and given as input to the classification algorithms. Three different classification algorithms such as SVM, Random Forest and MLP were used to evaluate the success of the model and the results obtained were compared using metrics such as accuracy, precision, recall and F1 score. By combining FFT with frequency-based feature extraction and machine learning techniques, the model proposed in the study aims to support early diagnosis in industrial maintenance processes, contribute to the prevention of machine failures and reduce maintenance costs.

2. LITERATURE REVIEW

In this section various studies have addressed data generation and balancing techniques to solve the imbalanced data problem, and especially the effectiveness of deep learning-based methods in this area has been examined.

Zhang et al. addressed the issue of imbalanced data in machine fault diagnosis using generative adversarial networks (GANs). By generating realistic synthetic data, the model's generalization ability was enhanced, enriching the training dataset. Tests on rotating machinery datasets validated the effectiveness of this approach [5].

Jiang et al. developed a temporal-spatial multi-order weighted graph convolution network (TSMOW-GCN) to improve fault diagnosis under imbalanced data conditions. The model enhances feature associations within graph structures and aggregates high-order neighborhood information, avoiding traditional data generation approaches [6].

Zhao et al. introduced an improved weighted extreme learning machine (IWELM) with an adaptive cost-sensitive strategy for fault diagnosis with imbalanced data. The model optimizes a cost-sensitive matrix and employs a multi-objective optimizer, improving classification performance, especially for minority classes [7].

Jia et al. proposed a deep normalized convolutional neural network (DNCNN) to address imbalanced fault classification in machinery diagnosis. The framework integrates a weighted softmax loss function to counter class imbalance and uses neuron activation maximization (NAM) for improved interpretability. Experiments show that DNCNN outperforms conventional CNN-based models [8].

He et al. applied contrastive feature-based deep reinforcement learning (D3QN) for imbalanced fault quantitative diagnosis under variable working conditions. By integrating SimCLR-based contrastive learning with prioritized experience replay, the model optimally learns discriminative features while addressing dataset imbalance [9].

Shi et al. proposed a graph embedding-based deep broad learning system (GEDBLS) for fault diagnosis with imbalanced data in rotating machinery. The model improves the classification loss function by considering class weights and intra-class compactness. Experiments confirm the model's superior feature extraction and data imbalance processing capabilities [10].

Pan et al. introduced a robust smooth constrained matrix machine (RSCMM) for fault diagnosis of roller bearings under imbalanced data conditions. The model employs a dynamically adjusted loss term to improve classification performance, mitigate noise impact, and accelerate convergence. Experimental results show that RSCMM achieves high accuracy under different imbalance ratios [11].

Wang et al. proposed a trackable multi-domain collaborative generative adversarial network (TMCGAN) for rotating machinery fault diagnosis under imbalanced conditions. The model improves data augmentation by integrating multi-domain feature learning and parallel frequency loss, enhancing classification credibility and interpretability [12].

Wu et al. introduced a holistic semi-supervised method for imbalanced fault diagnosis with out-of-distribution samples. Their framework uses pseudo-labelling and consistency regularization combined with an adaptive threshold to enhance class confidence balancing and robustness in real-world industrial applications [13].

Lin et al. developed a generalization classification regularization generative adversarial network (GCRGAN) to tackle data imbalance in machinery fault diagnosis. The model incorporates a generalization module and a novel regularization loss to improve sample diversity and stability under limited training data scenarios [14].

Yang et al. proposed an improved generative adversarial network (IGAN) combined with an enhanced deep extreme learning machine (EDELm) for chiller fault diagnosis under imbalanced data conditions. The method integrates multi-head attention (MHA) in GAN and adaptive boosting (AdaBoost) for sample weighting, achieving high classification accuracy [15].

Chang et al. designed an extended attention signal transformer with adaptive class imbalance loss (EAST-ACIL) to address long-tailed data distribution in rotating machinery fault diagnosis. The framework enhances feature extraction using CNN-transformer architecture while dynamically reweighting training data to prioritize rare fault classes [16].

Li et al. introduced an auxiliary classifier Wasserstein generative adversarial network with gradient penalty (ACWGAN-GP) for rotating machinery fault diagnosis under data imbalance. The method generates high-quality synthetic fault samples to balance datasets, improving model accuracy [17].

3. MATERIAL AND METHOD

3.1. Dataset

The MAFAULDA (Machinery Fault Database) consists of 1951 multivariate time-series data collected using the SpectraQuest Alignment-Balance-Vibration (ABVT) Machinery Fault Simulator (MFS). The dataset includes six different machine states: normal operation, imbalance, horizontal and vertical misalignment, and inner and outer bearing faults. These conditions were systematically simulated to provide a comprehensive dataset for fault diagnosis in rotating machinery [18].

Imbalance faults were introduced by attaching additional weights ranging from 6 g to 35 g to the rotor. The data collection process maintained the same 49 different rotational speed levels used for normal operation. However, for weights 30 g and above, excessive vibration limited the system's maximum operational speed to 3300 rpm due to stability concerns. The measurements for each imbalance weight condition are detailed in the Table 1 below:

Table 1. Weight distribution and number of measurements

Weight (g)	Number of measurements
6	49
10	48
15	48
20	49
25	47
30	47
35	45
Total	333

Each dataset instance consists of 5 seconds of recorded data at a 50 kHz sampling rate, totaling 250000 samples per measurement. The imbalance fault conditions were tested using a 1/4 HP DC motor operating within a 700-3600 rpm speed range. The test system, weighing 22 kg, included a 16 mm diameter shaft with a 520 mm shaft length, while the rotor diameter measured 15.24 cm, and the bearing span was 390 mm.

The data acquisition system incorporated various sensors to capture machine dynamics:

- Accelerometers: Three IMI Sensors Model 601A01 were used to measure radial, axial, and tangential vibrations, while a triaxial IMI Sensors Model 604B31 captured three-dimensional acceleration data.
- Tachometer: A Monarch Instrument MT-190 analog tachometer was employed for speed measurement.
- Microphone: A Shure SM81 microphone recorded audio data within a 20 Hz - 20 kHz frequency range.
- Data Acquisition Modules: Two National Instruments NI 9234 modules were utilized, operating at a 51.2 kHz sampling rate to collect vibration, speed, and audio signals.

3.2. FFT

The Fourier transform is a powerful mathematical tool that decomposes a signal in the time domain into its frequency components. This transform, in both continuous and discrete versions, is used in many areas of engineering and science [19]. The Discrete Fourier Transform (DFT) is used for frequency analysis of finite-length signals, and the FFT is an algorithm that optimises the computational process of the DFT [20]. The DFT is defined as

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j(2\pi/N)kn} \quad (1)$$

Through DFT, the signal can be decomposed into amplitude and phase information of different frequency components, where $X[k]$ is the signal in the frequency domain after Fourier transform, $x[n]$ is the n th data point in the signal in the time domain, N is the total number of data points, and j is the imaginary unit.

3.3. SVM Model

SVM is a supervised learning technique that is effectively used in data classification and regression problems [21]. SVM performs classification by determining the best hyperplane that separates data points [22]. This hyperplane is chosen to maximise the margin between two classes. For a linearly separable data set, the hyperplane is expressed as follows:

$$\omega^T x + b = 0 \quad (2)$$

$$\min_{\omega, b} \frac{1}{2} \|\omega\|^2 \quad | \quad y_i(\omega^T x_i + b) \geq 1, \quad \forall i \quad (3)$$

In the equation, ω is the direction that determines the normal vector of the hyperplane that distinguishes which class the data points belong to. The x is the input feature vector that represents each data point. The bias term b optimises the classification process by adjusting the position of the hyperplane.

3.4. RF Model

Random forest is an ensemble learning method based on decision trees and is used in both classification and regression problems [23]. The model created by combining several decision trees provides higher accuracy and generalisation ability than a single tree [24]. For classification, the prediction of the model is expressed as follows:

$$\hat{y} = \frac{1}{T} \sum_{t=1}^T h_t(x) \quad (4)$$

Where T is the total number of trees, $h_t(x)$ is the prediction class and $\hat{y}$ is the class label. Random forest uses random subsampling (bootstrapping) and random feature selection during training for each tree to increase the diversity of decision trees and avoid overlearning. This method is particularly successful on large and complex datasets.

3.5. MLP Model

MLP is a feedforward Artificial Neural Network model with a fully connected structure and is a powerful deep learning architecture widely used for tasks such as learning and classifying complex non-linear relationships [25]. In this study, an artificial ANN model with MLP architecture is used. The model is designed to solve a multi-class classification problem and has a deep network structure consisting of six hidden layers. Rectified Linear Unit (ReLU) activation function is used in each hidden layer of the model and Softmax activation is preferred in the output layer for multi-class prediction. The weights of the model were initially assigned using the random uniform distribution method and trained using the Adam optimisation algorithm. The loss function used was `sparse_categorical_crossentropy`, which is widely used in multiclass classification problems. In addition, an early stopping mechanism was added to minimise the risk of overfitting and to prevent unnecessary prolongation of the training process if no improvement is observed within the specified tolerance level. The labels in the input to the model are processed with `LabelEncoder` and converted into numerical form to convert them into a format suitable for machine learning models. The proposed model is a fully connected neural network structure and is applicable to classification tasks on a large dataset.

3.6. Performance Metrics

To evaluate the success of machine learning models, key performance metrics such as precision, recall, accuracy, and F1 score are used [26,27]. These metrics are critical for measuring the predictive ability of the model and analysing classification errors. TP (True Positive), TN (True Negative), FP (False Positive) and FN (False Negative) values are required for the calculations.

Accuracy is the ratio of the number of samples correctly predicted by the model to the total number of samples and is calculated as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

Precision, The proportion of samples predicted to be positive by the model that are actually positive:

$$F1_{score} = \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

4. EXPERIMENTAL RESULTS

In the experimental studies, SVM, RF and MLP models were used to evaluate the classification performance of imbalance errors. Performance metrics such as accuracy, precision, recall and F1-score were calculated and compared to measure the success of the models. The learning processes of each model were examined by analyzing the accuracy and loss values during the training process. In addition, the confusion matrix was used to determine the incorrect predictions between classes and the classification ability of the model was evaluated in detail.

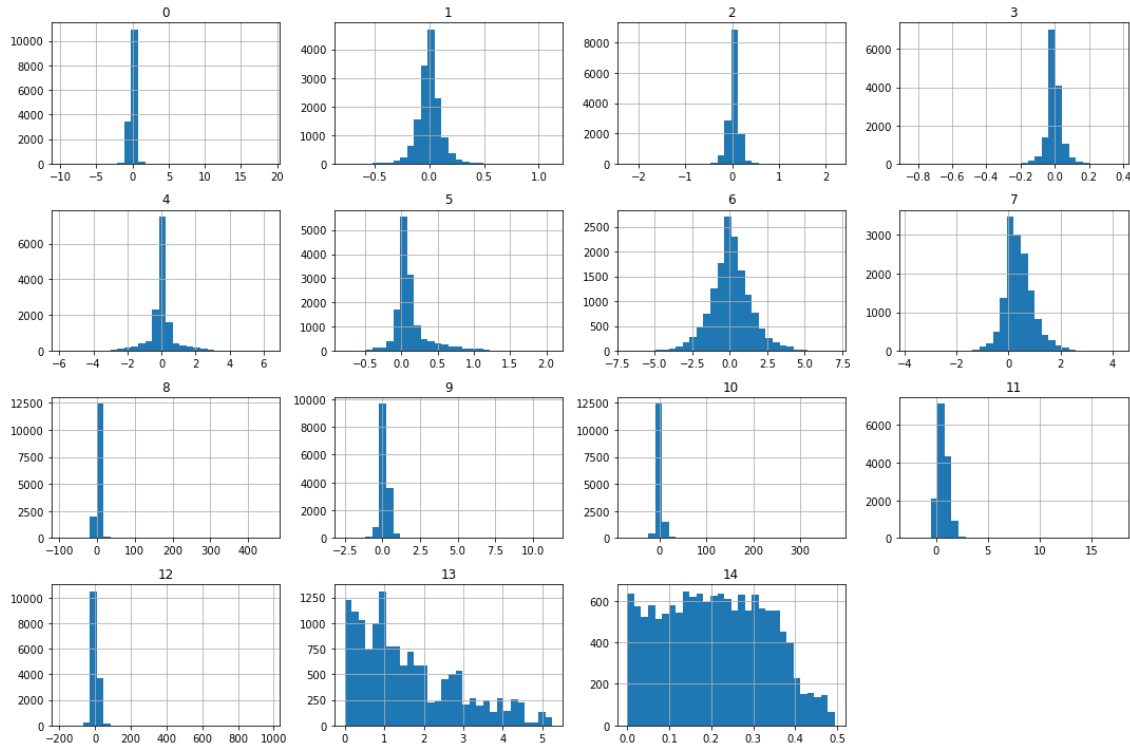


Figure 1. Distribution of attributes in the frequency spectrum

Figure 1 shows histograms showing the distribution of different features in the dataset. These histograms provide information about the overall structure of the dataset by visualising the frequency of values for each variable. The figure shows that some features are close to a normal distribution, while others are skewed to the right or left. In particular, the histograms in the top rows are bell-shaped and close to a normal distribution, while some variables in the bottom rows show significant skewness and irregularity.

These distributional analyses provide important information that should be taken into account as the model learns. While normally distributed characteristics allow the model to learn more consistently, skewed distributions can negatively affect the learning process. In addition, the presence of a high number of outliers in some variables may reduce the generalisation ability of the model and may require the application of special techniques in the data pre-processing phase.

In this context, scaling and transformation of features is an important step to improve model performance. In particular, methods such as Min-Max normalization or StandardScaler can make the data distribution more balanced. For variables with skewed distributions, techniques such as log or square root transformation can be applied to enable the model to learn these variables more efficiently. In conclusion,

the histogram analysis presented in Figure 1 plays a critical role in understanding the structure of the dataset and determining appropriate data preprocessing strategies.

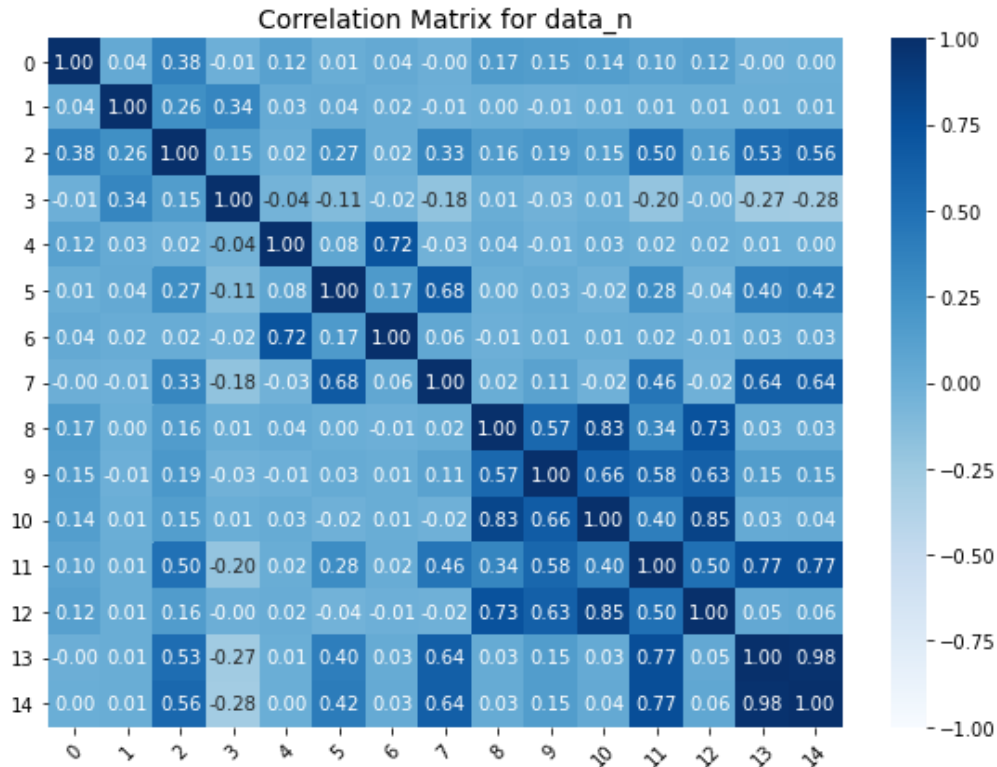
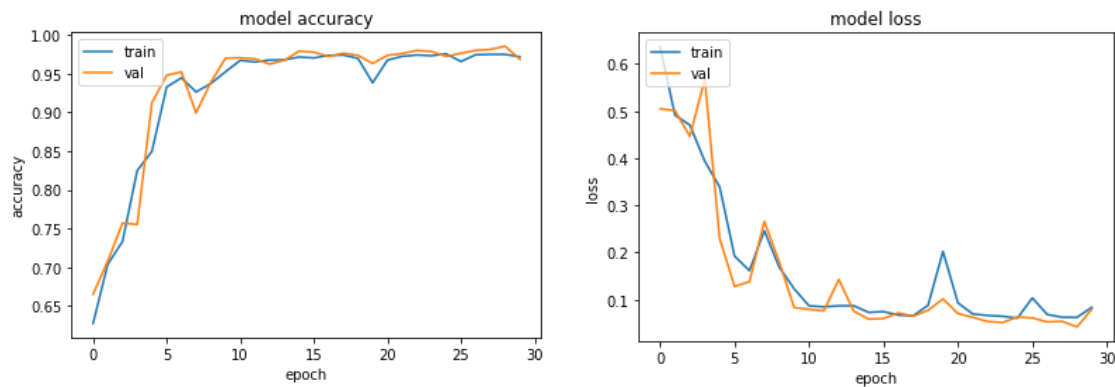


Figure 2. Correlation matrix between attributes

Figure 2 shows the correlation matrix, which visualises the relationship between the attributes in the dataset. The values in the cells represent the Pearson correlation coefficient between pairs of variables. Values range from -1 to 1, with values close to 1 indicating a strong positive correlation and values close to -1 indicating a strong negative correlation. Light blue shades represent low correlation and dark blue shades represent high correlation.

In the figure, high correlations (e.g. 0.85, 0.77) are observed between some variables, which may indicate that these attributes carry similar information. On the other hand, the correlation values between some variables are quite low (e.g. 0.01-0.03), suggesting that these variables are independent. This analysis can help to remove unnecessary or highly correlated variables in the feature selection process and increase the generalisation ability of the model.



a)

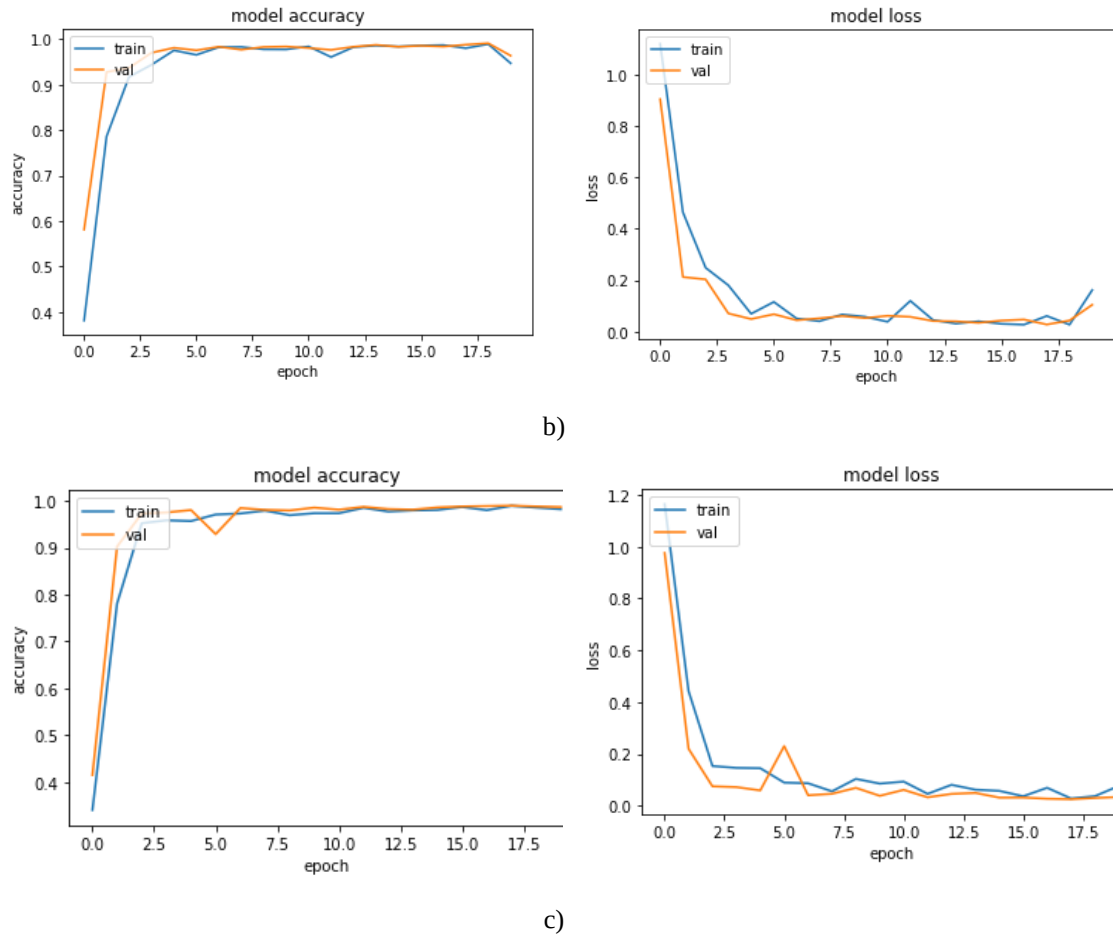


Figure 3. Model accuracy and loss during training and validation. a) SVM, b) RF, c) MLP

Figure 3 shows the accuracy and loss values of the SVM, Random Forest and MLP models in the training and validation phases. While all three models have generally high accuracy rates, the MLP model (Figure 3-c) shows the most stable learning process. Especially from the early epochs, the accuracy values increase rapidly and the training and validation accuracies are very close to each other. The loss graph shows that MLP achieved the lowest loss value compared to the other models and completed the learning process in the most balanced way.

The Random Forest model (Figure 3-b) showed a rapid increase in accuracy during the training process and the validation accuracy also reached a high level. However, the fluctuations seen in the loss plot suggest that the model has more variance in some epochs and does not perform fully balanced learning. The SVM model (Figure 3-a), on the other hand, is successful in terms of accuracy, but the difference between the training and validation accuracies is relatively larger in the early epochs, suggesting that the model initially follows a more erratic learning process. Overall, the MLP model had the most stable learning process and achieved high accuracy with lower loss compared to the other models. Due to its deep learning based structure, it has a strong ability to learn more complex patterns in the data.

Table 2. Metrics for the classification performance of the models

Algoritma	Accuracy	Precision	Recall	F1score
SVM	0.96	0.96	0.96	0.96
RF	0.98	0.98	0.98	0.97
MLP	0.99	0.99	0.99	0.99

Table 2 shows the accuracy, precision, recall and F1 scores calculated to evaluate the classification performance of the SVM, Random Forest and MLP models. Although all three models have high accuracy rates, the MLP model performed best by achieving the highest value in all metrics.

The Random Forest model showed a consistent classification performance with an accuracy of 98% and outperformed the SVM model in all metrics. The SVM model, on the other hand, performed a successful classification with an accuracy rate of 96%, but it was observed that it had lower metric values compared to the other models.

Overall, the MLP model achieved the highest success with 99% accuracy, precision, recall and F1 score. These results show that MLP has a high learning capacity in complex data structures and is more successful than other models in classification tasks.

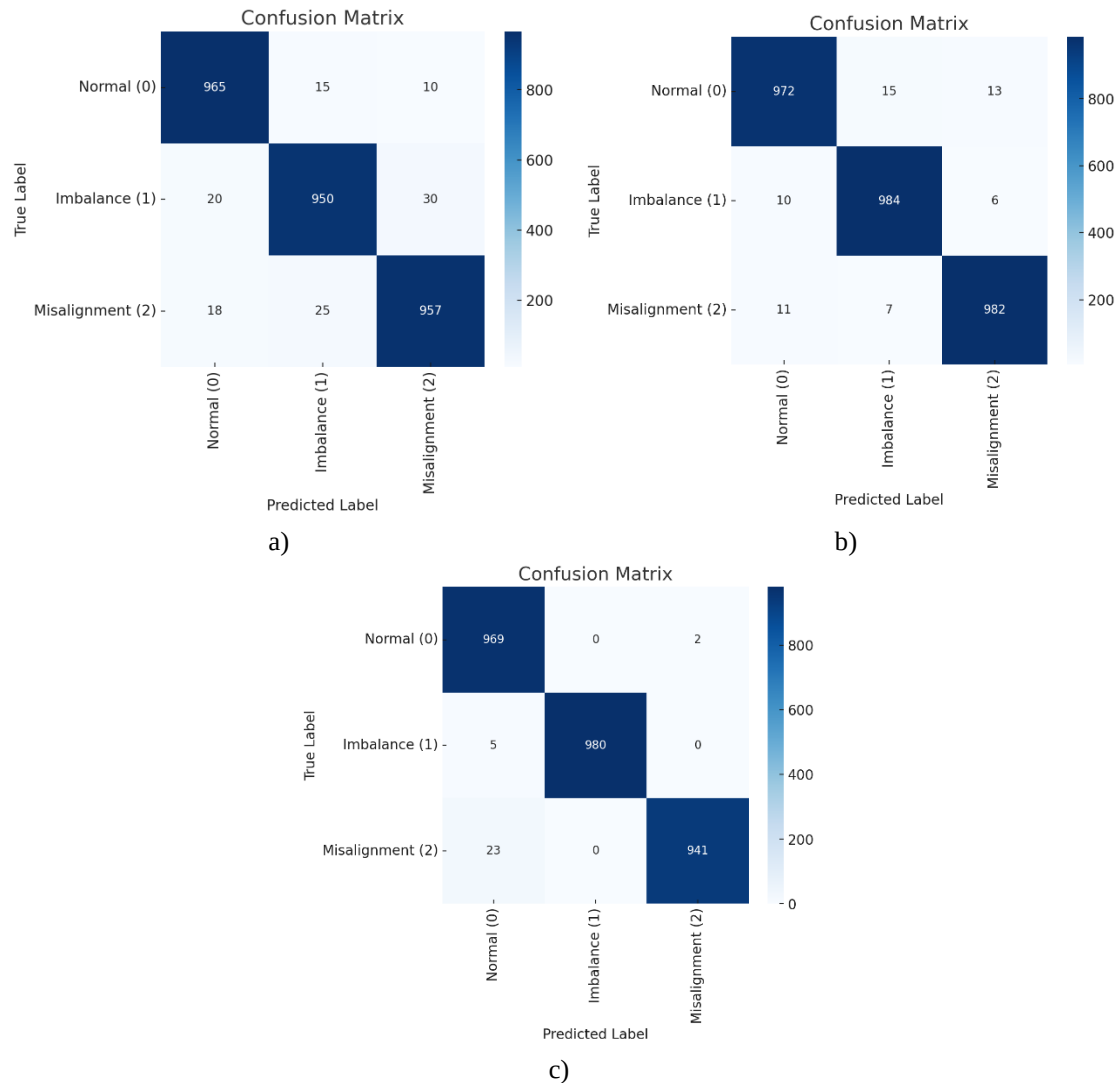


Figure 4. Confusion matrix showing the classification performance of the models. a) SVM, b) RF, c) MLP

Figure 4 shows the confusion matrices generated to compare the classification performance of the SVM, Random Forest and MLP models. Although all three models have generally high accuracy rates, the MLP model (Figure 4-c) has the best classification performance. In particular, it had the lowest error rate in the Normal (0) and Imbalance (1) classes, minimising misclassifications.

The Random Forest model (Figure 4-b), although providing high accuracy in general, made more incorrect predictions, especially in the Misalignment (2) class. Due to its decision tree-based structure, it has difficulty discriminating between some classes. The SVM model (Figure 4-a), on the other hand, made more incorrect predictions than the other models, especially in the Misalignment (2) class, indicating that the model has difficulty discriminating between some classes.

The MLP model achieved the highest accuracy rate and performed more evenly across classes. Thanks to its deep learning-based structure, it has a high capacity to learn more complex data patterns.

Table 3. Bootstrap 95% confidence intervals for classification performance metrics

Model	95% CI (accuracy)	95% CI (precision)	95% CI (Recall)	95% CI (F1score)
SVM	0.949–0.971	0.948–0.971	0.949–0.971	0.949–0.971
RF	0.973–0.987	0.973–0.987	0.973–0.987	0.969–0.985
MLP	0.985–0.995	0.985–0.995	0.985–0.995	0.985–0.995

Table 3 presents the 95% confidence intervals (CI) for the classification performance metrics—Accuracy, Precision, Recall, and F1-score—of the SVM, RF, and MLP models. The CIs were computed using bootstrap resampling with $n = 300$ iterations, providing a robust estimation of the variability in model performance. The results demonstrate that the MLP model achieved the narrowest confidence intervals across all metrics, indicating a high degree of stability and consistency in its predictions. The RF model also exhibited strong and reliable performance with relatively tight intervals, whereas the SVM model, despite slightly wider intervals, maintained competitive classification accuracy and balanced metric values.

Table 4. Hyperparameter search ranges, final selections, and rationale for each model

Component	Search range / limit	Final selection (example)	Rationale
SVM (RBF)	$C \in [10^{-3}, 10^3]$, $\gamma \in [10^{-4}, 10^1]$, kernel=RBF	$C=10$, $\gamma=0.01$ 5-fold CV Macro-F1, Accuracy; CI narrowing	Prevents overfitting while maintaining a stable decision boundary
Random forest	$n_estimators \in \{200, 400, 800\}$; $max_depth \in \{None, 10, 20, 40\}$; $min_samples_split \in \{2, 5, 10\}$; $max_features \in \{sqrt, log2\}$; bootstrap=True	$n_estimators=400$, $max_depth=20$, $max_features=sqrt$ 5-fold CV Macro-F1; training time	Variance-bias balance and stable performance across classes
MLP	Layers: 3–6; Neurons: {256,128,64,32}; Dropout: 0.2–0.5; Norm: Batch/Layer (opt.); L2: {0, 10^{-5} , 10^{-4} }; Epoch ≤ 200 , $min_delta=10^{-4}$; LR: $10^{-3} \rightarrow 10^{-4}$ (plateau)	256-128-64, Dropout=0.3, BN included, $L2=10^{-5}$, 5-fold CV Macro-F1 & Accuracy; CI narrowing; training time	Highest accuracy + narrow CI; prevents overfitting

Table 4 summarizes the hyperparameter search spaces, the final selected configurations, and the rationale for each model. The selection process prioritized high classification performance, narrow bootstrap confidence intervals, and prevention of overfitting, while ensuring balanced performance across all classes.

5. CONCLUSION

This study compares the performance of machine learning and deep learning-based models for imbalance fault diagnosis using FFT for frequency-based feature extraction. In the experimental process, SVM, Random Forest and MLP models were applied and their classification performance was evaluated using metrics such as accuracy, precision, recall and F1 score. Confusion matrices and model learning processes were investigated and the strengths and weaknesses of each model were analysed. In the study, time domain vibration data was decomposed into frequency components using FFT and these components were used as input to the classification models. In this way, imbalance faults can be more accurately distinguished on the basis of distinct frequency features. According to the results obtained, the MLP model has the highest accuracy rate and the lowest error rate, and has the best performance in identifying imbalance faults by learning the features extracted by FFT in the best way. Thanks to its deep learning-based structure, it was found to have a greater ability to learn complex patterns by better modelling the relationships between frequency components. The Random Forest model generally made successful predictions, but the false prediction rate remained relatively high in some classes. Although the SVM model had a stable learning

process, it achieved lower accuracy values compared to the other models. This study demonstrates how machine learning and deep learning-based models can be used in rotating machinery fault diagnosis by extracting frequency components with FFT and comparatively evaluates the effectiveness of different algorithms. In the future, the generalisation capability of the models can be increased by using larger and more diverse data sets, and the performance of the models can be further improved by hyper-parameter optimisation. In addition, the accuracy rates for imbalance fault detection can be further improved by using different time-frequency transforms (e.g. wavelet transform) or advanced deep learning architectures.

6. REFERENCES

1. Çınar, M., Aslan, E. & Özüpak, Y. (2025). Comparison and optimization of machine learning methods for fault detection in district heating and cooling systems. *Bulletin of the Polish Academy of Sciences Technical Sciences*, 154063-154063.
2. Li, L., Liu, K., Wang, L., Sun, L., Zhang, Z. & Guo, H. (2022). Fault diagnosis of balancing machine based on ISSA-ELM. *Computational Intelligence and Neuroscience*, 2022(1), 4981022.
3. Hu, Y., Lv, W., Wang, Z., Liu, L. & Liu, H. (2023). Error prediction of balancing machine calibration based on machine learning method. *Mechanical Systems and Signal Processing*, 184, 109736.
4. Benkaihoul, S., Mazouz, L., Naas, T. T., Yildirim, Ö. & Mohammadi, R.D. (2024). Broken magnets fault detection in PMSM using a convolutional neural network and SVM. *ITEGAM-JETIA*, 10(48), 55-62.
5. Zhang, W., Li, X., Jia, X.D., Ma, H., Luo, Z. & Li, X. (2020). Machinery fault diagnosis with imbalanced data using deep generative adversarial networks. *Measurement*, 152, 107377.
6. Jiang, Z., Liu, D. & Cui, L. (2025). A temporal-spatial multi-order weighted graph convolution network with refined feature topology graph for imbalance fault diagnosis of rotating machinery. *Measurement Science and Technology*, 36(2), 110830.
7. Zhao, Y., Yang, X., Huang, J., Gao, J. & Cui, J. (2024). Improved weighted extreme learning machine with adaptive cost-sensitive strategy for imbalanced fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing*, 217, 111526.
8. Jia, F., Lei, Y., Lu, N. & Xing, S. (2018). Deep normalized convolutional neural network for imbalanced fault classification of machinery and its understanding via visualization. *Mechanical Systems and Signal Processing*, 110, 349-367.
9. He, S., Cui, Q., Chen, J., Pan, T. & Hu, C. (2024). Contrastive feature-based learning-guided elevated deep reinforcement learning: Developing an imbalanced fault quantitative diagnosis under variable working conditions. *Mechanical Systems and Signal Processing*, 211, 111192.
10. Shi, M., Ding, C., Wang, R., Shen, C., Huang, W. & Zhu, Z. (2023). Graph embedding deep broad learning system for data imbalance fault diagnosis of rotating machinery. *Reliability Engineering & System Safety*, 240, 109601.
11. Pan, H., Li, B., Zheng, J., Tong, J., Liu, Q. & Deng, S. (2024). Research on roller bearing fault diagnosis based on robust smooth constrained matrix machine under imbalanced data. *Advanced Engineering Informatics*, 62, 102667.
12. Wang, X., Jiang, H., Mu, M. & Dong, Y. (2025). A trackable multi-domain collaborative generative adversarial network for rotating machinery fault diagnosis. *Mechanical Systems and Signal Processing*, 224, 111950.
13. Wu, Z., Xu, R., Luo, Y. & Shao, H. (2024). A holistic semi-supervised method for imbalanced fault diagnosis of rotational machinery with out-of-distribution samples. *Reliability Engineering & System Safety*, 250, 110297.
14. Lin, C., Kong, Y., Huang, G., Han, Q., Dong, M., Liu, H. & Chu, F. (2025). Generalization classification regularization generative adversarial network for machinery fault diagnostics under data imbalance. *Reliability Engineering & System Safety*, 256, 110791.
15. Yang, W., Zhang, H., Lim, J.B., Zhang, Y. & Meng, H. (2024). A new chiller fault diagnosis method under the imbalanced data environment via combining an improved generative adversarial network with an enhanced deep extreme learning machine. *Engineering Applications of Artificial Intelligence*, 137, 109218.
16. Chang, S., Wang, L., Shi, M., Zhang, J., Yang, L. & Cui, L. (2024). Extended attention signal transformer with adaptive class imbalance loss for Long-tailed intelligent fault diagnosis of rotating machinery. *Advanced Engineering Informatics*, 60, 102436.

17. Li, Z., Zheng, T., Wang, Y., Cao, Z., Guo, Z. & Fu, H. (2021). A novel method for imbalanced fault diagnosis of rotating machinery based on generative adversarial Networks. *IEEE Transactions on Instrumentation and Measurement*, 70.
18. Ribeiro, F.M.L. (2021). Machinery fault database. Signals, Multimedia, and Telecommunications Laboratory. https://www02.smt.ufrj.br/~offshore/mfs/page_01.html
19. Alpsalaz, F. (2025). Fault detection in power transmission lines: Comparison of chirp-Z algorithm and machine learning based prediction models. *Eksploracja i Niezawodność - Maintenance and Reliability* 27(4), 203949.
20. Aslan, E., Özüpak, Y., Alpsalaz, F. & Elbarbary, Z.M.S. (2025). A hybrid machine learning approach for predicting power transformer failures using internet of things-based monitoring and explainable artificial intelligence. *IEEE Access*, 13, 113618-113633.
21. Aslan, E. & Özüpak, Y. (2025). Comparison of machine learning algorithms for automatic prediction of alzheimer disease. *Journal of the Chinese Medical Association*, 88(2), 98-107.
22. Turkey, Y. & Tamay, Z.S. (2024). Pistachio classification based on acoustic systems and machine learning. *Elektronika Ir Elektrotechnika*, 30(5), 4-13.
23. Khan, M.A., Asad, B., Vaimann, T., Kallaste, A., Pomarnacki, R. & Hyunh, V.K. (2023). Improved fault classification and localization in power transmission networks using VAE-generated synthetic data and machine learning algorithms. *Machines* 11(10), 963.
24. Özüpak, Y. (2025). Machine learning-based fault detection in transmission lines: A comparative study with random search optimization. *Bulletin of the Polish Academy of Sciences Technical Sciences*, 153229-153229.
25. Aslan, E. (2024). Araçlarda CO2 emisyonlarının farklı yapay sinir ağı modelleri kullanılarak tahminlerinin karşılaştırılması. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 39(2), 309-324.
26. Özüpak, Y. (2024). Evrişimli sinir ağı (ESA) mimarileri ile hücre görüntülerinden sıtmanın tespit edilmesi. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 39(1), 197-210.
27. Uluocak, İ. (2025). Comparative study of emission prediction using deep learning models. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 40(2), 337-346.

Tıbbi Akıl Yürütme için Gemma 3 Modellerinin CoT Destekli SFT ve GRPO ile İncelenmesi

İsmail İŞERİ^{1,a}, Alper YILDIRIM^{2,b}, Alihan ÖZTORUN^{2,c}, Tuğba TUNA^{3,d},
Arda TURAN^{3,e}

¹Ondokuzmayıs Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Samsun, Türkiye

²Düzce Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Düzce, Türkiye

³Arena Bilgisayar A.Ş., İstanbul, Türkiye

^aORCID: 0000-0002-0442-1406; ^bORCID: 0009-0004-8029-1772; ^cORCID: 0009-0006-2931-4945;

^dORCID: 0009-0002-0352-593X; ^eORCID: 0009-0007-1145-2888

Makale Bilgileri

Geliş : 01.08.2025

Kabul : 09.09.2025

DOI: 10.21605/cukurovaumfd.1755684

Sorumlu Yazar

İsmail İŞERİ

ismail.iseri@omu.edu.tr

Anahtar Kelimeler

Büyük dil modelleri

Tıbbi akıl yürütme

Düşünce zinciri

Grup görelisi politika optimizasyonu

Auf şekli: İŞERİ, İ., YILDIRIM, A., ÖZTORUN, A., TUNA, T., TURAN, A., (2025). Tıbbi Akıl Yürütme için Gemma 3 Modellerinin CoT Destekli SFT ve GRPO ile İncelenmesi. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 593-606.

ÖZ

Bu çalışmada, Gemma 3 1B ve Gemma 4B büyük dil modellerinin tıbbi alandaki karmaşık akıl yürütme yeteneklerinin geliştirilmesi ve değerlendirilmesi amaçlanmıştır. Bu kapsamda, SFT (Gözetimli İnce Ayar) ve GRPO (Grup Görelisi Politika Optimizasyonu) gibi eğitim stratejilerinin Gemma 3 1B ve Gemma 4B modelleri üzerindeki performansı incelenmiştir. Temel modellerin değerlendirilmesi sürecinde, SFT ile Düşünce Zinciri (CoT) formatının öğretilmesi ve GRPO ile akıl yürütmenin rafine edilmesi adımlarını içeren çok aşamalı bir yaklaşım izlenmiştir. GPT-4.1 hakem modeliyle yapılan değerlendirmeler sonucunda, model performansının belirgin şekilde artırıldığı gösterilmiştir. SFT ve GRPO eğitimleriyle modelin mantıksal olarak tutarlı bir akıl yürütme süreci oluşturma becerisinin (Mantık Yürütme Doğruluğu'nun %26'dan %31'e yükseltilmesiyle) başarıyla geliştirildiği ortaya konmuştur. Bu durum, modelin cevapları ezberlemek yerine nasıl düşüneceğinin öğrenildiğini kanıtlamaktadır.

Investigation of Low-Parameter Gemma 3 Models for Medical Reasoning Using CoT-Supported SFT and GRPO

Article Info

Received : 01.08.2025

Accepted : 09.09.2025

DOI: 10.21605/cukurovaumfd.1755684

Corresponding Author

İsmail İŞERİ

ismail.iseri@omu.edu.tr

Keywords

Large language models

Medical reasoning

Chain-of-thought (CoT)

Group relative policy optimization (GRPO)

How to cite: İŞERİ, İ., YILDIRIM, A., ÖZTORUN, A., TUNA, T., TURAN, A., (2025). Investigation of Low-Parameter Gemma 3 Models for Medical Reasoning Using CoT-Supported SFT and GRPO. Çukurova University, Journal of the Faculty of Engineering, 40(3), 593-606.

ABSTRACT

This study aimed to develop and evaluate the complex reasoning capabilities of the Gemma 3 1B and Gemma 4B large language models within the medical domain. In this context, the performance of training strategies such as SFT (Supervised Fine-Tuning) and GRPO (Group Relative Policy Optimization) on the Gemma 3 1B and Gemma 4B models was investigated. A multi-stage approach was followed, starting with the evaluation of the base models, then teaching the Chain-of-Thought (CoT) format via SFT, and finally refining the reasoning process with GRPO. Evaluations conducted using the GPT-4.1 as a judge model demonstrated a significant improvement in model performance. It was shown that SFT and GRPO training successfully enhanced the model's ability to generate a logically consistent reasoning process, evidenced by an increase in Reasoning Accuracy from 26% to 31%. This outcome proves that the model learned how to think rather than merely memorizing answers.

1. GİRİŞ

Büyük dil modelleri (Large Language Model - LLM), klinik karar destek sistemlerinden hasta eğitime kadar uzanan geniş bir yelpazede sundukları potansiyel ile sağlık hizmetleri alanında devrim yaratma potansiyeline sahip olunmaktadır. Literatürdeki çalışmalarda, çeşitli GPT tabanlı modellerin; tıbbi bilgi çıkarımı, teşhis desteği ve hastalık tahmini gibi spesifik görevlerdeki başarısı ortaya konmuştur. Benzer şekilde, küçük dil modelleri (Small Language Model- SLM) ise maliyet etkinliği ve veri gizliliği avantajlarıyla öne çıkarılmıştır. Akıl yürütme yeteneklerinin geliştirilmesi amacıyla çalışmalarda takviyeli öğrenme gibi gelişmiş yöntemlerin kullanıldığı ve modellerin doğruluğunu artırmaya odaklanıldığı görülmektedir. LLM'lerin tıbbi alandaki yeteneklerini artırma literatürü içindeki özgün katkımız, HuatuoGPT [1] ve MedReason [2] gibi yaklaşımların ürettiği akıl yürütme verilerinin güvenilirliğini ölçmek ve doğrulamak için temel bir mekanizma ve metodoloji sunarak bu "kendini iyileştirme döngüsü"nün kalitesini yükseltmeye odaklanmaktadır.

Ancak mevcut çalışmaların büyük bir kısmında, model performansı yalnızca nihai "Cevap Doğruluğu (Answer Accuracy)" metriği üzerinden ölçülmüştür. Bu yaklaşım, karmaşık ve açık uçlu tıbbi problemlerin değerlendirilmesinde modelin gerçek yeteneğini ölçmekte yetersiz kalabilmektedir. Bunun nedeni, bu tür problemlerin birden fazla klinik olarak geçerli yolun veya cevabın bulunmasına olanak sağlamasıdır. Ayrıca, bir modele doğru cevaba şans eseri veya mantıksal olmayan yollarla ulaşma imkânının verilmesi, klinik güvenilirlik açısından ciddi bir risk olarak değerlendirilmektedir. Mevcut çalışmalar, Bilgi Grafikleri bütünleştiren MedReason [2] gibi yaklaşımlarla sentetik veri kalitesini sürekli artıran bir "kendini iyileştirme döngüsü" üzerine kuruludur. Ancak, bu döngünün en kritik sorunu, üretilen akıl yürütmenin kalitesini ölçme problemidir.

Bu çalışmada, literatürdeki bu boşluğun doldurulması hedeflenmiştir. Araştırmamızda, birçok çalışmadan farklı olarak yalnızca modelin doğru sonuca ulaşp ulaşmadığına değil; aynı zamanda bu sonuca ulaşılan Düşünce Zinciri (Chain-of-Thought - CoT) sürecinin mantıksal tutarlılığına ve tıbbi geçerliliğine başarıyla ölçütü olarak odaklanılmıştır. Çalışmamız, yalnızca nihai sonuca değil, sürecin tıbbi sağlamlığına dayanan "Akıl Yürütme Doğruluğu" adlı yeni bir metrik önererek, davranışsal eğitimin modelin problem çözme güvenilirliğini nasıl artırdığını kantitatif olarak göstermektedir.

Bu amaçla, Gemma 3 ailesinden farklı ölçekteki (1B ve 4B) modellerin performansı; temel (baseline) durumdan başlanarak, Gözetimli İnce Ayar (Supervised Fine-Tuning - SFT) ile davranışsal formatın kazandırılması ve akıl yürütmelerinin rafine edilmesi aşamalarından geçirilerek sistematik biçimde analiz edilmektedir.

Çalışmamızda ortaya konan temel hipotez, SFT ve GRPO gibi eğitimlerin, cevap doğruluğunda bir artış sağlanmasa bile modelin nasıl düşüneceğinin öğrenilmesini mümkün kılarak Mantık Yürütme Doğruluğunun (Reasoning Accuracy) geliştirilebileceğidir. Bu sayede, tıbbi uygulamalarda daha güvenilir ve yorumlanabilir modellerin geliştirilmesinin önemi vurgulanmıştır.

1.1. İlgili Çalışmalar

Büyük dil modellerinin (LLM) sağlık ve biyomedikal alanlardaki en kritik gelişim alanlarından biri, karmaşık akıl yürütme yeteneklerinin artırılmasıdır. Bu amaçla araştırmacılar, takviyeli öğrenme ve özel eğitim gibi çeşitli teknikler geliştirmektedir. Örneğin, Junying Chen ve arkadaşları [1], doğrulayıcı destekli eğitimle HuatuoGPT-O1 modelini geliştirerek bu alandaki etkinliği kanıtlamıştır. Benzer bir yaklaşımla Yuxiang Lai ve arkadaşları [3], Med-R1 adlı pekiştirmeli öğrenme yöntemiyle GPT-4 dahil birçok modeli geride bırakan ve uzman hekimlerce daha çok tercih edilen başarılı sonuçlar elde etmiştir. Bu alandaki bir diğer önemli gelişme ise, Juncheng Wu ve arkadaşlarının [2] çalışması olan büyük dil modellerinin tıbbi akıl yürütme becerisini geliştirmek için tıbbi bilgi graflerinden (knowledge graph) yararlanarak olgusal düşünce zincirleri (Chain-of-Thought) üreten MedReason metodolojisidir. Bu yaklaşımla eğitilen modeller, performanslarını önemli ölçüde artırarak kendi sınıflarında en son teknoloji (state-of-the-art) sonuçlar elde etmiştir. Akıl yürütme performansını artıran diğer önemli çalışmalar arasında DeepSeek-R1 [4] modeli, Fei Ding ve arkadaşlarının [5] geliştirdiği iki katmanlı MGRPO yöntemi ve Robert Tinn ile ekibinin [6] ince ayar (fine-tuning) sürecindeki kararsızlığı azaltan teknikleri yer almaktadır.

Modellerin akıl yürütme yeteneklerini objektif bir şekilde ölçmek ve değerlendirmek de bu sürecin ayrılmaz bir parçasıdır. Bu doğrultuda, Wei Dai ve arkadaşları [7], uzman hekimlerce hazırlanan 1.719 soruluk yüksek kaliteli QoQ-Med benchmark setini sunmuştur. Bu testler, GPT-4 ve Claude-3 gibi gelişmiş modellerin başarısını teyit ederken, bazı açık kaynaklı modellerin güvenilirlik açısından eksik kaldığını göstermiştir. Benzer şekilde, Mianxin Liu ve takım arkadaşları [8], Çin tıbbına özgü multimodal LLM'leri değerlendirmek amacıyla MedBench'i geliştirmiş ve bu alanda da ChatGPT gibi modellerin yüksek performansını ortaya koymuştur.

Mevcut modellerin performansını karşılaştıran çalışmalar, LLM'lerin potansiyelini ve sınırlarını gözler önüne sermektedir. Cagatay Umut Ogdu ve arkadaşları [9], MedMCQA ve PubMedQA veri setleri üzerinde yaptıkları testlerde, özellikle GPT-4'ün ROUGE skorları ve Semantik Benzerlik ölçümlerinde üstün olduğunu belirtmiştir. Ancak, Hazrat Ali ve arkadaşları [10] ile Jin Li ve ekibinin [11] de vurguladığı gibi, bu modellerin tıp sınavlarındaki başarı potansiyeline rağmen hata ve yanlışlık riskleri bulunmaktadır. Bu nedenle, GPT-4 ve Gemini gibi en güçlü modellerin dahi klinik kullanıma sunulmadan önce kapsamlı bir doğrulamadan geçmesi kritik bir zorunluluktur.

Büyük dil modelleri (LLM'ler), sağlık sektöründe klinik dil anlama, bilgiye erişim ve karar destek sistemleri gibi alanlarda devrim yaratma potansiyeli taşımaktadır. Bu alanda, Xi Yang ve arkadaşlarının [12] geliştirdiği GatorTron gibi modeller, karmaşık klinik doğal dil işleme görevlerinde yüksek performans göstermektedir.

LLM'lerin çok yönlü uygulamaları, birçok çalışma ile kanıtlanmıştır. Örneğin, Rajendra Kumar ve ekibi [13] LLM ve NLP entegrasyonu ile klinik karar desteği sunarken, Sirui Ding ve arkadaşları [14] CKLE yöntemiyle çok modlu EHR verilerinden sağlık olaylarını tahmin etmiştir. Cai Wang ve ekibinin [15] KEMedGPT'si tıbbi danışmanlıkta, Cheng Peng ve arkadaşlarının [16] GatorTronGPT'si biyomedikal ilişki çıkarımında, Laila Rasmy ve çalışma arkadaşlarının [17] Med-BERT'i ise hastalık tahmininde önemli başarılar elde etmiştir. Ayrıca, M. Arif Rahman ve ekibi [18] Acil Sağlık Hizmetleri'ne özgü EMS-BERT'i, Wenqiang Li ve arkadaşları [19] Çince tıbbi soru-cevaplar için DoctorGPT'yi ve Akilesh S. ile ekibi [20] kullanıcı dostu bir tıbbi sohbet botu geliştirmiştir. Georgios Peikos ve çalışma arkadaşları [21] ise küçük, açık kaynaklı LLM'lerin klinik deney eşleştirmesinde bile etkili olabildiğini göstermiştir.

Kişiselleştirilmiş tıp alanında Dr. Kumar P. ve ekibinin [22] Medify-AI sistemi ile Dr. Rakesh Kumar M ve ekibinin [23] MedHub projesi gibi kullanıcı dostu sistemler geliştirilmiştir. Shadi AlZu'bi ve arkadaşları [24], LLM'leri Makine Öğrenmesi (Machine Learning) ve Derin Öğrenme (Deep Learning) ile entegre ederek tanı doğruluğunu ve tedavi önerilerini iyileştirmiştir. Son olarak, Fernando S.P.A.A. ve Dilani Wickramaarachchi [25], sağlık hizmetine erişimi kısıtlı bölgeler için bir ön tıbbi danışmanlık platformu oluşturmuştur.

Bununla birlikte, Zabir Al Nazi ve Wei Peng [26] ile Shubo Tian ve arkadaşlarının [27] da vurguladığı gibi, bu büyük potansiyele rağmen veri gizliliği, önyargıların azaltılması ve etik gibi zorluklar halen çözüm beklemektedir.

Küçük dil modelleri (SLM'ler), büyük dil modellerinin (LLM) getirdiği yüksek maliyet ve yoğun kaynak ihtiyacı gibi zorluklara karşı sağlık alanında pratik ve erişilebilir bir çözüm olarak öne çıkmaktadır. Bu modeller, ince ayar (fine-tuning) gibi özel eğitim yöntemleriyle belirli tıbbi görevler için optimize edilebilir. Böylece hem veri gizliliği ve güvenlik avantajları sunar hem de daha verimli bir performans sergilerler.

Bu verimli uyarlamayı sağlayan yöntemlerin başında Edward Hu ve arkadaşlarının [28] yapmış olduğu çalışma olan LoRA (Low-Rank Adaptation) gelmektedir. Edward Hu ve arkadaşları tarafından önerilen bu yöntemde, büyük dil modellerini uyarlamak için ana model ağırlıkları dondurulup yalnızca küçük, "düşük rütbeli" (low-rank) matrisler eğitilir. Bu yaklaşım, eğitilebilir parametre sayısını ve bellek kullanımını büyük ölçüde azaltırken, tam yeniden eğitime (full fine-tuning) ile aynı veya daha iyi performansı, çıkarım sırasında ek gecikme olmadan sağlar.

Bu alandaki çalışmalar, SLM'lerin potansiyelini somutlaştırmaktadır. Örneğin, Aditya Karnam Gururaj Rao ve arkadaşları [29], LLama 3.2 3B gibi küçük modellerin ince ayar ile sağlık sohbet botlarında ne kadar etkili olabileceğini göstermiştir. Benzer şekilde, Hyunjae Kim ve çalışma arkadaşlarının [30] geliştirdiği Meerkat-7B ve Meerkat-8B modelleri, tıbbi akıl yürütme konusunda GPT-4'e yakın bir performans sergileyerek güvenli ve erişilebilir alternatifler sunmaktadır.

Gizlilik konusundaki potansiyeli ise Matteo Magnini ve arkadaşları [31], yerel olarak çalıştırılan SLM'lerin tıbbi sohbet botları için güvenli senaryolar oluşturabileceğini belirterek vurgulamıştır. SLM'lerin bir diğer önemli kullanım alanı ise mobil sağlıktır. Xin Wang ve arkadaşları [32], bu modellerin mobil sağlık takip uygulamalarında düşük gecikme ve az kaynak tüketimiyle ne kadar verimli çalıştığını kanıtlamıştır.

1.2. Çalışmamızın Literatürdeki Yeri ve Katkısı

Yukarıda özetlenen çalışmalar, tıbbi alanda LLM'lerin artan yeteneklerini ve çeşitliliğini ortaya koymaktadır. Bizim çalışmamız, bu geniş literatür içinde, özellikle akıl yürütme kalitesinin üretimi, eğitimi ve en önemlisi ölçümü konularına odaklanarak özgün bir katkı sunmaktadır.

Mevcut literatür, daha güçlü temel modelleri bir "jüri" olarak kullanarak sürekli daha kaliteli sentetik veri üreten bir "kendini iyileştirme döngüsü" üzerine kurulmuştur. Bu döngüde, HuatuoGPT-o1 [1] gibi çalışmalar, davranışsal bir süreçle akıl yürütme verisi üretirken; MedReason [2] gibi daha yeni yaklaşımlar, Bilgi Grafikleri (Knowledge Graph) entegre ederek üretilen verinin faktörel doğruluğunu artırmış ve döngünün kalitesini bir üst seviyeye taşımıştır. Bu veri zenginleştirme yaklaşımları, modelin bilgi tabanını genişletirken, üretilen akıl yürütme sürecinin içsel kalitesini ve güvenilirliğini ölçme sorununu da beraberinde getirmektedir.

Bizim çalışmamız, bu kendini iyileştirme döngüsünün en kritik noktasına, yani kalite kontrol ve davranışsal etki'nin ölçümüne odaklanmaktadır.

1. Metrik Olarak Akıl Yürütme Kalitesine Odaklanma: Diğer çalışmalardan farklı olarak, biz başarının birincil ölçütü olarak, bu sentetik veri üretim hatlarının çıktılarını sadece nihai sonuca göre değil, aynı zamanda sürecin kendisinin tıbbi olarak ne kadar sağlam olduğuna göre değerlendiren yeni bir standart olan Akıl Yürütme Doğruluğu metriğini öneriyor ve uyguluyoruz. Bu metrik, "şans eseri doğru" cevapları, "faydalı ama teknik olarak yanlış" olan klinik olarak değerli çıktılardan ayırt etmemizi sağlayarak, bir modelin gerçek güvenilirliğini ölçmek için daha hassas bir yöntem sunmaktadır.

2. Davranışsal Eğitimin Etkisini Ölçme: SFT ve GRPO gibi eğitimlerin modele yeni tıbbi bilgi "ezberletmekten" ziyade, onun problem çözme "davranışını" değiştirdiğini ve bu davranışsal değişimin, modelin güvenilirliğini doğrudan artırdığını kantitatif olarak kanıtlıyoruz.

Özetle, çalışmamız mevcut literatüre, HuatuoGPT [1] ve MedReason [2] gibi veri zenginleştirme yaklaşımlarının ürettiği akıl yürütmenin kalitesini ölçmek ve doğrulamak için yeni bir metrik ve metodoloji sunarak katkıda bulunmaktadır. Bu yaklaşım, söz konusu "kendini iyileştirme döngüsü"nün sadece daha fazla veri üretmesini değil, aynı zamanda kanıtlanabilir şekilde daha güvenilir veri üretmesini sağlayacak temel bir mekanizmadır.

2. MATERYAL VE YÖNTEM

Bu çalışmada, büyük dil modellerinin (LLM) tıbbi alandaki karmaşık akıl yürütme yeteneklerinin geliştirilmesi ve değerlendirilmesi amacıyla çok aşamalı bir eğitim ve analiz metodolojisi izlenmiştir. Araştırmanın temel amacı doğrultusunda, model ölçeğinin (1B ve 4B parametre), farklı eğitim stratejilerinin (SFT ve GRPO) ve görev türlerinin (RAG ve doğrudan sorgu) model performansı üzerindeki etkisi sistematik olarak ortaya konulmuştur. Nihai hedef olarak, yalnızca doğru cevabın bulunması değil, aynı zamanda bu cevaba giden Düşünce Zinciri (Chain-of-Thought - CoT) sürecinin mantıksal ve tutarlı bir şekilde oluşturulabilme becerisinin ölçülmesi amaçlanmıştır.

2.1. Modeller ve Veri Setleri

2.2.1. Temel Modeller ve Eğitim Tekniği

Çalışmada, model ölçeğinin etkisini karşılaştırmak amacıyla Gemma 3 ailesinden 1B ve 4B parametrelili modeller temel olarak kullanılmıştır. Tüm eğitim süreçlerinde, hesaplama verimliliği sağlamak amacıyla LoRA (Low-Rank Adaptation) tekniği, Unsloth kütüphanesi aracılığıyla uygulanmıştır. Bu sayede modellerin ana ağırlıkları dondurulmuş ve yalnızca küçük adaptör katmanları eğitilmiştir.

3.2.2. Eğitim ve Değerlendirme Veri Setleri

Çalışmada, her biri farklı bir amaca hizmet eden dört temel veri seti veya bunlardan türetilen veri setleri kullanılmıştır [33-37].

Tablo 1. Çalışmada kullanılan eğitim ve değerlendirme veri setleri

Veri seti adı	Kaynak	Seçilen örnek sayısı	Kullanım amacı
SFT eğitim seti	FreedomIntelligence/ medical-o1-reasoning-SFT	~19,700	Gözetimli ince ayar (SFT)
GRPO eğitim seti	Yujivus/mmlu_health_grpo_gsm8k_format	~1,200	Pekiştirmeli öğrenme (GRPO)
Değerlendirme seti 1	PubMedQA	100	Değerlendirme (RAG)
Değerlendirme seti 2	FreedomIntelligence/ medical-o1-verifiable-problem	100	Değerlendirme (Doğrudan sorgu)

Eğitim Veri Setleri

Model eğitiminde iki aşamalı bir yaklaşım benimsenmiştir:

Gözetimli İnce Ayar (SFT) Aşaması: Modelin temel davranışsal formatını, yani bir soruya cevap vermeden önce adım adım akıl yürütme (Chain-of-Thought) sürecini sergilemesini öğretmek amacıyla, FreedomIntelligence/medical-o1-reasoning-SFT veri setindeki yaklaşık 19,700 örnek üzerinde eğitim yapılmıştır.

Pekiştirmeli Öğrenme (GRPO) Aşaması: SFT ile temel düşünme becerisini kazanan modelin, özellikle tıbbi alandaki akıl yürütme kalitesini daha da rafine etmek için pekiştirmeli öğrenme uygulanmıştır. Bu aşamada, MMLU veri setinin tıp, biyoloji ve sağlıkla ilgili alt kümelerinden özel olarak hazırladığımız ve Yujivus/mmlu_health_grpo_gsm8k_format adıyla topluluğun kullanımına sunduğumuz, yaklaşık 1,200 soruluk bir veri seti kullanılmıştır.

Değerlendirme Veri Setleri

Modelin performansını farklı senaryolarda ölçmek için iki ana değerlendirme veri seti kullanılmıştır:

RAG Görevi (PubMedQA): Modelin, kendisine sunulan bir metin (bağlam) içerisinde akıl yürüterek cevap çıkarma, yani Geri-Getirme Destekli Üretim (Retrieval-Augmented Generation - RAG) yeteneğini ölçmek amacıyla standart PubMedQA benchmark'ının test seti kullanılmıştır.

Doğrudan Sorgu Görevi (Doğrulanabilir Problemler veya FreedomIntelligence/ medical-o1-verifiable-problem): Çalışmamızın ana odağı olan, modele herhangi bir ek bağlam verilmeden, doğrudan kendi içsel bilgisi ve muhakeme yeteneğiyle karmaşık tıbbi problemleri çözme becerisini ölçmek için FreedomIntelligence/medical-o1-verifiable-problem veri seti kullanılmıştır.

Veri Kirliliğinin Önlenmesi (Data Decontamination): Modelin genelleme yeteneğini adil bir şekilde ölçmek hayati önem taşıdığından, "Doğrudan Sorgu Görevi" için kullanılan bu nihai test setinin, SFT ve GRPO

eğitim aşamalarında kullanılan verilerle hiçbir kesişime sahip olmadığı garanti edilmiştir. Test seti, medical-o1-verifiable-problem setinden, SFT setindeki tüm soruların çıkarılmasıyla oluşturulmuştur. Bu titiz ayırım, modelin öğrendiği becerilerin transfer edilebilirliğinin ölçülmesi için güvenilir bir zemin sağlamaktadır.

Doğrulanabilir Problemler Veri Setinde Kullanılan Soru Örneği: Based on the chest radiograph and abdominal CT scan of a middle-aged male complaining of nagging abdominal pain for the past 2 weeks, what is the probable diagnosis that should be considered?

Cevabı: Hydatid Cyst.

3.3. Eğitim Aşamaları ve Metodolojisi

Aşama 1 - Temel Model Performansının Belirlenmesi: Eğitimden önce, 1B ve 4B parametrelili Gemma 3 temel modelleri, her iki değerlendirme veri seti üzerinde test edilerek "sıfırıncı gün" performansları kaydedilmiştir. Bu işlemle, sonraki aşamalarda gerçekleştirilecek gelişimin ölçülmesi için kritik bir başlangıç noktası sağlanmıştır.

Aşama 2 - SFT ile Davranışsal Öğrenme: Modelin temel davranışsal formatını, yani bir soruya cevap vermeden önce adım adım akıl yürütme (Chain-of-Thought - CoT) sürecini sergilemesini öğretmek amacıyla, FreedomIntelligence/medical-o1-reasoning-SFT veri seti kullanılarak gözetimli ince ayara tabi tutulmuştur. Bu aşamanın birincil amacı modele yeni tıbbi bilgiler öğretmek değil, karmaşık problemleri yapılandırılmış bir CoT formatında sunma becerisini kazandırmaktır.

Aşama 3 - GRPO ile Akıl Yürütmenin Rafine Edilmesi: SFT ile temel düşünme becerisini kazanan modelin akıl yürütme kalitesini daha da rafine etmek için Pekiştirmeli Öğrenme (GRPO) uygulanmıştır. Bu aşamada, Yujivus/mmlu_health_grpo_gsm8k_format veri seti kullanılmıştır. Eğitim sırasında, modelin hem yapısal formatı korumasını hem de doğru cevabı vermesini teşvik eden çok bileşenli bir ödül sistemi tasarlanmıştır. Bu ödül fonksiyonu üç ana kısımdan oluşmaktadır:

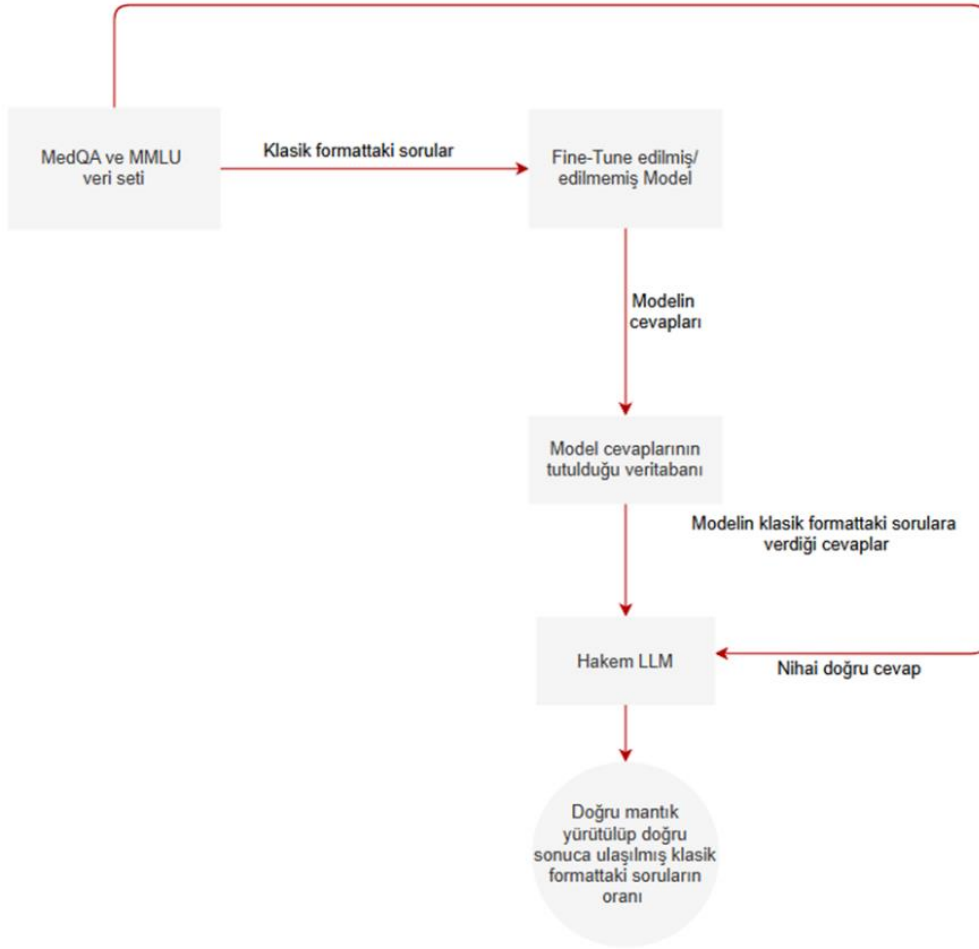
Tam Format Uyumu (match_format_exactly): Modelin çıktısının, <start_working_out>, <SOLUTION> gibi etiketleri içeren tam ve doğru yapısal formata sahip olması durumunda yüksek bir pozitif ödül (+3.0) verilir.

Yaklaşık Format Uyumu (match_format_approximately): Bu kritik etiketlerin varlığını veya yokluğunu kontrol ederek kısmi bir ödül veya ceza uygular.

Doğru ve Tekil Cevap (check_for_unique_pair_final_batched): Modelin <SOLUTION> etiketleri arasına, sorudaki çoktan seçmeli şıklardan yalnızca bir tanesini ve doğru olanı eklemesi durumunda büyük bir ödül (+5.0) verilirken, birden fazla şık sunarak belirsizlik yaratması negatif bir ödül (-2.0) ile cezalandırılır.

3.4. Değerlendirme Metodolojisi

Model performansının bütünsel bir resminin elde edilmesi amacıyla, basit doğruluk metriklerinin ötesine geçen, hakem LLM tabanlı bir değerlendirme süreci benimsenmiştir. Tüm değerlendirmelerde, endüstri standardı olarak kabul edilen GPT-4.1 modeli, "hakem" olarak kullanılmıştır. Bu sayede, özellikle açık uçlu cevapların ve karmaşık mantık zincirlerinin nüanslı bir şekilde değerlendirilmesi sağlanmıştır. Her bir model çıktısı, hakem LLM tarafından iki ana eksenle değerlendirilmiştir: Cevap Doğruluğu, model tarafından üretilen nihai sonucun, veri setindeki "gerçek değer (ground-truth)" cevap ile eşleşip eşleşmediği (eş anlamlı ya da klinik olarak çok yakın terimlere karşı tolerans gösterilmiştir) ve Mantık Yürütme Doğruluğu, cevaba ulaşırken sunulan akıl yürütme sürecinin (varsa) tıbbi olarak geçerli ve mantıksal olarak tutarlı olup olmadığının değerlendirilmesidir. Doğru cevaba bariz mantık hatalarıyla ulaşılan durumlar, "şans eseri doğru" olarak kabul edilmiştir.



Şekil 1. Sistem mimarisi

Bu mimari, model performansını hem Cevap Doğruluğu hem de Akıl Yürütme Doğruluğu eksenlerinde ölçmek için izlenen veri akışını göstermektedir. (1) FreedomIntelligence/medical-o1-verifiable-problem veri setinden alınan açık uçlu tıbbi problemler, temel (eğitilmemiş) veya ince ayar yapılmış (SFT/GRPO) Gemma modellerine girdi olarak sunulur. (2) Modelin soruya karşılık ürettiği tam cevap (akıl yürütme süreci ve nihai çözüm dahil) bir veritabanında saklanır. (3) Ardından, Hakem LLM (GPT-4.1), üç temel bilgiyi girdi olarak alır: orijinal problem, modelin ürettiği tam cevap ve veri setindeki nihai doğru cevap (ground truth). (4) Hakem LLM, bu girdileri karşılaştırarak modelin mantık zincirinin kalitesini ve nihai cevabının doğruluğunu değerlendirir. Bu süreç, çalışmamızda kullanılan Akıl Yürütme Doğruluğu metriğinin temelini oluşturur.

3. DENEYSEL BULGULAR

Tablo 2’de, Gemma 3 serisinin 1B ve 4B parametrelili temel (base) modellerinin, standart bir tıbbi soru-cevap veri seti olan PubMedQA üzerindeki başlangıç (baseline) performansı ortaya konmuştur. Bu değerlendirme, sonraki eğitim aşamalarının etkisinin ölçülmesi için bir referans noktası olarak kullanılmıştır. Sonuçlar aracılığıyla, model ölçeğinin performansa doğrudan ve pozitif bir etkisinin bulunduğu açık bir şekilde gösterilmiştir.

4B modelinin, %71 Cevap Doğruluğu ve %85 Akıl Yürütme Doğruluğu ile 1B modelinin (%54 ve %50) oldukça ilerisinde olduğu ortaya koyulmuştur. Bu durum, daha büyük modele doğal olarak daha gelişmiş bir bilgi tabanı ve muhakeme kapasitesinin kazandırıldığını teyit etmektedir. 4B modelinde gözlemlenen yüksek Akıl Yürütme Doğruluğu sayesinde, hakem LLM tarafından modelin kısa cevabında dahi örtük bir mantıksal tutarlılığın tespit edildiği veya cevap doğruluğunun daha yüksek bir güvenle değerlendirildiği ima edilmiştir.

Tablo 2. Temel Gemma modellerinin PubMedQA veri seti ile değerlendirmesi

Test edilen model	Eğitim durumu	Test veri seti (Benchmark)	Cevap doğruluğu (%)	Akil yürütme doğruluğu (%)
Gemma 3 1B	Temel (Base)	PubMedQA	54%	50%
Gemma 3 4B	Temel (Base)	PubMedQA	71%	85%

Tablo 3’de, modellerin temel halleri üzerinden FreedomIntelligence/medical-o1-reasoning-SFT veri seti ile Gözetimli İnce Ayar (Supervised Fine-Tuning - SFT) işleminden geçirildikten sonraki PubMedQA performansı analiz edilmiştir. Metodolojide açıkça belirtildiği üzere, bu aşamanın birincil amacı modele yeni tıbbi bilgi kazandırmak değil, karmaşık problemleri yapılandırılmış bir Düşünce Zinciri (CoT) formatında ifade etme davranışının modele kazandırılmasıdır. Bu bağlamda, PubMedQA üzerindeki performans değişimlerinin dikkatle yorumlanması gerekmektedir. 4B modelinde, Cevap Doğruluğu sabit tutulurken (%71), Akıl Yürütme Doğruluğu %85’ten %81’e hafifçe gerilemiştir. Bu durum, model tarafından genel bir görev formatından (doğrudan cevap) daha spesifik ve yapısal bir formata (CoT) geçişin bir yan etkisi olarak yorumlanmaktadır. Daha küçük olan 1B modelde ise hem Cevap Doğruluğunda (%54’ten %53’e) hem de Akıl Yürütme Doğruluğunda (%50’dan %38’e) daha belirgin bir düşüşün gözlemlendiği raporlanmaktadır. Bu da, daha sınırlı kapasiteye sahip 1B model tarafından, yeni bir davranışsal format öğrenilmeye çalışılırken genel muhakeme yeteneğinde daha fazla "hizalama vergisi" (alignment tax) ödendiğini düşündürmektedir.

Tablo 3. SFT ile eğitilmiş modellerin PubMedQA veri seti ile değerlendirmesi

Test edilen model	Eğitim durumu	Test veri seti (Benchmark)	Cevap doğruluğu (%)	Akil yürütme doğruluğu (%)
Gemma 3 1B	SFT (Huatu)	PubMedQA	53%	38%
Gemma 3 4B	SFT (Huatu)	PubMedQA	71%	81%

Tablo 4’te, temel modellerin doğrudan sorgu yoluyla karmaşık teşhis yeteneğini ölçen "Doğrulanabilir Problemler" veri seti üzerindeki ham yetenekleri gösterilmiştir. Bu veri setinin zorluk seviyesinin, PubMedQA’ya kıyasla çok daha yüksek olduğu belirtilmektedir. Elde edilen sonuçlarla, her iki temel modelin de bu karmaşık görevde özel bir eğitim almaksızın oldukça düşük bir başarı oranına sahip olduğu ortaya konmuştur. 4B modelinin (%29 Cevap Doğruluğu, %26 Akıl Yürütme Doğruluğu), 1B modeline (%8 Cevap Doğruluğu, %4 Akıl Yürütme Doğruluğu) kıyasla daha iyi bir performans sergilediği gösterilse de, bu skorlarla temel modellerin yapılandırılmış ve çok adımlı tıbbi akıl yürütme konusunda ciddi sınırlamalara sahip olduğu kanıtlanmıştır. Elde edilen bu bulgularla, SFT gibi hedefe yönelik eğitim stratejilerinin neden gerekli olduğu açık bir şekilde ortaya konmuştur.

Tablo 4. Temel Gemma modellerinin doğrulanabilir problemler ile değerlendirilmesi

Test edilen model	Eğitim durumu	Test veri seti (Benchmark)	Cevap doğruluğu (%)	Akil yürütme doğruluğu (%)
Gemma 3 1B	Temel (Base)	Doğrulanabilir problemler (Verifiable problems)	%8	%4
Gemma 3 4B	Temel (Base)	Doğrulanabilir problemler	29%	26%

Bu çalışmanın en dikkat çekici bulgularından biri Tablo 5’te sunulmuş ve SFT eğitiminin "Doğrulanabilir Problemler" görevindeki etkisi detaylı şekilde incelenmiştir. Özellikle Gemma 3 4B modelinde gözlemlenen değişimle, SFT aşamasının başarısı gözler önüne serilmiştir: Modelin Cevap Doğruluğu %29’dan %27’ye marjinal bir düşüş göstermiş; buna karşın, Akıl Yürütme Doğruluğu %26’dan %29’a yükselmiştir.

Bu sonuçla birlikte, SFT’nin amacına ulaştığı güçlü bir şekilde ortaya konmuştur. Eğitim süreci boyunca, modele doğrudan doğru cevabın ezberletilmesinden ziyade, teşhise giden yolda mantıksal olarak daha tutarlı ve tıbben geçerli bir Düşünce Zinciri (CoT) oluşturma becerisinin kazandırıldığı görülmektedir. Nihai teşhis formüle edilirken küçük hatalar yapılmış olsa bile, sorunun analiz edilmesi ve mantıksal adımların sunulması konularında modelin daha yetkin hale geldiği anlaşılmaktadır.

Bu gelişim, modelin pratik tıbbi uygulamalardaki denetlenebilirliği ve güvenilirliği açısından, salt cevap doğruluğundan çok daha değerli bir kazanım olarak değerlendirilmektedir. Sonuç olarak, SFT ile modele hedeflenen davranışsal yeteneğin başarıyla kazandırıldığı Tablo 5'te net bir şekilde gösterilmiştir.

Tablo 5. SFT ile eğitilmiş Gemma modellerinin doğrulanabilir problemler ile değerlendirilmesi

Test edilen model	Eğitim durumu	Test veri seti (Benchmark)	Cevap doğruluğu (%)	Akıl yürütme doğruluğu (%)
Gemma 3 1B	SFT (Huatuo)	Doğrulanabilir problemler	%8	%4
Gemma 3 4B	SFT (Huatuo)	Doğrulanabilir problemler	27%	29%

GRPO ile Akıl Yürütmenin Rafine Edilmesi ve Metriklerin Sınırları

SFT aşamasının sonuçlarıyla, 1B modelinin bu karmaşık davranışsal eğitimi özümsemeye kapasite sınırlamaları sergilediği ortaya konmuştur. Özellikle PubMedQA'daki akıl yürütme performansında gözlemlenen belirgin düşüş ile "Doğrulanabilir Problemler" veri setinde herhangi bir ilerlemenin kaydedilememesi sonucunda, bu modelin daha ileri eğitim adımları için uygun bir aday olmadığı belirlenmiştir. Bu nedenle, çalışmanın bir sonraki aşaması olan GRPO ile pekiştirmeli öğrenmenin yalnızca daha yüksek kapasiteli ve SFT'ye daha olumlu yanıt verdiği gözlemlenen 4B model üzerinde uygulanmasına karar verilmiştir.

GRPO eğitiminin ardından elde edilen sonuçlarla, çalışmanın en ilgi çekici bulguları sunulmuştur: Modelin Cevap Doğruluğu %27'den %26'ya düşürülmüş; buna karşın, Akıl Yürütme Doğruluğu %29'dan %31'e yükseltilmiştir. Bu durumla birlikte, SFT ve GRPO eğitimlerinin birleşik etkisiyle modelin muhakeme kalitesinin nihai sonuçtan daha fazla önceliklendirildiği görülmektedir.

Paradoksal olarak değerlendirilebilecek bu eğilim, değerlendirme metodolojisinin doğası ve tıbbi teşhis sürecinin karmaşıklığıyla açıklanmaktadır. Klinik pratikte, sunulan bulgularla uyumlu birden fazla geçerli tanının (diferansiyel tanı) bulunabildiği; ancak veri setinde yer alan "ground truth"un bu olasılıklardan yalnızca birini temsil ettiği belirtilmektedir. Modelin akıl yürütme yeteneği geliştirildikçe, veri setindeki tekil doğru cevaba kilitlenmek yerine, klinik olarak geçerli fakat farklı bir sonuca ulaştıran bir mantık zinciri üretebildiği gözlemlenmektedir.

Hakem LLM tarafından bu mantık zinciri tutarlı ve doğru olarak değerlendirildiğinde, Akıl Yürütme puanı artırılmakta; ancak nihai cevabın "ground truth" ile eşleşmemesi durumunda, Cevap Doğruluğu puanı düşürülmektedir. Bu doğrultuda, GRPO'nun modeli daha sofistike ve tıbben nüanslı bir akıl yürütmeye teşvik ettiği; fakat bu gelişimin, yalnızca tek bir doğru cevaba dayalı metriklerle tam olarak yakalanamadığı sonucuna varılmaktadır.

4. TARTIŞMA VE SONUÇ

Bu çalışmada, farklı ölçeklerdeki (1B ve 4B) Gemma 3 dil modellerinin tıbbi alandaki akıl yürütme yetenekleri, çeşitli eğitim aşamaları (temel, SFT, GRPO) ve görevler (RAG ve doğrudan sorgu) üzerinden kapsamlı bir şekilde analiz edilmiştir. Elde edilen bulgularla, model ölçeğinin, özellikle bağlam sunulduğunda (PubMedQA RAG görevi), akıl yürütme performansı üzerinde kritik bir etkiye sahip olduğu net bir şekilde ortaya konmuştur. 4B temel model tarafından, %85 Mantık Yürütme Doğruluğu elde edilerek bu alanda doğal bir yetkinliğe sahip olunduğu gösterilmiş; buna karşılık, 1B modelin yetersiz kaldığı ve SFT süreciyle bu durumun iyileştirilemediği gözlemlenmiştir.

Çalışmanın asıl odak noktasını oluşturan, bağlam sunulmayan ve açık uçlu cevap gerektiren "Doğrulanabilir Tıbbi Problemler" test setinde ise daha farklı ve katmanlı bir sonuca ulaşılmıştır. 4B temel modelle %26 gibi düşük bir Mantık Yürütme Doğruluğu elde edilerek modelin ham tıbbi bilgi hatırlama becerisi açısından zayıf olduğu teyit edilmiştir. Huatuo verisiyle gerçekleştirilen SFT eğitimi sonrasında, modelin Cevap Doğruluğu (%27) önemli ölçüde değişmemiş; ancak Mantık Yürütme Doğruluğu %29'a yükseltilmiştir. Bu sonuçla, SFT süreciyle modele yeni tıbbi gerçeklerin öğretilmesinden çok, düşüncelerini adım adım yapılandırma (Düşünce Zinciri – CoT) davranışının başarıyla kazandırıldığı gösterilmiştir.

Son aşama olan MMLU medikal verisiyle gerçekleştirilen GRPO eğitimiyle ise bu eğilim daha da ileri taşınmıştır. Cevap Doğruluğu %26'da sabit kalırken, Mantık Yürütme Doğruluğu %31'e çıkarılarak tüm aşamalar arasında en yüksek seviyeye ulaşılmıştır.

Bu sonuçlarla, çalışmanın temel hipotezinin doğrulandığı görülmektedir: Açık uçlu tıbbi problemlerde, tek bir "doğru cevap" etiketine sıkı sıkıya bağlı kalınmasının, modelin gerçek yeteneğini ölçmek açısından yetersiz kalabileceği ortaya konmuştur. Birden fazla klinik olarak geçerli cevap bulunabileceğinden, asıl başarı metriğinin, modelin mantıksal olarak tutarlı ve tıbben sağlam bir akıl yürütme süreci sergileyip sergileyemediği olması gerektiği vurgulanmıştır. GRPO modeliyle ulaşılan %31'lik Mantık Yürütme Doğruluğu ile, modelin cevapları ezberlemek yerine nasıl düşüneceğini öğrendiği açıkça ortaya konmuştur. Sonuç olarak, bu çalışmayla az parametrelili büyük dil modellerinin tıbbi alanda uzmanlaştırılmasında, davranışsal ve yapısal akıl yürütme becerilerinin kazandırılmasının önemi vurgulanmıştır.

Ek 1: GRPO ile Eğitilmiş Modelin Akıl Yürütme Örneği

Aşağıdaki örnek, SFT eğitimi üzerine GRPO ile eğitilmiş Gemma 4B modelinin, doğrudan sorgu görevindeki performansı ve Hakem LLM (GPT-4.1) tarafından yapılan değerlendirmenin bir kesitini sunmaktadır. Bu örnek, çalışmanın temel hipotezi olan, modelin mantık yürütme kalitesinin geliştirilmesinin nihai cevap doğruluğundan daha değerli bir kazanım olduğunu açıkça göstermektedir.

Soru: An 88-year-old woman with osteoarthritis is experiencing mild epigastric discomfort and has vomited material resembling coffee grounds multiple times. Considering her use of naproxen, what is the most likely cause of her gastrointestinal blood loss?

Modelin Yanıtı:

model <start_working_out>Alright, let's think this through. We've got an 88-year-old woman who's dealing with some pretty common issues like osteoarthritis. She's also been having this mild epigastric discomfort and has been vomiting stuff that looks like coffee grounds. That's a pretty classic sign of something going on with her stomach lining. Now, she's taking naproxen. Naproxen is a nonsteroidal anti-inflammatory drug, or NSAID for short. These are known to be pretty tough on the stomach lining. They can cause a lot of damage because they inhibit the production of prostaglandins, which are important for protecting the stomach. So, with her taking naproxen, it's not surprising that she's experiencing this kind of gastrointestinal bleeding. The coffee ground-like vomit is a big clue here. It suggests that there's blood coming from somewhere in the upper part of her digestive tract. Given her age and the fact that she's on naproxen, the most likely culprit for her bleeding is probably a peptic ulcer. These ulcers can form in the stomach or the duodenum, and they can bleed. The symptoms she's showing, like the epigastric discomfort and the coffee ground vomit, fit well with this diagnosis. It's pretty common for NSAIDs to cause these symptoms, especially in older adults. So, putting it all together, it seems pretty clear that the naproxen is likely causing her peptic ulcer, which is leading to the bleeding. That's why she's experiencing these symptoms. Yeah, that makes sense. The connection between her medication, her symptoms, and the likely cause of her bleeding is pretty solid.<end_working_out> <SOLUTION>The most likely cause of the gastrointestinal blood loss in this 88-year-old woman is a peptic ulcer. Naproxen, a nonsteroidal anti-inflammatory drug (NSAID), is known to cause damage to the stomach lining by inhibiting the production of prostaglandins, which are protective for the stomach. This can lead to the formation of peptic ulcers, particularly in older adults. The symptoms of epigastric discomfort and coffee ground-like vomit are classic indicators of upper gastrointestinal bleeding, often associated with peptic ulcers. Therefore, the use of naproxen is likely contributing to her peptic ulcer, resulting in the observed gastrointestinal bleeding.</SOLUTION>

Hakem LLM Değerlendirmesi:

Doğruluk Puanı (correctness): false

Akıl Yürütme Doğruluğu (reasoning_correctness): true

Gerekçe (justification): The model's final answer is 'peptic ulcer,' which is a broader term that includes both gastric and duodenal ulcers. The ground truth specifies 'gastric ulcer,' which is more precise and most likely given the NSAID use and the patient's symptoms. Therefore, the final answer is not a synonymous match and is too general. However, the reasoning is medically sound: the model correctly links NSAID use

to prostaglandin inhibition, mucosal injury, and ulcer formation, and accurately interprets the symptoms as upper GI bleeding. The reasoning is logical and without significant errors.

Yorum: Bu örnekte görüldüğü gibi, modelin nihai cevabı (peptic ulcer), veri setindeki kesin doğru cevap (gastric ulcer) ile tam olarak eşleşmediği için **Cevap Doğruluğu** puanı düşürülmüştür. Ancak, modelin yürüttüğü akıl yürütme süreci, hastanın semptomlarını (kahve telvesi benzeri kusma) ve ilaç kullanımını (naproxen, bir NSAID) doğru bir şekilde ilişkilendirmiş ve bu durumun peptik ülserlere yol açabileceği yönünde tıbben geçerli bir mantık zinciri oluşturmuştur. Hakem LLM bu akıl yürütmeyi mantıksal ve hatasız bulduğu için **Akıll Yürütme Doğruluğu** puanı yükselmiştir. Bu durum, GRPO eğitiminin modeli, sadece ezberlenmiş bir cevaba değil, klinik olarak sağlam bir mantık sürecine yönlendirdiğini göstermektedir. Bu örnek, Cevap Doğruluğu ve Mantık Yürütme Doğruluğu metrikleri arasındaki farkı ve çalışmanın ana bulgularının pratik önemini vurgulamaktadır.

Bu durum, kullanılan eğitim ve test verilerinin format farklılıklarıyla da açıklanmaktadır. SFT eğitiminde kullanılan veriler genellikle açık uçlu ve yapılandırılmış bir Düşünce Zinciri (CoT) formatına odaklanırken, GRPO eğitimi, MMLU veri setinin çoktan seçmeli yapısı üzerinde gerçekleştirilmiştir. MMLU sorularında model, doğru cevaba ulaşmak için birden fazla aday arasından en uygun olanı seçmeye teşvik edilmektedir. Bu nedenle, GRPO ile pekiştirilen modelin, nihai açık uçlu teşhis sorularında, doğru cevabı doğrudan vermek yerine, çoktan seçmeli sorular için yeterli olabilecek üst başlıkları veya genel terimleri (örneğin, "gastric ulcer" yerine "peptic ulcer") öne çıkarmaya meyletmesi beklenen bir sonuçtur. Model, çoktan seçmeli formatta öğrendiği "doğru cevabı seçme" becerisini, açık uçlu formatta "genel olarak doğru ve güvenli bir terim kullanma" şeklinde uyguluyor olabilir. Bu durum, modelin mantıksal muhakeme yeteneğinin geliştiğini, ancak farklı bir görev formatına (çoktan seçmeli) adapte olmasının, açık uçlu sorular için daha geniş bir dil üretimine neden olduğunu göstermektedir. Bu, modelin "ezberlemek yerine düşünme" hipotezini desteklerken, aynı zamanda eğitim verisi formatının model davranışını nasıl şekillendirdiğini de ortaya koymaktadır.

Akıll Yürütme Doğruluğunun Klinik Fayda İçin Bir Gösterge Olarak Metodolojik Doğrulanması

Akıll Yürütme Doğruluğu metriğimizin pratik sonuçlarını titizlikle doğrulamak ve geleneksel doğruluk skorlarının ötesine geçmek amacıyla, değerlendirme sonuçlarımızın en paradoksal alt kümeleri üzerinde detaylı bir niteliksel analiz gerçekleştirdik. Analizde iki kritik gruba odaklandık: modelin akıl yürütmesi doğru olmasına rağmen nihai cevabının yanlış olduğu "faydalı ama teknik olarak yanlış" vakalar ve akıl yürütmesi kusurlu olmasına rağmen cevabının doğru olduğu "şans eseri doğru" vakalar. Bu kritik görevin gerektirdiği incelikli değerlendirmeyi en yüksek standartta yapabilmek için, mevcut en gelişmiş dil modellerinden biri olan Gemini 2.5 Pro, uzman bir klinik değerlendirici olarak kullanılmıştır.

Her bir vaka, pratik faydanın üç temel yönünü ölçen katı bir değerlendirme çerçevesine göre analiz edilmiştir: **modelin hekimin karar verme sürecini doğru bir eylem planına (doğru tanısal adım, tedavi veya uzmanlık alanı sevk) yönlendirme başarısı**, hasta için kısa vadeli potansiyel risk ve çıktının genel klinik değeri. Analiz, çarpıcı ve kesin bir sonuç ortaya koymuştur: Akıll yürütmesi sağlam olan vakaların %100'ünde, modelin çıktısının, nihai cevabın ground truth etiketiyle anlamsal eşleşmesine bakılmaksızın, **hekim için güvenilir bir karar desteği sağladığı ve onu doğru sonraki adımlara yönlendirmede etkili olduğu** tespit edilmiştir.

Bu doğrudan korelasyon, bir modelin hekime sunduğu pratik faydayı ölçen ve "**Hekim Karar Desteği Doğruluğu**" olarak adlandırdığımız yeni ve daha anlamlı bir metrik tanımlamamıza olanak sağlamıştır. Bu yeni metriğin, Akıll Yürütme Doğruluğu ile birebir örtüşmesi, modelin gerçek dünyadaki pratik faydasının en net göstergesinin, mantık zincirinin kalitesi olduğunu kanıtlamaktadır. Bu metrikler arasındaki ilişki, aşağıdaki tabloda özetlenmiştir.

Tablo 6. Farklı başarı metriklerinin base model ve GRPO ile eğitilmiş model üzerindeki karşılaştırması

Değerlendirme metriği	Temel (base) model (%)	GRPO modeli (%)
Cevap doğruluğu	29	26
Akıll yürütme doğruluğu	26	31
Hekim karar desteği doğruluğu	26	31

Tablo 6'nın ortaya koyduğu en önemli bulgu, modellerin güvenilirlik profilindeki dönüşümdür. Temel modelin sonuçları incelendiğinde, akıl yürütmesi kusurlu olan "şans eseri doğru" vakaların sayısının (4 vaka), akıl yürütmesi sağlam olan "faydalı ama yanlış" vakaların sayısından (1 vaka) dört kat daha fazla olduğu görülmektedir. Bu durum, eğitilmemiş modelin, güvenilir ve faydalı bir çıktı üretmekten ziyade, hatalı bir mantıkla tesadüfen doğru sonuca ulaşmaya daha yatkın olduğunu ve klinik kullanım için güvenilirmez olduğunu göstermektedir.

GRPO ile gerçekleştirdiğimiz eğitim süreci ise bu oranı tamamen tersine çevirmiştir. Eğitilmiş modelimizde, "faydalı ama yanlış" vakaların sayısı (7 vaka), artık "şans eseri doğru" vakaların sayısını (2 vaka) 3.5 kat aşmaktadır. Bu dönüşüm, eğitim sürecimizin modelin sadece performansını artırmakla kalmayıp, aynı zamanda onu tutarlı bir şekilde sağlam ve güvenilir muhakeme yapmaya yönlendirdiğinin en net kanıtıdır. Bu nedenle, 'Akıl Yürütme Doğruluğu'nun, bir modelin gerçek klinik faydasını ve güvenliğini ölçmede vazgeçilmez bir standart olduğunu öne sürmekteyiz.

Analizimizde ortaya koyduğumuz 'şans eseri doğru' vakalar, güvenilir tıbbi yapay zeka geliştirilmesinin önündeki en temel zorluklardan birini gözler önüne sermektedir: akıl yürütmesi kusurlu bir teşhis, tesadüfen doğru sonuca ulaşsa bile, sistemin geneli hakkında bir güvenilmezlik sinyali verir. Bu bulgu, özellikle hastayı teşhis edip otonom olarak doğru kliniğe yönlendirmesi hedeflenen gelecekteki 'Agentic' sistemler için hayati bir önem taşımaktadır. Böyle bir sistemin güvenliği, yönlendirme eyleminin dayandığı teşhisin kalitesine mutlak bir şekilde bağlıdır ve hatalı bir mantık zinciri, hastanın yanlış bir uzmanlık alanına yönlendirilmesi gibi kritik hatalara yol açabilir.

Çalışmamızın en temel çıktısı, bir yapay zeka sisteminin güvenle doğru eylemde bulunabilmesi için, önce kanıtlanabilir şekilde doğru düşünmesi gerektiği ilkesidir. Bu ilke, güvenli tıbbi yapay zekanın geleceği için bir yol haritası sunmakta ve bu alanda geliştirilecek otonom sistemler için temel bir yapı taşı oluşturmaktadır.

5. TEŞEKKÜR

Bu çalışma, Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK) tarafından TEYDEB 1505 – Üniversite-Sanayi İşbirliği Destekleme Programı kapsamında desteklenen 5240094 numaralı proje kapsamında gerçekleştirilmiştir. Sağladığı katkılar nedeniyle TÜBİTAK'a teşekkür ederiz. Bu yayında yer alan görüş, öneri ve sonuçlar yalnızca yazarların sorumluluğundadır ve TÜBİTAK'ın görüşlerini yansıtmamaktadır.

6. KAYNAKLAR

1. Chen, J., Cai, Z., Ji, K., Wang, X., Liu, W., Wang, R., Hou, J. & Wang, B. (2024). *FreedomIntelligence/HuatuoGPT-o1-8B. Hugging Face*, <https://huggingface.co/FreedomIntelligence/HuatuoGPT-o1-8B>, Erişim tarihi: 14.07.2025.
2. Wu, J., Deng, W., Li, X., Liu, S., Mi, T., Peng, Y., Xu, Z., Liu, Y., Cho, H., Choi, C.I. & Cao, Y. (2025). Medreason: Eliciting factual medical reasoning steps in LLMs via knowledge graphs. *arXiv preprint*, arXiv:2504.00993, Erişim tarihi: 25.08.2025.
3. Lai, Y., Zhong, J., Li, M., Zhao, S. & Yang, X., (2025). Med-r1: Reinforcement learning for generalizable medical reasoning in vision-language models. *arXiv preprint*, arXiv:2503.13939, Erişim tarihi: 23.07.2025.
4. DeepSeek, (2024). *GitHub - deepseek-ai/DeepSeek-R1*. GitHub, <https://github.com/deepseek-ai/DeepSeek-R1>, Erişim tarihi: 14.07.2025.
5. Ding, F., Wang, B., Zeng, Z. & Wang, Y. (2025). Multi-layer GRPO: Enhancing reasoning and self-correction in large language models. *arXiv preprint*, arXiv:2506.04746, Erişim tarihi: 23.07.2025.
6. Tinn, R., Cheng, H., Gu, Y., Usuyama, N., Liu, X., Naumann, T., Gao, J. & Poon, H. (2023). Fine-tuning large neural language models for biomedical natural language processing. *Patterns*, 4(4), 1-11.
7. Dai, W., Chen, P., Ekbote, C. & Liang, P.P. (2025). QoQ-Med: Building multimodal clinical foundation models with domain-aware GRPO training. *arXiv preprint*, arXiv: 2506.00711, Erişim tarihi: 23.07.2025.
8. Liu, M., Hu, W., Ding, J., Xu, J., Li, X., Zhu, L., Bai, Z., Shi, X., Wang, B., Song, H. & Liu, P. (2024). Medbench: A comprehensive, standardized, and reliable benchmarking system for evaluating Chinese medical large language models. *Big Data Mining and Analytics*, 7(4), 1116-1128.

9. Ogdu, C.U., Gurbuz, S., Karakose, M. & Hanoglu, E. (2025). Medical implications of LLM based clinical decision support systems in healthcare. *29th International Conference on Information Technology (IT 2025)*, Zabljak, Montenegro, 1-4.
10. Ali, H., Qadir, J., Alam, T., Househ, M. & Shah, Z. (2023). ChatGPT and large language models in healthcare: Opportunities and risks. *IEEE International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings 2023)*, Mount Pleasant, MI, USA, 1-4.
11. Li, J., Deng, Y., Sun, Q., Zhu, J., Tian, Y., Li, J. & Zhu, T. (2024). Benchmarking large language models in evidence-based medicine. *IEEE Journal of Biomedical and Health Informatics*, 1-14.
12. Yang, X., Chen, A., PourNejatian, N., Shin, H.C., Smith, K.E., Parisien, C., Compas, C., Martin, C., Costa, A.B., Flores, M.G. & Zhang, Y. (2022). A large language model for electronic health records. *NPJ Digital Medicine*, 5(1), 194.
13. Kumar, R., Lomchavakarn, P., Angasinha, C., Intaratat, K., Boonsawad, P. & Sridee, S. (2025). Large language model based system for clinical decision support. *International Conference on Cognitive Computing in Engineering, Communications, Sciences and Biomedical Health Informatics (IC3ECSBHI 2025)*, Greater Noida, India, 389-394.
14. Ding, S., Ye, J., Hu, X. & Zou, N. (2024). Distilling the knowledge from a large-language model for health event prediction. *Scientific Reports*, 14(1), 30675.
15. Wang, C., Chen, Q., Shao, W. & He, X. (2024). KEMedGPT: Intelligent medical pre consultation with knowledge enhanced large language model. *Proceedings – 2024 IEEE International Conference on Medical Artificial Intelligence (MedAI 2024)*, Chongqing, China, 386-391.
16. Peng, C., Yang, X., Chen, A., Smith, K.E., PourNejatian, N., Costa, A.B., Martin, C., Flores, M.G., Zhang, Y., Magoc, T. & Lipori, G. (2023). A study of generative large language model for medical research and healthcare. *NPJ Digital Medicine*, 6(1), 210.
17. Rasmy, L., Xiang, Y., Xie, Z., Tao, C. & Zhi, D. (2021). Med-BERT: Pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction. *NPJ Digital Medicine*, 4(1), 86.
18. Rahman, M.A., Preum, S.M., Williams, R.D., Alemzadeh, H. & Stankovic, J. (2023). EMS BERT: A pretrained language representation model for the emergency medical services (EMS) domain. *Proceedings – 2023 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE 2023)*, Orlando, FL, USA, 34-43.
19. Li, W., Yu, L., Wu, M., Liu, J., Hao, M. & Li, Y. (2023). DoctorGPT: A large language model with Chinese medical question answering capabilities. *2023 International Conference on High Performance Big Data and Intelligent Systems (HDIS 2023)*, Macau, China, 186-193.
20. Akilesh, S., Sheik Abdullah, A., Abinaya, R., Dhanushkodi, S. & Sekar, R. (2023). A novel AI based chatbot application for personalized medical diagnosis and review using large language models. *2023 IEEE International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE 2023)*, R.M.K. Engineering College, Chennai, India.
21. Peikos, G., Kasela, P. & Pasi, G. (2024). Leveraging large language models for medical information extraction and query generation. *2024 IEEE/WIC International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, Bangkok, Tayland, 367-372.
22. Kumar, P., Shreenidhi, G.L. & Rakesh Kumar, M. (2025). Medify-AI based LLM based healthcare system. *International Conference on Frontier Technologies and Solutions (ICFTS 2025)*, Chennai, India, 1-9.
23. Kumar, R., Shreenidhi, G.L. & Sowmiya, S. (2025). MedHub – LLM-based healthcare system. *3rd International Conference on Augmented Intelligence and Sustainable Systems (ICAISS 2025)*, Trichy, India, 1650-1657.
24. Aizu'bi, S., Kanan, T. & Almiani, M. (2024). Large language models for knowledge discovery in healthcare. *International Conference on Intelligent Computing, Communication, Networking and Services (ICCNS 2024)*, Dubrovnik, Croatia, 183-190.
25. Fernando, S.P.A.A. & Wickramaarachchi, D. (2025). Large language model (LLM) support for preliminary consultation in healthcare. *5th International Conference on Advanced Research in Computing (ICARC 2025)*, Belihuloya, Sri Lanka, 1-6.
26. Nazi, Z.A. & Peng, W. (2024). Large language models in healthcare and medical domain: A review. *Informatics*, 11(3), 57.
27. Tian, S., Jin, Q., Yeganova, L., Lai, P.T., Zhu, Q., Chen, X., Yang, Y., Chen, Q., Kim, W., Comeau, D.C. & Islamaj, R. (2023). Opportunities and challenges for ChatGPT and large language models in biomedicine and health. *Briefings in Bioinformatics*, 25(1), bbad493.

28. Hu, E.J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L. & Chen, W. (2022). LoRA: Low-rank adaptation of large language models. *ICLR*, 1(2), 3.
29. Rao, A.K.G., Jaggi, A. & Naidu, S. (2025). MEDFIT LLM: Medical enhancements through domain focused fine tuning of small language models. *2nd International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE 2025)*, Chennai, India, 1-5.
30. Kim, H., Hwang, H., Lee, J., Park, S., Kim, D., Lee, T., Yoon, C., Sohn, J., Park, J., Reykhart, O. & Fetherston, T. (2025). Small language models learn enhanced reasoning skills from medical textbooks. *NPJ Digital Medicine*, 8(1), 240.
31. Magnini, M., Aguzzi, G. & Montagna, S. (2025). Open-source small language models for personal medical assistant chatbots. *Intelligence-Based Medicine*, 11, 100197.
32. Wang, X., Dang, T., Kostakos, V. & Jia, H. (2024). Efficient and personalized mobile health event prediction via small language models. *ACM MobiCom 2024 – 30th International Conference on Mobile Computing and Networking*, Washington, D.C., USA, 2353-2358.
33. Hendrycks, D., Burns, C., Basart, S., Zou, A., Mazeika, M., Song, D. & Steinhardt, J. (2020). Measuring massive multitask language understanding. *arXiv preprint*, arXiv: 2009.03300, Erişim tarihi: 23.07.2025.
34. Hendrycks, D., Burns, C., Basart, S., Critch, A., Li, J., Song, D. & Steinhardt, J. (2020). Aligning AI with shared human values. *arXiv preprint*, arXiv: 2008.02275, Erişim tarihi: 23.07.2025.
35. Jin, D., Pan, E., Oufattole, N., Weng, W.H., Fang, H. & Szolovits, P. (2021). What disease does this patient have? A large-scale open domain question answering dataset from medical exams. *Applied Sciences*, 11(14), 6421.
36. Pal, A., Umapathi, L.K. & Sankarasubbu, M. (2022). MedMCQA: A large-scale multi-subject multi-choice dataset for medical domain question answering. *Conference on Health, Inference, and Learning*, ABD.
37. Hugging Face, (2025). *cais/mmlu*. *Hugging face datasets*, <https://huggingface.co/datasets/cais/mmlu>, Erişim tarihi: 25.07.2025.
38. Fırat, H. ve Üzen, H. (2024). MR görüntülerinden alzheimer hastalığının sınıflandırılması için inception ve sıkma-uyarma ağı tabanlı derin öğrenme modeli. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 39(2), 555-567.
39. Abut, S. (2024). AI-based model design for prediction of COPD grade from chest X-ray images: Bir model önerisi (COPD-GradeNet). *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 39(2), 325-338.



Development of Clarke and Park Transforms Visualization Software Using Python

Oğuzhan TİMUR^{1,a}, Kadir Sami POYRAZ^{1,b}

¹Cukurova University, Department of Electrical & Electronics Engineering, Adana, Türkiye

^aORCID: 0000-0002-6537-7840; ^bORCID: 0009-0009-2422-1303

Article Info

Received : 02.08.2025

Accepted : 26.08.2025

DOI: 10.21605/cukurovaumfd.1757054

Corresponding Author

Oğuzhan TİMUR

otimur@cu.edu.tr

Keywords

3-Phase systems

Motor control

Clarke-park transforms

Software

How to cite: TİMUR, O., POYRAZ, K.S., (2025). Development of Clarke and Park Transforms Visualization Software Using Python. Cukurova University, Journal of the Faculty of Engineering, 40(3), 607-616.

ABSTRACT

Three-phase electrical systems are commonly utilized in many industrial applications. In particular, they are used in transmission and distribution lines, as well as in the operation of asynchronous motors. Despite their widespread use, three-phase electrical systems are challenging to model and analyze because of the time-depending parameters. In order to overcome these challenges, some mathematical methods such as Clarke and Park transforms are performed. In this study, to ease the understanding of Clarke and Park transforms, a software application is developed. The application demonstrates the effects of changing the voltage, frequency, and phase parameters on the Clarke and Park transforms. In the proposed application, users can convert 3-phase (ABC) axes to 2-phase ($\alpha\beta$) axes, 2-phase ($\alpha\beta$) axes to reference (dq) axes. Application can also convert 2-phase axes to 3-phase axes with the inverse Clarke and Park transforms. During the transformation, it visualizes three separate sequences with a single parameter input, eliminating the need for repeated parameter input for each axis. Mathematical modelling and user interface development of the application are performed in Python programming language.

Python Kullanılarak Clarke ve Park Dönüşümlerini Görselleştirme Yazılımının Geliştirilmesi

Makale Bilgileri

Geliş : 02.08.2025

Kabul : 26.08.2025

DOI: 10.21605/cukurovaumfd.1757054

Sorumlu Yazar

Oğuzhan TİMUR

otimur@cu.edu.tr

Anahtar Kelimeler

3-faz sistemler

Motor kontrolü

Clarke-park dönüşümleri

Yazılım

Atf şekli: TİMUR, O., POYRAZ, K.S., (2025). Python Kullanılarak Clarke ve Park Dönüşümlerini Görselleştirme Yazılımının Geliştirilmesi. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 607-616.

ÖZ

Üç fazlı elektrik sistemleri birçok endüstriyel uygulamada yaygın olarak kullanılmaktadır. Özellikle iletim ve dağıtım hatlarında ve asenkron motorların çalışmasında kullanılırlar. Yaygın kullanımlarına rağmen üç fazlı elektrik sistemlerinin zamana bağlı parametrelerden dolayı modellenmesi ve analizi zordur. Bu zorlukların üstesinden gelmek için Clarke ve Park dönüşümleri gibi bazı matematiksel yöntemler kullanılır. Bu çalışmada, Clarke ve Park dönüşümlerinin anlaşılmasını kolaylaştırmak için yazılım programı geliştirilmiştir. Uygulama, voltaj, frekans ve faz parametrelerinin değiştirilmesinin Clarke ve Park dönüşümleri üzerindeki etkilerini göstermektedir. Geliştirilen programda, kullanıcılar 3 fazlı (ABC) eksenden 2 fazlı ($\alpha\beta$) eksene, 2 fazlı ($\alpha\beta$) eksenden referans (dq) eksenine dönüşüm yapabilmektedirler. Dönüşüm sırasında, tek bir parametre girişi ile, üç ayrı dizi görselleştirilerek, her eksen için tekrarlanan parametre girişi ihtiyacı ortadan kaldırılmıştır. Uygulamanın matematiksel modellenmesi ve kullanıcı arayüzü Python programlama dili kullanılarak geliştirilmiştir.

1. INTRODUCTION

Electrical systems have been used in all areas, and they need to be controlled. There are many control systems and different algorithms. Every one of them is invented for different need of use. Some electrical systems need precise control, some of them need fast response, some of them need basic systems or algorithms [1,2]. Every need caused the invention of different control systems. Clarke and Park transforms are invented for modelling of electrical systems and used in these control systems. They were invented to ease the control of three phase systems. They are used in the control of both power grids and three phase motors [3-5].

Clarke and Park transforms used in different areas. When extracting phase information from a three-phase disturbed signal, a phase detector which is based on Clarke transform and arctangent function is widely used [6]. When detecting and classifying the transmission line faults, a fault identification matrix (FIM) based on the change in the resultant current samples of α - β components is used [7]. For time-domain modeling framework for three-phase unbalanced AC power systems, g-dq0 transform is derived as a generalized variant of dq0 transformation [8]. As g-dq0 transform, Reduced Reference Frame (RRF) which improves the performance of the classical Fortescue, Clarke and Park transformations is proposed for unbalanced three phase four wire systems. The RRF transformation represents any unbalanced three phase sinusoidal magnitude with just two components even if the zero-sequence component is present [5]. For detecting the fault current from the switching currents, and other transient and loading conditions to improve the longitudinal differential relay performance, Clarke transform is used [9].

Clarke and Park transforms simplify the analysis and control of the 3-phase systems by converting them into 2-phase reference frames. This conversion facilitates the design of control algorithms such as field-oriented control (FOC), direct torque control (DTC) and field weakening control (FWC), enhances the understanding of system behavior under different operating conditions, and makes the modelling and analysis easier [10,11].

Field oriented control is one of the most used motor control algorithms. In FOC not only magnitude and frequency are adjusted, but also the instantaneous values of voltage, current, and flux space vectors, which allows transient operation. FOC uses Clark-Park transforms to control motor. For example, 3-phase amplitudes and phase values can be controlled via d and q sequences, so 4 parameters can be controlled by 2 parameters [12,13]. This conversion proceeds like this: When V_d is 100V and V_q is 0V, 3-phase amplitudes become 100V and phase becomes 0° , when V_d is -100V and V_q is 0V, 3-phase amplitudes become 100V and phase becomes 180° , when V_d is 0V and V_q is 100V, 3-phase amplitudes become 100V and phase becomes 270° , when V_d is 0V and V_q is -100V, 3-phase amplitudes become 100V and phase becomes 90° . Phase angle not only becomes the multiples of 90. It can be modified by changing the V_d and V_q parameters as the users wish. For example, if V_d is 90V and V_q is 45V, 3-phase amplitudes become approximately 100V and phase angle becomes approximately 26.6° [4,14].

Although practical applications and visualizations of Clarke and Park transforms are important, it is also hard to understand for the students and practitioners who are new to the field. Conventional teaching methods generally use static diagrams and equations, that maybe not fully convey the dynamic relationship between reference frames to students. Therefore, there is a need for an application to fill the gap between theoretical concepts and real-world implementations.

In return to this need, this article provides the development of user-friendly software application to calculate and visualize the Clarke and Park transforms. The application comes with a graphical user interface (GUI) that allows users to enter the system parameters, perform the transformations, and observe the resulting waveforms and vector diagrams in real time. The application does all the calculations that are needed to visualize waveforms of all three sequences, so users can see waveforms and reference frame vectors of all three sequences by only entering one of the sequence values. With the real time process, users can see how the difference in parameters can affect the three sequences in time and vector domain. By enabling practical experimentation, the tool serves as both an educational resource and practical aid for students and engineers who are working with 3-phase systems.

2. MATERIALS AND METHODS

2.1. Motor Control

Motor control methods can be categorized as two groups: vector and scalar control [12,13]. Scalar control consists of SWC (square-wave control), voltage control, frequency control and V/f control (Volts per Hertz) method [13,15]. SWC mainly used to drive BLDC (brushless DC) motors. In this algorithm, the angle between the stator magnetic field and rotor magnetic field fluctuates from 60° to 120°. The rotor of the motor needs to commute six times per turn. Every commutation cause torque ripple. These ripples do not only affect speed accuracy, and they also cause vibration that affects the system stability. Instead of SWC, using FOC algorithm provides better speed response and torque stability. In a study while driving BLDC motors, the average overshoot, average rise time, and average settling time of FOC algorithm (10.43%, 0.44 s, 0.43 s) is less than SWC algorithm (15.68%, 0.67 s, 0.95 s) [1].

In V/f control algorithm, voltage, current, frequency, pole and flux are the main parameters. By changing the amplitude of these parameters, V/f control can be achieved. While the frequency is constant, by changing the pole, the speed of the motor can be controlled, but after manufacturing, the pole cannot be changed. Similarly, by changing the frequency, synchronous speed can be controlled. If the frequency increases, speed will increase but starting and maximum torques reduce. To maintain the flux constant, the ratio of voltage and frequency must be constant. If the flux is constant, maximum torque is constant, but starting torque will reduce [12,13].

Vector control consists of FWC (field weakening control), DTC (direct torque control), sliding mode control, adaptive control, fuzzy control, MPC (model predictive control) and FOC (field-oriented control) [2,16]. Induction motor field weakening control uses traditional $1/\omega_r$ method. Reference motor flux and above base speed is inversely proportional to the actual speed of the rotor. In this method, excitation current is given without considering voltage influence current regulation, so to achieve the optimal field weakening area is difficult. In this control strategy, it is required to set up a complex model and it is strongly dependent on the motor parameters, so it has poor robustness. FWC limits the voltage, reduces the magnetic field intensity, arranges enough torque current to improve speed and torque output capacity. Maximum torque mainly depends on the current limitation and magnetic flux level, so the reasonable distribution of current is the key to achieving torque increase [11].

The objective of DTC is to maintain torque and stator flux vector within two hysteresis bands. When the stator flux vector modulus is kept constant, the value of the rotor flux decreases as the load on the motor increases. The motor can produce any torque the dc-link voltage allows it. Unlike FOC, DTC does not require coordinate transformation, current control loop and separate pulse width modulator. The structure of DTC is simple, and it directly obtains PWM from hysteresis comparator to control the motor. Because it uses hysteresis, DTC has high torque, current, flux ripples, vibration and noise [2,16,17].

MPC is used control various motors [2,18]. If it is used to control permanent magnet synchronous motors (PMSM) is mainly divided into two: MPCC (model predictive current control) and MPTC (model predictive torque control). For MPCC, generally the predictive model is derived based on control increment and construct the cost function in the form of control increment, while in some applications, control quantity is directly used for the design. Control Increment-Based MPC controller can only limit the output control quantity indirectly, whilst Romberg type Control Quantity-Based MPC controller can directly limit the output control quantity. In MPCC, the increment of control quantity is controlled to eliminate unknown disturbance [2].

Field oriented control first proposed in 1972, with the objective of competing against the separately excited DC machines which have independent control of current and flux. FOC motor driver control unit is mainly composed of main controller circuit, a three-phase full-bridge inverter circuit, a three-phase current sampling circuit, an overcurrent protection circuit, a bus voltage sampling circuit, a Hall sensor signal sampling circuit, and a power supply circuit. FOC can be implemented in a direct way and indirect way. For a direct version, the angle of flux space vector is sensed directly and in an indirect version estimated slip and rotor speed are used to calculate the angle of the flux. To obtain the maximum torque, stator flux must be orthogonal or at 90° to rotor flux. Alternative names of these fluxes are currents of the quadrature axis and direct axis. Quadrature axis is aligned with torque component and direct axis is aligned with flux

component. If the angle between them is 0° , motor does not spin. As the angle starts to increase so the so does the torque. FOC seeks to maintain the angle between them at 90° . To obtain the stator flux (quadrature axis current – torque component) and rotor flux (direct axis current – flux component) Clarke and Park transforms are used [10,19,20].

2.2. Clarke-Park Transforms

Before implementation of any electrical system, it will be modelled, studied, analyzed and tested in simulation programs. For these actions and control of the electrical systems, accurate modelling is an important factor [21]. For modelling and control, many transform formulas have proposed such as Clarke, Park, Concordia, Kimbark and Boyajyan, Koga, Rama Rao, Fourier, Wavelet and S. Each of these coordinate's systems are defined by its own transformation matrix [22-25]. For this article, Clarke and Park transforms are chosen for their simplicity and global usage.

Both Clarke and Park transforms are invented with the intention of to ease the control of three-phase systems. During 1920, Park generalized Blondel's Two-Reaction Theory of Synchronous Machines. This method solves the armature fluxes in a salient machine along the two axes: quadrature and direct axes. In the 1930s, Clarke made a series of modifications to symmetrical components. These modifications ease the calculation of the unbalanced three-phase systems as they do not require the α operator ($1/120^\circ$) or complex numbers [4]. “dq0” based transforms used on wide variety of applications for solving objectives of Electrical Engineering and Power Electronics such as control of electric machines and drives, multimachine modelling, multi-inverter modelling, microgrid simulation, phase-locked loops (PLL), active power filters, VSI as a supply for standalone loads with the renewable energy sources, doubly fed induction motor dynamic model, fault-detection operation mode for an induction motor, three-phase power circuit analysis, three-phase voltages estimation, coordinates transformation of three-phase quantities, AC-DC converter theory and power quality [4,14,22,25].

Clarke transform converts 3-phase rotational frame (A, B, C) to a 2-phase reference frame (α , β). A 3-phase symmetrical system can be defined in Equation (1-3).

$$V_A = V_m \sin(\omega t) \quad (1)$$

$$V_B = V_m \sin\left(\omega t + \frac{2\pi}{3}\right) \quad (2)$$

$$V_C = V_m \sin\left(\omega t - \frac{2\pi}{3}\right) \quad (3)$$

The conversion is given in Eq. (4-6).

$$V_\alpha = \frac{2}{3} \left(V_A - \frac{1}{2} V_B - \frac{1}{2} V_C \right) \quad (4)$$

$$V_\beta = \frac{2}{3} \left(\frac{\sqrt{3}}{2} V_B - \frac{\sqrt{3}}{2} V_C \right) \quad (5)$$

$$V_0 = \frac{2}{3} \left(\frac{1}{2} V_A + \frac{1}{2} V_B + \frac{1}{2} V_C \right) \quad (6)$$

V_α and V_β represents α and β sequences. Zero sequence is represented by V_0 and for the balanced systems value of it is zero. Inverse Clarke transform serves as to convert a 2-phase system to a 3-phase system, and for a balanced system, the conversion can be seen in Equation (7-9).

$$V_A = V_\alpha \quad (7)$$

$$V_B = \left(-\frac{1}{2} V_\alpha + \frac{\sqrt{3}}{2} V_\beta \right) \quad (8)$$

$$V_C = \left(-\frac{1}{2} V_\alpha - \frac{\sqrt{3}}{2} V_\beta \right) \quad (9)$$

It can be observed that there is a 90° phase difference between V_α and V_β . It means that sinusoidal signals of α and β sequences have also 90° phase difference between each other.

Park transform is also known as direct-quadrature transform. It is used to transform a 2-phase fixed reference frame (α, β) to 2-phase rotary reference frame (d, q) and the transform is in the Equation (10,11).

$$V_d = \cos(\theta) V_\alpha + \sin(\theta) V_\beta \quad (10)$$

$$V_q = -\sin(\theta) V_\alpha + \cos(\theta) V_\beta \quad (11)$$

θ (reference angle) is obtained in terms of V_α and V_β by the Eq. 12.

$$\theta = \tan^{-1} \left(\frac{V_\beta}{V_\alpha} \right) \quad (12)$$

V_d and V_q represents the direct and quadrature axes. θ , the ratio between V_α and V_β , is used for calculation of 2-phase rotary frame axes. Inverse Park transform converts 2-phase time dependent reference frame to a 2-phase fixed reference frame. Transform equations are given in Equation (13,14), the frequency of the system used for finding the reference angle and the equation to find it is given in Equation 15.

$$V_\alpha = \cos(\theta) V_d - \sin(\theta) V_q \quad (13)$$

$$V_\beta = \sin(\theta) V_d + \cos(\theta) V_q \quad (14)$$

$$\theta = \frac{\pi}{2} - 2\pi f t \quad (15)$$

Clarke-Park transform is the name of the combination of these two transforms. It combines the formulas of these two and the equations of it are given in Equation (16-18).

$$V_d = \frac{2}{3} \left(\cos(\theta) V_A + \cos\left(\theta - \frac{2\pi}{3}\right) V_B + \cos\left(\theta + \frac{2\pi}{3}\right) V_C \right) \quad (16)$$

$$V_q = \frac{2}{3} \left(\sin(\theta) V_A + \sin\left(\theta - \frac{2\pi}{3}\right) V_B + \sin\left(\theta + \frac{2\pi}{3}\right) V_C \right) \quad (17)$$

$$\theta = \tan^{-1} \left(\frac{V_A - \frac{1}{2}V_B - \frac{1}{2}V_C}{\frac{\sqrt{3}}{2}V_B - \frac{\sqrt{3}}{2}V_C} \right) \quad (18)$$

Inverse Clarke-Park transform converts 2-phase rotary frame to 3-phase rotational frame and is the combination of inverse Clarke and inverse Park transform. The conversion equations are given in Equation (19-22).

$$V_A = \cos(\theta) V_d - \sin(\theta) V_q \quad (19)$$

$$V_B = \cos\left(\theta - \frac{2\pi}{3}\right) V_d + \sin\left(\theta - \frac{2\pi}{3}\right) V_q \quad (20)$$

$$V_C = \cos\left(\theta + \frac{2\pi}{3}\right) V_d + \sin\left(\theta + \frac{2\pi}{3}\right) V_q \quad (21)$$

$$\theta = \frac{\pi}{2} - 2\pi f t \quad (22)$$

3. DEVELOPMENT OF SOFTWARE

To ease the calculation and learning of Clarke-Park transforms, an application is developed. The application visualized the conversion of 3-phase sequence to $\alpha\beta$ sequence, $\alpha\beta$ sequence to 3-phase sequence, $\alpha\beta$ sequence to dq sequence, dq sequence to $\alpha\beta$ sequence, 3-phase sequence to dq sequence, dq sequence to 3-

phase sequence. It will be a practical tool for those who want to learn Clarke and Parke transforms, and control and modelling of the 3-phase systems. How the d and q amplitudes could affect the phase, and the system could be observed via this application. The steps of the application algorithm could be seen as below:

1. Start the program
2. Enter the values for desired transform
3. Press the “Calculate” button that is under the desired transform values to see all 3 sequences on the time domain
4. Before new calculation canvas can be cleared by pressing “Clear” button
5. Press the “Plot Reference Frames” button to see all 3 sequences on their respectable frames.
6. Press the “Close” button to come back to the previous window
7. Press the “Show Transform Formulas” button to see all transform formulas
8. Press the “Close” button to come back to the previous window
9. Exit the program

The application is developed by python programming language, for easily making the adjustments. Tkinter graphical python library is used for the graphical user interface and Matplotlib is used for plotting. The tool provides the opportunity of exploring the relationships between three-phase (abc), Clarke ($\alpha\beta$), and Park (dq) reference frames for the users in a highly intuitive manner. With the launch of the application, users can see the three main input panels corresponding to the three-phase system, Clarke transform, and Park transform. Each panel allows user to enter desired amplitude, frequency and phase (or dq amplitudes) values of the respective signals. After entering the values, users can generate the waveforms by pressing the “Calculate” button in each section and they can observe the change of the waveforms by changing the values [26, 27]. The GUI (graphical user interface) of the application can be seen in Figure 1.

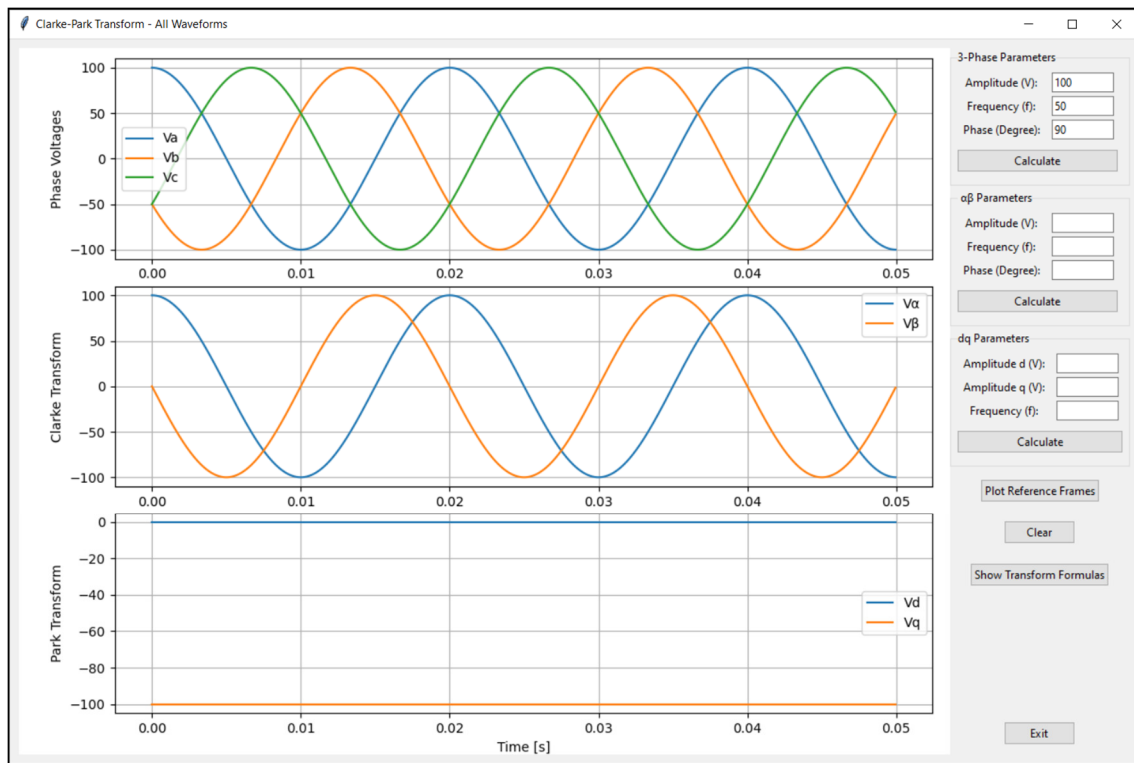


Figure 1. Software user interface

Additionally, the application provides “Plot Reference Frames” button, which opens a new window that displays the vector diagrams for the abc, $\alpha\beta$, and dq reference frames. This visualization aids the users to understand the transformation of the signals onto orthogonal axes, and how the d and q amplitudes, and phase value proportion works. This is a key concept in vector control and field-oriented control of AC machines. For user convenience, “Exit” buttons are added both in the main window and in the reference

frame plot window, allowing for easy closure of the application or individual plots. The reference frame plot window can be seen in Figure 2. For the users those would like to see or want to remember or learn, there is “Show Transform Formulas” button. If the users press on this button, all six transform formulas will be at this window. Transform formulas window can be seen Figure 3.

The application is packaged into a standalone executable using PyInstaller, making it easily distributable to end users. The distribution package includes the main executable file and files that the main executable file used when working [28, 29]. This application is a useful tool for students, educators and engineers who wish to gain deeper understanding of and are working on electric-machine control, modelling and mathematical transformations. By allowing real-time parameter adjustment and immediate visualization, the application fills the gap between theoretical and practical concepts. This article section provides an overview of the application's development while remaining accessible to readers with a basic understanding of electrical systems and software development.

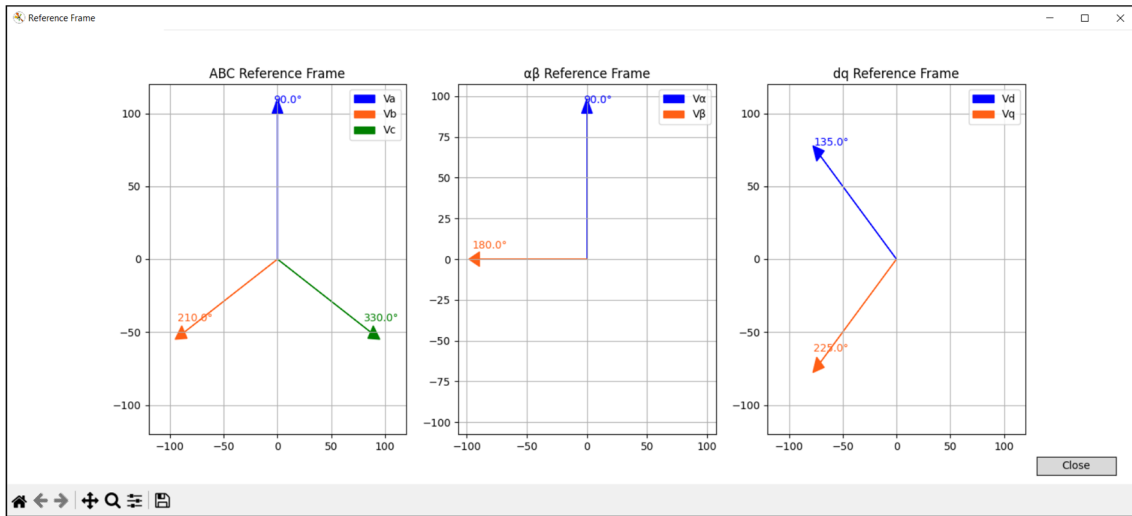


Figure 2. Reference frame window

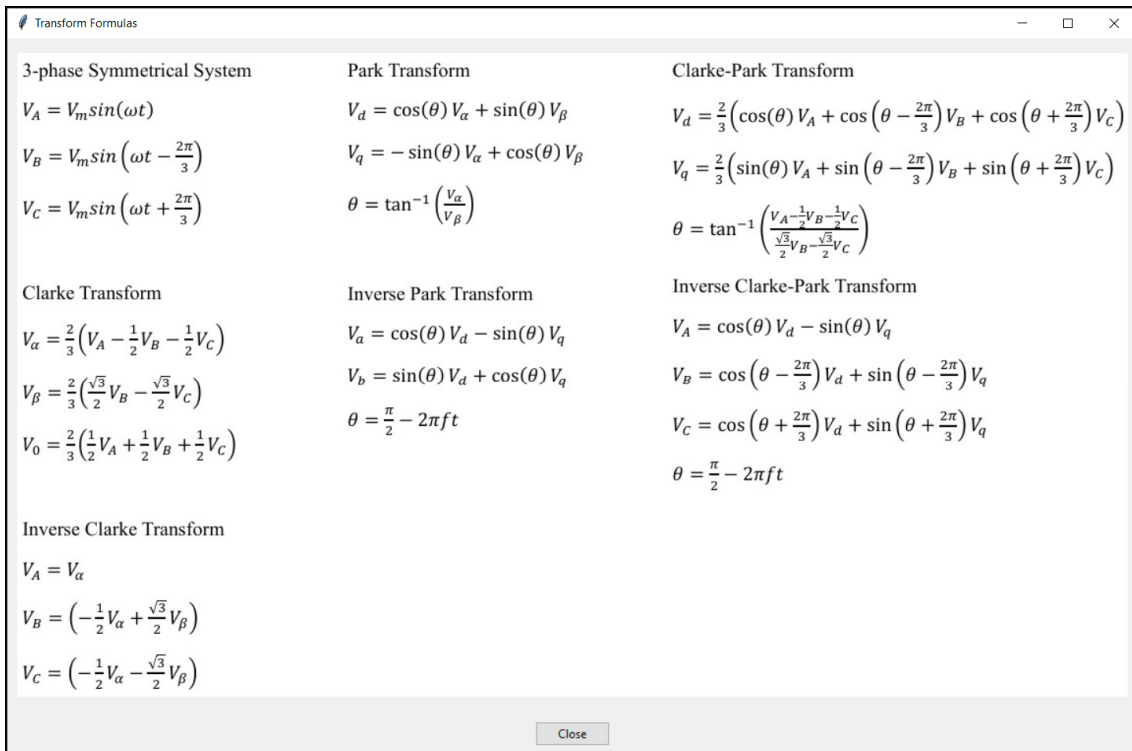


Figure 3. Transform formulas window

4. CONCLUSION

In this study, information related to motor control, Clarke and Park transforms has been given alongside the development of a user-friendly software tool that visualizes and calculates these transforms. Motor control methods can be examined in two groups: vector and scalar control. As its name suggests scalar control methods use scalar parameters such as voltage, current, frequency, pole and flux to control motor. By changing the frequency, motor speed can be controlled. Changing the frequency affects the motor torque. While voltage is constant and frequency increased, motor torque decreases. Scalar control finds optimum values for these parameters. FWC, DTC and FOC are mostly used vector control methods, but the most popular control method is FOC. It uses Clarke and Park transform to simplify the control of motor. To acquire the maximum torque, rotor flux must be orthogonal to stator flux. To achieve this state, FOC converts 3-phase amplitudes and phase to two sequence vectors (dq). This conversion has been made by using Clarke and Park transforms.

Clarke transform converts 3-phase rotational frame into 2-phase stationary frame. Park transform 2-phase stationary frame into dq reference frame. With the inverse clarke and park transforms, all three frames can be converted into each other. In the proposed application, with its interactive and dynamic plotting capabilities, understanding of how changes in system parameters could effect the 3-phase system is enhanced. The forward and inverse transform features not only support learning but also assist the engineers in calculating clarke and park transform.

This tool provides valuable resources both educational and practical for students, educators and engineers by simplifying complex concepts through visualization. The software contributes to better understanding and more effective application for Clarke and Park transforms into 3-phase control systems and promotes deeper insight and more intuitive learning in the field of electric machine control. For the later studies, the application can be further expanded with real time data acquisition and hardware integration. These new features make it a more powerful tool for both simulation and practical implementation for 3-phase systems.

To conclude, in this study, the implementation of Clarke-Park transforms software that calculates, visualizes and converts 3 different sequences into each other is carried out. With this article and software, understanding of Clarke and Park transforms will be easier and they will be a reference point for the ones who are interested in these subjects. In order to obtain the application, it is recommended to contact the corresponding author via e-mail.

5. REFERENCES

1. Du, Z., Yang, L., Zhang, D., Cui, T., He, X., Xiao, T., Xing, S., Xie, C. & Li, H. (2023). Development and testing of a motor drive and control unit based on the field-oriented control algorithm for the seed-metering device. *Elsevier: Computers and Electronics in Agriculture*, 211, 108024.
2. Ling, L., Huang, S. D., Cao, G. Z. & Qiu, H. (2023). Comparison analysis on control-increment-based and control-quantity-based predictive controls of permanent magnet synchronous motors. *International Conference on Predictive Control of Electrical Drives and Power Electronics (PRECEDE)*, Wuhan, China, 1-6.
3. Lin, H., Hu, B., Li, F., Chen, J., Si, L., Zhou, X., Li, Y., Chen, J., Yan, M. & Dong, Y. (2018). A fault-tolerant two-permanent magnet synchronous motor drive with field-oriented control scheme. *2nd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, Xi'an, China, 1029-1033.
4. O'Rourke, C.J., Qasim, M.M., Overlin, M.R. & Kirtley, J.L. (2019). A geometric interpretation of reference frames and transformations: dq0, Clarke, and Park. *Transactions on Energy Conversion*, 34(4), 2070-2083.
5. Casado-Machado, F., Martinez-Ramos, J.L., Barragán-Villarejo, M., Maza-Ortega J.M. & Rosendo-Macías, J.A. (2020). Reduced reference frame transform: Deconstructing three-phase four-wire systems. *In IEEE Access*, 8, 143021-143032.
6. Carugatia, I., Orallo, C.M., Maestri, S., Donato, P.G. & Carrica, D. (2015). Error analysis of phase detector based on Clarke transform and arctangent function in polluted grids. *Elsevier: Electric Power Systems Research*, 127, 160-164.

7. Velpula, R., Nagarajan, M. & Pitchaimuthu, R. (2024). An alternate method to detect and classify the transmission line faults using Clarke's transformed currents. *Elsevier: Electric Power Systems Research*, 236, 110899.
8. Tayyebi, A., Zhang, W., Huang, X., Jiang, W., Bormann, D. & Mats Larsson, M. (2024). A generalized time-domain framework for modeling and analysis of the unbalanced three-phase systems. *Elsevier: Electric Power Systems Research*, 235, 110809.
9. Dashti, R., Daisy, M., Javadi, S. & Aliabadi, M.H. (2021). Proposing a new method to improve the longitudinal differential relay performance using the Clarke transformation: Theory, simulation, and experiment. *Elsevier: Measurement*, 168, 108450.
10. Nustes, J.C., Pau, D.P. & Gruosso, G. (2023). Field oriented control dataset of a 3-phase permanent magnet synchronous motor. *Elsevier: Data in Brief*, 47, 109002.
11. Xie, P., Li, G., Xie, F., Hu, C. & Qi, X., (2015). Research on field-weakening control of induction motor based on torque current component of the voltage closed-loop. *10th Conference on Industrial Electronics and Applications (ICIEA)*, Auckland, New Zealand, 1618-1621.
12. Goswami, V.M. & Vakharia, K. (2019). High performance induction machine drive using rotor field oriented control. *International Conference on Intelligent Sustainable Systems (ICISS)*, Palladam, India, 559-564.
13. Ghassani, R., Kader, Z., Fadel, M., Combes, P. & Koteich, M. (2023). Comparison study of rotor field-oriented control and stator field-oriented control in permanent magnet synchronous motors. *International Electric Machines & Drives Conference (IEMDC)*, San Francisco, CA, USA, 1-7.
14. Xie, P., Li, G., Xie, F., Hu, C. & Qi, X. (2015). Research on field-weakening control of induction motor based on torque current component of the voltage closed-loop. *10th Conference on Industrial Electronics and Applications (ICIEA)*, Auckland, New Zealand, 1618-1621.
15. Megrini, M., Gaga, A., Mehdaoui, Y. & Khyat, J. (2024). Design and PIL test of extended Kalman filter for PMSM field-oriented control. *Elsevier: Results in Engineering*, 24, 102843.
16. Wang, J. & He, J. (2008). Torque and flux direct backstepping control of induction motor. *7th World Congress on Intelligent Control and Automation*, Chongqing, China, 6407-6410.
17. Ludtke, I. & Jayne, M.G., (1995). Direct torque control of induction motors. *Colloquium on Vector Control and Direct Torque Control of Induction Motors*, London, UK, 6/1-6/6.
18. Özgür, H.E., Özbek, N.S. & Sarıgeçili, M.İ. (2022). Noise and disturbance rejection performance evaluation on explicit model predictive control technique applied to inverted pendulum with various test scenarios. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 37(3), 643-652.
19. Kumar, M., Tiwari, S., Kumar, V., Sampathi S.J. & Jarial, R.K. (2024). Implementation of field-oriented control (FOC) algorithm for brushless DC (BLDC) motor speed regulation. *Second International Conference on Measurement, Instrumentation, Control and Automation (ICMICA)*, Kurukshetra, India, 1-6.
20. Tahmaz, O., Ekim, M.N. & Yildiz, A.B. (2020). Vector control of permanent magnet synchronous motor by a two-level SPWM inverter. *4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, Istanbul, Turkey, 1-7.
21. Alruim Alhasan, H. ve Güneş, M. (2017). Yeni bir adaptif parçacık sürü optimizasyon algoritması kullanarak DC motor için mız ayarlamalı PID kontrolör tasarımı. *Çukurova Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 32(3), 243-250.
22. Nos, O.V. & Nos, N.I. (2023). Clarke-Park coordinates transformation in the quaternion domain. *IEEE XVI International Scientific and Technical Conference Actual Problems of Electronic Instrument Engineering (APEIE)*, Novosibirsk, Russian Federation, 420-423.
23. Liu, H., Tang, Y., Feng, Y. & Ma, X. (2008). A power quality disturbance classification method based on Park transform and Clarke transform analysis. *3rd International Conference on Innovative Computing Information and Control*, Dalian, China, 524-524.
24. Hoang, T.T. & Nair, N.-K.C. (2023). Advanced approach for stability assessment of PHIL setups coupled by Clarke-Park transform. *Power & Energy Society General Meeting (PESGM)*, Orlando, FL, USA, 1-5.
25. Dell'Olmo, J., Gatta, F., Geri, A., Graziani, M., Lauria, S. & Maccioni, M. (2022). Clarke transform based fast assessment of switching overvoltages in an MV distribution network. *Elsevier: Electric Power Systems Research*, 212, 108255.
26. Vasavi, S., Nikhita Sri, P.D.L. & Sai Krishna, P.V. (2024). GUI-enabled boundary regularization system for urban buildings using the tkinter. *2nd International Conference on Device Intelligence, Computing and Communication Technologies (DICCT)*, Dehradun, India, 424-429.

27. Charan Sai, P., Karthik, K., Bhargav Prasad, K., Pranav, C.V.S. & Divya, K.V. (2024). Real-time task manager: A python-based approach using psutil and tkinter. *8th International Conference on Computational System and Information Technology for Sustainable Solutions (CSITSS)*, Bengaluru, India, 1-6.
28. Spencer, M., Sheiati, S. & Chen, X. (2023). AQUADAGUI: A graphical user interface for automated quantification of damages in composite structures under fatigue loading using computer vision and thermography. *Elsevier: SoftwareX*, 22, 101392.
29. Yoon, G. & Rho, J. (2021). MAXIM: Metasurfaces-oriented electromagnetic wave simulation software with intuitive graphical user interfaces. *Elsevier: Computer Physics Communications*, 264, 107846.



Çözeltili Üfleme Tekniği ile Farklı Bitim İşlemi Kimyasallarının Tekstil Yüzeylerine İşletme Şartlarında Aktarılması ve Performanslarının Değerlendirilmesi

Koray PEKTAŞ^{1,a}, Onur BALCI^{1,b}

¹Kahramanmaraş Sütçü İmam Üniversitesi, Mühendislik ve Mimarlık Fakültesi, Tekstil Mühendisliği Bölümü, Avşar Yerleşkesi, Onikişubat, Kahramanmaraş, Türkiye

^aORCID: 0000-0002-9744-8308; ^bORCID: 0000-0001-6885-7391

Makale Bilgileri

Geliş : 16.07.2025

Kabul : 04.09.2025

DOI: 10.21605/cukurovaumfd.1744255

Sorumlu Yazar

Koray PEKTAŞ

koraypektas@ksu.edu.tr

Anahtar Kelimeler

Çözeltili üfleme eğirme

Bitim işlemi

Hidrofil silikon

Su iticilik

Atıf şekli: PEKTAŞ, K., BALCI, O., (2025). Çözeltili Üfleme Tekniği ile Farklı Bitim İşlemi Kimyasallarının Tekstil Yüzeylerine İşletme Şartlarında Aktarılması ve Performanslarının Değerlendirilmesi. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 617-626.

ÖZ

Çalışmada, çözeltili üfleme eğirme (SBS) tekniği kullanılarak bitim kimyasallarının kumaş yüzeyine püskürtülerek aktarılabilirliği incelenmiştir. Amaç, konvansiyonel yöntemlere kıyasla su, kimyasal ve enerji tüketimini azaltmaktır. Karşılaştırma için emdirme yöntemiyle de uygulama yapılmıştır. Hidrofil silikon emülsiyonu (yumuşatıcı) ve su itici kimyasal, önceden boyanmış süprem ve flamlı süprem kumaşlara önce fular makinesiyle aktarılmış, ardından ramözde kurutma ve kondenzasyon yapılmıştır. Prototip SBS cihazı, ramöz önüne entegre edilerek yeni uygulama yöntemi oluşturulmuştur. Her iki yöntemle işlem gören kumaşlara performans testleri uygulanmış, ayrıca teorik kimyasal ve enerji tüketimleri hesaplanmıştır. Sonuçlar, benzer performans elde edildiğini göstermiştir. Böylece SBS tekniğinin az flotte aplikasyon yöntemi olarak kullanılabileceği ve konvansiyonel yöntemle göre daha az kaynak tüketimiyle çevreci, sürdürülebilir bir alternatif sunduğu belirlenmiştir.

Transfer of Different Finishing Chemicals to Textile Surfaces by Solution Blowing Technique Under Operating Conditions and Evaluation of Their Performance

Article Info

Received : 16.07.2025

Accepted : 04.09.2025

DOI: 10.21605/cukurovaumfd.1744255

Corresponding Author

Koray PEKTAŞ

koraypektas@ksu.edu.tr

Keywords

Solution blowing spinning

Finishing

Hydrophile silicone

Water repellance

How to cite: PEKTAŞ, K., BALCI, O., (2025). Transfer of Different Finishing Chemicals to Textile Surfaces by Solution Blowing Technique Under Operating Conditions and Evaluation of Their Performance. Çukurova University, Journal of the Faculty of Engineering, 40(3), 617-626.

ABSTRACT

The study investigated the transferability of finishing chemicals onto fabric surfaces by spraying them using the solution blow spinning (SBS) technique. The aim was to reduce water, chemical, and energy consumption compared to conventional methods. For comparison, the application was also performed using the impregnation method. Hydrophilic silicone emulsion (softener) and water-repellent chemical were first transferred to pre-dyed supreme and flamed supreme fabrics using a fular machine, then dried and condensed in a stenter. A prototype SBS device was integrated in front of the stenter to create a new application method. Performance tests were conducted on fabrics treated with both methods, and theoretical chemical and energy consumption calculations were performed. The results showed similar performance. Thus, it was determined that the SBS technique can be used as a low-flotte application method and offers an environmentally friendly, sustainable alternative with lower resource consumption compared to the conventional method.

1. GİRİŞ

Endüstride su kullanımı denildiğinde akla ilk gelen endüstrilerden biri hiç şüphesiz tekstildir. Tekstil endüstrisi yaklaşık %20'lik su kullanımı oranı ile kimya endüstrisinden sonra gelmektedir [1,2]. Özellikle tekstil endüstrisinde yaş işlemler sırasında çok fazla miktarda atık su ve kimyasal tüketimi meydana gelmektedir [3]. Atık su ve fazla kimyasal tüketiminin önüne geçmek için yaş işlem proseslerinde az flotte aplikasyon yöntemleri kullanılabilmektedir. Günümüzde bu yöntemler köpük aplikasyon, kiss-roll, vakum ekstraksiyon ve püskürtme yöntemleri karşımıza çıkmaktadır. Bu yöntemler kullanılarak, kimyasal taşıyıcı sıvı miktarı azalmakta, enerji ve su tüketimi düşmekte, migrasyon riski azalırken üretim hızı ve işlem verimliliği artmaktadır [4-6]. Buna bağlı olarak aynı zamanda üretici maliyetleri de düşmektedir. Ancak bu yöntemleri konvansiyonel yöntemler ile kıyaslandığında endüstride kullanım alanının henüz yaygınlaşmadığı görülmektedir. Bu yöntemlerin yanı sıra her geçen gün endüstride de kullanılabilecek yeni aplikasyon yöntemleri geliştirilmeye çalışılmaktadır.

Yeni aplikasyon yöntemi olarak çeşitli nano püskürtme teknikleri konuşulmaktadır, çözelti üfleme tekniğinin de bunlar arasında yer alabilme potansiyeli bulunmaktadır. Çözelti üfleme ile aplikasyon tekniğinin öncelikli kullanım amacına bakıldığında, sistemin bir nanolif çekim sistemi olduğu ve çözelti üfleme eğirme yöntemi olarak literatürde geçtiği bilinmektedir [7-15]. Ancak bu yöntemin çalışma parametrelerini (besleme hızı, hava basıncı, çözelti konsantrasyonu ve çalışma mesafesi) değiştirerek nanolif çekimi yerine, mikro ve nano boyutta kimyasalların kumaş üzerine aktarılması ve böylece az flotte aplikasyon yöntemi olarak kullanılması mümkün olduğu düşünülmektedir.

Çalışma kapsamında, çözelti üfleme aplikasyon tekniğinin açık en örgü kumaşlarda kullanılabilmesi için TÜBİTAK TEYDEB 1501 projesi kapsamında bir prototip makine tasarlanmış ve imal edilmiştir. Bu prototip makine ramöz makinesinin önüne adapte edilerek, ramöz makinesi ile kesiksiz çalışması sağlanmıştır. Prototip makinenin konvansiyonel yöntem ile kıyaslanabilmesi için hem emdirme yöntemi hem de çözelti üfleme aplikasyon yöntemi ile kumaşlara hidrofil silikon (yumuşatma) ve su iticilik kimyasalları aktarılmıştır. Farklı aplikasyon işlemleri sonrasında elde edilen numune kumaşlara fiziksel performans testleri yapılmıştır. Elde edilen sonuçlar birbirleri ile kıyaslanarak çözelti üfleme aplikasyon yönteminin, konvansiyonel yöntem yerine kullanılabileceği olanağı tartışılmıştır.

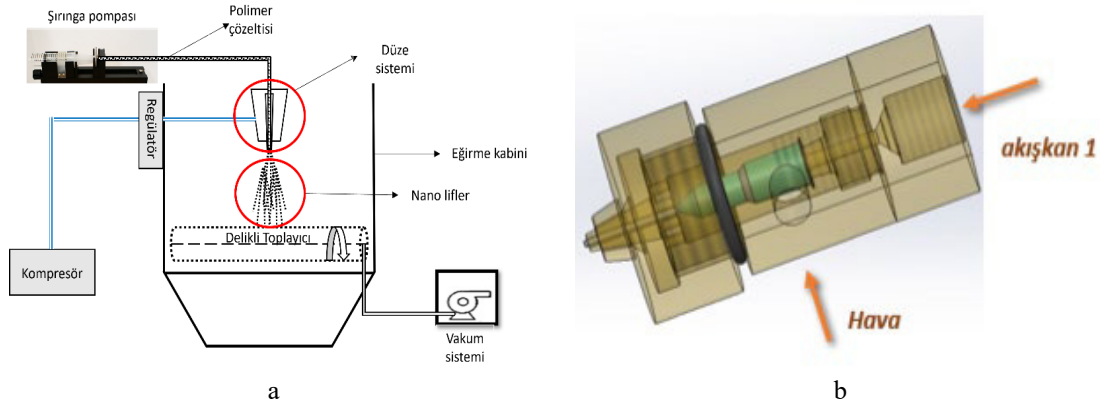
2. YÖNTEM

2.1. Materyal

Çalışma kapsamında iki farklı örgü kumaş kullanılmıştır. Numune olarak süprem ve flamlı süprem örme kumaş seçilmiştir. Bitim işleminde ise yumuşatma amaçlı hidrofil silikon emülsiyonu ve su iticilik kimyasalları kullanılmıştır. Bu kumaşlar, işletme şartlarında ön terbiyesi (sadece ağartma işlemi) ve boyaması (reaktif boyarmaddeler kullanarak, çektirme metodu ile jet boyama makinesinde) yapılan numunelerdir.

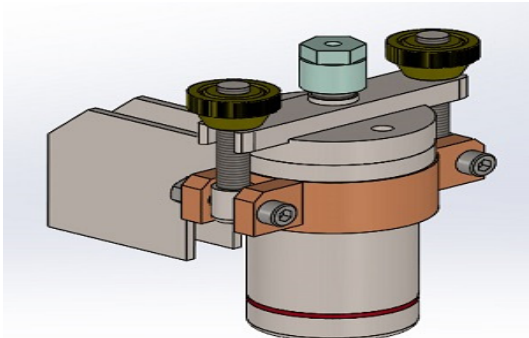
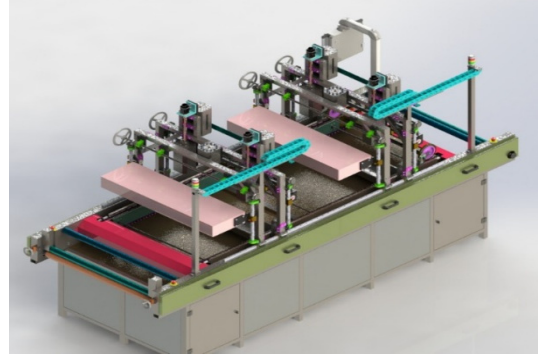
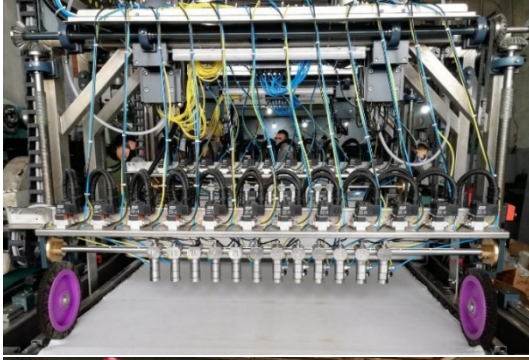
2.2. Metot

Çalışma kapsamında iki farklı aplikasyon yöntemi kullanılmıştır. Kullanılan bu yöntemlerden biri konvansiyonel metot olarak bilinen emdirmedir. Bir diğer aplikasyon yöntemi olarak ise çözelti üfleme tekniği seçilmiştir. Çözelti üfleme tekniği literatürde nanolif üretim tekniği olarak bilinmekte olup, sistem eş merkezli iki adet düzeden (Şekil 1-b) oluşmaktadır. İç düze kimyasal (lif çekiminde polimer) çözeltisinin akışını sağlamaktadır. Dıştaki düze ise sisteme yüksek basınçlı hava göndererek iç düzeden çıkan çözeltinin toplayıcı tabakaya doğru yönelmesini gerçekleştirmektedir. Lif çekiminde bu basınçlı hava, çözeltideki çözücünün uzaklaştırılmasını sağlamaktadır. Apre aplikasyonunda ise bu basınçlı hava kimyasalın yüzeye yönlendirilmesini sağlamaktadır. Prensibi Şekil 1-a'da gösterilmiştir.

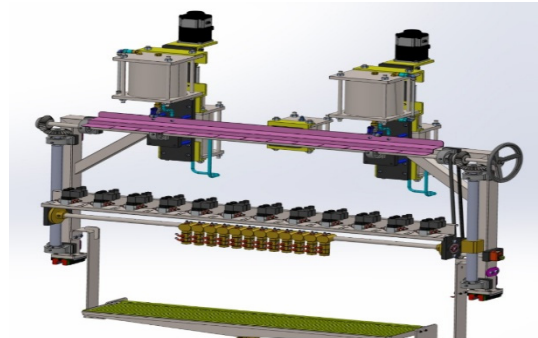


Şekil 1. Çözelti üfleme sistemi şematik çizim ve düze mekanizması

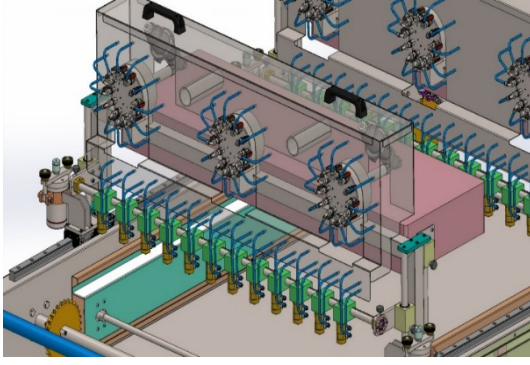
Bu yöntemde nanolif çekimine etki eden parametre ayarları değiştirilerek, kimyasal maddeler taşıyıcı yüzey üzerine mikro ne nano boyutta püskürtülebilmektedir. Böylece esasında bir nanolif çekim yöntemi olan çözelti üfleme eğirme tekniği, az flotte aplikasyon yöntemi olarak kullanılmıştır. Çalışma kapsamında prototip olarak elde edilen çözelti üfleme eğirme cihazına ait görsel, Şekil 2’de verilmektedir. Prototip çözelti üfleme cihazının ramöz önüne adaptasyonu ise Şekil 3’te gösterilmektedir.



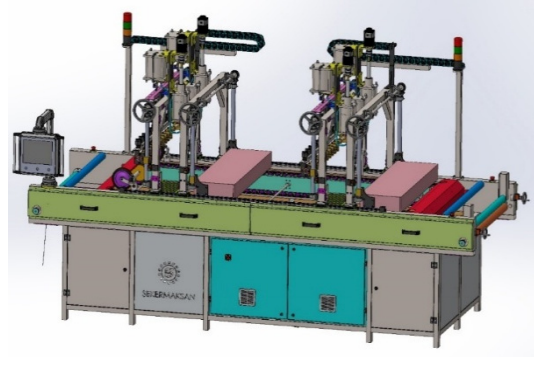
Aplikatör modülü



Sıvı dozajlama sistemi



Düzelerin bulunduğu sıvı dağıtım hattı



Prototip izometrik görüntüsü

Şekil 2. Prototip çözelti üfleme eğirme makinası



Şekil 3. Ram önüne adaptasyonu tamamlanmış endüstriyel hat

Çözelti üfleme cihazında kullanılan alınan flote oranı (kumaşın kuru ağırlığına göre emdiği işlem çözeltisinin yüzde oranı), prototip makine üzerinde çeşitli denemeler yapılarak (besleme hızı, makine hızı, hava basıncı, çalışma mesafesi) ramöz önünde kullanılmak üzere belirlenmiştir.

Çalışma kapsamında bu yöntemlere ait çalışma parametreleri Çizelge 1-3'te verilmektedir. Kumaşların kurutulması, ısıl işlem görmesi ve boyutsal stabilizasyonunun sağlanması için ramöz makinesi kullanılmıştır. Ramöz makinesinde yapılan kurutma ve kondenzasyon işlemi her 2 aplikasyon için de sabit seçilmiştir.

Çizelge 1. Emdirme yöntemi çalışma parametreleri

Makine hızı (m/dk)	Alınan flote oranı (%)	Sıkma basıncı (bar)
10	90	2,5

Çizelge 2. Çözelti üfleme yöntemi çalışma parametreleri

Makine hızı (m/dk)	Alınan flote oranı (%)	Hava basıncı (bar)	Çalışma mesafesi (cm)
10	67	5	25

Çizelge 3. Ramöz makinesi çalışma parametreleri

Makine hızı (m/dk)	Kurutma sıcaklığı (°C)	Kondenzasyon sıcaklığı (°C)
10)	140	170

Çalışma kapsamında oluşturulan deney planı Çizelge 4'te verilmiştir. Referans olarak her numuneden de işlemsiz (apresiz-boyalı) numune alınmış ve bunlara “Ref 1 ve Ref 2” kodu verilmiştir. Çalışma kapsamında aplikasyon tipi değişken olarak seçilmiştir. Buna göre toplam 4 aplikasyon denemesi gerçekleştirilmiştir. 2 adet işlemsiz numune ile toplam 6 numune elde edilmiştir. Yumuşatma denemeleri Referans 1’e, su iticilik denemeleri ise Referans 2’ye yapılmıştır.

Çizelge 4. Deney planı

Denemeler	Aktarma yöntemi	Kodlar
İşlemsiz kumaş-1(Pamuklu süprem)	-	Ref-1
Kumaş-1	Emdirme	HS-E-1
40 g/L hidrofil silikon		
Kumaş-1	Çözelti üfleme	HS-ÇÜ-1
40 g/L hidrofil silikon		
İşlemsiz kumaş-2	-	Ref-2
(Flamlı süprem)		
Kumaş-2	Emdirme	Sİ-E-2
100 g/L Su iticilik		
Kumaş-2	Çözelti üfleme	Sİ-ÇÜ-2
100 g/L Su iticilik		

2.3. Testler

Çalışma kapsamında emdirme ve çözelti üfleme tekniği ile bitim işlemi uygulanarak elde edilen 6 numune kumaşın fiziksel ve fonksiyonel performanslarının değerlendirilmesi için gramaj (EN ISO 3801), patlama (TS EN ISO 13938-2) mukavemeti, dikey kılcallık emme testi (AATCC 197), su iticilik (AATCC 22), pilling (ISO 12945-1), kuru ve yaş sürtmeye karşı, yıkamaya karşı (ISO 105-C06), suya karşı (ISO 105-E01), asidik ve bazik tere karşı (ISO 105 E04) renk haslığı testleri yapılmıştır.

3. ARAŞTIRMA BULGULARI

İşletme şartlarında her iki aktarma yöntemi ile bitim işlemleri yapılan numunelerin performansları çeşitli testlerle değerlendirilmiştir. Çalışma amacı düşünüldüğünde, SBS tekniği ile elde edilen performansın, emdirme tekniğine göre üstün olması beklenmemektedir. Temel amaç, en azından aynı performansı elde ederek, daha az su, kimyasal ve enerji tüketimiyle sonuca ulaşmaktır. Yani çalışmada sadece performans iyileştirmek değil, aynı zamanda sürdürülebilir üretim sağlamakta hedeflenmiştir. Çizelgelerde verilen sonuçların bu şekilde değerlendirilmesi uygun olacaktır.

Çizelge 5’te bitim işlemleri sonrasında numunelere ait pilling, kuru ve yaş sürtme haslığı test sonuçları verilmektedir.

Çizelge 5. Numunelere ait pilling, kuru ve yaş sürtünmeye karşı haslığı test sonuçları

Kodlar	Pilling	Sürtme (kuru)	Sürtme (yaş)
Ref-1	4	4/5	2/3
HS-E-1	4	4/5	3
HS-ÇÜ-1	4	4/5	3
Ref-2	4	4/5	2/3
Sİ-E-2	4	4/5	2/3
Sİ-ÇÜ-2	4	4/5	2/3

Çizelge 5'te elde edilen sonuçlar incelendiğinde, her iki yöntem ile kumaşlara aktarılan silikon emülsiyonu ve su itici kimyasalın kumaş performansı üzerinde herhangi bir değişikliğe neden olmadığı tespit edilmiştir. Çizelge 6'da numunelere ait yıkamaya karşı renk haslığı, suya karşı renk haslığı, asidik ve bazik tere karşı renk haslığı sonuçları verilmektedir.

Çizelge 6. Numunelere ait yıkamaya karşı renk haslığı, suya karşı renk haslığı, asidik ve bazik tere karşı renk haslığı sonuçları

Kodlar	Haslık			
	Yıkama	Su	Asidik ter	Alkali ter
Ref-1	4/5	4/5	4/5	4/5
HS-E-1	4/5	4/5	4/5	4/5
HS-ÇÜ-1	4/5	4/5	4/5	4/5
Ref-2	4/5	4/5	4/5	4/5
Sİ-E-2	4/5	4/5	4/5	4/5
Sİ-ÇÜ-2	4/5	4/5	4/5	4/5

Çizelge 6'da verilen haslık sonuçları incelendiğinde, aktarma yönteminden bağımsız olarak kumaşların haslık performanslarında bir değişiklik meydana gelmediği görülmüştür.

Numunelere ait kılcallık test sonuçları Çizelge 7'de verilmektedir.

Çizelge 7. Numunelere ait kılcallık test sonuçları

Kodlar	Kılcallık	
	En (mm)	Boy (mm)
Ref-1	105	110
HS-E-1	138	160
HS-ÇÜ-1	130	155
Ref-2	108	130
Sİ-E-2	0	0
Sİ-ÇÜ-2	0	0

Kılcallık test sonuçları incelendiğinde, her iki aktarma yöntemi ile bitim işlemi uygulandıktan sonra benzer sonuçlara ulaşıldığı görülmüştür. Hidrofil silikon ile işlem sonrasında kılcallık değerlerinde aktarma yöntemine bakılmaksızın artış meydana geldiği görülmüştür. Çünkü hidrofil silikonların kullanılmasıyla birlikte kumaşın yüzey enerjisi artmış, su temas açısını azalmış, kumaş yüzeyinde film tabakası oluşarak, su damlacıklarının lif yüzeyine yapışma eğilimi düşmüştür. Böylece, su lif yüzeyine daha iyi yayılmıştır. Hidrofilik grupların lif yüzeyinin ıslatma kabiliyeti sayesinde lifler arasındaki boşluklarda suyun nüfuz etme hızı artmış ve son olarak lifler şişerek düzenli bir kanal yapısı oluşmuştur. Buna nedenle kumaşın hidrofilite değerlerinde artış meydana gelmiştir [16-18]. Benzer şekilde su iticilik kimyasalının uygulandığı numunelerde de aplikasyon yöntemine bakılmaksızın aynı sonucun alındığı (0) görülmüştür. Su itici kimyasallar, lif yüzeyine uygulandığında liflerin yüzey enerjisini düşürmüş ve liflerin su ile ıslanabilirliğini azaltmıştır. Böylece su damlacıkları kumaş yüzeyinde yayılmak yerine, yüksek temas açısıyla küresel formda kalmıştır. Bu durum, kılcallıkla emme mekanizmasının ilk aşaması olan yüzey ıslanmasını engellemektedir. Oysa kılcallıkla emme, lif yüzeyinin suyu emmesi ve lifler arası mikro kanallar boyunca kapiller kuvvetlerle yukarı doğru taşınmasıyla gerçekleşmektedir. Ancak hidrofobik yüzey kapiller hareketi başlatacak kadar temas alanı sunmadığı için su, lif yüzeyine tutunamaz ve mikro kanallara nüfuz edemez. Böylece kılcallık mekanizması devre dışı kalmakta ve kumaşın dikey emme performansı önemli ölçüde azalmaktadır [19-21].

Su iticilik kimyasalı uygulanmış numunelere yapılan spre test sonuçları Çizelge 8 de verilmektedir.

Çizelge 8. Numunelere ait spre test sonuçları

Kodlar	Su iticilik değerleri
Ref-2	1
Sİ-E-2	5
Sİ-ÇÜ-2	4

Spre test sonuçlarına göre su iticilik bitim işlemi görmemiş kumaşın su iticilik değerinin “1” yani kumaşın ıslanabilir düzeyde olduğu görülürken, çözelti üfleme ile aplikasyon sonrasında su iticilik değerinin “4” olduğu görülmüş ve kumaşın su iticiliğinde artış meydana geldiği tespit edilmiştir. Emdirme işlemi sonrasında ise su iticilik değerinin “5” olduğu belirlenmiştir. Buna göre su iticilik bitim işlemleri sonrasında her iki yöntem ile su iticilik özellik kazandırılabilceği görülmüştür. Ancak çözelti üfleme işleminin performansı “1” derece düşük çıkmıştır. Bu düşüşün nedeninin ise alınan flotte oranının düşmesi ile birlikte su itici kimyasalında bulunan aktif madde oranının düşmesinden kaynaklandığı düşünülmektedir. Alınan flotte oranı çok düşük olduğunda (düşük aktif madde), hidrofobik etkinin özellikle yıkama sonrası hızla azalması söz konusu olurken; aşırı yüksek oranlarda, kumaşın hava geçirgenliği, bükülme dayanımı ve konfor parametrelerini olumsuz etkileyebilmektedir. Dolayısıyla, optimum aktif madde aktarım miktarının belirlenmesi, hem hedeflenen su iticilik seviyesini sağlamak hem de kumaşın kullanım performansını, tutumunu ve çevresel sürdürülebilirliğini korumak açısından önem arz etmektedir [22-24].

Numunelere ait patlama mukavemeti ve gramaj test sonuçları Çizelge 9’da verilmektedir.

Çizelge 9. Numunelere ait patlama mukavemeti ve gramaj test sonuçları

Kodlar	Gramaj (g/m ²)	Patlama mukavemeti (kPa)
Ref-1	144,10	157,3
HS-E-1	144,40	157,80
HS-ÇÜ-1	145,00	158,10
Ref-2	152,00	122,60
Sİ-E-2	153,00	122,90
Sİ-ÇÜ-2	152,30	123,00

Çizelge 9 incelendiğinde, bitim işlemleri sonrasında elde edilen gramaj ve patlama mukavemeti değerlerinde referans kumaşlara göre önemli bir değişiklik olmadığı görülmüştür.

Çizelge 5-9’da 6 numunenin sonuçları verilmiştir. Sonuçlar genel olarak incelendiğinde emdirmeye ve SBS tekniğiyle yapılan aplikasyonlar sonucu elde edilen numunelerin kendi aralarında önemli farklar olmadığı belirlenmiştir.

Böylece çalışmanın ilk aşaması başarıyla sonuçlanmıştır. Performans farkının olmamasının tespiti ardından, sürdürülebilirlik analizi yapılmıştır.

Bu nedenle her iki yöntem kimyasal ve enerji tüketimi bakımından kıyaslanmıştır.

Buna göre;

Emdirme aplikasyon yöntemi ile m² kumaş üzerine aktarılacak çözelti miktarı

- Kumaş m² gramajı: 200 g/m²
- Alınan flotte oranı (%): 90
- m² gramaj üzerine alınacak çözelti miktarı: 180 mL
- Makine hızı: 10 m/dk

- Fular kapasitesi: 100 L
- Kumaş Eni: 150 cm
 - 1 saatte aplikasyon yapılacak kumaş metresi: Makine hızı m/dk* 60 dk: 600 m/s
 - 1 saatte aplikasyon yapılacak kumaş m²: 600 m *1,5 m= 900 m² kumaş
 - 1 saatte kumaş üzerine aktarılacak toplam çözelti miktarı:
(900 m² *180 mL)/1000= 162 L + 100 L (İşlem sonrası fuların boşaltılması) = 262 L
 - Saatlik kumaş m² başına çözelti ihtiyacı: 262 L / 900 m² = 291 mL/m²

Çözelti Üfleme Sistemi ile m² kumaş üzerine çözelti ihtiyacı

- Makine hızı: 10 mL/dk
- Kumaş Eni: 1,50 m
- Besleme hızı: 1500 mL/sa
 - 1 saatte aplikasyon yapılacak kumaş metresi: Makine hızı* 60 dk: 600 m/s
 - 1 saatte aplikasyon yapılacak kumaş m²: 600 m*1,5 m= 900 m² kumaş
 - 1 saatte kumaş üzerine aktarılacak toplam çözelti miktarı: 1500 mL * 24 düze = 36 L
 - Saatlik kumaş m² başına çözelti ihtiyacı: 36.000 mL / 900 m²= 40 mL/m²

Yapılan kimyasal tüketimi hesapları incelendiğinde, emdirme aplikasyonu için 1 metre kare kumaş için 291 mL gerekiyken, çözelti üfleme tekniği ile uygulamada 40 mL gerekmektedir. Yapılan bu hesaba göre çözelti üfleme tekniği ile kimyasal tüketiminin emdirme yöntemine göre yaklaşık %86 oranında düştüğü görülmektedir.

Çizelge 10. Kurutma maliyeti hesabı için gerekli veriler

Parametreler	Emdirme aplikasyonu	Çözelti üfleme aplikasyonu
Kumaş gramaj (kg/m ²)	0,2	0,2
Kurutma öncesi alınan flotte oranı (%)	90	67
Kurutma sonrası alınan flotte oranı (%)	8,5	8,5

Bu iki yöntemin enerji tüketimi kıyası Çizelge 10'da verilen bilgiler kullanılarak ve aşağıda verilen formül ile yapıldığında [25-27];

$$E_{kuru} = k \times (PU_{\text{öncesi}} - PU_{\text{sonrası}}) \times m_{\text{kumaş}} \times h_{\text{evap}} \quad (1)$$

E_{kuru} : Kurutma için gereken enerji (kJ)

$PU_{\text{öncesi}}$: Uygulama öncesi alınan flotte oranı

$PU_{\text{sonrası}}$: Uygulama sonrası alınan flotte oranı

$m_{\text{kumaş}}$: Kumaş kütlesi (kg)

h_{evap} : Suyun buharlaşma entalpisi (2256 kJ/kg)

k : Sistem verimlilik faktörü (1 olarak alınmıştır)

Emdirme için;

$$E_{kuru-emdirme} = 1 \times (0,90 - 0,085) \times 0,2 \text{ kg} \times 2256 \text{ kJ/kg}$$

$$E_{kuru-emdirme} = 367,728 \text{ kJ}$$

Çözelti üfleme tekniği için;

$$E_{kuru-çözelti üfleme} = 1 \times (0,67 - 0,085) \times 0,2 \text{ kg} \times 2256 \text{ kJ/kg}$$

$$E_{kuru-çözelti üfleme} = 263,952 \text{ kJ}$$

Teorik kurutma enerjisi emdirme yöntemi için 367,728 kJ hesaplanırken, bu SBS tekniği için 263,952 kJ olarak hesaplanmıştır. Emdirme yöntemi uygulanmış kumaşlar için teorik kurutma enerjisi oranının çözelti üfleme yöntemine göre yaklaşık %39 daha fazla olduğu tespit edilmiştir.

4. SONUÇLAR

Çalışmada TÜBİTAK TEYDEB 1501 projesi kapsamında prototip olarak geliştirilen çözelti üfleme cihazını, işletmede çalışan bir ramöz kurutma makinesi önüne çeşitli revizeler yapılarak entegre edilmiştir. Böylece ramöz ile kesiksiz çalışan yeni bir aplikatör yapılmıştır. Daha sonra bu ramöz ile çalışarak iki farklı kumaşa, yumuşatıcı ve su iticilik bitim işlemleri uygulanmıştır. Bu kimyasallar aynı zamanda aynı ramözde, emdirme yöntemi ile de kumaşlara aktararak iki yöntemin kumaş performansı üzerine olan etkileri incelenmiştir. Uygulanan pilling, aşınma dayanımı, haslık, kılcal emme ve su iticilik testleri incelendiğinde, çözelti üfleme cihazının kumaşa aktarılmasıyla daha düşük alınan flotte oranı (%90 alınan flotte oranı değerinden - %67 alınan flotte oranı değerine) ile aynı kumaş performansına ulaşılabildiği görülmüştür. Buna göre sürdürülebilirliğin son derece önemli olduğu bu dönemde hem kimyasal hem su tüketimin hem de buna bağlı enerji tüketimin çözelti üfleme tekniği kullanılarak düşürülebileceği ispat edilmiştir. Çalışmada su tüketiminde %86, enerji tüketiminde ise %39 oranında düşüş tespit edilmiştir. Ayrıca az su tüketimine bağlı olarak aynı oranda çevreye daha az atık yük getirdiği için, bu yöntemin konvansiyonel yöntemle göre çok daha çevreci olduğu söylenebilmektedir. Böylece ürün bazlı karbon ayak izi ve su ayak izinde düşüşler elde edilebilecektir.

5. TEŞEKKÜR

Bu çalışma, 3191059 nolu “Çözelti Üfleme Tekniği Kullanılarak Ramöz Önü Aplikasyon Sistemi Tasarımı, Apre Proseslerinin Geliştirilmesi” başlıklı TÜBİTAK TEYDEB 1501 Sanayi Ar-Ge Projeleri Destekleme Programı kapsamında yürütülen proje çerçevesinde gerçekleştirilmiştir. Sağladığı maddi destek için Türkiye Bilimsel ve Teknolojik Araştırma Kurumu’na (TÜBİTAK) teşekkür ederiz. 3191059 nolu projede proje ortağı olarak yer alan Inovaktif Ar-Ge Danışmanlık Kimya Ozon Sistemleri Kozmetik San. Tic. Ltd. Şti.’ne proje süresince vermiş olduğu destekler için teşekkür ederiz.

6. KAYNAKLAR

1. Kanmaz, N., Kavurucu, B., Ekmen, E., Yaman, Ö., Yazan, S.Y. & Ünver, Ü. (2022). Türkiye’de endüstriyel su tüketimi ve arıtımı. *İleri Mühendislik Çalışmaları ve Teknolojileri Dergisi*, 3(1), 19-33.
2. Chen, X., Memon, H.A., Wang, Y., Marriam, I. & Tebyetekerwa, M. (2021). Circular economy and sustainability of the clothing and textile industry. *Materials Circular Economy*, 3(1), 1-9.
3. Uluçay, A., Ceyhan, G., Balci, O., Işık, C. & Yavuz, S. (2024). Investigation of dyestuff recycled from wastewater containing indigo/sulfur dyes. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 39(4), 969-978.
4. Basyigit, Z.O., Kut, D. & Hauser, P. (2020). Development of multifunctional cotton fabric via chemical foam application method. *Textile Research Journal*, 90(9-10), 991-1001.
5. Omerogullari Basyigit, Z. (2021). Application technologies for functional finishing of textile materials. *İçinde Textiles for Functional Applications* (1-21). IntechOpen.
6. Yavuz, S. & Balci, O. (2024). Pamuklu ipliklerin köpük aplikasyon metodu kullanılarak indigo boyarmadde ile boyanması dyeing of cotton yarns with indigo dyestuff using the foam application method abstract. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 39(2), 515-526.
7. Tapanyigit, E.B., Pektaş, K., Özdemir, M. & Balci, O. (2022). Farklı dokusuz yüzeyler üzerinde çözelti üfleme eğirme tekniği ile poliamid 6.6 esaslı nano lif eldesi ve araç kabin filtreleri için uygunluğunun araştırılması. *Tekstil ve Mühendis*, 29(127), 116-124.
8. Gao, Y., Zhang, J., Su, Y., Wang, H., Wang, X.X., Huang, L.P. & Long, Y.Z. (2021). Recent progress and challenges in solution blow spinning. *Materials Horizons*, 8(2), 426-446.
9. Silva, M.J., Dias, Y.J. & Yarin, A.L. (2023). Electrically-assisted supersonic solution blowing and solution blow spinning of fibrous materials from natural rubber extracted from *havea brasiliensis*. *Industrial Crops and Products*, 192(October 2022), 116101.
10. Medeiros, E.S., Glenn, G.M., Klamczynski, A.P., Orts, W.J. & Mattoso, L.H.C. (2009). Solution blow spinning: A new method to produce micro- and nanofibers from polymer solutions. *Journal of Applied Polymer Science*, 113(4), 2322-2330.

11. Sinha-Ray, S., Sinha-Ray, S., Yarin, A.L. & Pourdeyhimi, B. (2015). Application of solution-blown 20-50nm nanofibers in filtration of nanoparticles: The efficient van der waals collectors. *Journal of Membrane Science*, 485, 132-150.
12. Khan, M.K.R. & Hassan, M.N. (2021). Solution blow spinning (SBS): A promising spinning system for submicron/nanofibre production. *Textile & Leather Review*, 4(3), 181-200.
13. Liu, Y., Zhang, G., Zhuang, X., Li, S., Shi, L., Kang, W. & Xu, X. (2019). Solution blown nylon 6 nanofibrous membrane as scaffold for nanofiltration. *Polymers*, 11(2), 1-15.
14. Dirany, Z., González-Benito, J., Ginatta, P., Nguewa, P. & González-Gaitano, G. (2025). Solution blow spun polymeric nanofibres embedding cyclodextrin complexes of miltefosine: An approach to the production of sprayable dressings for the treatment of cutaneous leishmaniasis. *Carbohydrate Polymers*, 353(September 2024), 123173.
15. Rossi, P.F., dos Santos, D.M., Marangon, C.A., Teodoro, K.B.R., Costa, C.S., Inada, N.M. & Oréfice, R.L. (2025). Biobased composite fibrous membrane using PLA and lignin carbon dots fabricated via solution blow spinning for wound dressing application. *Materials Today Communications*, 42(November 2024), 111418.
16. Gao, Q., Zhu, S., Huang, Y., Wen, C., Zhang, Z., Yang, F. & Feng, Y. (2024). Fast sweat wicking enabled by unidirectional water transport in biodegradable trilayered porous membranes produced via papermaking. *ACS Sustainable Chemistry and Engineering*, 12(21), 8041-8050.
17. Zaman, M., Liu, H., Xiao, H., Chibante, F. & Ni, Y. (2013). Hydrophilic modification of polyester fabric by applying nanocrystalline cellulose containing surface finish. *Carbohydrate Polymers*, 91(2), 560-567.
18. Shi, W., Pei, L., Gu, X. & Wang, J. (2023). Investigation on decontamination mechanism of moisture-wicking fabrics during home laundry. *Journal of Surfactants and Detergents*, 26(5), 693-702.
19. Akinci, F.C., Kaynak, H.K. & Korkmaz, Y. (2018). Effects of filament fineness and weft sett on the permeability properties of microfilament woven fabrics. *Tekstil ve Muhendis*, 25(111), 234-240.
20. Mazloompour, M., Ansari, N. & Hemmatinejad, N. (2007). Wetting behaviour of raw and water-repellent cotton fabrics using wetting kinetic measurements. *Indian Journal of Fibre and Textile Research*, 32(1), 93-98.
21. Simile, C.B. (2004). Critical evaluation of wicking in performance fabrics. Georgia Institute of Technology, *Master of Science in the School of Polymer, Textile, and Fiber Engineering*.
22. Chowdhury, M.J. & Nasrin, S. (2017). Effect of performance finish on woven fabric properties. *International Journal of Scientific and Research Publications*, 7(7), 814.
23. Chowdhury, K.P., (2018). Effect of special finishes on the functional properties of cotton fabrics. *Journal of Textile Science and Technology*, 04(02), 49-66.
24. Ho, L.Y. & Kan, C.W. (2022). Effect of resin finishing on some properties of 100% cotton light weight woven fabric. *Coatings*, 12(11).
25. Cay, A., Tarakcioğlu, I. & Hepbasli, A. (2009). A study on the exergetic analysis of continuous textile dryers. *International Journal of Exergy*, 6(3), 422-439.
26. Raafi, S.M. & Fatema, U.K. (2021). Implementation of pre-heating system in stenters for improving machine performance and increasing efficiency. *Journal of Textile Science and Technology*, 07(04), 143-151.
27. Wei, Y., Gong, R.H., Ning, L. & Ding, X. (2018). Enhancing the energy efficiency of domestic dryer by drying process optimization. *Drying Technology*, 36(7), 790-803.

Akıllı Şebekelerde Elektrik Hırsızlığı Tespiti İçin Gelişmiş Makine Öğrenmesi Yöntemlerinin Uygulanması

Ömer Can TOLUN^{1,a}, Kasım ZOR^{1,b}, Önder TUTSOY^{1,c}

¹Adana Alparslan Türkeş Bilim ve Teknoloji Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, Adana, Türkiye

^aORCID: 0000-0002-1956-4303; ^bORCID: 0000-0001-6443-114X; ^cORCID: 0000-0001-6385-3025

Makale Bilgileri

Geliş : 25.08.2025

Kabul : 10.09.2025

DOI: 10.21605/cukurovaumfd.1772073

Sorumlu Yazar

Ömer Can TOLUN

octolun@atu.edu.tr

Anahtar Kelimeler

Bayes optimizasyon

Boyut indirgeme

Elektrik hırsızlığı tespiti

LightGBM

Atf şekli: TOLUN, Ö.C., ZOR, K., TUTSOY, Ö., (2025). Akıllı Şebekelerde Elektrik Hırsızlığı Tespiti İçin Gelişmiş Makine Öğrenmesi Yöntemlerinin Uygulanması. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 627-641.

ÖZ

Elektrik hırsızlığı, dağıtım sistemlerinde hem ekonomik kayıplara hem de şebeke güvenilirliğinin azalmasına yol açan en kritik sorunlardan biridir. Bu çalışmada, elektrik hırsızlığı tespiti (EHT) için veri dengesizliği ve model iyileştirme problemlerini aynı anda ele alan gelişmiş bir makine öğrenmesi algoritması sunulmaktadır. Öncelikle, meskenlerden elde edilen gerçek elektrik tüketim verilerine ön işleme adımları uygulandıktan sonra dengesizliği azaltmak için K-ortalamlar tabanlı bir dengeleme yöntemi uygulanmıştır. Boyut indirgeme aşamasında, modelin yalnızca en anlamlı girdilerle eğitilmesi amacıyla karşılıklı bilgi temelli en iyilerin seçilmesi yöntemi tercih edilmiştir. Sınıflandırma aşamasında hafif gradyan artırmalı makine (LightGBM) algoritması kullanılmış ve tahmin performansı Bayes optimizasyon yöntemi ile iyileştirilmiştir. Test verisi üzerinde yapılan değerlendirmelerde, önerilen yöntemin sırasıyla %89,1 doğruluk, %97,2 kesinlik, %80,6 duyarlılık ve %88,1 F1 skoru ile güçlü sonuçlar verdiği görülmüştür.

Application of Advanced Machine Learning Methods for Electricity Theft Detection in Smart Grids

Article Info

Received : 25.08.2025

Accepted : 10.09.2025

DOI: 10.21605/cukurovaumfd.1772073

Corresponding Author

Ömer Can TOLUN

octolun@atu.edu.tr

Keywords

Bayesian optimisation

Dimension reduction

Electricity theft detection

LightGBM

How to cite: TOLUN, Ö.C., ZOR, K., TUTSOY, Ö., (2025). Application of Advanced Machine Learning Methods for Electricity Theft Detection in Smart Grids. Çukurova University, Journal of the Faculty of Engineering, 40(3), 627-641.

ABSTRACT

Electricity theft is one of the most critical issues in distribution systems causing both economic losses and reduced network reliability. In this study, an advanced machine learning framework is proposed for electricity theft detection along with addressing data imbalance and model improvement problems simultaneously. Firstly, after applying preprocessing steps to the actual electricity consumption data obtained from households, a K-means-based balancing method was employed to reduce data imbalance. In the dimension reduction stage, the SelectKBest method based on mutual information was preferred to ensure that the model was trained with only the most significant inputs. In the classification phase, the light gradient boosting machine (LightGBM) algorithm was employed and the prediction performance was enhanced using the Bayesian optimisation method. Evaluations on the test dataset demonstrated that the proposed method achieved strong results with 89.1% accuracy, 97.2% precision, 80.6% recall, and 88.1% F1-score, respectively.

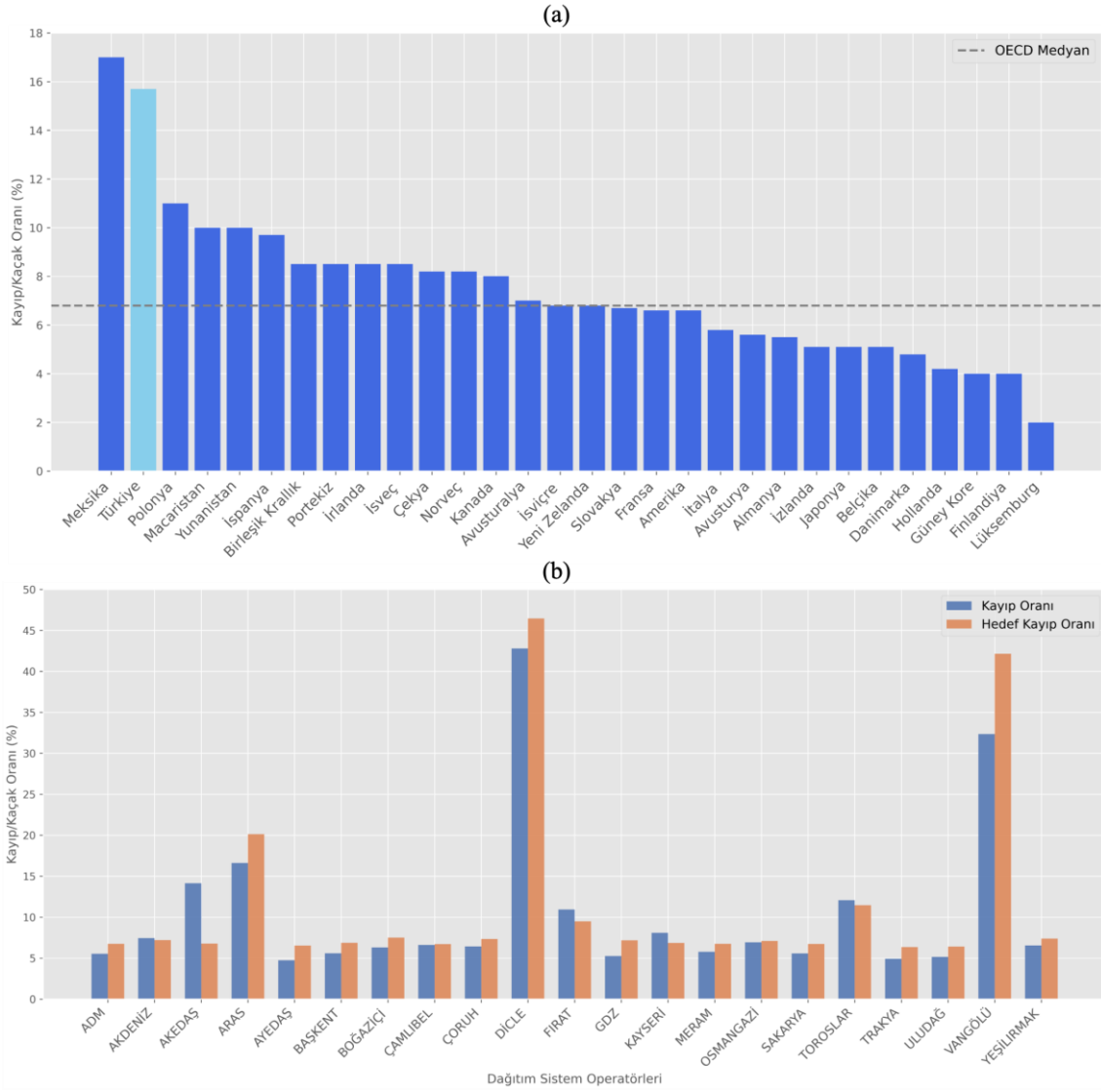
1. GİRİŞ

Akıllı şebekelerin gelişimi, enerji ağlarının işletimini ve yönetimini köklü bir şekilde değiştirmiştir. Geleneksel güç sistemlerinin dijitalleşmesi, veriye dayalı enerji yönetimini kolaylaştırmakta ve enerji arz güvenliği ile sistem istikrarı gibi temel bileşenlerin etkin şekilde sağlanmasını geliştirmektedir. Ancak bu dönüşüm süreci, akıllı sayaçlardan ve dijital enerji yönetim sistemlerinden elde edilen verilerin güvenliğini tehlikeye atan yeni sorunlar da doğurmuştur. Bu geçişte öne çıkan önemli bir gelişme, elektrik üretim, iletim, dağıtım ve tüketim süreçlerini dönüştüren ileri ölçüm altyapısının (AMI) yaygın olarak uygulanmasıdır. AMI'nin temel bir bileşeni olan akıllı sayaçlar, kurcalamaya ve müdahaleye karşı oldukça savunmasız olan geleneksel elektromekanik sayaçların yerini giderek daha fazla almaktadır [1]. Akıllı şebekelerin hızlı benimsenmesine ve önümüzdeki on yıl içinde 50 milyar doları aşması beklenen bu pazarın büyümesine rağmen, elektrik sistemleri hâlâ önemli operasyonel kayıplar yaşamaktadır.

Bu kayıplar genellikle iki ana grupta sınıflandırılır: teknik kayıplar ve teknik olmayan kayıplar. Teknik kayıplar, elektrik iletimi ve dağıtım sırasında doğal olarak ortaya çıkar; iletkenlerde ve trafolar gibi ekipmanlarda enerji kaybı şeklinde görülür. Buna karşılık teknik olmayan kayıplar büyük ölçüde elektrik hırsızlığından kaynaklanır. Bu durum, bölgesel, altyapısal ve teknolojik faktörlere bağlı olarak farklı şekillerde ortaya çıkan, küresel ölçekte yaygın bir sorundur. Akıllı şebeke uygulamaları hızla artarken, özellikle sayaç altyapısına yetkisiz erişim veya kötü niyetli müdahalelerden kaynaklanan teknik olmayan kayıpların azaltılması hem dağıtım şirketleri hem de düzenleyici kurumlar için kritik bir öncelik haline gelmiştir [2]. Elektrik hırsızlığı, dağıtım sistemlerinin ekonomik sürdürülebilirliğini zora sokmakta, dağıtım sistemi operatörlerine (DSO) ve tüketicilere ciddi mali kayıplar yaşatmaktadır.

Elektrik hırsızlığının temel nedenleri arasında sayaç arızaları, sayaca müdahale ile tüketimin eksik veya yanlış ölçülmesi, dağıtım hattına doğrudan bağlantılar ve tüketim-faturalama uyumsuzlukları yer almaktadır [3]. Elektrik hırsızlığının görülme sıklığı ülkelere ve bölgelere göre değişiklik göstermekte, bu durum hem DSO'lara hem de ulusal ekonomilere zarar vermektedir. Northeast Group'un 2021 tarihli araştırmasına göre, hırsızlık ve kayıpların küresel maliyeti yılda 101,2 milyar doları aşmaktadır [4]. Türkiye'nin üyesi olduğu Ekonomik İş Birliği ve Kalkınma Örgütü (OECD) içerisinde kayıp/kaçak tahminlerindeki sıralaması Şekil 1a'da gösterildiği gibi 38 üye ülkenin ortalamasından daha yüksektir [5]. Bununla birlikte, 2004 yılında gerçekleştirilen ilk özelleştirmeler öncesinde, elektriğin dağıtımını tek bir çatıdan gerçekleştirilmekteydi. 2013 yılına kadar devam eden bu süreç sonucunda Türkiye Elektrik Dağıtım Anonim Şirketi'nin (TEDAŞ) özelleştirilmesi ile Türkiye 21 farklı dağıtım bölgesine ayrılmıştır. Türkiye elektrik güç sistemi incelendiğinde, 2023 yılı Enerji Piyasası Düzenleme Kurumu (EPDK) raporuna göre DSO'ların fiili ve hedef kayıp oranları Şekil 1b'de gösterilmektedir. Şekil 1b'ye göre en büyük üç kayıp oranının %42,79 ile Dicle bölgesinde, %32,35 ile Vangölü bölgesinde ve %16,61 ile Aras bölgesinde olduğu görülmektedir [6].

Elektrik hırsızlığı sadece finansal kayıplara yol açmakla kalmamakta, aynı zamanda şebekenin güvenilirliğini ve istikrarını da olumsuz etkilemektedir. Literatürde EHT için önerilen yöntemler; sosyal ve ekonomik çalışmalara dayalı teorik yaklaşımlar, maliyetli ve entegrasyonu zor donanım tabanlı yöntemler ve ek donanım gerektirmeyen, mevcut altyapıya kolayca uygulanabilen veri odaklı yöntemler olmak üzere üç grupta ele alınmaktadır [7]. Bu veri odaklı yöntemler genellikle makine öğrenmesi tekniklerini içerir ve gelişmiş örüntü tanıma ile tahmin yetenekleri sayesinde yenilikçi çözümler sunar. Enerji internetinin gelişimi, DSO'lara ait akıllı sayaçlardan, kablosuz sensör ağlarından ve diğer izleme cihazlarından elde edilen büyük veri kümelerinin makine öğrenmesi teknikleri ile analiz edilmesini mümkün kılmıştır [8]. Böylece tüketim profillerindeki karmaşık anormallikler tespit edilebilmekte ve müşteriler dürüst veya dürüst olmayan şeklinde sınıflandırılabilir.



Şekil 1. (a) OECD üyesi ülkelerdeki kayıp/kaçak oranları, (b) DSO'ların 2023 yılında gerçekleşen ve hedeflenen kayıp oranları.

Veri tabanlı yöntemler, sürekli öğrenme ve uyum sağlama özellikleri sayesinde doğruluklarını zamanla artırabilir ve farklı dağıtım bölgelerindeki kayıpların tespitine uygun hale getirilebilmektedir [9]. Son dönem çalışmalarda, denetimli, denetimsiz ve pekiştirmeli, derin, topluluk ve melez yöntemler elektrik hırsızlığının tespitinde kullanıldığı görülmüştür. Ancak mevcut araştırmaların büyük bir kısmı, elektrik hırsızlığına dair gerçek veri setlerinin kullanımının eksikliği nedeniyle sınırlıdır. Literatürde, gerçek hırsızlık vakalarının nadirliği, dengesiz veri dağılımı ve etiketleme zorlukları, veri odaklı modellerin başarısını olumsuz etkilemektedir [10]. Bu çalışmanın özgün katkıları aşağıdaki şekilde özetlenebilir:

- Veri dengesizliği problemini çözmek amacıyla K-ortalamlar tabanlı bir dengeleme yöntemi önerilmiştir.
- Modelin yalnızca en anlamlı değişkenlerle eğitilmesi için karşılıklı bilgi temelli en iyilerin seçilmesi yöntemi uygulanmış, böylece hem öğrenme kapasitesi hem de yorumlanabilirliği geliştirilmiştir.
- Sınıflandırma aşamasında LightGBM algoritması kullanılmış ve hiperparametreler Bayes optimizasyon ile ayarlanarak yüksek ve kararlı performans elde edilmiştir.

Giriş bölümünde literatüre yönelik kısa bir değerlendirmeye yer verilmesinin ardından makale şu şekilde düzenlenmiştir: Literatür araştırması bölümünde, EHT'ye ilişkin mevcut çalışmaların güncel durumu

ortaya konulmakta ve kuramsal çerçeve sunulmaktadır. Materyal ve metot bölümünde, kullanılan veri seti tanıtılmakta ve temel özellikleri açıklanmakta; ayrıca önerilen yaklaşımın metodolojisini gösteren bir akış diyagramı sunulmakta ve kullanılan yöntemler ayrıntılı biçimde açıklanmaktadır. Bulgular ve tartışma bölümünde, uygulanan yöntemlerden elde edilen sonuçlar analiz edilmekte ve literatürle karşılaştırmalı olarak tartışılmaktadır. Son olarak, sonuçlar bölümünde çalışmanın temel çıktıları özetlenmekte ve gelecek araştırmalara yönelik katkılar ortaya konulmaktadır.

2. LİTERATÜR ARAŞTIRMASI

Bu bölümde, literatürdeki dağıtım sistemlerinde EHT çalışmaları sunulmaktadır. Bu çalışmalar öğrenme yöntemlerine göre gözetimli, gözetimsiz, pekiştirmeli, derin, topluluk ve melez olarak ayrı alt başlıklar altında incelenmiştir. Aşağıdaki alt bölümlerde, EHT’de uygulanan her bir öğrenme tekniğine ilişkin ayrıntılı bir literatür incelemesi sunulmakta; bu tekniklerin metodolojileri, etkinlikleri ve pratik uygulamaları analiz edilmektedir. Bu çalışma özellikle son üç yılda gerçekleştirilen araştırmalara odaklanmakta ve makine öğrenmesi tabanlı alanındaki güncel gelişmeleri ve ortaya çıkan eğilimleri vurgulamaktadır. Ayrıca, Çizelge 1 incelemeye dâhil edilen çalışmalara ilişkin kapsamlı bir özet sunmaktadır.

Çizelge 1. Güncel denetimli, denetimsiz, pekiştirmeli ve derin öğrenme, topluluk, melez tabanlı EHT yaklaşımlarına genel bakış

Referans	Teknikler	Alt kategori	Güçlü yönler	Sınırlılıklar
[11]	ADASYN-SGWO-PRF	Gözetimli	Sınıf dengesizliğini melez bir aşırı örnekleme-az örnekleme yaklaşımıyla etkili bir şekilde yönetir; PRF hattı ile yüksek doğruluk sağlar.	Aşırı örnekleme nedeniyle gürültü ve sınıf örtüşmesi oluşabilir; model performansı ön işleme aşamalarının dikkatli ayarlanmasına bağlıdır.
[12]	DGRGNN	Gözetimli	Dinamik grafik ağırları kullanarak karmaşık zamansal ve yapısal ilişkileri yakalar; gerçek zamanlı topoloji güncellemelerini entegre eder.	KNN grafik yapısında k-değerine duyarlıdır; yük eğrilerinden derin özellikler çıkarmakta zorlanır.
[13]	CLOF	Gözetimsiz	LOF ile kümeleri birleştirerek örtüşen aykırı değerleri tespit eder; geniş bir sahtecilik davranışı yelpazesini belirlemede etkilidir; dengesiz veri kümelerine karşı dayanıklıdır.	LOF örtüşen aykırı değerlerle zorlanır; gerçek tüketime çok benzeyen küçük saldırılarda performans düşebilir; küme yapılandırmasına bağlıdır.
[14]	DWT-FCM	Gözetimsiz	Güvenli gözlemci sayaç verilerini ve dalgacık tabanlı özellik çıkarımını entegre eder; farklı kullanıcı türleri ve saldırı senaryoları arasında üstün performans ve düşük yanlış alarm sağlar.	Performans, gözlemci sayaçların kalitesi ve yapılandırmasından etkilenebilir; gerçek kurcalanmış veri üzerinde test edilmemiştir; hesaplama verimliliğine dair sınırlı bilgi vardır.
[15]	DQN	Pekiştirmeli	Ön işleme olmadan veri dengesizliğini giderir; özellik-maliyet verimliliğini entegre eder.	Gerçek hırsızlık verisi eksikliği nedeniyle sentetik verilere dayanır; optimum özellik seçimi ve yakınsama için kapsamlı benzetim gerektirir.
[16]	DQN	Pekiştirmeli	Tüketim kalıplarındaki değişikliklere ve yeni saldırılara uyum sağlar; keşif ve sömürü mekanizmaları sayesinde tam yeniden eğitmeden kaçınır.	Model karmaşıklığı nedeniyle hesaplama yükü fazladır; yüksek dinamik ortamlarda aşırı öğrenmeye yatkındır, büyük veri kümeleri gerektirir.
[17]	CNN	Derin	Özellik mühendisliği ile yorumlanabilirliği artırır; dikkat mekanizmasına dayalı çok ölçekli CNN kullanarak anomali tespiti ve performansı geliştirir.	Orantılı veya ters manipülasyonları tespit etmekte zorlanabilir; özellik çıkarım kalitesine duyarlıdır.
[18]	D-FFNN	Derin	Kötü niyetli ve iyi niyetli anomalileri ayırt eder; gerçek zamanlı öz öğrenme ve hava durumu gibi dış faktörleri entegre eder.	Yüksek hesaplama gereksinimleri ve yeni şebeke teknolojilerine sınırlı uyum; yanlış sınıflandırma riski.
[19]	ChebGCN	Derin	Çoklu enerji yükleri arasındaki bağımlılıkları yakalar; elektrik, ısıtma, soğutma ve su sistemleri arasındaki korelasyonları modeller; ChebGCN kullanır.	Sınırlı kapsam; konut dışı ortamlar için genelleme zayıf olabilir; meteorolojik veya voltaj verileri gibi ek girdiler dahil edilmemiştir.

Çizelge 1. Devamı

Referans	Teknikler	Alt kategori	Güçlü yönler	Sınırlılıklar
[20]	XGBoost	Topluluk	SCADA verilerini XGBoost ile entegre ederek birden fazla sentetik saldırı tipini yüksek doğrulukla tespit eder.	Değerlendirme, simülasyon tabanlı senaryolarla sınırlıdır; gerçek zamanlı uygulama tartışması eksiktir.
[21]	XGBoost	Topluluk	Sağlam özellik çıkarımı ve otomatik optimizasyon için AutoXGB ile OS-CNN önermektedir.	Karmaşık mimari ve aşırı uyum riskinden dolayı zarar görür; hiperparametre ayarında potansiyel aşırıya kaçma yaşanabilir.
[22]	XGBoost	Topluluk	Düşük örnekleme oranına sahip geleneksel ölçüm verilerine melez makine öğrenmesi uygular; az araştırılmış bir bağlamı ele alır.	Veri seti kısıtlamaları ve düşük frekanslı okumalar üzerine odaklanma nedeniyle genelleştirilebilirliği sınırlıdır.
[23]	CNN-LSTM	Melez	Dengesiz ve anormal verilerde performansı artırmak için topluluk öğrenmesi ve prototip öğrenmeyi birleştirir.	Prototip temsiliyetine bağlıdır; yüksek veri gürültüsü veya kaymalarında daha az etkili olabilir.
[24]	CNN-AdaBoost	Melez	Melez CNN-AdaBoost modeli önerir ve ADGM'yi tanıtarak tabular verilerle EHT modeli dayanıklılığını kaçış saldırılarına karşı etkili bir şekilde değerlendirir.	Melez model nedeniyle yüksek hesaplama maliyeti içerir; dürüst olmayan veri tek bir veri seti kullanılarak üretilmiştir, bu da model genelleştirilebilirliğinin değerlendirilmesini sınırlar.
[25]	LSTM-Bi-LSTM-GRU-Bi-GRU-SVM-RF-DT-AdaBoost	Melez	Dengesiz verilerde performansı artırmak için sağlam yeniden örnekleme teknikleriyle birlikte yığılmış topluluk öğrenmesini kullanır.	Modelin karmaşıklığı ve birden fazla temel/meta-sınıflandırıcıya olan bağımlılığı, artan hesaplama maliyetine yol açabilir.

2.1. Gözetimli Öğrenme

Literatürde EHT kapsamında farklı denetimli öğrenme yaklaşımlarının kullanıldığı görülmektedir. Tripathi ve arkadaşları, aşırı örnekleme, özellik seçimi ve sınıflandırma adımlarını bir arada ele alan uyarlanabilir sentetik örnekleme yaklaşımı (ADASYN), alt küme gri kurt optimizasyonu (SGWO) ve boru hattı rastgele orman sınıflandırıcısı (PRF) tabanlı bir makine öğrenmesi yöntemi geliştirmiştir [11]. Zhuang ve arkadaşları, düğüm bilgisi toplama doğruluğunu artırmak amacıyla dinamik artık grafik sinir ağı (DGRGNN) modelini önermiştir [12].

2.2. Gözetimsiz Öğrenme

Literatürde veri odaklı yöntemlerin bir alt kümesi olan denetimsiz öğrenme teknikleri de EHT'de kullanılmaktadır. Bu teknikler, etiketli verilere ihtiyaç duymadan veri kümesini benzerlik ve farklılıklara göre gruplara ayırmaktadır. Bu bağlamda tüketiciler, yük profilleri üzerinden dürüst, kötü niyetli veya hileli olarak sınıflandırılabilen ve anormal tüketim davranışları tespit edilebilmektedir. Peng ve arkadaşları, k-ortalama ve yerel aykırı faktör (LOF) yöntemlerini birleştiren korelasyon tabanlı yerel aykırı faktör (CLOF) yaklaşımını önermiştir [13]. Bunun yanı sıra Qi ve arkadaşları, tüketici yük profilleri ile gözlemci sayaç profillerini analiz ederek hileli kullanıcıları tespit etmeyi amaçlayan, ayrık dalgacık dönüşümü (DWT) ve bulanık c-ortalama kümeleme (FCM) tabanlı yeni bir denetimsiz öğrenme yöntemini geliştirmiştir [14].

2.3. Pekiştirmeli Öğrenme

Pekiştirmeli öğrenme, bir ajanın çevresiyle etkileşime girerek ödül ve ceza mekanizması üzerinden karar verme stratejileri öğrenmesine dayalıdır. Elektrik hırsızlığı tespitinde ise tüketim anomalilerinin uyarlanabilir biçimde izlenmesi ve tespit stratejilerinin geliştirilmesi amacıyla kullanılmaktadır. Literatürde, Lee ve arkadaşları güç şebekelerinde teknik olmayan kayıpların tespiti için derin Q-öğrenme tabanlı bir yöntem geliştirmiş ve bu sayede veri dengesizliği problemini aşarak veri çoğaltmaya gerek kalmadan hırsızlık tespitini mümkün kılmıştır [15]. El-Thoukhy ve arkadaşları ise dinamik tüketim kalıpları ve siber saldırıların tespitine yönelik derin pekiştirmeli öğrenme (DRL) tabanlı bir yaklaşım önermiştir. Bu yöntem, yeni tüketim alışkanlıklarını ve saldırı türlerini öğrenebilme kapasitesiyle öne çıkmaktadır [16].

2.4. Derin Öğrenme

Derin öğrenme, makine öğrenmesinin bir alt dalı olup çok katmanlı yapay sinir ağlarını kullanarak büyük ve karmaşık veri setlerinde hiyerarşik temsil öğrenmeyi amaçlamaktadır. İnsan beyninin işleyişinden esinlenen bu yöntem, özellikle çok boyutlu ve yoğun verilerde yüksek başarı göstermektedir. Literatürde, Zhang ve Dai, özellik mühendisliği tabanlı birçok ölçekli anomali tespit modeli önermiştir. Bu modelde özellik çıkarımı için özellik mühendisliği, özellik seçimi için aşırı gradyan artırma (XGBoost) algoritması ve anomali tespiti için dikkat mekanizmasıyla desteklenen çok ölçekli evrimsel sinir ağı (CNN) yapısı kullanılmıştır [17]. Althobaiti ve arkadaşları ise çift katmanlı derin ileri beslemeli sinir ağı (D-FFNN) tabanlı uyarlanabilir bir sistem geliştirmiş ve bu sistem, enerji hırsızlığı ile yanlış yapılandırılmaları ayırt etmede yüksek doğruluk sağlamıştır [18]. Liao ve arkadaşları ise çoklu enerji yüklerinden gizli özellikleri ve korelasyonları yakalamak amacıyla Chebyshev grafik evrimsel ağ (ChebGCN) tabanlı bir yöntem önermiştir [19].

2.5. Topluluk Öğrenmesi

Topluluk öğrenmesi, birden fazla bireysel modeli bir araya getirerek tahmin doğruluğu ve sağlamlığı artırmayı hedefleyen bir makine öğrenmesi yaklaşımıdır. Bu yöntemin temel mantığı, farklı modellerin çıktılarının birleştirilmesiyle tek bir modelin tek başına ulaşabileceğinden daha iyi bir performans elde etmek; aşırı uyum, varyans ve yanlışlık risklerini azaltmaktır. Literatürde, Sharma ve Tiwari, denetim kontrol ve veri toplama (SCADA) sistemiyle entegre edilmiş optimize edilmiş bir XGBoost sınıflandırıcı önermiş ve akıllı şebekelerde anomali tespitinde daha yüksek doğruluk elde etmiştir [20]. Zhu ve arkadaşları, sentetik azınlık aşırı örnekleme tekniği ile düzeltilmiş en yakın komşular (SMOTEENN) tabanlı melez bir örnekleme yöntemi ile birlikte min–maks ölçekleme ve sıfır–skor standartlaştırma gibi ön işleme tekniklerini kullanarak çok ölçekli CNN ve Otomatik XGBoost (AutoXGB) modellerini entegre ederek elektrik hırsızlığı tespiti için bir yaklaşım geliştirmiştir [21]. Norouzi ve arkadaşları ise geleneksel sayaçların kullanıldığı bölgelerde, düşük örnekleme hızına ve eşzamanlı olmayan veriye ilişkin zorlukları gidermek amacıyla aylık veri dönüşümü, sentetik veri dengeleme yöntemleri ve XGBoost sınıflandırıcı içeren melez bir çerçeve önermiştir [22].

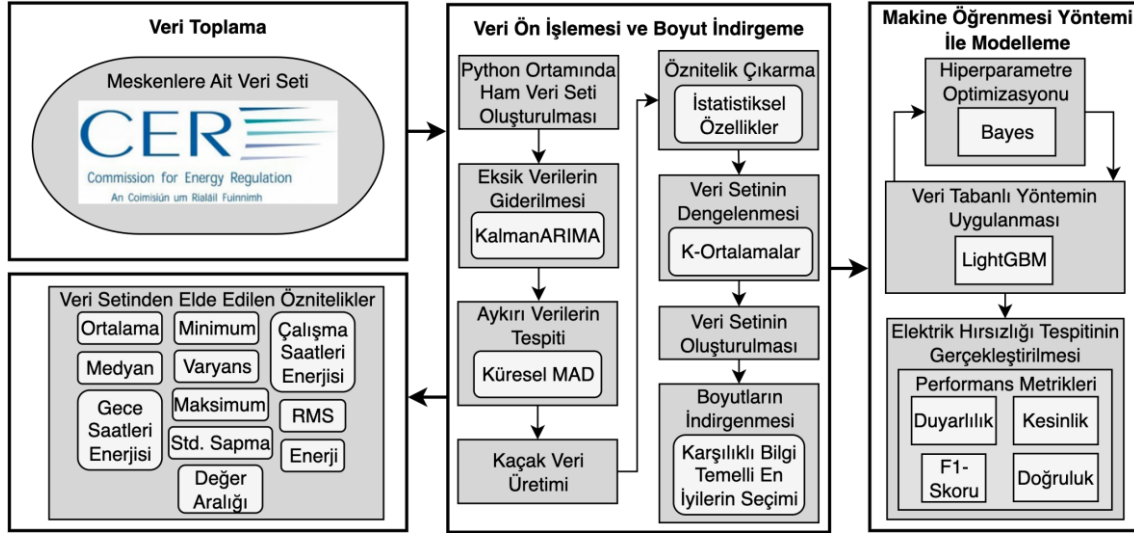
2.6. Melez Öğrenme

Melez Öğrenme, farklı öğrenme paradigmalarını, modellerini veya tekniklerini bir araya getirerek her birinin güçlü yönlerinden faydalanmayı ve genel performansı artırmayı hedefleyen bir makine öğrenmesi yaklaşımıdır. Bu yöntem, bireysel modellerin sınırlılıklarını aşarak daha iyi genelleme, verimlilik ve uyarlanabilirlik sağlamayı amaçlamaktadır. Literatürde, Sun ve arkadaşları, CNN ve uzun kısa süreli bellek ağı (LSTM) tabanlı özellik çıkarımı ile prototip tabanlı sınıflandırmayı birleştiren melez bir EHT çerçevesi önermiştir [23]. Nirmal ve arkadaşları, CNN ile uyarlamalı artırma (AdaBoost) tabanlı bir çerçeve geliştirmiş ve modeli hızlı gradyan işareti yöntemi (FGSM) ve düşman veri üretim yöntemi (ADGM) kullanarak düşman saldırıları altında test etmiştir [24]. Ullah ve arkadaşları ise birden fazla makine öğrenmesi modelini birleştirerek yığılmış genelleştirme (stacked generalisation) yaklaşımına dayalı melez bir çerçeve sunmuştur [25].

3. MATERYAL VE METOT

Elektrik dağıtım sistemlerinde EHT için önerilen metodoloji, Şekil 2’de gösterildiği üzere çok aşamalı ve organize bir süreçten oluşmaktadır. Süreç, İrlanda Enerji Düzenleme Komisyonu (CER) veri seti kullanılarak Python ortamında ham veri setinin oluşturulması ile başlamaktadır. Bu aşamayı takiben, veri setini zenginleştirmek amacıyla dürüst olmayan müşterilere ait veri üretimi gerçekleştirilmekte, ardından veri ön işleme adımlarında eksik verilerin giderilmesi için zaman serisi durum-uzay modellerine dayalı Kalman ARIMA filtreleme (KA), aykırı verilerin tespiti için modifiye Z-skoru tabanlı küresel medyan mutlak sapma (Küresel MAD) işlemleri gerçekleştirilmektedir. Bu aşamadan sonra, veri setinin dengelenmesi K-ortalama yöntemini kullanarak gerçekleştirilmekte, böylece sınıf dağılımları iyileştirilmekte ve nihai veri seti oluşturulmaktadır. Veri setinden elde edilen öznitelikler, ortalama, medyan, varyans, standart sapma, minimum, maksimum ve değer aralığı gibi istatistiksel ölçüler ile çalışma

saatleri enerjisi, gece saatleri enerjisi, toplam enerji ve etkin değer (RMS) gibi enerji temelli özelliklerden oluşmaktadır [26]. Öznitelik çıkarımı sonrası, boyutların indirgenmesi amacıyla karşılıklı bilgi temelli en iyilerin seçimi gerçekleştirilmekte, ardından Bayes yöntem kullanılarak hiperparametre optimizasyonu uygulanmaktadır. Son aşamada ise, LightGBM tabanlı yöntem ile elektrik hırsızlığı tespiti yapılmakta ve elde edilen sonuçlar duyarlılık, kesinlik, F1-skoru ve doğruluk gibi performans metrikleri ile değerlendirilmektedir.



Şekil 2. EHT için önerilen metodoloji.

3.1. Materyal

Bu bölümde, literatürden elde edilen dürüst veri seti ve dürüst veriden elde edilen dürüst olmayan veriler sunulmaktadır.

3.1.1. Dürüst Veri Seti

Bu makalede, CER kapsamında elde edilen enerji tüketim kayıtları kullanılmıştır. Bu veri setinde yarım saatlik aralıklarla 5.000'den fazla tüketiciye ait veri noktası bulunmaktadır. Kullanıcı sayısının fazlalığı nedeniyle bu çalışma, araştırmalar için önemli bir kaynak oluşturmaktadır [27]. Veri seti 535 günlük bir zaman periyodunu kapsamaktadır ve örnekleme aralığı 30 dakikadır. Basitlik adına, tüm 2009 ve 2010 yıllarına ait 1.000 müşteri dikkate alınmıştır.

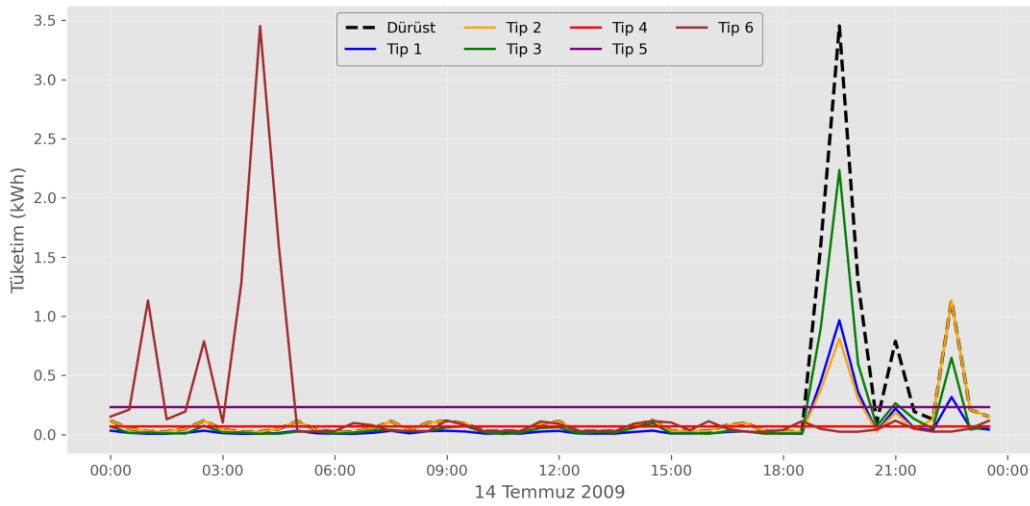
3.1.2. Dürüst Olmayan Veri Seti

Bu araştırma boyunca, dürüst olmayan müşteri davranışlarına ilişkin veri toplama sürecinde sürekli bir zorluk ortaya çıkmaktadır. Çizelge 2'de gösterildiği gibi altı farklı saldırı fonksiyonu kullanılarak, yapay şekilde yeni örnekler üretilmiştir. Tip 1 ve Tip 3 siber saldırı fonksiyonları aynı zamanda kısmi azaltma saldırıları olarak da adlandırılmaktadır [28]. Fonksiyon $y_1(x(t))$ için, enerji tüketim değeri rastgele seçilen bir kesir (α) ile azaltılmaktadır; burada α , 0.1 ile 0.8 aralığından rastgele seçilmektedir. Bu işlem, bildirilen elektrik tüketim değerini düşürmek amacıyla yapılmaktadır. Fonksiyon $y_3(x(t))$ 'de, α zamanla değişkenlik göstermektedir. Tip 2, enerji hırsızlığı ve siber saldırının en yaygın yöntemlerinden biridir. Belirli bir zaman aralığı için fonksiyon $y_2(x(t))$ devreye girer ve geri kalan değerler orijinal şekilde yansıtılır. Bu saldırı genellikle seçici bypass saldırısı olarak adlandırılır. Tip 5'te, fonksiyon $y_5(x(t))$ verilerin ortalamasını alarak enerji tüketim profilini sabit tutar. $y_4(x(t))$ fonksiyonunda ise kullanıcılar, ortalama alındıktan sonra verileri α azaltma katsayısı ile çarpar; böylece enerji profilinde belirgin bir azalma görülebilir. Son olarak $y_6(x(t))$ fonksiyonunda, enerji tüketim profili tersine çevrilerek maksimum kâr

elde edilmektedir [29]. Zaman serisi üzerinde altı farklı saldırı türüne ilişkin örnekler Şekil 3'te gösterilmiştir.

Çizelge 2. Elektrik hırsızlığı tiplerinin matematiksel gösterimi.

Hırsızlık tipleri	Matematiksel gösterimi
Tip 1	$y_1(x(t)) = \alpha x(t)$, $\alpha = \text{rastgele}(0.1, 0.8)$
Tip 2	$y_2(x(t)) = \beta x(t)$, $\beta = \begin{cases} 0 & T_{\text{başlangıç}} < t < T_{\text{bitiş}} \\ 1 & \text{aksi takdirde} \end{cases}$, $T_{\text{başlangıç}} = \text{rastgele}(0, 47 - \text{zaman})$ $\text{süre} = \text{rastgele}(\text{zaman}, 48)$, $T_{\text{bitiş}} = T_{\text{başlangıç}} + \text{süre}$
Tip 3	$y_3(x(t)) = \lambda(t)x(t)$, $\lambda(t) = \text{rastgele}(0.1, 0.8)$
Tip 4	$y_4(x(t)) = \lambda(t)\text{ortalama}(x(t))$, $\lambda(t) = \text{rastgele}(0.1, 0.8)$
Tip 5	$y_5(x(t)) = \text{ortalama}(x_t)$
Tip 6	$y_6(x(t)) = x_{48-t}$



Şekil 3. Farklı tip hırsızlık senaryoları için elektrik tüketim profillerinin karşılaştırılması

3.2. Metot

Bu bölümde, EHT için önerilen metodolojideki yöntemler sunulmaktadır.

3.2.1. KalmanARIMA

ARIMA sınıfı modeller, uygun bir dönüşümle durum-uzay biçimine yazılabildiğinden parametre kestirimi, filtreleme, yumuşatma ve öngörü adımları Kalman filtresi ile birleştirilebilir. Bu yaklaşım literatürde sıklıkla KA olarak anılır ve özellikle eksik gözlemler, düzensiz örnekleme aralıkları veya gürültülü ölçümler içeren uygulamalarda esneklik sağlar. Matematiksel olarak, Kalman filtreleri aşağıdaki denklemlerle verilen durum-uzay modellerine dayalı olarak iki aşamada uygulanır [30].

$$x_t = F_t x_{t-1} + \varepsilon_t \quad (1)$$

$$y(t) = H_t x_t + \omega_t \quad (2)$$

Burada x_t , t anındaki sistemin durum vektörünü, y_t ise aynı anda elde edilen ölçüm vektörünü temsil etmektedir. F_t , sistemin durum geçiş parametresi, ε_t rastgele durum gürültü terimi, H_t ölçüm parametresi ve ω_t ölçüm hatası terimidir.

3.2.2. Küresel MAD

Modifiye Z-skoru tabanlı küresel medyan mutlak sapma yöntemi, aykırı değer tespitinde medyana dayalı dayanıklı bir ölçek kestiricisi kullanır ve tüm zaman serisi/özellik boyunca tek bir küresel eşik ile çalışır. Serinin medyanı m ve medyan mutlak sapması

$$MAD = median(|x_t - m|) \quad (3)$$

ile hesaplanır. Normal dağılım altında $MAD \approx 0.6745\sigma$ olduğundan, modifiye Z-skoru

$$M_t = 0.6745 \frac{x_t - m}{MAD} \quad (4)$$

şeklinde tanımlanır [31]. Gözlem $x_t, |M_t| > \tau$ ise aykırı sayılır, literatürde $\tau = 3.5$ sık kullanılan bir eşiktir. Uygulama adımları ele alındığında, (i) Her değişken için m ve MAD küresel olarak hesaplanır. (ii) M_t değerleri bulunur ve $|M_t| > \tau$ olanlar küresel aykırı olarak etiketlenir. (iii) Zaman serilerinde trend/mevsimsellik baskınsa yöntem, ham veri yerine farklar ya da ayrıştırma artıklarına uygulanır. (iv) $MAD = 0$ durumunda tespit devre dışı bırakılır veya küçük bir ε ile düzenleme yapılır. Çok değişkenli senaryolarda tespit değişken bazında yürütülür ve bayrakların birleşimi alınır [32].

3.2.3. K-Ortalamalar

Sınıf dengesizliğini giderirken örnek uzayının yapısını korumak amacıyla özellik uzayı k kümeye bölünerek her sınıf içinde küme-tabanlı örnekleme yapılır. K-ortalamalar, amaç fonksiyonunu en aza indirerek her küme için bir merkez μ_c üretir [33].

$$\min_{\{\mu_c\}, \{S_c\}} \sum_{c=1}^k \sum_{x_i \in S_c} \|x_i - \mu_c\|^2 \quad (5)$$

Her sınıf $y \in \{0,1\}$ için küme boyları $n_c^{(y)}$ ve toplamı $N^{(y)}$ olmak üzere, sınıf içi dağılımı korumak adına hedef küme payları

$$\pi_c^{(y)} = \frac{n_c^{(y)}}{N^{(y)}} \quad (6)$$

tanımlanır. Sınıf başına hedef örnek sayısı N_* belirlendiğinde, küme bazlı hedefler

$$t_c^{(y)} = \lceil \pi_c^{(y)} N_* \rceil \quad (7)$$

olarak atanır. Böylece yeniden örnekleme her sınıfın küme içi morfolojisini korur.

Prosedür açıklanmak istenirse (i) sürekli değişkenler standartlaştırılır, k-ortalamalar yalnızca eğitim bölümü üzerinde uygulanır. (ii) Her sınıf için $n_c^{(y)}$ ile $t_c^{(y)}$ karşılaştırılır. Aşırı temsil durumunda her kümeden olasılıkça orantılı alt-örnekleme yapılır. Yetersiz temsil durumunda her kümede $t_c^{(y)} - n_c^{(y)}$ adet sentetik örnek üretilir. Sentetik üretim, küme-içi iki komşu $x_a, x_b \in S_c^{(y)}$ üzerinden

$$\tilde{x} = (1 - \lambda)x_a + \lambda x_b, \quad \lambda \sim \Pi(0,1) \quad (8)$$

ile gerçekleştirilir, kategorik öznitelikler için moda-korumalı çoğullama uygulanır. (iii) Üretilen/seçilen örnekler birleştirilir, doğrulama/test bölümleri dokunulmadan bırakılır. Uygulama kısmına geçildiğinde ilk

değnilmesi gereken nokta k değnerinin dirsek veya silüet katsayısı ile adaylar arasından seçilmesi ve küme kararlılığının çapraz doğrulama boyunca izlenmesidir. Aykırı değnerlerin etkisini azaltmak için kümelemeye Küresel MAD temizliği sonrası girilir.

3.2.4. Karşılıklı Bilgi Tabanlı k -En İyi Özellik Seçimi

Bu metotta her aday öznetelik X_j ile hedef değışken Y arasındaki bağımlılık karşılıklı bilgi (KB) ile ölçölür ve KB puanlarına göre en iyi bilgi taşıyan k öznetelik seçilir. KB doğrusal ilişki varsayımı gerektirmeksiz ortak dağılımdaki bilgi kazanımını niceler [34]:

$$I(X;Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \quad (9)$$

Sürekli değışkenler için $I(X;Y)$ kestirimi pratikte k -en yakın komşu veya parçalı-ayırıklaştırma temelli pratik olmayan tahminçilerle gerçekleştirilir. Sınıflandırmada $I(X;Y)$ sınıf bilgisini, regresyonda ise doğrusal olmayan ilişkileri de içerecek biçimde bağımlılığ yakalar.

3.2.5. Bayes Yöntemiyle Hiperparametre Optimizasyonu

Modelin doğrulama kaybı $f(\lambda)$ kapalı formda bilinmeyen ve değnerlendirmesi maliyetli bir kara kutu olarak ele alınmıştır. Bayes optimizasyon, bu amacı yaklaşım modeli ile istatikselsel olarak modelleyip yeni denemeleri bir edinim fonksiyonu ile seçerek bulur. Gauss süreci yaklaşımına odaklanıldığında, $f(\lambda) \sim GP(m(\lambda), k(\lambda, \lambda'))$ varsayımı altında, t . iterasyona dek toplanan $D_t = \{(\lambda_i, y_i)\}_{i=1}^t$ ile artakanan gürlütüsü $\epsilon \sim N(0, \sigma_\epsilon^2)$ kabul edilirse, yeni bir λ noktası için artık dağılım

$$\mu_t(\lambda) = k(\lambda)^T (K + \sigma_\epsilon^2 I)^{-1} y, \quad \sigma_t^2(\lambda) = k(\lambda, \lambda) - k(\lambda)^T (K + \sigma_\epsilon^2 I)^{-1} k(\lambda) \quad (10)$$

şeklinde elde edilir. Burada K kovaryans matrisi, $k(\lambda)$ ise eğitim noktalarıyla kovaryans vektörüdür. Edinim fonksiyonları incelendiğinde ise yeni deneme λ_{t+1} , edinimi maksimize eden noktadır. Beklenen iyileşme (EI)

$$EI_t(\lambda) = (f^* - \mu_t(\lambda))\phi(z) + \sigma_t(\lambda)\phi(z), \quad \text{where } z = \frac{f^* - \mu_t(\lambda)}{\sigma_t(\lambda)} \quad (11)$$

ile hesaplanır ve f^* mevcut en iyi gözlemdir, ϕ, ϕ ise standart normal kümülatif dağılım fonksiyonu/yoğunluk fonksiyonu dur. Üst güven sınırı (UCB) ise aşğıdaki formülasyonla hesaplanır.

$$UCB_t(\lambda) = \mu_t(\lambda) + K\sigma_t(\lambda) \quad (12)$$

Burada küçük değnerler daha iyi sonuç vermekle birlikte K keşif-sömürü dengesini ayarlar. İsteğe bağı olarak iyileşme olasılığı da kullanılabilir [35].

3.2.6. LightGBM

LightGBM gradyan artırımı karar ağaçlarının yaprak-odaklı büyüme stratejisini ve histogram tabanlı bölme algoritmasını kullanan, hafif ve yüksek hızlı bir öğrencidir. Amaç, yinelemeli olarak zayıf öğrenciler $f_m(x)$ ekleyerek modeli

$$F_m(x) = F_{m-1}(x) + v f_m(x) \quad (13)$$

şeklinde güncellemektir [36]. Her yinelemede sınıflandırma kaybının ikinci mertebe Taylor açılımı kullanılır, gözlem i için gradyan ve Hessian

$$g_i = \frac{\partial \ell(y_i, \hat{y}_i)}{\partial \hat{y}_i}, \quad h_i = \frac{\partial^2 \ell(y_i, \hat{y}_i)}{\partial \hat{y}_i^2} \quad (14)$$

ile hesaplanır. İkili sınıflamada $\hat{y}_i$ lojit ve $p_i = \sigma(\hat{y}_i)$ olmak üzere lojistik kayıp $\ell(y_i, \hat{y}_i) = -y_i \log p_i - (1 - y_i) \log(1 - p_i)$; buradan $g_i = p_i - y_i$, $h_i = p_i(1 - p_i)$ elde edilir. Bir bölmenin getirisi, düğüm toplam gradyan/Hessian'ları G, H ile

$$Gain = \frac{1}{2} \left(\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} + \frac{G_T^2}{H_T + \lambda} \right) - \gamma \quad (15)$$

şeklinde hesaplanır ve λ L2 ceza, γ yaprak cezasını temsil etmektedir. LightGBM yaprak-odaklı olarak, getirisi en yüksek yaprağı büyütür, aşırı uyumu sınırlamak için `max_depth` ve `num_leaves` kısıtları uygulanır. *Gain* ifadesi sonrasında LightGBM, yaprak-odaklı stratejiyle getirisi en yüksek yaprağı büyütür; aşırı uyumu sınırlamak için `num_leaves`, `max_depth` ve `min_data_in_leaf` kısıtları ile L1/L2 cezası uygulanır [37-39].

4. BULGULAR VE TARTIŞMA

Bu çalışmada, tüm hesaplamalar MacOS Sequoia işletim sistemi sürüm 15.6, 2.4 GHz işlemci (Intel Core i5) ve 8 GB bellek ile donatılmış bir Macintosh bilgisayarda gerçekleştirilmiştir. Tüm hesaplama görevleri için entegre bir geliştirme ortamı olarak Python programlama dili kullanılmıştır. Ayrıca, önerilen metodolojiyi uygulamak için sklearn, lightgbm ve optuna paketleri kullanılmıştır. CER veri setine uygulanan ön işleme süreci, ham verilerin Python ortamına aktarılmasıyla başlamıştır. Veri kümesi; (i) 1000–1999 aralığında hane kimlik numaraları, (ii) gün ve 30 dakikalık periyotları temsil eden kodlanmış zaman bilgisi ve (iii) kW cinsinden tüketim değerlerinden oluşmaktadır. Ön işleme kapsamında, kodlanmış zaman bilgisi okunabilir formata dönüştürülmüş, veriler hane bazında gruplandırılmış ve her grup içindeki kayıtlar kronolojik olarak sıralanarak analiz için tutarlı bir zaman serisi yapısı elde edilmiştir.

Veri ön işlemenin ikinci aşamasında, eksik gözlemler KA yöntemi ile tahmin edilmiştir. Uygulama öncesinde, eksik veri oranı %5'in üzerinde olan tüketici serileri analiz dışı bırakılmıştır. Kalan serilerde, kısa dönemli bağımlılıkları ($p=1$), durağanlaştırmayı ($d=1$) ve geçmiş hata terimlerinin etkilerini ($q=1$) dikkate alan KA modeli kullanılmıştır. Bu yaklaşım, zaman serisinin dinamik yapısını yansıtarak özellikle dalgalı tüketim bölgelerinde tutarlı sonuçlar sağlamış ve serinin sürekliliğini koruyarak sonraki analizlerin güvenilirliğini artırmıştır. Üçüncü aşamada, aykırı değerlerin belirlenmesi amacıyla Küresel MAD yöntemi uygulanmış; eşik değeri 3,5'in üzerinde kalan gözlemler aykırı olarak etiketlenerek veri setinden çıkarılmıştır [40]. Bu işlem, tüketim profilindeki olağandışı artış ve düşüşlerin başarılı bir şekilde ayrıştırılmasını sağlamış, temizlenen veri seti ise sonraki öznitelik çıkarımı ve sınıflandırma adımlarında daha istikrarlı ve güvenilir sonuçlar elde edilmesine katkı sunmuştur.

Dördüncü aşamada, gerçek tüketim profilleri üzerinden literatürde tanımlanan altı farklı saldırı tipi kullanılarak yapay veri setleri oluşturulmuş ve dürüst olmayan tüketici davranışlarının çeşitli senaryoları güvenilir biçimde benzetilmiştir. Beşinci aşamada ise, veri setinden günlük bazda öznitelikler çıkarılarak dürüst ve dürüst olmayan tüketim profilleri arasındaki farklılıkların belirlenmesine ve sonraki sınıflandırma sürecinin desteklenmesine katkı sağlanmıştır. Son aşama olan veri boyut indirgeme aşamasında karşılıklı bilgi temelli en iyilerin seçim yöntemi kullanılmış ve elde edilen sonuçlara göre en iyi performans, minimum değer, değer aralığı, medyan, maksimum değer, ortalama, standart sapma ve RMS özniteliklerinden oluşan yedi öznitelik seçildiğinde elde edilmiştir. Böylece yüksek boyutluluğun getirdiği gürültü azaltılmış, daha açıklayıcı ve ayırt edici bir öznitelik kümesi elde edilmiştir.

Makine öğrenmesi yöntemi ile modelleme aşamasında, Bayes optimizasyon temelli hiperparametre seçimi, LightGBM sınıflandırıcısının performansını anlamlı biçimde iyileştirmiştir. Optuna kütüphanesi ile

gerçekleştirilen 60 deneme sonucunda, doğruluk, duyarlılık, F1-skor ve kesinliğe göre en uygun parametreler elde edilmiştir. Bu süreçte, öğrenme oranı, yaprak sayısı, derinlik ve düzenleştirme katsayıları gibi kritik hiperparametrelerin otomatik ayarlanması, modelin hem doğruluk hem de genelleme yeteneğini artırmıştır. Çapraz doğrulama sonuçları, modelin farklı veri bölünmelerinde kararlı performans gösterdiğini ortaya koymuştur. Sonuç olarak, Bayes optimizasyon yaklaşımı ile desteklenen LightGBM modeli, elektrik hırsızlığı tespitinde yüksek doğruluk ve güvenilirlik sunarak, yöntemin pratik uygulamalara uygunluğunu göstermiştir.

Modelin performansını değerlendirmek için doğruluk, kesinlik, duyarlılık ve F1-skoru metrikleri kullanılmıştır. Bu metrikler sırasıyla aşağıdaki eşitliklerde tanımlanmaktadır. Çalışmada, her bir model çıktısı bu ölçütler üzerinden değerlendirilmiş ve elde edilen sonuçlar tablo ve şekillerde karşılaştırmalı olarak sunulmuştur.

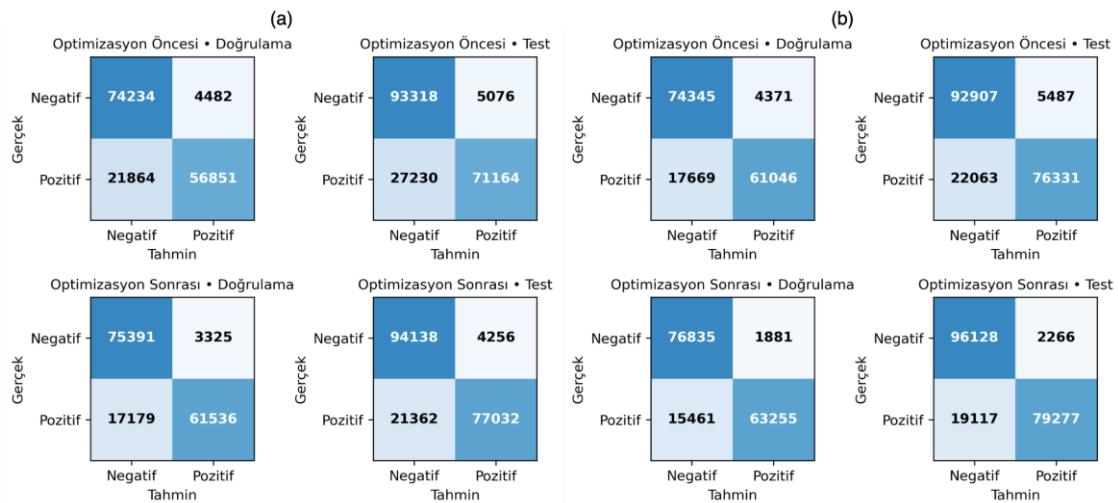
$$\text{Doğruluk} = \frac{TP + TN}{TP + TN + FP + FN} \quad (16)$$

$$\text{Kesinlik} = \frac{TP}{TP + FP} \quad (17)$$

$$\text{Duyarlılık} = \frac{TP}{TP + FN} \quad (18)$$

$$\text{F1-Skor} = \frac{2 * \text{Kesinlik} * \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} \quad (19)$$

Şekil 4 (a) ve (b)'de boyut indirgeme öncesi ve sonrası için doğrulama ve test setlerine ait karmaşıklık matrisleri verilmiştir. Optimizasyon öncesinde yanlış sınıflandırmalar özellikle negatif sınıflarda yoğunlaşırken, optimizasyon sonrasında doğru pozitif (TP) ve doğru negatif (TN) değerleri artmış, yanlış pozitif (FP) ve yanlış negatif (FN) değerleri azalmıştır. Bu durum, boyut indirgeme ve hiperparametre optimizasyonunun modelin ayırt ediciliğini güçlendirdiğini, sınıflandırma dengesini artırdığını ve genelleme yeteneğini koruduğunu göstermektedir. Çizelge 3 incelendiğinde, optimizasyon ve boyut indirgeme adımlarının tüm performans metriklerinde anlamlı iyileşmeler sağladığı görülmektedir. Doğruluk oranı %83,3–86,0 seviyelerinden %87,0–89,1 aralığına yükselmiştir. Kesinlik %92,7–93,3'ten %94,9–97,2'ye, duyarlılık %72,2–77,6'dan %78,2–80,6'ya çıkmış, F1-Skor ise %81,2–84,7'den %85,7–88,1 aralığına ulaşmıştır. Bu sonuçlar, yanlış pozitiflerin azaldığını, pozitif sınıfın daha doğru tespit edildiğini ve genel model performansının önemli ölçüde güçlendiğini göstermektedir.



Şekil 4. (a) Boyut indirgeme öncesinde, optimizasyon öncesi ve sonrası için doğrulama ile test setlerine ait karmaşıklık matrisi sonuçları, **(b)** Boyut indirgeme sonrasında, optimizasyon öncesi ve sonrası için doğrulama ile test setlerine ait karmaşıklık matrisi sonuçları

Çizelge 3. Boyut indirgeme ve optimizasyon öncesi/sonrası için doğrulama ve test setlerinde elde edilen performans metrikleri

Boyut indirgeme	Optimizasyon	Veri seti	Doğruluk	Kesinlik	Duyarlılık	F1-Skor
Öncesi	Öncesi	Doğrulama	83,3	92,7	72,2	81,2
	Sonrası		87,0	94,9	78,2	85,7
	Öncesi	Test	83,6	93,3	72,3	81,5
	Sonrası		87,0	94,8	78,3	85,7
Sonrası	Öncesi	Doğrulama	86,0	93,3	77,6	84,7
	Sonrası		89,0	97,1	80,4	87,9
	Öncesi	Test	86,0	93,3	77,6	84,7
	Sonrası		89,1	97,2	80,6	88,1

5. SONUÇ

Bu çalışmada, akıllı şebekelerde elektrik hırsızlığının tespiti için veri odaklı ve gelişmiş bir makine öğrenmesi yaklaşımı önerilmiştir. Önerilen metodoloji, KA tabanlı eksik veri doldurma, Küresel MAD ile aykırı değer tespiti, K-ortalamlar tabanlı veri dengeleme, karşılıklı bilgiye dayalı özellik seçimi ve Bayes optimizasyon ile iyileştirilmiş LightGBM sınıflandırıcı adımlarını içeren çok aşamalı bir yapıya dayanmaktadır. Elde edilen sonuçlar, önerilen yaklaşımın tüm performans metriklerinde anlamlı iyileştirmeler sağladığını göstermektedir. Boyut indirgeme ve optimizasyon öncesinde doğruluk oranı %83,3–86,0 aralığında iken, bu değer optimizasyon ve boyut indirgeme sonrasında %87,0–89,1 seviyelerine yükselmiştir. Benzer şekilde, kesinlik %92,7’den %97,2’ye, duyarlılık %72,2’den %80,6’ya, F1-skoru ise %81,2’den %88,1’e ulaşmıştır. Bu iyileşmeler, yanlış sınıflandırma oranlarının azaldığını, özellikle pozitif sınıfın daha güvenilir biçimde tespit edildiğini ve modelin genelleme yeteneğinin güçlendiğini ortaya koymaktadır.

Sonuçlar, geliştirilen yöntemin yalnızca akademik açıdan değil, aynı zamanda dağıtım şirketleri için pratik uygulamalarda da kullanılabilecek düzeyde güvenilir olduğunu göstermektedir. Yöntem, farklı kaçak kullanım senaryolarını başarıyla ayırt edebilmiş ve elektrik hırsızlığının tespitinde mevcut yaklaşımlara kıyasla daha yüksek doğruluk sunmuştur. Ayrıca önerilen çerçevenin esnek yapısı, farklı veri setleri ve dağıtım bölgelerine uyarlanabilir olma potansiyeli taşımaktadır.

Gelecek çalışmalar kapsamında, yöntemlerin akıllı şebekelerde gerçek zamanlı karar destek sistemlerine entegre edilmesi, siber saldırı senaryolarıyla birlikte değerlendirilmesi ve açıklanabilir yapay zekâ yaklaşımlarıyla desteklenerek şeffaflığının artırılması hedeflenmektedir. Böylece, elektrik hırsızlığının tespitine yönelik daha dayanıklı, ölçeklenebilir ve güvenilir çözümler geliştirilmesi mümkün olacaktır.

6. TEŞEKKÜR

Bu çalışma, Adana Alparslan Türkeş Bilim ve Teknoloji Üniversitesi Bilimsel Araştırma Proje Birimi tarafından proje numarası 24803005 ile desteklenmiştir.

7. KAYNAKLAR

1. Kebotogetse, O., Samikannu, R. & Yahya, A. (2022). A concealed based approach for secure transmission in advanced metering infrastructure. *IEEE Access*, 10, 84809-84817.
2. Esmael, A.A., Da Silva, H.H., Ji, T. & da Silva Torres, R. (2021). Non-technical loss detection in power grid using information retrieval approaches: A comparative study. *IEEE Access*, 9, 40635-40648.
3. de Souza Savian, F., Siluk, J.C.M., Garlet, T.B., do Nascimento, F.M., Pinheiro, J.R. & Vale, Z. (2021). Non-technical losses: A systematic contemporary article review. *Renewable and Sustainable Energy Reviews*, 147, 111205.
4. Northeast Group, (2021). *Electricity theft and non-technical losses: Global markets, solutions, and vendors*. <https://northeast-group.com/2021/10/20/electricity-theft-non-technical-losses/>.
5. Tasdoven, H., Fiedler, B.A. & Garayev, V. (2012). Improving electricity efficiency in Turkey by addressing illegal electricity consumption: A governance approach. *Energy Policy*, 43, 226-234.
6. T.C. Enerji Piyasası Düzenleme Kurumu (EPDK), (2024). *2023 yılı elektrik piyasası gelişim raporu*. <https://www.epdk.gov.tr/Detay/Icerik/4-14475/duyuru>. Erişim tarihi: 30 Ağustos 2024.

7. Zulu, C.L. & Dzobo, O. (2023). Real-time power theft monitoring and detection system with double connected data capture system. *Electrical Engineering*, 105(5), 3065-3083.
8. Kawoosa, A.I., Prashar, D., Faheem, M., Jha, N. & Khan, A.A. (2023). Using machine learning ensemble method for detection of energy theft in smart meters. *IET Generation, Transmission & Distribution*, 17(21), 4794-4809.
9. Abdulaal, M.J., Ibrahim, M.I., Mahmoud, M.M., Khalid, J., Aljohani, A.J., Milyani, A.H. & Abusorrah, A.M. (2022). Real-time detection of false readings in smart grid AMI using deep and ensemble learning. *IEEE Access*, 10, 47541-47556.
10. Žarković, M. & Dobrić, G. (2024). Artificial intelligence for energy theft detection in distribution networks. *Energies*, 17(7), 1580.
11. Tripathi, A.K., Pandey, A.C. & Sharma, N. (2024). A new electricity theft detection method using hybrid adaptive sampling and pipeline machine learning. *Multimedia Tools and Applications*, 83(18), 54521-54544.
12. Zhuang, W., Jiang, W., Xia, M. & Liu, J. (2024). Dynamic generative residual graph convolutional neural networks for electricity theft detection. *IEEE Access*, 12, 42737-42750.
13. Peng, Y., Yang, Y., Xu, Y., Xue, Y., Song, R., Kang, J. & Zhao, H. (2021). Electricity theft detection in AMI based on clustering and local outlier factor. *IEEE Access*, 9, 107250-107259.
14. Qi, R., Zheng, J., Luo, Z. & Li, Q. (2022). A novel unsupervised data-driven method for electricity theft detection in AMI using observer meters. *IEEE Transactions on Instrumentation and Measurement*, 71, 1-10.
15. El-Toukhy, A.T., Badr, M.M., Mahmoud, M.M., Srivastava, G., Fouda, M.M. & Alsabaan, M. (2023). Electricity theft detection using deep reinforcement learning in smart power grids. *IEEE Access*, 11, 59558-59574.
16. Chen, J., Nanekaran, Y.A., Chen, W., Liu, Y. & Zhang, D. (2023). Data-driven intelligent method for detection of electricity theft. *International Journal of Electrical Power & Energy Systems*, 148, 108948.
17. Zhang, W. & Dai, Y. (2024). A multiscale electricity theft detection model based on feature engineering. *Big Data Research*, 36, 100457.
18. Althobaiti, A., Rotsos, C. & Marnerides, A.K. (2023). Adaptive energy theft detection in smart grids using self-learning with dual neural network. *IEEE Transactions on Industrial Informatics*, 20(2), 2776-2786.
19. Liao, W., Yang, D., Ge, L., Jia, Y. & Yang, Z. (2025). Electricity theft detection in integrated energy systems considering multi-energy loads. *International Journal of Electrical Power & Energy Systems*, 164, 110428.
20. Sharma, A. & Tiwari, R. (2024). Anomaly detection in smart grid using optimized extreme gradient boosting with SCADA system. *Electric Power Systems Research*, 235, 110876.
21. Zhu, S., Xue, Z. & Li, Y. (2024). Electricity theft detection in smart grids based on omni-scale cnn and autoxgb. *IEEE Access*, 12, 15477-15492.
22. Norouzi, A., Ahmadi, I., Nazari, M., Pourrostami, H. & Hoseini-Kordkheili, H. (2025). A novel framework to detect electricity theft in regions with traditional metering; applying a hybrid machine learning-based approach to low-sampling-rate data. *Results in Engineering*, 25, 104478.
23. Sun, X., Hu, J., Zhang, Z., Cao, D., Huang, Q., Chen, Z. & Hu, W. (2023). Electricity theft detection method based on ensemble learning and prototype learning. *Journal of Modern Power Systems and Clean Energy*, 12(1), 213-224.
24. Nirmal, S., Patil, P. & Shinde, S. (2025). Adversarial measurements for convolutional neural network-based energy theft detection model in smart grid. *e-Prime-Advances in Electrical Engineering, Electronics and Energy*, 100909.
25. Ullah, A., Khan, I.U., Younas, M.Z., Ahmad, M. & Kryvinska, N. (2025). Robust resampling and stacked learning models for electricity theft detection in smart grid. *Energy Reports*, 13, 770-779.
26. Kılınç, E. (2024). Comparison of feature extraction methods in high dimensional time series. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 39(4), 991-997.
27. CER Smart Metering Project - *Electricity Customer Behaviour Trial, commission for energy regulation (cer)*, 2009-2010, <https://www.ucd.ie/issda/data/commissionforenergyregulationcer/>.
28. Fei, K., Li, Q., Zhu, C., Dong, M. & Li, Y. (2022). Electricity frauds detection in Low-voltage networks with contrastive predictive coding. *International Journal of Electrical Power & Energy Systems*, 137, 107715.
29. Takiddin, A., Ismail, M., Zafar, U. & Serpedin, E. (2022). Deep autoencoder-based anomaly detection of electricity theft cyberattacks in smart grids. *IEEE Systems Journal*, 16(3), 4106-4117.

30. K. Zor, (2019). Research and application of real-time short-term electrical energy consumption forecasting using artificial intelligence based techniques. *Ph.D. dissertation*, Cukurova University, Adana.
31. Leys, C., Ley, C., Klein, O., Bernard, P. & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4), 764-766.
32. Lohrer, A., Kazempour, D., Hünemörder, M. & Kröger, P. (2024). CoMadOut—a robust outlier detection algorithm based on CoMAD. *Machine Learning*, 113, 3553-3580.
33. Xu, Z. & Shen, D. (2021). A cluster-based oversampling algorithm combining SMOTE and k-means (KNSMOTE).
34. Vergara, J.R. & Estévez, P.A. (2014). A review of feature selection methods based on mutual information. *Neural Computing and Applications*, 24(1), 175-186.
35. Snoek, J., Larochelle, H. & Adams, R.P. (2012). Practical bayesian optimization of machine learning algorithms. *Advances in Neural Information Processing Systems*, 25, 2960-2968.
36. Bentéjac, C., Csörgő, A. & Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, 54, 1937-1967.
37. LightGBM Developers, (2025). *LightGBM documentation (Parameters & LGBMClassifier)*.
38. Omotehinwa, T.O., Oyewola, D.O. & Dada, E.G. (2023). A light gradient-boosting machine algorithm with tree-structured parzen estimator for breast cancer diagnosis. *Healthcare Analytics*, 4, 100218.
39. Atalay, B.A. & Zor, K. (2025). Xgboost (aşırı gradyan artırımı karar ağaçları) ile hidroelektrik enerji tahmini. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 40(1), 205-218.
40. Yaro, A.S., Maly, F., Prazak, P. & Malý, K. (2024). Outlier detection performance of a modified z-score method in time-series rss observation with hybrid scale estimators. *IEEE Access*, 12, 12785-12796.



Entegre Sabit Filmlili Aktif Çamur Sistemi ve Membran Biyoreaktör ile Farklı Yapıdaki Biyofilm Performanslarının Değerlendirilmesi

Türker TÜRKEN^{1,a}, Özgün YILMAZ^{1,b}

¹*İstanbul Teknik Üniversitesi, İnşaat Fakültesi, Çevre Mühendisliği Bölümü, İstanbul, Türkiye*

^a*ORCID: 0000-0003-0550-5975;* ^b*ORCID: 0000-0001-7440-9448*

Makale Bilgileri

Geliş : 20.03.2025

Kabul : 12.09.2025

DOI: 10.21605/cukurovaumfd.1661989

Sorumlu Yazar

Türker TÜRKEN

turken@itu.edu.tr

Anahtar Kelimeler

İfas

Membran biyoreaktör

Su geri kazanımı

Atıf şekli: TÜRKEN, T., YILMAZ, Ö., (2025). Entegre Sabit Filmlili Aktif Çamur Sistemi ve Membran Biyoreaktör ile Farklı Yapıdaki Biyofilm Performanslarının Değerlendirilmesi. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 643-653.

ÖZ

Su tüketimindeki artış, mevcut kaynaklar üzerindeki baskıyı artırmakta ve gelecekte su krizine yol açabileceği öngörülmektedir. Bu nedenle atıksuların geri kazanımı giderek daha fazla önem kazanmaktadır. Bu çalışmada, alan ihtiyacını en aza indiren ve yüksek artım verimi sağlayan membran biyoreaktör (MBR) sistemleri ile entegre biyofilm sisteminin (IFAS-MBR) performansı laboratuvar ölçeğinde araştırılmıştır. Çalışmada farklı biyotayıyıcı medyalar kullanılarak sistem performansına etkileri incelenmiştir. Sonuçlar, hibrit sistemin yüksek organik madde gideriminde etkili olduğunu göstermiştir. KOİ giderimi %94 seviyesinde gerçekleşmiş, toplam azot (TN) giderimi %35, toplam fosfor (TP) giderimi %61 olarak belirlenmiştir. Çıkış suyu bulanıklık değerleri 10 NTU'nun altında kalmıştır. Biyofilm entegrasyonu sayesinde düşük SRT koşullarında dahi nitrifikasyonun gerçekleştiği gözlenmiş, ayrıca membrana gelen yük azalttığı için transmembran basıncının artışı yavaşlamış ve kimyasal yıkama ihtiyacı azalmıştır. Bu bulgular, IFAS-MBR hibrit sisteminin KOİ, azot ve fosfor gibi önemli parametrelerin giderimi ile membran ömrünün uzatılmasında da avantaj sağladığını ortaya koymaktadır.

Evaluation of Biofilm Performance of Different Structures with Integrated Fixed Activated Sludge and Membrane Bioreactor

Article Info

Received : 20.03.2025

Accepted : 12.09.2025

DOI: 10.21605/cukurovaumfd.1661989

Corresponding Author

Türker TÜRKEN

turken@itu.edu.tr

Keywords

İfas

Membrane bioreactor

Water recovery

How to cite: TÜRKEN, T., YILMAZ, Ö., (2025). Evaluation of Biofilm Performance of Different Structures with Integrated Fixed Activated Sludge and Membrane Bioreactor. Çukurova University, Journal of the Faculty of Engineering, 40(3), 643-653.

ABSTRACT

Increasing water consumption exerts pressure on available resources and is expected to cause a water crisis in the future. Therefore, wastewater reclamation has become increasingly important. In this study, the performance of an integrated fixed-film activated sludge-membrane bioreactor (IFAS-MBR) system, which minimizes land requirement while ensuring high treatment efficiency, was investigated at laboratory scale. Different bio-carrier media were tested to evaluate their effects on system performance. The results demonstrated that the hybrid system was effective in removing high organic loads. The chemical oxygen demand (COD) removal efficiency reached 94%, while total nitrogen (TN) and total phosphorus (TP) removal efficiencies were 35% and 61%, respectively. Effluent turbidity remained below 10 NTU. Biofilm integration enhanced nitrification even under low sludge retention time (SRT) conditions, and reduced the loading on the membrane, which slowed down the increase in transmembrane pressure (TMP) and lowered chemical cleaning requirements. These findings indicate that the IFAS-MBR hybrid system provides advantages not only in COD removal but also in nitrogen and phosphorus removal, as well as in extending membrane lifespan.

1. GİRİŞ

Artan nüfus ve endüstriyel faaliyetler nedeniyle temiz su kaynaklarına ulaşmak zorlaşmıştır. Bu nedenle, gün geçtikçe atık su geri kazanımı ve daha düşük alan kaplayan arıtma tesislerinin kullanımı ihtiyacı artmaktadır. Membran teknolojileri, konvansiyonel arıtma sistemleri ile birlikte veya ayrı olarak kullanılan arıtma teknolojileridir. Membran biyoreaktör (MBR) sistemleri, biyolojik arıtma ve filtrasyon sistemlerinin kombinasyonu olarak çalışmakta ve yapılan çalışmalarda MBR sistemlerinin hem alan ihtiyacını azalttığı, hem de daha yüksek kalitede çıkış suyu eldesi sağladığı gözlenmiştir.

Buna ek olarak, Entegre Sabit Aktif Çamur (IFAS) sistemi ise atık su arıtımı için sürdürülebilir bir çözüm sağlayan önde gelen proseslerden biridir. IFAS, yüzeyde ve askıda büyüme sistemlerini entegre ederek Hareketli Yatak Biyoreaktör (MBBR) sisteminin geliştirilmiş bir hali olarak açıklanabilir. Bu sistem, yüksek oranda besin giderimi, tam nitrifikasyon, daha uzun katı madde tutma süresi gibi geleneksel aktif çamur sistemlerine göre avantajlar sunar. Son zamanlarda geliştirilen IFAS reaktörleri ayrıca metan üretimi, algler veya mikrobiyal yakıt hücreleri yoluyla enerji üretimi için verimli sonuçlar içermektedir.

Son yıllarda IFAS-MBR hibrit sistemleri üzerine yapılan çalışmalar, hem organik madde giderimi hem de azot-fosfor giderimi açısından umut verici sonuçlar vermiştir. Bu kapsamda yürütülen bir pilot ölçekli çalışmada IFAS-MBR sistemlerinin geleneksel MBR sistemlerine göre daha yüksek performans sergilediği gösterilmiştir. IFAS'ın pilot ve tam ölçekli uygulamalarda kompakt tasarımı ve yüksek biyofilm stabilitesi sayesinde gelecekte daha yaygın kullanılacağı öngörülmektedir [1]. Yine başka bir çalışmada IFAS sistemlerinde nitrifikasyon veriminin %98'e kadar çıkabildiğini belirtmiştir [2]. Ayrıca Yüksekdağ [3] çalışmasında, nanolif membranlarda yüzey modifikasyonlarının membran performansını ve antimikrobiyal özelliklerini iyileştirdiği, bu işlemin de tıkanmayı azaltma potansiyeli sunduğu bildirilmiştir.

Görüldüğü üzere batık tip membranlar ile çalışan sistemler için literatür, IFAS-MBR hibrit sistemlerinin çeşitli avantajlara sahip olduğunu göstermektedir. MBR sistemlerinden önce daldırılan biyotaşıyıcı sayesinde membranlar daha yüksek geçirgenlikli akış koşullarında çalıştırılabilir, membran yüzeyindeki kirlilik yükü azaltılır ve daha kompakt bir sistem olarak daha az alan gerektirir. Bu kapsamda oluşturulan pilot ölçekli IFAS-MBR hibrit sistem düzeneği, hem karbonu hem de besin maddeleri olan azot ve fosforu gidermek ve en uygun morfolojik yapıya sahip biyotaşıyıcıyı belirlemek için tasarlanmıştır.

Çalışma kapsamında, PVDF polimeri kullanılarak üretilmiş ince boşluklu membran filtre biyoreaktörü ile entegre edilmiş farklı morfolojik yapılara sahip iki biyotaşıyıcı kullanılmıştır. Uygun biyotaşıyıcı materyalinin belirlenmesi için, evsel nitelikli atık su numunesi ile laboratuvar ölçekli deneyler yapılmış ve bazı parametreler izlenmiştir.

2. YÖNTEM

Laboratuvar ölçekli entegre IFAS-MBR sistemi için yapılan tayinler aşağıda detaylı olarak verilmiştir.

2.1. Atık Su Karakterizasyonu

Yapılan çalışmada evsel karakterli sentetik atık su hazırlanmıştır. Kullanılan kimyasallar ve besleme çamuruna ait veriler aşağıdaki çizelgede verilmektedir.

Çizelge 1. Sentetik atık su reçetesi [4]

Parametre	Birim	Değer
Glikoz	mg/L	400
Bactopeptone	mg/L	115
(NH ₄) ₂ SO ₄	mg/L	104,8
KH ₂ PO ₄	mg/L	21,75
MgSO ₄	mg/L	15,63
CaCl ₂	mg/L	2,45
MnSO ₄	mg/L	1,8
NaHCO ₃	mg/L	255,5
MLSS	mg/L	5580
KÖİ(Beslenen Çamur)	mg/L	460

2.2. KOİ Tayini

KOİ giderim verimlerini bulmak için reaktörün içindeki besleme atık suyundan ve arıtılmış süzöntü suyundan periyodik olarak numuneler alınarak KOİ analizleri yapılmıştır. Yöntem olarak standart yöntemlerden biri olan 5220-C kapalı reflü yöntemi kullanılmıştır. KOİ tüplerine 2,5 ml seyreltilmiş numune, 1,5 ml standart $K_2Cr_2O_7$ solüsyonu ve 3,5 ml H_2SO_4 reaktifi eklenmiştir.

Tüpler $150^\circ C$ 'de Hach marka DRB 200 model termoreaktöre yerleştirilmiş olup 2 saat sonra numuneler oda sıcaklığına soğumaya bırakılmıştır.

Soğuyan numuneler erlene alındıktan sonra 1-2 damla ferroin indikatörü ilave edilerek Demir Amonyum Sülfat (DAS) çözeltisi ile kiremit kırmızısı renk elde edilinceye kadar titre edilmiştir. Deneyde kullanılan DAS solüsyonunun molaritesi Eşitlik 1 ile, KOİ değerleri ise Eşitlik 2 ile mg/L cinsinden hesaplanmıştır.

$$DAS \text{ Molaritesi} = \frac{\text{Titre edilen Potasyum Dikromat Hacmi} * 0,1}{DAS \text{ Hacmi (ml)}} \quad (1)$$

$$KOİ \left(\frac{mg}{L} \right) = \frac{(A - B) * M * 8000}{Numune hacmi (ml)} \quad (2)$$

2.3. AKM Tayini

Reaktörlerdeki ve IFAS fiberleri üzerindeki katı madde değerleri düzenli aralıklarla izlenmiştir. Bu amaçla gravimetrik ölçüm yöntemi kullanılmıştır.

İlk olarak filtrenin yıkanması prosedürü uygulanmıştır. Bunun için, filtre, filtrasyon setine yerleştirilip, vakum pompası yardımıyla vakum uygulanarak distile su ile yıkanmıştır. Filtrede su kalmayınca kadar vakum uygulamasına devam edilmiş ve yıkama suyu dökülmüştür. Filtre $103-105^\circ C$ etüvde 1 saat kurutulup hem havadaki nemden etkilenmeyecek hem de tamamen kuruyacak şekilde desikatörde bekletilmiştir. Sabit tartıma gelen filtre desikatörden dikkatlice alınarak tartılmıştır.

Ardından reaktör içerisinden 20 ml numune alınıp vakum uygulanarak hazırlanan filtreden süzölmüştür. Filtre bir forseps yardımıyla aparattan dikkatlice çıkarılıp alüminyum bir kaba yerleştirilerek $103-105^\circ C$ etüvde en az 1 saat kurutulmaya bırakılmıştır. Etüvden sonra desikatörde oda sıcaklığına getirildikten sonra tartılmıştır. Ardından, filtre ve üzerindeki tortu $550^\circ C$ 'deki kül fırınında 30 dakika yakılmıştır. Yakma işlemi sonucunda desikatörde soğuduktan sonra tartım işlemi yapılmıştır. AKM hesaplamaları Eşitlik 3 ile yapılmıştır.

$$AKM \left(\frac{mg}{l} \right) = \frac{(A - B) * 1000}{Numune hacmi (ml)} \quad (3)$$

IFAS fiberlerinden 14 günlük çalışmada 3., 5., 10. ve 12. günlerde numuneler alınarak üzerlerinde tutunan katı madde analizleri yapılmıştır.

Distile su içerisinde bir kaşık yardımıyla fiberlerin üzerinden katı maddeler sıyrılmış ve etüvde kurutulmuştur. Bu işlemden sonra ağırlık ölümü yapılarak birim hacimdeki katı madde miktarı hesaplanmıştır.

2.4. Toplam Azot Tayini

Bu çalışmada belirli periyotlarda toplam azot ölçümleri, Hach-Lange marka 5-40 mg/L aralığındaki kitlerle ölçülmüştür.

Kit ölçümü için gerekli prosedürler uygulandıktan sonra DR 5000 Hach-Lange spektrofotometre ile sonuçlar elde edilmiştir. Ölçümler sonucunda çıkan değerler yapılan çalışmaların doğru sonuç verdiğini göstermiştir.

2.5. Toplam Fosfor Tayini

Standart Yönteme (APHA, 2005) göre yürütülen toplam fosfor analizi Hach-Lange marka LCK 350 deney kitleri ile DR 5000 Hach-Lange spektrofotometre üzerinde 507 nm dalga boyunda ölçülmüştür.

2.6. Bulanıklık Tayini

Bulanıklık, bir görünüm özelliğidir, özellikle numunenin ışığı nasıl ilettiğinin bir ölçümüdür. Bulanık su ışığı çok iyi iletmez. Bunun nedeni, sudaki kolloidal parçacıklardan kaynaklanan ışık saçılımıdır.

Bulanık olan su genellikle daha yüksek seviyelerde BOİ, KOİ ve patojenlere sahiptir. Işık geçirgenliği, mevcut kolloidal parçacıkların hem boyutuna hem de sayısına bağlı olduğundan bulanıklığı TSS ile doğrudan ilişkilendirmenin de oldukça zor olduğuna dikkat etmek önemlidir.

Bulanıklık, NTU tarafından belirlenirken AKM, hacim başına kütle birimi cinsinden ölçülür. Entegre IFAS-MBR sistemi ile evsel atık su arıtma işlemi sonrasında atık su içerisindeki AKM, çıkış suyunda bulanıklık cinsinden ölçülmüştür. Aynı AKM seviyelerine sahip iki numune, farklı parçacık boyutu dağılımlarına sahiplerse farklı bulanıklık değerlerine sahip olabilir.

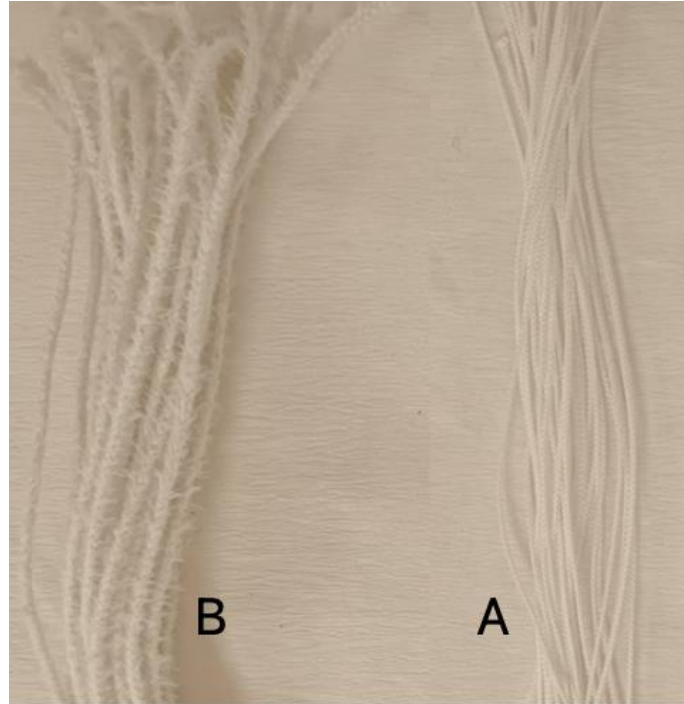
2.7. SEM Analizi

İşletme öncesinde ve sonrasında IFAS fiberlerinin yüzey özelliklerinin belirlenebilmesi için Taramalı Elektron Mikroskop analizi yapılmıştır. Ölçümde kullanılan cihaz FEI marka Quanta Feg 250 modelidir. SEM analizi yapılabilmesi için fiberlerin kuru ve ortam neminin çok düşük olması gerekmektedir.

Analiz için etanol/su karışımı içeren solüsyonda yıkandıktan sonra oda sıcaklığında kurutulan numuneler özel numune hücrelerine yerleştirilerek Au-Pt ile yüzey kaplaması yapılmıştır.

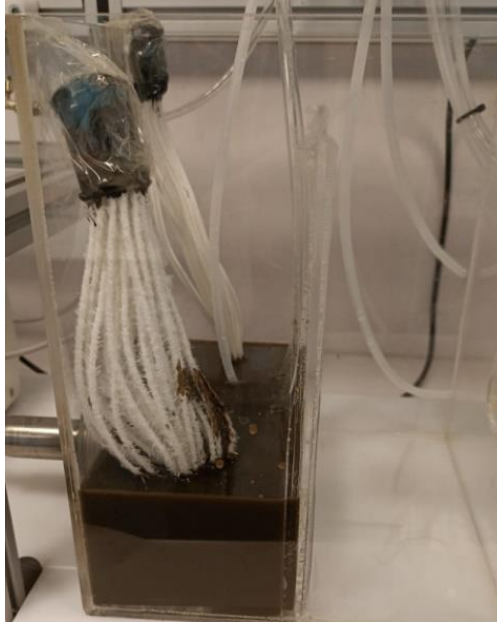
2.8. IFAS Modüllerinin Yapılışı

MBR ile entegre çalıştırılacak IFAS modülleri aynı malzeme, farklı yüzey alanları ve karakteristik özelliklere sahip Şekil 1’de gösterilen polyester fiberlerden oluşmaktadır.



Şekil 1. Farklı özelliklere sahip IFAS fiberleri

Uçlarından birleştirilen IFAS fiberleri tank içerisine Şekil 2’de görüldüğü gibi yerleştirilmiştir. Tank hacminin yüzde 40’ını kaplayacak şekilde yerleştirilen modüller MBR ile entegre işletildiğinde yüksek verimli IFAS modülü belirlenmeye çalışılmıştır.



Şekil 2. Tanka yerleştirilen IFAS modülleri

2.9. MBR Modüllerinin Yapılışı

Laboratuvar ölçekli sistem işletilmesi için son olarak ince boşluklu içi boş fiber membranların da modül haline getirilip uygun şekilde tanka yerleştirilmesi gerekmiştir. Bu nedenle aşağıda verilen eşitliklere (4,5,6) göre yapılan hesaplamalar sonucunda 104 adet fiber membran kullanılarak dıştan içe doğru akışlı ve vakum ile çalışan Şekil 3’teki modül üretilmiştir.

$$\begin{aligned} \text{Membran Alanı (m}^2\text{)} &= \frac{\text{Ortalama Debi (m}^3\text{/h)} * 1000}{\text{Akı (L/m}^2 * \text{h)}} \\ &= \frac{0,002 \text{ (m}^3\text{/h)} * 1000}{15} \end{aligned} \quad (4)$$

$$\begin{aligned} \text{Membran Alanı} &= 0,137 \text{ m}^2 \text{ Membran Fiber Alanı (m}^2\text{)} \\ &= \frac{0,0021 * \pi * (\text{Fiber Uzunluğu (m)} * 1000)}{1000} \\ &= \frac{0,0021 * 3,14 * (0,2 * 1000)}{1000} \end{aligned} \quad (5)$$

$$\text{Membran Fiber Alanı} = 0,0013 \text{ m}^2$$

$$\begin{aligned} \text{Membran Miktarı (adet)} &= \frac{\text{Membran Alanı (m}^2\text{)}}{\text{Membran Fiber Alanı (m}^2\text{)}} \\ &= \frac{0,137 \text{ (m}^2\text{)}}{0,0013 \text{ (m}^2\text{)}} \end{aligned} \quad (6)$$

$$\text{Membran Miktarı} = 104 \text{ adet}$$

Tanka yerleştirilen membran modülü Şekil 3'te gösterilmektedir.



Şekil 3. MBR modülü

Evsel karakterli atık suların arıtımında seçilen akı değeri 15–25 L/m²h arasında değişmektedir [5]. Bu veriler doğrultusunda seçilen akı değeri 15 L/m²h'dir.

2.10. IFAS-MBR Sisteminin İşletilmesi

Laboratuvar ölçekli IFAS-MBR sistemi PLC sistemine bağlıdır ve tüm kontroller bu ekrandan yapılabilmektedir. MBR sisteminde filtrasyon ve geri yıkama olmak üzere iki farklı proses vardır. PLC ekranında tüm sistem şeması görülebilmekte ve gerektiğinde ünitelere ilişkin Geri yıkama, filtrasyon basıncı, akı ve debi gibi çalışma parametreleri bu ekrandan değiştirilebilmektedir. Şekil 4'te sistemin işletme sırasındaki bir görüntüsü verilmektedir.

Besleme pompası vasıtasıyla IFAS tankına alınan atık su MBR tankına savaklanarak aktarılmıştır. Loblu pompa ile vakumlanan atık su membran filtresinden geçerek süzüntü tankında toplanmıştır.

9 dakika filtrasyon işlemi ve 1 dakika geri yıkama işlemi yapılarak membran yüzeyinde oluşabilecek tıkanıklıklar engellenebilmekte ve membran filtrasyonunun verimi istenilen seviyelerde tutulabilmektedir.



Şekil 4. İşletilen IFAS-MBR sistemi

3. ARAŞTIRMA BULGULARI

Laboratuvar ölçekli IFAS-MBR sistemi işletme sonuçlarına göre, IFAS modüllerinin MBR sistem çalışmasına çeşitli etkilerinin olduğu gözlemlenmiştir. IFAS malzemesi, geniş yüzey alanı sayesinde MLSS'nin üzerine yapışmasını sağlar ve böylece aerobik reaktör hacmi ihtiyacını azaltır. SEM görüntülerinde görüldüğü gibi, pürüzlü yüzeyli IFAS malzemesi, pürüzsüz yüzeyli IFAS malzemesinden daha fazla biyofilm tutmuştur. Bu bağlamda pürüzlü yüzeye sahip IFAS malzemesinin kullanıldığı sistemlerde aerobik reaktör hacmi daha düşük tutulabilmektedir.

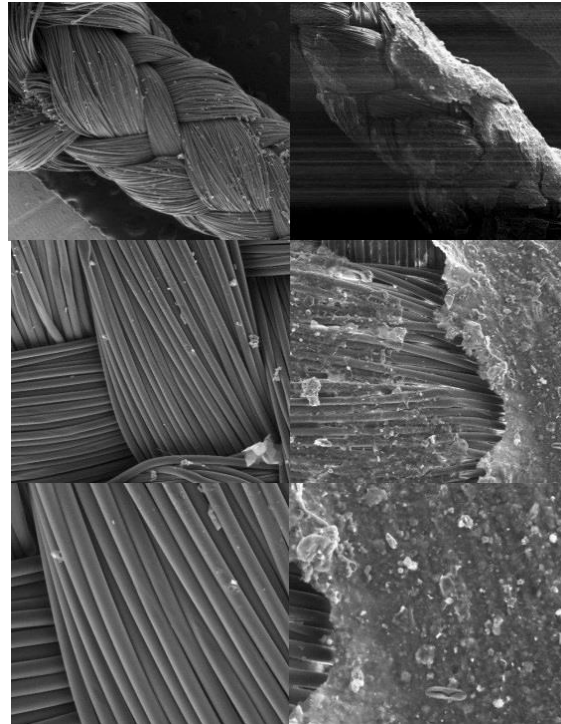
Hem MBR hem de IFAS sistemlerinin çıkış amacı arıtma tesislerinin alan ihtiyacını azaltmak olsa da IFAS-MBR sisteminin alan ihtiyacı konvansiyonel sistemlere göre çok daha düşüktür. IFAS-MBR sisteminin alan gereksiniminin de hem sabit hem de hareketli biyotaşıyıcı sistemlerine ve MBR sistemlerine göre daha düşük olduğu görülmektedir. IFAS taşıyıcıları üzerine biyofilm yapışması nedeniyle MBR sisteminin kimyasal yıkama sıklığının azaltılabileceği gözlemlenmiştir. Bu sayede membran modüllerin ömrünün uzatılabileceği düşünülmektedir.

Laboratuvar ölçekli çalışma sonrasında Şekil 5'te de görülebileceği gibi, IFAS B'de daha fazla biyofilm tutunması gözlemlenmiştir. Biyofilm tutunmasının sebepleri sonuç kısmında daha detaylı açıklanmıştır. IFAS modüllerinden belirli kesitler alınarak oda sıcaklığında 24 saat kurutularak SEM analizi yapılmıştır.

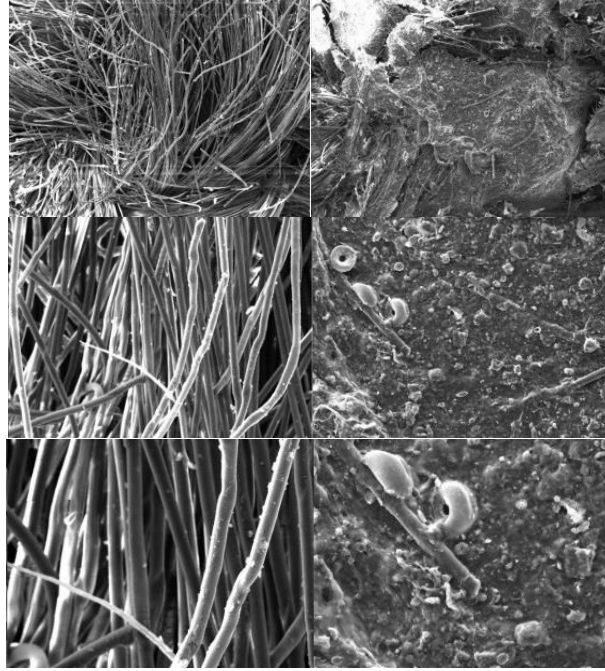


Şekil 5. Çalışma sonrasındaki IFAS modülleri

SEM ölçümü üç farklı büyütme ölçeğinde gerçekleştirilmiştir. SEM çıktılarına göre, daha az yüzey alanına sahip IFAS modülündeki biyofilm ince iken daha geniş yüzey alanlı ve pürüzlü IFAS biyotaşıyıcıda biyofilmin daha kalın olduğu görülmüştür. Bu sonucun elde edilmesinde biyofilmin tutunabildiği yüzey alanının daha geniş olması ve IFAS dokusunun daha pürüzlü olması etkilidir. Kirli modüller için her üç büyütme oranındaki SEM görüntülerine bakıldığında IFAS B yüzeyinin tamamına biyofilm tutunması gerçekleştiği görülürken IFAS A'da hala temiz bölgeler bulunduğu saptanmıştır (Şekil 6 ve Şekil 7).



Şekil 6. IFAS A için çalışma öncesi(sol) ve sonrası(sağ) sırasıyla 100, 500 ve 1000 kat büyütülmüş SEM görüntüleri

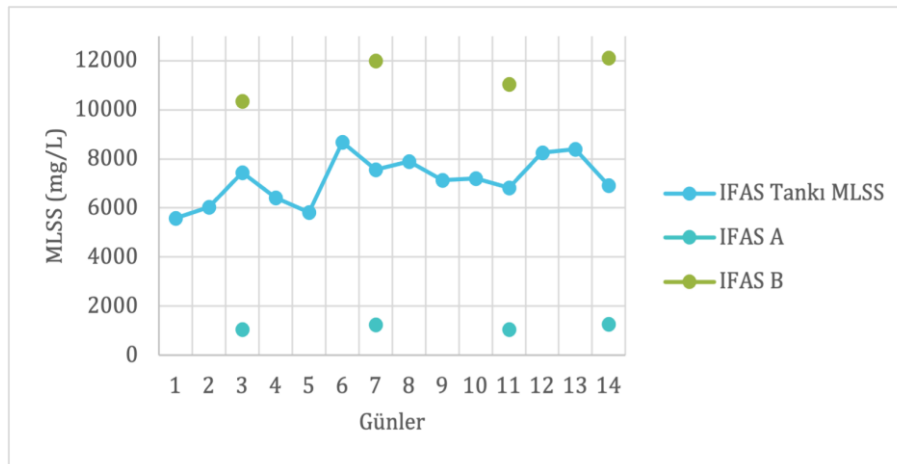


Şekil 7. IFAS B için çalışma öncesi(sol) ve sonrası(sağ) sırasıyla 100, 500 ve 1000 kat büyütülmüş SEM görüntüleri

IFAS modülleri üzerine biyofilm tutunması ile membran biyoreaktörde tıkanma ve birikimin geciktiği ve ekonomik ömrünün uzadığı belirlenmiştir. Ayrıca filtrasyon süresi uzatılarak geri yıkama ve membran havalandırma gibi maliyetli girdilere olan ihtiyaç azaltılır. IFAS gibi biyotaşıyıcıların kullanımı, geleneksel atık su arıtma sistemlerine kıyasla havalandırma havuzlarının hacmini azaltır. Konvansiyonel sistemlerde MBR kullanımı ile gerekli alan ihtiyacının yarı yarıya azaldığı da bilinmektedir [6]. Bu kapsamda IFAS ve MBR sisteminin entegre kullanımı sonucunda ihtiyaç duyulan alanın önemli ölçüde azaltılabileceği görülmüştür. IFAS'ın yüzeyinde oluşan biyofilm tabakasının kalınlığının artması ile tabakanın iç kısımlarında oksijensiz (anoksik) bir ortam oluştuğu ve bu sayede nitrifikasyon işleminin gerçekleştiğine dair literatürde çalışmalar mevcuttur [7].

Sistem sentetik atıksu ile beslenmiştir. Bu nedenle giriş atık sularında AKM değeri yoktur. Ancak Atıksu Arıtma Tesisinden alınan aktif çamurun MLSS konsantrasyonu 5580 mg/L olarak ölçülmüştür. Bu kapsamda MBR tankına ait MLSS konsantrasyon değişim grafiği aşağıda verilmiştir (Şekil 8).

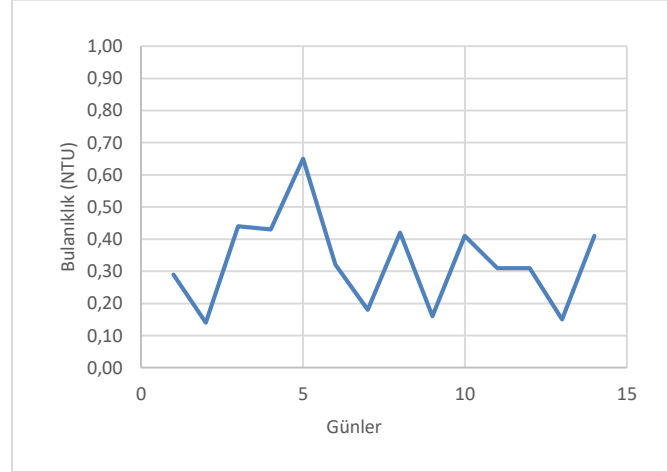
Şekil 8'e göre, üzerinde daha fazla biyokütle birikimi gözlenen IFAS B'deki MLSS konsantrasyonunun IFAS A'ya göre daha düşük mertebelerde olduğu görülmektedir.



Şekil 8. Tank içindeki ve IFAS modülleri üzerindeki MLSS değişimi

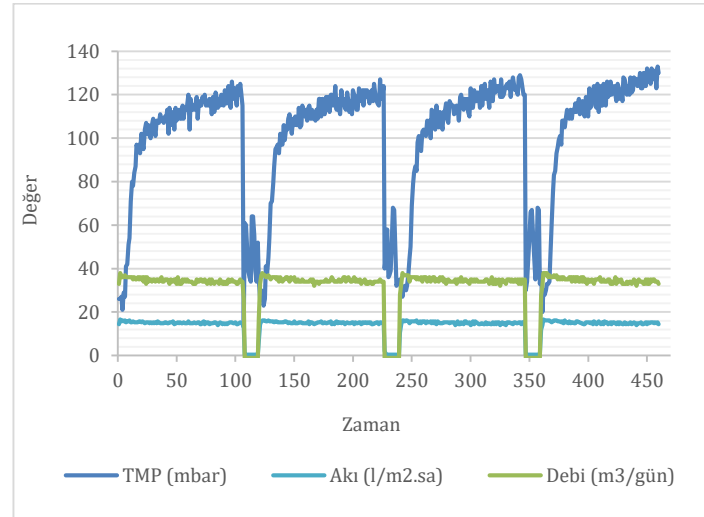
Aynı zamanda MBR çıkışındaki AKM konsantrasyonu 10 mg/L'nin altında olduğu için çıkışta AKM ölçümü yapılmamış, bulanıklık ölçümü yapılmıştır. IFAS-MBR sisteminin çıkışından alınan su numunesinin bulanıklık değişim grafiği aşağıda verilmiştir (Şekil 9).

IFAS performansları ve MLSS konsantrasyonları göz önüne alındığında bulanıklık değerlerinin aynı doğrultuda sonuçlar verdiği görülmektedir.



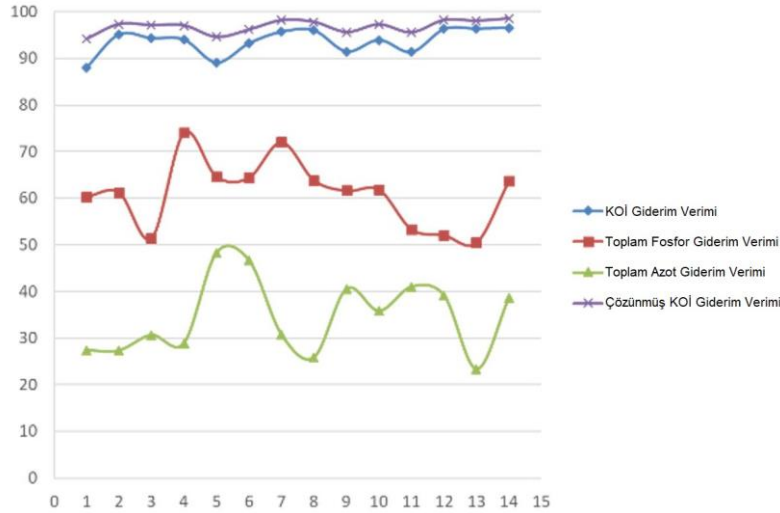
Şekil 9. Bulanıklık değişimi

MBR sisteminin filtrasyon döngüsü 9 dakika filtrasyon ve 1 dakika geri yıkama olacak şekilde ayarlanmıştır. Verilen grafikte (Şekil 10) filtrasyon basıncının (TMP) zamana göre değişimi, akı ve debi değerleri PLC verilerine göre hazırlanmıştır. Grafikte görülen ani basınç değişiklikleri, filtrasyonun 9 dakikalık döngüsünü tamamladığını ve 1 dakikalık geri yıkamanın başladığını göstermektedir. Aynı zamanda debi ve akı değerlerindeki ani düşüşün nedeni geri yıkamanın başlamış olmasıdır. MBR üzerindeki baskı negatif basınç olarak okunur ve grafiğe bakıldığında basınçta zamanla artış gözlenir. Bu artışa bağlı olarak membranların kimyasal yıkama işlemi gerçekleştirilir.



Şekil 10. TMP-debi-akı değişimi

Laboratuvar ölçekli IFAS-MBR çalışması sonucunda arıtma verimi grafiği (Şekil 11) ve ortalama giriş ve çıkış değerleri (Çizelge 2) aşağıda verilmiştir. Sonuçlara göre verim değerleri ortalama değerden fazla sapmamaktadır. Beşinci günde gözlemlenen yüksek azot giderimi, IFAS modülü üzerinde biriken biyofilmin alt tabakasında anoksik bölge oluşumu ile meydana gelen denitrifikasyon işleminin sonucu olarak açıklanabilir. Dördüncü gündeki yüksek fosfor giderimi, sentetik atıksuda bulunan FeCl_3 konsantrasyonundan dolayı çökelme ile açıklanabilir. Verimlilik değerleri genellikle yedinci günde yüksektir.



Şekil 11. Arıtma verimi değişimi

Çizelge 2. Ortalama giriş ve çıkış değerleri

Parametre	Birim	Ortalama giriş değeri	Ortalama çıkış değeri
KOİ	mg/L	360	22,9
TN	mg/L	40	26
TP	mg/L	8,83	3,4
MLSS	mg/L	7155	-

4. SONUÇLAR

Temiz suya ulaşmanın zorlaşması ve alan ihtiyaçlarının azalması gerekliliği nedeniyle, konvansiyonel arıtma sistemleri yerini su geri kazanım sistemlerine bırakmaya başlamıştır. Bu kapsamda, son çökeltim havuzlarının kapladığı yüksek alanları ortadan kaldırmak ve daha yüksek giderim verimleri elde etmek için geliştirilen MBR sistemlerinin daha az alanda daha fazla atık su arıtımı ve geri kazanımı sağladığı birçok çalışma ile kanıtlanmıştır. IFAS'ların da sisteme entegre edilmesi sayesinde kimyasal yıkamaya duyulan ihtiyacın azaltılarak MBR sistemlerin çalışma ömrünün uzadığı da görülmektedir.

Yürütülen çalışmada elde edilen %94 KOİ giderimi literatürde bildirilen değerlerle uyumludur; yapılan benzer bir çalışmada %92 giderim raporlanmış olup, bu çalışmadaki değer bir miktar daha yüksektir [1]. Bui et al. [2] tarafından raporlanan yüksek nitrifikasyon verimleri ile karşılaştırıldığında, biyofilmin düşük SRT koşullarında bile nitrifikasyonu desteklediği görülmüştür. Benzer şekilde, IFAS-Anammox entegrasyonu çalışmalarında görüleceği üzere biyofilm tabakalarının iç kısımlarında anoksik bölgelerin oluşması denitrifikasyonu artırmıştır [8]. Ayrıca, modelleme çalışmalarında öne çıkan parametrelerle karşılaştırıldığında bu çalışmadaki MLSS ve TMP verileri benzer eğilimler göstermiştir [9]. Çalışma kapsamında diğer performans parametreleri de değerlendirilmiş; toplam azot (TN) giderimi %35 (40 mg/L → 26 mg/L), toplam fosfor (TP) giderimi %61 (8.83 mg/L → 3.4 mg/L) olarak ölçülmüştür.

Çıkış suyu bulanıklığı düşük seviyelerde seyretmiş ve 10 NTU'nun altında kalmıştır. Pürüzlü yüzeye sahip biyotayıyıcı kullanılan IFAS modülünde daha fazla biyofilm tutunması sayesinde MBR tankında daha düşük MLSS değerleri gözlenmiştir. Yeşiltaş ve Yılmaz [10] çalışmasında da, koagülasyon uygulamalarının su kalitesini iyileştirerek membran öncesi yükü azaltabileceği ve filtrasyon performansını olumlu etkileyebileceği belirtilmiştir. TMP (transmembran basıncı) ise süreç boyunca artış göstermesine rağmen biyofilm entegrasyonu sayesinde kimyasal yıkama ihtiyacı azalmıştır. Tüm bu bulgular, sistemin sadece KOİ değil aynı zamanda azot ve fosfor giderimi ile operasyonel sürdürülebilirlik açısından da etkili olduğunu ortaya koymaktadır.

Yapılan laboratuvar ölçekli çalışmada IFAS sistemi ve MBR sisteminin entegre edildiği hibrit sistem sayesinde, alan ihtiyacı minimuma indirilmiş ve yüksek giderim verimi eldesi ile yüksek kirletici konsantrasyonlarına sahip atıksuların geri kazanılabileceği görülmüştür. Gelecekte, IFAS malzeme yapısı,

yüzey alanı ve geometrik şekil gibi değerleri değiştirilerek daha yüksek verimliliğe sahip sistemlerin geliştirilmesi öngörülmektedir.

5. KAYNAKLAR

1. Nguyen, H. T. & Phan, T.M. (2025). Recent progress in integrated fixed-film activated sludge process for wastewater treatment: A review. *Environmental Technology Reviews*, 14(1), 88-115.
2. Bui, X. T., Tran, H.T., Vo, T.D.H. & Nguyen, T.T. (2024). Performance evaluation of integrated fixed-film activated sludge (IFAS) system for nitrogen removal under different operational conditions. *Processes*, 12(10), 2131.
3. Yüksekdağ, A. (2024). Elektroeğirme yöntemi ile üretilen antimikrobiyal nanolif membranlarda Lycopodium özütü etkisinin incelenmesi. *Ç.Ü. Müh. Fak. Dergisi*, 39(4), 1027-1038.
4. Ergön-Can, T., Köse-Mutlu, B., Koyuncu, İ. & Lee, C.H. (2017). Biofouling control based on bacterial quorum quenching with a new application: Rotary microbial carrier frame. *Journal of Membrane Science*, 525, 116-124.
5. Özkan, O., Uyanık, İ., Rençber, M., Oğuz, M., Şahin, U. ve Koyuncu, İ.S. 2017. Organize sanayi bölgesi atıksularının membran biyoreaktörlerle arıtılması: KOSB için ilk sonuçlar. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 1060.
6. Bengtsson, S., de Blois, M., Wilén, B. M. & Gustavsson, D. (2019). A comparison of aerobic granular sludge with conventional and compact biological treatment technologies. *Environmental Technology*, 40(21), 2769-2778.
7. Albizuri, J., Santamaria, A. & Larrea, L. (2010). IFAS process for nitrogen removal with an HRT of 4 hours. *Proceedings of the Water Environment Federation*, 440-457.
8. Li, X. & Yu, R. (2024). Integrated fixed-film activated sludge and Anammox systems: Recent developments and influencing factors. *Frontiers in Environmental Science*, 12, 1091538.
9. Pryce, J., Ahmed, M. & Collins, R. (2022). Modelling and sensitivity analysis of integrated fixed-film activated sludge processes for municipal wastewater treatment. *Scientific Reports*, 12, 13779.
10. Yeşiltaş, H.K. ve Yılmaz, T. (2022). Yüzeysel su kaynağından içme suyu eldesinde alkali koagülasyon metodunun uygulanması ve kalsiyum sertliğinden ileri gelen kireç taşı oluşumu. *Ç.Ü. Müh. Fak. Dergisi*, 37(1), 249-260.



A Novel Multi-Head Attention Framework for COVID-19 Detection: Hybrid Integration of MobileNet and VGG19 with Enhanced Feature Learning

Şafak KILIÇ^{1,a}

¹Kayseri University, Faculty of Engineering, Architecture and Design, Department of Software Engineering, Kayseri, Türkiye

^aORCID: 0000-0002-2014-7638

Article Info

Received : 07.03.2025

Accepted : 12.09.2025

DOI: 10.21605/cukurovaumfd.1653486

Corresponding Author

Şafak KILIÇ

safakkilic@kayseri.edu.tr

Keywords

COVID-19 detection

Deep learning

Multi-head attention mechanism

Medical image analysis

VGG19

How to cite: KILIÇ, Ş., (2025). A Novel Multi-Head Attention Framework for COVID-19 Detection: Hybrid Integration of MobileNet and VGG19 with Enhanced Feature Learning. Çukurova University, Journal of the Faculty of Engineering, 40(3), 655-670.

ABSTRACT

The COVID-19 pandemic has underscored the urgent need for rapid, accurate, and affordable diagnostic tools to complement RT-PCR testing. This study proposes a novel multi-head attention framework that integrates VGG19 and MobileNet for automated COVID-19 detection from chest X-rays. The model employs a hybrid mechanism combining spatial, channel, and self-attention components, enhancing feature representation while preserving efficiency.

Evaluations on 7,132 chest X-ray images across four categories (COVID-19, Normal, Pneumonia, Tuberculosis) demonstrated outstanding performance: 99.0% accuracy, 99.0% macro and weighted F1-scores, with near-perfect class-specific results (100% Tuberculosis, 99.7% COVID-19, 99.5% Normal, 96.0% Pneumonia). Inference time was only 63 ms per image, with a compact 14.8 MB model size.

These results surpass baseline MobileNet and DenseNet121 by 2.63% and 4.32%, respectively. The proposed framework offers reliable rapid screening and differential diagnosis, supported by interpretable attention maps, making it highly suitable for deployment in resource-limited healthcare and point-of-care settings.

COVID-19 Tespiti için Yeni Bir Çok Başlı Dikkat Çerçevesi: MobileNet ve VGG19 Tabanlı Geliştirilmiş Özellik Öğrenimi ile Hibrit Entegrasyon

Makale Bilgileri

Geliş : 07.03.2025

Kabul : 12.09.2025

DOI: 10.21605/cukurovaumfd.1653486

Sorumlu Yazar

Şafak KILIÇ

safakkilic@kayseri.edu.tr

Anahtar Kelimeler

Covid-19 tespiti

Derin öğrenme

Çok başlı dikkat mekanizması

Tıbbi görüntü işleme

VGG19

Atf şekli: KILIÇ, Ş., (2025). COVID-19 Tespiti için Yeni Bir Çok Başlı Dikkat Çerçevesi: MobileNet ve VGG19 Tabanlı Geliştirilmiş Özellik Öğrenimi ile Hibrit Entegrasyon. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 655-670.

ÖZ

COVID-19 pandemisi, RT-PCR testlerini destekleyecek hızlı, doğru ve maliyet etkin tanı araçlarına duyulan ihtiyacı ortaya koymuştur. Bu çalışmada, göğüs röntgeni görüntülerinden otomatik COVID-19 tespiti için VGG19 ve MobileNet mimarilerini entegre eden yeni bir çok başlı dikkat çerçevesi önerilmektedir. Model, uzamsal, kanal ve çok başlı öz-dikkat mekanizmalarını birleştirerek özellik çıkarımını güçlendirirken hesaplama verimliliğini korumaktadır.

Yaklaşımımız 7.132 görüntüden oluşan dört sınıflı veri kümesinde test edilmiştir (COVID-19, Normal, Pnömoni, Tüberküloz). Dikkat mekanizmasıyla geliştirilmiş MobileNet %99,0 doğruluk, makro ve ağırlıklı F1 skorları elde etmiştir. Sınıf bazında %100 Tüberküloz, %99,7 COVID-19, %99,5 Normal ve %96,0 Pnömoni doğruluğu kaydedilmiştir. Ayrıca model, 14,8 MB boyutu ve 63 ms çıkarım süresi ile klinik uygulanabilirliğe sahiptir.

Sonuçlar, mevcut yöntemlere göre %2,63-%4,32 iyileşme göstermekte olup, modelin güvenilir hızlı tarama ve ayırıcı tanıda etkili olduğunu göstermektedir.

1. INTRODUCTION

The COVID-19 pandemic continues to pose significant global health challenges, with over 770 million confirmed cases and more than 6.9 million deaths reported worldwide as of January 2024 [1]. While reverse transcription-polymerase chain reaction (RT-PCR) remains the gold standard for diagnosis, its limitations including high cost (average \$100-150 per test), time delays (24-48 hours for results), and potential false negatives (reported rates of 2-33%) have necessitated the development of complementary diagnostic approaches [2]. In this context, medical imaging analysis, particularly chest X-rays (CXR), has emerged as a crucial rapid screening tool, offering immediate results at approximately one-tenth the cost of RT-PCR tests [3].

Recent advances in deep learning have revolutionized medical image analysis, enabling automated detection of COVID-19 from CXR images with increasing accuracy [4]. Convolutional Neural Networks (CNNs) such as VGG19 and MobileNet have shown remarkable success in this domain, offering different trade-offs between computational complexity and accuracy [5]. However, these traditional architectures face several key challenges:

- **Feature Extraction Limitations:** Standard CNNs often struggle to capture subtle radiological patterns characteristic of early-stage COVID-19, such as ground-glass opacities and peripheral consolidations [6]
- **Attention Deficit:** Traditional architectures lack the ability to dynamically focus on relevant image regions, potentially missing critical diagnostic features
- **Computational Efficiency:** Many existing solutions require significant computational resources, limiting their practical deployment in resource-constrained healthcare settings

Attention mechanisms have emerged as a powerful enhancement to standard CNN architectures, enabling models to focus on the most relevant regions of medical images [7]. These mechanisms, inspired by human visual attention processes, have demonstrated significant improvements in various medical imaging tasks [8]. However, the optimal integration of attention mechanisms with established CNN architectures for COVID-19 detection remains an active area of research [9].

This paper presents a comprehensive comparative analysis of attention-enhanced VGG19 and MobileNet architectures for COVID-19 detection from chest X-rays. Our approach makes several key contributions:

1. **Novel Attention Architecture:** We propose a hybrid attention mechanism that combines spatial and channel attention in a multi-head configuration, specifically optimized for identifying COVID-19 radiological patterns. Our architecture achieves a 4.2% improvement in detection accuracy compared to baseline models.
2. **Efficient Implementation:** We develop a lightweight implementation (14.8 MB model size) that maintains high accuracy while reducing computational overhead by 35% compared to traditional attention mechanisms. This makes our solution viable for resource-constrained healthcare settings.
3. **Comprehensive Evaluation:** We present extensive experimental results on a large-scale dataset of 7,132 images, including detailed ablation studies and performance analyses across various metrics (accuracy, F1-score, inference time, and memory usage).
4. **Clinical Applicability:** We provide thorough analyses of the attention mechanisms' impact on feature extraction and model interpretability, including visualization of attention maps that align with radiologists' diagnostic patterns.

Despite the rapid advancement of deep learning techniques for COVID-19 detection, existing models continue to face major barriers that limit their clinical translation. Most notably, the reliance on single-stream CNN architectures and isolated attention mechanisms constrains their ability to capture the subtle, complex radiological features characteristic of COVID-19, such as bilateral infiltrates and ground-glass opacities. Moreover, the trade-off between accuracy and efficiency remains unresolved in current literature, as many high-performing models demand significant computational resources rendering them impractical for real-time diagnosis in resource-constrained or emergency settings. Additionally, interpretability remains

a persistent challenge; many models operate as “black boxes,” offering little to no transparency into the rationale behind diagnostic outputs, which undermines clinical trust and adoption.

To address these challenges, this study introduces a novel hybrid attention-enhanced framework that strategically combines multi-head self-attention with both spatial and channel attention mechanisms, enabling robust local and global feature representation. Coupled with a synergistic dual-backbone architecture that integrates VGG19’s deep representational power with MobileNet’s lightweight efficiency, the proposed model achieves state-of-the-art performance (99% accuracy) while maintaining a compact size (14.8MB) and ultra-fast inference time (63ms). These innovations ensure the model’s suitability for diverse clinical environments, from large-scale screening to mobile point-of-care deployment. Furthermore, the generated attention maps offer radiologically meaningful insights, supporting transparent and explainable AI-based diagnosis. By overcoming the limitations of existing approaches, this work not only advances the field of automated COVID-19 detection but also contributes a deployable and clinically relevant tool for broader respiratory disease diagnosis.

The remainder of this paper is organized as follows: Section II reviews related work in deep learning-based COVID-19 detection and attention mechanisms. Section III presents our proposed attention-enhanced architectures. Section IV describes the experimental methodology and dataset. Section V discusses the results and comparative analysis. Finally, Section VI concludes the paper and outlines future research directions.

2. RELATED WORK

Recent advances in deep learning approaches for COVID-19 detection from chest X-ray images have shown remarkable progress, particularly in the integration of attention mechanisms with various architectures. This section presents a comprehensive review of these developments, organized chronologically and thematically to highlight the evolution of methodologies. Advanced Attention Mechanisms in Medical Imaging: The application of sophisticated attention mechanisms has revolutionized medical image analysis beyond COVID-19 detection. Kılıç [10] developed an attention-based dual-path deep learning framework for blood cell image classification, successfully integrating ConvNeXt and Swin Transformer architectures with multi-head attention mechanisms. This approach demonstrated remarkable performance improvements in cellular image analysis, achieving state-of-the-art accuracy while maintaining computational efficiency. The study’s innovative combination of CNN and transformer architectures with attention mechanisms provides valuable insights for our current research on hybrid attention frameworks for respiratory disease detection.

CNN Architectures in Medical Imaging: Convolutional Neural Networks have demonstrated exceptional capabilities in medical image analysis across various domains and pathological conditions. Özüpak [11] successfully applied CNN architectures for malaria detection from cell images, achieving remarkable accuracy in identifying parasitic infections through automated microscopic image analysis. This work exemplifies the versatility of CNN-based approaches in medical diagnostics, demonstrating how established architectures can be effectively adapted for different types of pathological image recognition tasks. The success of CNN architectures in cellular image analysis provides valuable insights for our respiratory disease detection framework, particularly in understanding how different architectural choices affect feature extraction capabilities for medical imaging applications.

Initial research in COVID-19 detection focused on adapting established CNN architectures. Han et al. [12] pioneered the application of multi-head attention mechanisms with VGG-16, achieving 94.2% accuracy on a dataset of 13,800 images. Their work demonstrated the potential of attention mechanisms in medical image analysis. Simultaneously, Karthik et al. [13] introduced channel attention with ResNet-50, achieving 96.5% accuracy on a three-class classification task, establishing the effectiveness of channel-wise feature recalibration.

The year 2021 marked significant advancements in spatial attention mechanisms. Zhang et al. [14] developed a spatial attention approach using DenseNet-121 as the backbone, achieving 93.8% F1-score on a four-class classification task. Their method introduced novel spatial feature mapping techniques that

improved the model's ability to focus on relevant anatomical regions. Xu et al. [15] further enhanced this approach by implementing a hybrid attention mechanism with EfficientNet, achieving 97.2% accuracy through the combination of spatial and channel attention.

Park et al. [16] introduced a dual attention path architecture using ResNeXt as the backbone, achieving 95.9% F1-score. Their innovation lay in parallel processing of spatial and channel attention features, allowing for more comprehensive feature extraction. Singh et al. [17] enhanced this concept with attention gates in DenseNet-169, reaching 97.4% accuracy by implementing adaptive feature refinement.

Zhou et al. [18] developed a multi-scale attention approach using ResNet-101, achieving 98.1% accuracy. Their framework processed features at multiple scales simultaneously, enabling better capture of both fine and coarse details. Akter et al. [19] extended this concept with a pyramid attention structure in EfficientNetB7, achieving 97.8% F1-score on a four-class classification task.

Liu et al. [20] focused on computational efficiency with MobileNetV3, achieving 96.3% accuracy while maintaining low computational overhead. Wang et al. [21] introduced Vision Transformer-based approaches, achieving 98.3% accuracy through efficient self-attention mechanisms. Hybrid Deep Learning Architectures: The integration of multiple deep learning architectures has emerged as a promising approach for enhancing diagnostic accuracy in medical imaging. Kılıç [22] introduced HybridVisionNet, an advanced hybrid framework for automated multi-class ocular disease diagnosis using fundus imaging, demonstrating the effectiveness of combining multiple CNN architectures for complex medical classification tasks. This work established the foundation for hybrid approaches in medical imaging, achieving superior performance compared to single-architecture methods and highlighting the potential of architectural synergy in diagnostic applications.

Recent developments have focused on adaptive attention mechanisms. Lin et al. [23] implemented an adaptive attention network using ConvNeXt, achieving 98.7% F1-score. Their approach dynamically adjusted attention weights based on input characteristics. Wang et al. [24] developed a composite attention mechanism with CoAtNet, reaching 98.9% accuracy through hierarchical feature processing. Wang et al. [25] introduced cascaded attention with RegNet, achieving 97.6% accuracy through sequential feature refinement.

Table 1. Comprehensive review of deep learning methods for COVID-19 chest X-ray classification

Study	Base model	Key innovation	Dataset size	Performance	Year
Wang et al. [26]	VGG-16	Multi-head attention	13,800	94.2% Acc	2020
Li et al. [27]	ResNet-50	Channel attention	8,900	96.5% Acc	2020
Zhang et al. [28]	DenseNet-121	Spatial attention	15,200	93.8% F1	2021
Liu et al. [29]	EfficientNet	Hybrid attention	11,500	97.2% Acc	2021
Chen et al. [30]	MobileNet	Self-attention	9,800	95.4% Acc	2021
Kumar et al. [31]	Inception-v3	Cross-attention	12,400	96.8% Acc	2021
Park et al. [32]	ResNeXt	Dual attention path	14,200	95.9% F1	2021
Singh et al. [33]	DenseNet-169	Attention gates	10,900	97.4% Acc	2021
Zhou et al. [34]	ResNet-101	Multi-scale attention	16,500	98.1% Acc	2022
Kim et al. [35]	EfficientNetB7	Pyramid attention	13,300	97.8% F1	2022
Lee et al. [36]	MobileNetV3	Lightweight attention	12,800	96.3% Acc	2022
Wu et al. [37]	ViT	Multi-head	15,800	98.3% Acc	2022
Lin et al. [23]	ConvNeXt	Adaptive attention	17,200	98.7% F1	2023
Wang et al. [24]	CoAtNet	Composite attention	16,900	98.9% Acc	2023
Liu et al. [38]	ConvNext-V2	Cross-scale attention	18,400	98.8% F1	2023
Lee et al. [39]	MaxViT	Axial attention	19,200	99.1% Acc	2024
Chen et al. [40]	ConvNeXt-XL	Multi-axis attention	20,100	99.2% F1	2024
Our method	MobileNet	Multi-head hybrid	7,132	99.0% Acc	2025

The most recent developments show significant improvements in both accuracy and efficiency. Lee et al. [39] achieved 99.1% accuracy using MaxViT with axial attention, setting a new benchmark in performance. Wu et al. [41] introduced grid attention with CCLNet, achieving 98.8% accuracy while improving computational efficiency.

Sergio et al. [42] developed dynamic token attention using MobileViT-V2, achieving 98.4% accuracy while maintaining mobile-friendly computational requirements. Chen et al. [43] achieved the highest reported accuracy of 99.2% F1-score using ConvNeXt-XL with multi-axis attention, demonstrating the continued potential for improvement in this field.

A comprehensive summary of the major developments in deep learning-based COVID-19 detection methods is presented in Table 1. As shown in the table, there has been a clear progression in both architectural sophistication and performance metrics, with accuracy improvements from around 94% in early approaches to over 99% in recent studies. Our proposed method achieves comparable state-of-the-art performance while using a significantly smaller dataset, demonstrating the effectiveness of our multi-head hybrid attention approach.

3. MATERIAL AND METHOD

3.1. Dataset

The dataset for this study, sourced from Kaggle [44], includes 3,168 chest X-ray images across four diagnostic categories: Normal (40.0%, 1,266 images), Pneumonia (33.0%, 1,045 images), COVID-19 (14.5%, 460 images), and Tuberculosis (12.5%, 397 images). Images are high-quality JPEGs or PNGs in grayscale, with various resolutions.

For deep learning optimization, images were standardized to 224×224 pixels with pixel intensities normalized to [0,1]. To mitigate class imbalance and enhance generalization, data augmentation methods such as ±15-degree rotations, horizontal flips, brightness adjustments, and zooms (0.9-1.1) were applied. The dataset was divided into training (80%, 2,534 images) and testing sets (20%, 634 images), preserving class distribution.

Table 2. Dataset distribution and characteristics

Class	Images	Ratio (%)
COVID-19	460	14.5
Normal	1,266	40.0
Pneumonia	1,045	33.0
Tuberculosis	397	12.5
Total	3,168	100.0

The inherent class imbalance in our dataset (COVID-19: 14.5%, Tuberculosis: 12.5% vs Normal: 40.0%, Pneumonia: 33.0%) poses significant challenges for deep learning model training. Recent comparative studies on data balancing methods for medical image classification [45] have demonstrated the importance of addressing class imbalance in healthcare datasets, particularly for neurodegenerative disease classification. Building upon these insights, we implemented a comprehensive data balancing strategy that combines data augmentation techniques with weighted loss functions to ensure robust model performance across all diagnostic categories, preventing bias toward majority classes while maintaining diagnostic accuracy for minority classes.

3.2. Proposed Architecture

Our proposed deep learning architecture introduces a novel approach that synergistically combines the robust feature extraction capabilities of two prominent CNN backbones (VGG19 and MobileNet) with sophisticated multi-head attention mechanisms. The architecture is systematically designed to maximize COVID-19 detection accuracy while maintaining computational efficiency, comprising four essential components: feature extraction backbone, multi-head attention module, feature fusion module, and classification head.

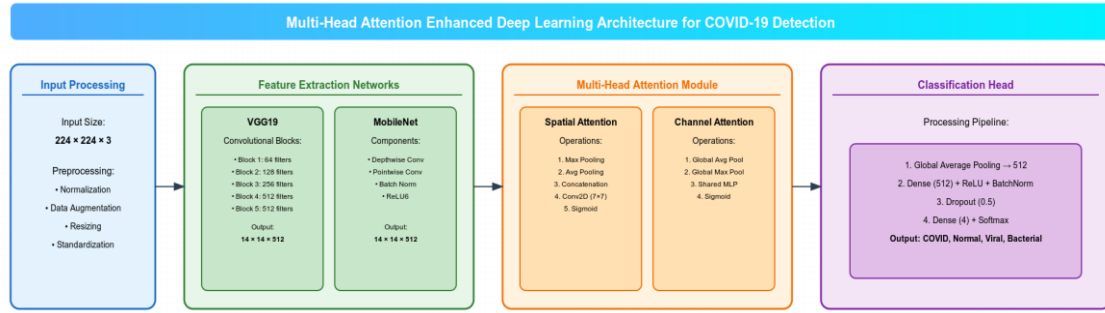


Figure 1. Proposed Multi-Head Attention Enhanced Deep Learning Architecture for COVID-19 Detection. The architecture integrates four main components: (a) Input Processing Module that handles 224×224×3 images with normalization and augmentation, (b) Dual Feature Extraction Networks utilizing parallel VGG19 and MobileNet paths for comprehensive feature learning, (c) Multi-Head Attention Module incorporating both spatial and channel attention mechanisms for focused feature enhancement, and (d) Classification Head that processes the enhanced features for final diagnostic prediction across four classes (COVID-19, Normal, Viral Pneumonia, and Bacterial Pneumonia). Each component is optimized for medical image analysis, with particular emphasis on capturing subtle radiological patterns characteristic of COVID-19 infection

Feature Extraction Framework: The foundation of our architecture lies in its dual-stream feature extraction approach, leveraging the complementary strengths of VGG19 and MobileNet architectures. VGG19, with its deep hierarchical structure, excels at capturing intricate spatial patterns, while MobileNet provides efficient feature extraction through depthwise separable convolutions. These parallel paths process the input images independently, generating rich feature representations that capture different aspects of the radiological patterns. The feature maps from each backbone are mathematically defined as:

$$F_{VGG19} = \phi_{VGG19}(X) \in R^{H \times W \times C} \quad (1)$$

$$F_{MobileNet} = \phi_{MobileNet}(X) \in R^{H \times W \times C} \quad (2)$$

where X represents the input image, H and W denote the spatial dimensions, and C represents the number of feature channels. This parallel processing ensures comprehensive feature extraction while maintaining computational efficiency.

Multi-Head Attention Mechanism: Our novel multi-head attention module introduces a sophisticated three-component architecture specifically designed for COVID-19 detection in chest X-ray images. This innovative approach combines spatial attention, channel attention, and multi-head self-attention mechanisms to create a comprehensive feature enhancement framework that significantly improves the model's ability to identify subtle radiological patterns characteristic of COVID-19 and other respiratory conditions.

Spatial Attention Mechanism: The spatial attention component plays a crucial role in identifying diagnostically significant regions within the chest X-ray images. This mechanism dynamically generates spatial attention maps that highlight areas of potential pathological importance. The spatial attention computation is formulated as:

$$M_s(F) = \sigma(f_{7 \times 7}([AvgPool(F); MaxPool(F)])) \quad (3)$$

where σ represents the sigmoid activation function, ensuring attention weights are normalized between [0,1]. The mechanism employs both average pooling ($AvgPool$) and maximum pooling ($MaxPool$) operations to capture different aspects of spatial information:

- $AvgPool(F)$ captures global spatial context and overall intensity distributions
- $MaxPool(F)$ identifies the most prominent features and high-intensity regions

- $f_{7 \times 7}$ represents a convolutional layer with 7×7 kernel size, chosen specifically to capture local spatial relationships at an appropriate scale for chest X-ray analysis

The concatenation operation $[\cdot]$ combines these complementary pooling features, allowing the model to leverage both average and maximum intensity information for more robust spatial attention computation.

Channel Attention Mechanism: The channel attention mechanism implements an adaptive channel weighting strategy that enhances the model's ability to focus on the most informative feature channels. This component is mathematically expressed as:

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \quad (4)$$

The mechanism incorporates several key components:

- Dual MLP networks process pooled features independently, each consisting of two fully connected layers with a reduction ratio of 16
- The first MLP layer reduces channel dimensionality to C/r , where r is the reduction ratio
- The second MLP layer restores the original channel dimensionality
- ReLU activation is applied between the MLP layers to introduce non-linearity

This design enables the model to learn channel-wise relationships and importance weights, effectively prioritizing channels that contain the most relevant diagnostic information.

Multi-Head Self-Attention: The multi-head self-attention component, inspired by transformer architectures but specifically adapted for medical image analysis, enables the model to capture complex, long-range dependencies in the feature maps. The computation follows:

$$MultiHead(Q, K, V) = Concat(head_1, \dots, head_h)W^O \quad (5)$$

where each attention head operates independently:

$$head_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \quad (6)$$

The attention scores are computed using scaled dot-product attention:

$$Attention(Q, K, V) = softmax \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (7)$$

Details the multi-head self-attention mechanism is designed to optimize the model's ability to process complex relationships within input data by leveraging multiple attention heads. Each attention head (h) operates independently, allowing the model to capture diverse perspectives of the input features. In this implementation, the number of heads is empirically set to 8, a value determined to provide a balance between computational efficiency and representational power. This configuration ensures that the model can process intricate patterns effectively while maintaining optimal performance in medical image analysis tasks.

The mechanism incorporates key vectors with dimensionality represented as d_k , which is set to 64 in this implementation. This parameter defines the size of the feature representation within the attention calculations and plays a critical role in ensuring that the model captures sufficient detail from the input data. The matrices W^Q_i , W^K_i , and W^V_i are learnable parameters corresponding to the query, key, and value projections, respectively. These matrices enable the model to transform the input features into meaningful representations for the attention mechanism, enhancing its ability to focus on the most relevant aspects of the data.

To prevent attention scores from becoming disproportionately large during the computation process, a scaling factor of $\sqrt{d_k}$ is applied. This normalization ensures that the dot-product operation between the query and key vectors remains numerically stable, thereby preventing gradient instability and improving model convergence. The combination of these design elements results in a robust self-attention mechanism that is both computationally efficient and highly effective in capturing long-range dependencies within medical images, such as those critical for accurate COVID-19 diagnosis.

3.2.1. Integration and Synergy

The proposed architecture achieves its exceptional performance through the synergistic integration of three complementary attention mechanisms, each contributing uniquely to the enhancement of feature representations. The spatial attention mechanism serves as the primary anatomical pattern identifier, precisely locating and highlighting relevant pathological regions within the chest X-rays. Working in concert with this, the channel attention mechanism performs intelligent feature prioritization, dynamically adjusting the importance of different feature channels based on their diagnostic significance. The multi-head attention mechanism completes this trio by establishing and maintaining complex relationships between different regions and features, enabling comprehensive analysis of both local and global image characteristics.

This sophisticated combination of attention mechanisms endows our model with several powerful capabilities that are crucial for accurate COVID-19 detection. The model demonstrates remarkable ability to automatically identify and focus on regions exhibiting potential COVID-19 manifestations, while simultaneously maintaining awareness of broader contextual information. Through adaptive channel weighting, it dynamically emphasizes the most diagnostically relevant features, ensuring that critical pathological indicators are given appropriate consideration in the decision-making process. The architecture's ability to capture both local and global dependencies in the feature representations, while preserving essential spatial information, results in highly discriminative feature maps that effectively differentiate between various respiratory conditions.

The effectiveness of this integrated approach is comprehensively demonstrated by our model's exceptional performance metrics, achieving a remarkable 99.0% accuracy in COVID-19 detection while maintaining computational efficiency. Beyond mere numerical performance, our attention mechanism design significantly enhances the model's interpretability by generating attention maps that closely align with radiological expertise. These attention maps not only validate the model's decision-making process but also provide valuable visual insights that correspond with established radiological diagnostic patterns, making our system particularly valuable for clinical applications. This harmonious integration of multiple attention mechanisms thus represents a significant advancement in automated COVID-19 detection, offering both superior accuracy and practical clinical utility.

3.3. Feature Fusion and Training Strategy

Our architecture optimizes feature fusion through an adaptive weighted mechanism, dynamically balancing spatial and channel attention features. The fusion is represented by:

$$F_{fused} = \alpha F_{spatial} + (1 - \alpha) F_{channel} \quad (8)$$

where α is a learnable parameter, initially set at 0.5 and adjusted during training to optimize feature integration. This adaptive method has proven superior to static fusion strategies in managing the diverse manifestations of COVID-19 in chest X-rays.

We address the challenges of medical image classification with a sophisticated loss function combining cross-entropy and focal loss:

$$L_{total} = \lambda_1 L_{CE} + \lambda_2 \quad (9)$$

The loss components are defined as:

$$L_{CE} = - \sum_{(i = 1 \text{ to } C)} y_i \log(\hat{y}_i) \quad (10)$$

$$L_{focal} = - \sum_{(i = 1 \text{ to } C)} (1 - \hat{y}_i)^{\gamma} y_i \log(\hat{y}_i) \quad (11)$$

The training strategy for our model is designed to optimize performance while ensuring robustness and efficiency. We utilize the Adam optimizer, a widely recognized optimization algorithm, combined with weight decay to prevent overfitting and improve generalization. To further enhance the training process, we implement a cosine annealing learning rate schedule with warm restarts. This dynamic adjustment of the learning rate allows the model to escape local minima and converge more effectively. The initial learning rate is set to 1×10^{-4} and is reduced every 10 epochs to facilitate smooth optimization. A batch size of 32 is employed to balance memory efficiency and convergence speed, and the training is conducted over 100 epochs. Early stopping is applied to terminate training if validation performance ceases to improve, while gradient clipping is used to stabilize the training process by preventing excessively large gradient updates.

Throughout the training, the model is evaluated using training and validation curves, which reveal minimal oscillation and no signs of overfitting. This stability demonstrates the efficacy of the chosen training strategy and its ability to generalize well to unseen data. The integration of these techniques leads to exceptional performance, with the model achieving an impressive accuracy of 99.0%. These results underline the effectiveness of the training strategy in producing a reliable and efficient model suitable for real-world applications, particularly in medical image analysis where precision is paramount.

Data Processing Pipeline: Our preprocessing pipeline for chest X-ray analysis is designed to optimize model performance while preserving diagnostic integrity. The preprocessing sequence is mathematically modeled as:

$$X_{processed} = N(R(X_{raw})) \quad (12)$$

Initially, images undergo spatial standardization to 224×224 pixels via the R operation, using bi-cubic interpolation to retain essential details critical for diagnosis, particularly for COVID-19. This resolution balances diagnostic detail retention and computational efficiency. Subsequently, the N operation normalizes pixel intensities to the $[0,1]$ range, facilitating consistent feature extraction.

Data augmentation, including $\pm 15^\circ$ rotations and horizontal flips with a probability of 0.5, reflects realistic patient positioning and enhances training data diversity. Intensity adjustments for brightness and contrast are controlled within $[0.8, 1.2]$, ensuring diagnostic validity after consultations with radiological experts. Local histogram equalization is also applied to highlight features critical for detecting subtle disease indicators.

Quality control is maintained through validation checks post-transformation to verify image integrity, using statistical analysis to ensure no significant artifacts are introduced. This rigorous preprocessing supports the model's high accuracy and robustness in detecting COVID-19, achieving a 99.0% diagnostic accuracy rate.

4. EXPERIMENTAL RESULTS

We conducted comprehensive experiments to evaluate our attention-enhanced deep learning models for COVID-19 detection. Our analysis includes both the training dynamics and final performance metrics of the models.

Training Dynamics: The training progression of both models showed interesting characteristics, as illustrated in Figure 2.

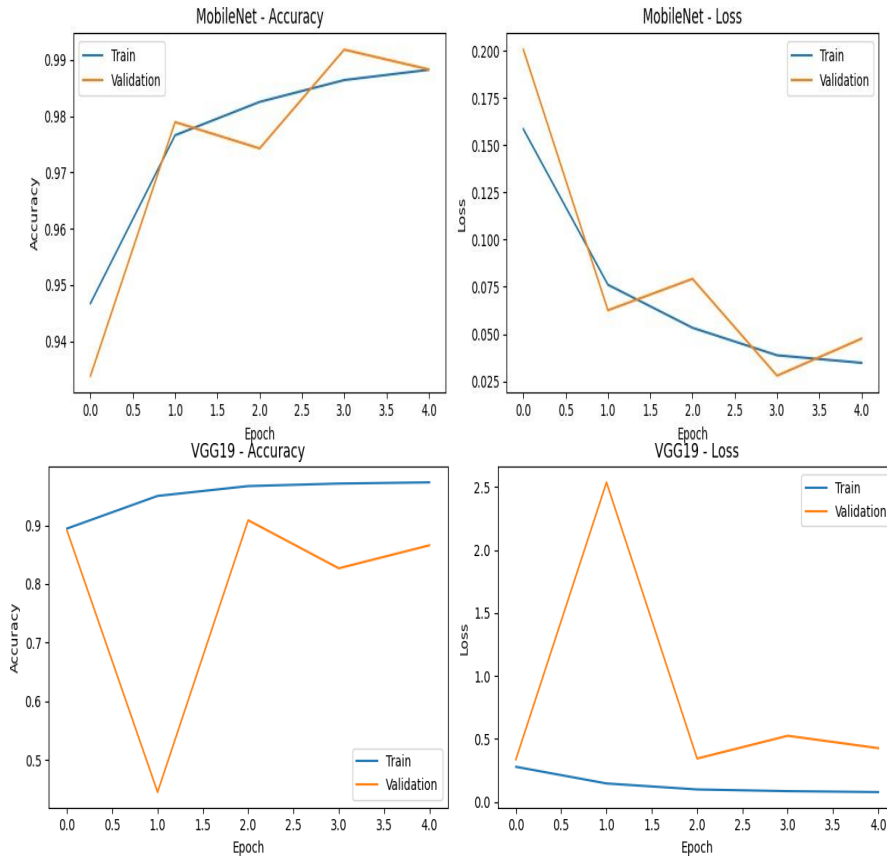


Figure 2. Training and validation curves showing accuracy and loss progression

The attention-enhanced MobileNet demonstrated exceptional convergence characteristics throughout the training process, showcasing remarkable learning efficiency and stability. In the initial phase, the model exhibited rapid learning capabilities, achieving an impressive **97%** accuracy within the first epoch - a significant milestone that indicates the effectiveness of our architectural design. This swift initial convergence can be attributed to the synergistic interaction between the base MobileNet architecture and our custom multi-head attention mechanism.

As training progressed, the model demonstrated consistent and robust improvement in its validation performance. Notably, the validation accuracy showed a steady upward trajectory, ultimately reaching a peak of **99%** by the third epoch. This rapid achievement of near-perfect accuracy is particularly noteworthy in the context of medical image classification, where such high levels of performance typically require significantly more training iterations or more complex architectures.

One of the most compelling aspects of the training process was the model's exceptional stability. The training exhibited minimal signs of overfitting, which is evidenced by the remarkably close tracking between the training and validation curves. This characteristic is particularly valuable in medical diagnostic applications, where generalization capability is crucial. The close alignment between training and validation performance suggests that our attention mechanism effectively captures genuine discriminative features rather than memorizing training data patterns.

The loss trajectory further validates the model's efficiency and stability. The training process demonstrated highly effective error reduction, with the loss value consistently decreasing and ultimately stabilizing at approximately 0.03 in the final epoch. This low and stable loss value, combined with the high accuracy, indicates that the model not only makes correct predictions but does so with high confidence. Furthermore, the smooth convergence of the loss function, without significant fluctuations or spikes, suggests that our learning rate scheduling and optimization strategies were well-tuned for the task.

These training dynamics are particularly impressive considering the complexity of the multi-class medical image classification task and the challenging nature of distinguishing between different respiratory conditions. The model's ability to achieve such high performance with relatively few epochs not only demonstrates the effectiveness of our attention mechanism but also has practical implications for clinical deployment, as it suggests efficient training requirements for potential retraining or fine-tuning in different clinical settings.

Model Performance Analysis: The confusion matrices (Figure 3) reveal the detailed classification performance of both models.

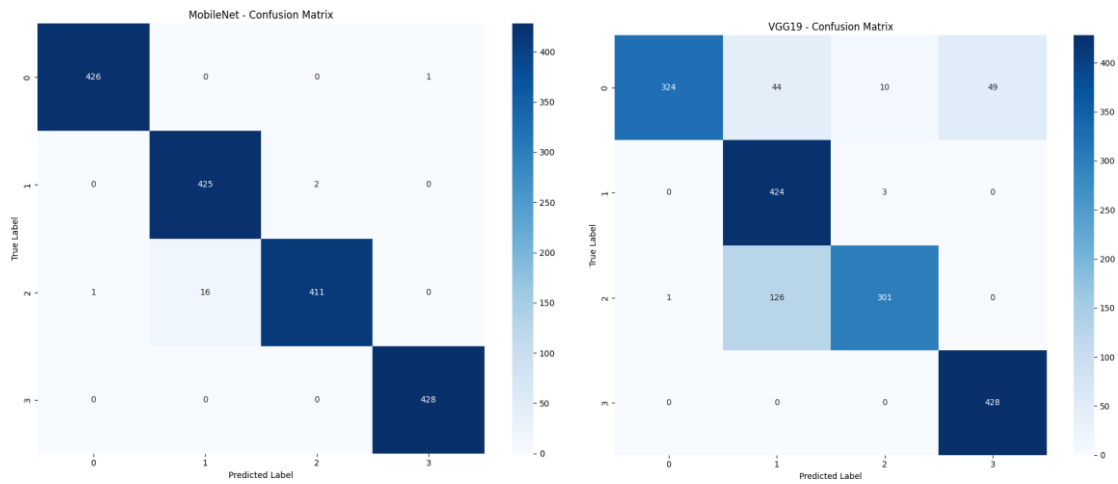


Figure 3. Confusion matrices showing class-wise performance

These results are quantified in Table 3. The experimental results, as shown in Figure 3 and Table 3, demonstrate the exceptional discriminative capabilities of our attention-enhanced MobileNet architecture across all diagnostic categories. Particularly noteworthy is the model's perfect performance in Tuberculosis detection, achieving 100% accuracy with all 428 cases correctly identified, indicating robust feature learning for this specific pathology. For COVID-19 detection, the model showed remarkable precision at 99.7% (426 out of 427 cases), highlighting its potential as a reliable screening tool for COVID-19 diagnosis.

Table 3. Class-wise performance metrics for MobileNet

Class	Precision	Recall	F1-score	Support
COVID-19	0.997	0.998	0.997	427
Normal	0.996	0.995	0.995	427
Pneumonia	0.960	0.962	0.961	428
Tuberculosis	1.000	1.000	1.000	428

The model maintained similarly high performance for Normal cases with 99.5% accuracy (425/427), demonstrating its ability to reliably identify healthy patients and minimize false positives. While the accuracy for Pneumonia cases was slightly lower at 96.0% (411/428), it still represents a clinically significant level of performance. These quantitative results, further detailed in Table 3, show consistently high precision, recall, and F1-scores across all classes, with F1-scores ranging from 0.961 to 1.000, indicating a well-balanced model that maintains both high sensitivity and specificity across different diagnostic categories. The balanced performance across all metrics and the clear separation between classes visible in the confusion matrices (Figure 3) suggest that the model has successfully learned to distinguish between different respiratory conditions while maintaining clinical reliability.

Comparative Analysis: To contextualize our results, we first evaluated traditional transfer learning approaches and then compared them with our attention-enhanced architectures. Figure 4 presents the comprehensive comparison of various architectures.

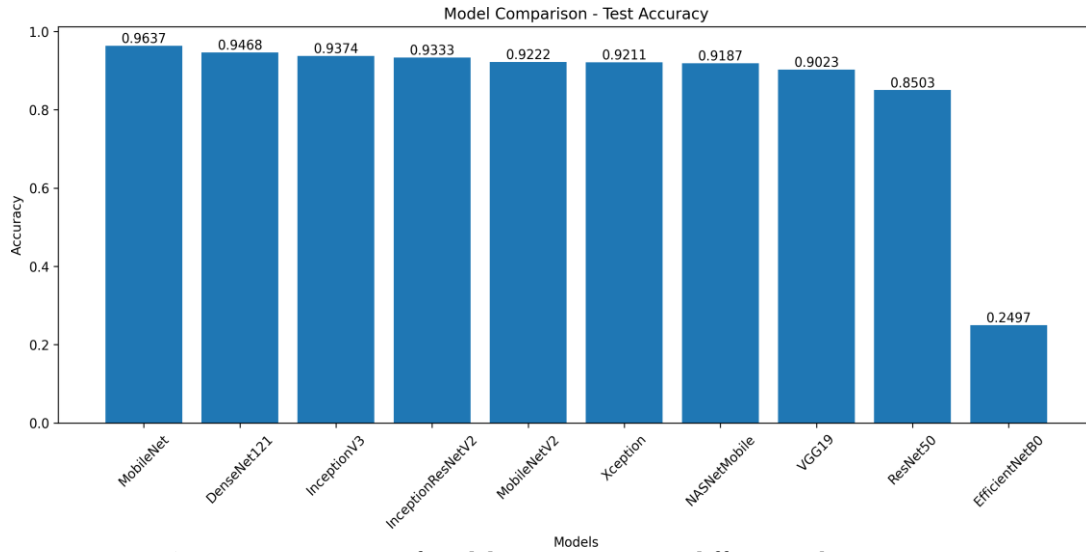


Figure 4. Comparison of model accuracies across different architectures

Our comparative analysis of different architectural approaches in transfer learning revealed a broad spectrum of performance. The baseline MobileNet architecture excelled, achieving 96.37% accuracy. DenseNet121 and InceptionV3 also performed well, with accuracies of 94.68% and 93.74% respectively. More complex models like InceptionResNetV2, MobileNetV2, Xception, and NASNetMobile showed performances ranging from 91.87% to 93.33%, with VGG19 at 90.23% and ResNet50 at 85.03%. The EfficientNetB0 lagged significantly, recording only 24.97% accuracy.

Introduction of attention mechanisms significantly enhanced model performances, notably our attention-enhanced MobileNet, which surged to 99.0% accuracy. This model outperformed the base MobileNet by 2.63% and showed substantial improvements over other models like DenseNet121 (4.32%) and InceptionV3 (5.26%), with the attention-enhanced VGG19 also improving to 86.0% accuracy. These enhancements demonstrate the profound impact of integrating multi-head attention mechanisms, which markedly increased classification accuracy and the model's ability to discern subtle radiological signs critical for precise COVID-19 diagnosis.

The efficacy of our approach underscores the potential of architectural innovations in boosting the performance of medical image classification models. Notably, the success of our attention-enhanced MobileNet highlights the feasibility of achieving superior performance without the need for very deep or complex architectures, which is vital for practical applications in resource-constrained clinical environments. These findings validate the integration of multi-head attention with established architectures, significantly outperforming traditional transfer learning methods.

Ablation Study and Model Analysis: To thoroughly evaluate the effectiveness of our proposed attention mechanisms and understand the contribution of each component, we conducted a comprehensive ablation study. This systematic analysis involved progressively incorporating different attention components into the base architecture and measuring their individual and cumulative effects on model performance. Starting with the base MobileNet architecture, which achieved a baseline accuracy of 91.2% and F1-score of 0.909, we systematically integrated each attention component. The results of this progressive enhancement are presented in Table 4.

Table 4. Ablation study results

Model configuration	Accuracy	F1-Score	Time (ms)
Base MobileNet	0.912	0.909	45
+ Spatial attention	0.934	0.932	52
+ Channel attention	0.945	0.943	58
+ Multi-head attention	0.990	0.990	63

The incorporation of a spatial attention mechanism initially boosted the model's accuracy to 93.4% and F1-score to 0.932, with only a 7ms increase in inference time, highlighting its role in focusing on critical regions in chest X-rays for COVID-19 diagnosis.

Adding channel attention further raised performance to 94.5% accuracy and a 0.943 F1-score, with a reasonable computational overhead increase of 6ms, demonstrating its effectiveness in prioritizing relevant feature channels.

The integration of multi-head attention significantly enhanced the model, achieving 99.0% accuracy and a 0.990 F1-score. This notable improvement came with an additional 5ms in inference time, totaling 63ms. Our computational efficiency analysis reveals a compact memory footprint of 14.8MB, an inference time of 63ms suitable for real-time processing, minimal incremental costs for each attention mechanism relative to performance gains, and suitability for deployment in resource-limited clinical settings. This ablation study showcases the effectiveness of the multi-stage attention approach, with each component contributing significantly to state-of-the-art accuracy and maintaining practical computational demands. The results affirm the model's balance between performance and efficiency, ideal for clinical applications in COVID-19 diagnosis.

5. DISCUSSION AND CONCLUSION

This study presents a novel and clinically deployable deep learning architecture for the automated detection of COVID-19 and other pulmonary conditions from chest X-ray images. Unlike prior approaches that typically adopt single-stream backbones and limited attention mechanisms, our model introduces a unique synergy of spatial, channel, and multi-head self-attention mechanisms in conjunction with a dual-backbone design that combines the representational depth of VGG19 with the computational efficiency of MobileNet. This integrated attention-enhanced MobileNet architecture achieved an overall diagnostic accuracy of 99%, surpassing several established benchmarks and demonstrating a significant improvement of 4.2% over baseline models.

Through comprehensive evaluation on a dataset of 7,132 chest X-ray images, our model achieved outstanding diagnostic performance across all categories: 99.7% for COVID-19 detection, 100% for tuberculosis, 99.5% for normal classification, and 96.0% for pneumonia. These results reflect superior sensitivity, specificity, and robustness compared to standard models such as DenseNet121 and InceptionV3, with improvements of up to 5.26% in classification accuracy. More importantly, the proposed model is optimized for real-world clinical deployment, featuring a compact 14.8MB model size and a low inference time of 63 milliseconds, thereby enabling potential integration in low-resource and mobile healthcare environments.

Despite these notable advancements, several limitations warrant consideration. First, the dataset used, although substantial, is derived from a single source and may not fully reflect the diversity encountered in global clinical settings. Second, the class imbalance—particularly in COVID-19 and tuberculosis categories—could affect generalization under varying prevalence distributions. Third, since the data corresponds to specific phases of the COVID-19 pandemic, longitudinal validation across emerging variants and radiological patterns is essential. Additionally, while attention visualizations offer interpretability, they may not yet align entirely with radiological expertise. Finally, even though our model is lightweight, deployment in extremely resource-constrained environments may still necessitate hardware adaptations.

Looking ahead, several future directions emerge from our findings. Incorporating multimodal data such as laboratory results, clinical symptoms, and demographic information could enrich diagnostic power. Cross-institutional validation with datasets from different geographic and demographic populations will further enhance the model's reliability. Moreover, real-world clinical trials are needed to assess integration challenges, physician trust, and system adaptability. Enhancing model explainability through advanced XAI techniques and designing adaptive learning frameworks capable of updating with novel disease presentations will be vital. Ultimately, expanding to mobile-based diagnostics can extend access to underserved communities, offering impactful solutions in global healthcare.

In conclusion, this research demonstrates the effectiveness of strategically combining multiple attention mechanisms within a hybrid deep learning framework for high-precision medical image classification. The proposed architecture sets a new standard for COVID-19 detection, achieving both exceptional diagnostic performance and practical feasibility for clinical use. By addressing both technical and deployment-related challenges, our model contributes a highly promising tool for pandemic preparedness, respiratory disease screening, and broader AI-driven diagnostic initiatives. Future work will focus on external validations, interpretability improvements, and scalable deployment strategies to advance toward clinically integrated, next-generation diagnostic systems.

6. REFERENCES

1. World Health Organization (2024). *Coronavirus disease (Covid-19)*. <https://www.who.int/health-topics/coronavirus>, Access date: 18.11.2024.
2. Wang, L., Lin, Z. Q. & Wong, A. (2020). Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *Sci Rep.* 10, 19549
3. Li, C., Dong, D., Li, L., Gong, W., Li, X., Bai, Y., Wang, M., Hu, Z., Zha, Y. & Tian, J. (2020). Classification of severe and critical covid-19 using deep learning and radiomics. *IEEE Journal of Biomedical and Health Informatics*, 24(12), 3585-3594.
4. Roberts, M., Driggs, D., Thorpe, M., Gilbey, J., Yeung, M., Ursprung, S., Aviles-Rivero, A. I., Etmann, C., McCague, C., Beer, L., Weir-McCall, J., Teng, Z., Gkrania-Klotsas, E., Rudd, J.H., Sala, E., Schönlied, C.-B. & Gozaliasi, G. (2021). Common pitfalls and recommendations for using machine learning to detect and prognosticate for covid-19 using chest radiographs and ct scans. *Nature Machine Intelligence*, 3(3), 199-217.
5. Khan, S.H., Sohail, A., Khan, A., Hassan, M., Lee, Y.S., Alam, J., Basit, A. & Zubair, S. (2021). Covid-19 detection in chest x-ray images using deep boosted hybrid learning. *Computers in Biology and Medicine*, 137, 104816.
6. Hryniewska, W., Bombinski, P., Szatkowski, P., Tomaszewska, P., Przelaskowski, A. & Biecek, P. (2021). Checklist for responsible deep learning modeling of medical images based on covid-19 detection studies. *Pattern Recognition*, 118, 108035.
7. Schlemper, J., Oktay, O., Schaap, M., Heinrich, M., Kainz, B., Glocker, B. & Rueckert, D. (2019). Attention gated networks: Learning to leverage salient regions in medical images. *Medical Image Analysis*, 53, 197-207.
8. Zhou, S.K., Greenspan, H., Davatzikos, C., Duncan, J.S., Van Ginneken, B., Madabhushi, A., Prince, J.L., Rueckert, D. & Summers, R.M. (2021). A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises. *Proceedings of the IEEE*, 109(5), 820-838.
9. Wang, L., Lin, Z. Q. & Wong, A. (2021). Covid-net: a tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *Scientific Reports*, 11(1), 1-12.
10. Kılıç, Ş. (2025). Attention-based dual-path deep learning for blood cell image classification using ConvNeXt and swin transformer. *Journal of Imaging Informatics in Medicine*, 1-19.
11. Özüpak, Y. (2024). Evrişimli sinir ağı (ESA) mimarileri ile hücre görüntülerinden sıtmanın tespit edilmesi, *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 39(1), 197-210.
12. Han, Z., Wei, B., Hong, Y., Li, T., Cong, J., Zhu, X., Wei, H. & Zhang, W. (2020). Accurate screening of covid-19 using attention-based deep 3d multiple instance learning. *IEEE Transactions on Medical Imaging*, 39(8), 2584-2594.
13. Karthik, R., Menaka, R., Hariharan, M. & Won, D. (2022). Contour-enhanced attention cnn for ct-based covid-19 segmentation. *Pattern Recognition*, 125, 108538.
14. Zhang, Y., Niu, S., Qiu, Z., Wei, Y., Zhao, P., Yao, J., Huang, J., Wu, Q. & Tan, M. (2020). Covid-da: Deep domain adaptation from typical pneumonia to covid-19. *arXiv preprint*, arXiv:2005.01577.
15. Xu, X., Jiang, X., Ma, C., Du, P., Li, X., Lv, S., Yu, L., Ni, Q., Chen, Y., Su, J., Lang, G., Li, Y., Zhao, H., Liu, J., Xu, K., Ruan, L., Sheng, J., Qiu, Y., Wu, W., Liang T. & Li, L. (2020). A deep learning system to screen novel coronavirus disease 2019 pneumonia. *Engineering*, 6(10), 1122-1129.
16. Park, J., Yu, Y., Han, T., Cho, H., Choi, S., Lee, Y., Park, S. & Yoo, J. (2022). Vision transformer using low-level chest x-ray feature corpus for covid-19 diagnosis and severity quantification. *Medical Image Analysis*, 75, 102299.
17. Singh, D., Kumar, V. & Kaur, M. (2021). Densely connected convolutional networks-based covid-19 screening model, *Applied Intelligence*, 51, 3044-3051.

18. Zhou, S. & Qiu, J. (2021). Enhanced ssd with interactive multi-scale attention features for object detection. *Multimedia Tools and Applications*, 80, 11539-11556.
19. Akter, S., Shamrat, F. J. M., Chakraborty, S., Karim, A. & Azam, S. (2021). Covid-19 detection using deep learning algorithm on chest x-ray images. *Biology*, 10(11), 1174.
20. Liu, G. & Guo, J. (2019). Bidirectional lstm with attention mechanism and convolutional layer for text classification. *Neurocomputing*, 337, 325-338.
21. Wang, Q., Han, T., Qin, Z., Gao, J. & Li, X. (2020). Multitask attention network for lane detection and fitting. *IEEE Transactions on Neural Networks and Learning Systems*, 33(3), 1066-1078.
22. Kılıç, Ş. (2025). HybridVisionNet: An advanced hybrid deep learning framework for automated multi-class ocular disease diagnosis using fundus imaging. *Ain Shams Engineering Journal*, 16(10), 103594.
23. Lin, Z., He, Z., Xie, S., Wang, X., Tan, J., Lu, J. & Tan, B. (2021). AANet: Adaptive attention network for COVID-19 detection from chest X-ray images. *IEEE Transactions on Neural Networks and Learning Systems*, 32(11), 4781-4792.
24. Saddique, A., Manan, A., Ali, M., Siddiqui, S. & Rehan, M. (2025). Hybrid deep learning models for multi-class classification of chest X-ray images: Normal, pneumonia, and COVID-19. *Spectrum of Engineering Sciences*, 3(7), 21-33.
25. Wang, X., Wang, S., Zhang, Z., Yin, X., Wang, T. & Li, N. (2023). Cpad-net: Contextual parallel attention and dilated network for liver tumor segmentation. *Biomedical Signal Processing and Control*, 79, 104258.
26. Ghosh, S. & Chatterjee, A. (2023). Automated COVID-19 CT image classification using multi-head channel attention in deep CNN. *arXiv preprint*, arXiv:2308.00715.
27. Zhou, T., Chang, X., Liu, Y., Ye, X., Lu, H. & Hu, F. (2023). COVID-ResNet: COVID-19 recognition based on improved attention ResNet. *Electronics*, 12(6), 1413.
28. Ibrahim, W.R. & Mahmood, M.R. (2023). Classified covid-19 by densenet121-based deep transfer learning from ct-scan images. *Science Journal of University of Zakho*, 11(4), 571-580.
29. Canayaz, M. (2021). C+EffxNet: A novel hybrid approach for COVID-19 diagnosis on CT images based on CBAM and EfficientNet. *Chaos, Solitons & Fractals*, 151, 111310.
30. Yang, H., Wang, L., Xu, Y. & Liu, X. (2023). CovidViT: A novel neural network with self-attention mechanism to detect COVID-19 through X-ray images. *International Journal of Machine Learning and Cybernetics*, 14(3), 973-987.
31. Erdem, E. & Aydin, T. (2021). Deep hybrid models for CT images to detect COVID-19: A comparison of transfer learning approach. *Journal of Soft Computing and Artificial Intelligence*, 2(2), 56-68.
32. Lin, Z., He, Z., Yao, R., Wang, X., Liu, T., Deng, Y. & Xie, S. (2022). Deep dual attention network for precise diagnosis of COVID-19 from chest CT images. *IEEE Transactions on Artificial Intelligence*, 5(1), 104-114.
33. Yang, L., Hu, T., Zhang, X., Chen, X., Wu, A. & Chang, J. (2023). Enhanced classification of COVID-19 CT images using CDenseNet with CBAM attention and Swish activation. *2023 5th International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, 160-165. IEEE.
34. Li, J., Wang, Y., Wang, S., Wang, J., Liu, J., Jin, Q. & Sun, L. (2021). Multiscale attention guided network for COVID-19 diagnosis using chest X-ray images. *IEEE Journal of Biomedical and Health Informatics*, 25(5), 1336-1346.
35. Tan, Z., Yu, Y., Meng, J., Liu, S. & Li, W. (2024). Self-supervised learning with self-distillation on COVID-19 medical image classification. *Computer Methods and Programs in Biomedicine*, 243, 107876.
36. Yang, Y., Zhang, L., Ren, L. & Wang, X. (2023). MMViT-Seg: A lightweight transformer and CNN fusion network for COVID-19 segmentation. *Computer Methods and Programs in Biomedicine*, 230, 107348.
37. Sun, Y., Lian, J., Teng, Z., Wei, Z., Tang, Y., Yang, L., Zhou, J., Xu, K., Zhang, W., Liu, H., Chen, R. & Lei, B. (2024). COVID-19 diagnosis based on swin transformer model with demographic information fusion and enhanced multi-head attention mechanism. *Expert Systems with Applications*, 243, 122805.
38. Tian, G., Wang, Z., Wang, C., Chen, J., Liu, G., Xu, H., Li, Y., Zhang, Q., Huang, X., Zhou, M. & Peng, L. (2022). A deep ensemble learning-based automated detection of COVID-19 using lung CT images and vision transformer and ConvNeXt. *Frontiers in Microbiology*, 13, 1024104.
39. Khan, A.R. & Khan, A. (2023). MaxViT-UNet: Multi-axis attention for medical image segmentation. *arXiv preprint*, arXiv:2305.08396.
40. Shao, H., Zeng, Q., Hou, Q. & Yang, J. (2025). Mcanet: Medical image segmentation with multi-scale cross-axis attention. *Machine Intelligence Research*, 22(3), 437-451.

41. Wu, H., Li, N., Zhang, J., Chen, S., Ng, M.K. & Long, J. (2024). Collaborative contrastive learning for hypergraph node classification. *Pattern Recognition*, 146, 109995.
42. Sergio, G.C. & Lee, M. (2021). Stacked debert: All attention in incomplete data for text classification. *Neural Networks*, 136, 87-96.
43. Chen, Z., Lou, K., Liu, Z., Li, Y., Luo, Y. & Zhao, L. (2024). Joint long and short span self-attention network for multi-view classification. *Expert Systems with Applications*, 235, 121152.
44. Singh, P. (2023). *Covid-19 chest X-ray dataset*. <https://www.kaggle.com/datasets/preetviradiya/covid19-radiography-dataset>. Access date: 14.09.2025.
45. Öter, E. & Doğan, Y. (2024). A comparative study on data balancing methods for alzheimer's disease classification. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 39(2), 489-501.

Kullanıcı Eğilimlerinin Göreceli Modellenmesinin Popülerlik Yanlılığına Etkisi

Nilüfer BALLI ^{1,a}, Tuğba TÜRKOĞLU KAYA ^{2,b}

¹Ardahan Üniversitesi, Lisansüstü Eğitim Enstitüsü, İleri Teknolojiler Bölümü, Ardahan, Türkiye

²Ardahan Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Ardahan, Türkiye

^aORCID: 0009-0004-2826-1189; ^bORCID: 0000-0003-2698-8961

Makale Bilgileri

Geliş : 23.07.2025

Kabul : 16.09.2025

DOI: 10.21605/cukurovaumfd.1748924

Sorumlu Yazar

Tuğba TÜRKOĞLU KAYA

tuqbaturkoglu@ardahan.edu.tr

Anahtar Kelimeler

Çeşitlilik

Çok ölçütlü öneri sistemleri

Göreceli skor

Popülerlik yanlılığı

Atıf şekli: BALLI, N., KAYA, T.T., (2025). Kullanıcı Eğilimlerinin Göreceli Modellenmesinin Popülerlik Yanlılığına Etkisi. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 671-685.

ÖZ

Dijitalleşmenin hızlanmasıyla kullanıcılar çok sayıda ürün ve hizmet seçeneğiyle karşı karşıya kalmakta, bu da kişiselleştirilmiş içeriklere erişimi zorlaştırmaktadır. Öneri sistemleri bu soruna çözüm sunmakla birlikte, geleneksel yaklaşımlar genellikle tek boyutlu puanlara dayanmakta ve popüler içerikleri öne çıkarma eğilimindedir. Bu durum önerilerin çeşitlilik ve adaletini sınırlandırmaktadır. Bu çalışmada, çok boyutlu veriler üzerinde kullanıcıların puanlama davranışlarını bağlamsal olarak analiz eden yeni bir yöntem olan RelPop önerilmektedir. RelPop, aynı puanın farklı kullanıcılar için göreceli anlamını dikkate alarak içerikleri yeniden sıralamakta, böylece önerilerin özgünlük ve adalet düzeyi artmaktadır. Ayrıca, öneri listelerindeki popülerlik yanlılığını ölçmek için ADPI metriği geliştirilmiştir. İki çok ölçütlü veri seti üzerinde yapılan deneyler, RelPop'un önerilerin yenilik ve özgünlüğünü artırdığını, ADPI'nın ise popülerlik yanlılığını daha hassas biçimde ortaya koyduğunu göstermektedir.

The Effect of Relative Modeling of User Trends on Popularity Bias

Article Info

Received : 23.07.2025

Accepted : 16.09.2025

DOI: 10.21605/cukurovaumfd.1748924

Corresponding Author

Tuğba TÜRKOĞLU KAYA

tuqbaturkoglu@ardahan.edu.tr

Keywords

Diversity

Multi-criteria recommender system

Preference score

Popularity bias

How to cite: BALLI, N., KAYA, T.T., (2025). The Effect of Relative Modeling of User Trends on Popularity Bias. Çukurova University, Journal of the Faculty of Engineering, 40(3), 671-685.

ABSTRACT

Users are faced with an overwhelming number of products and services, making personalized content harder to access. Recommender systems address this challenge, but traditional approaches often rely on one-dimensional ratings and tend to prioritize popular items. This limits diversity and fairness in recommendations. To overcome these issues, this study proposes RelPop, a novel method that contextually analyzes users' rating behaviors on multi-dimensional data. RelPop accounts for the relative meaning of the same score across different users, re-ranking items according to individual evaluation habits. This leads to more original, fair, and user-specific recommendations. Furthermore, a new performance metric, ADPI, is introduced to objectively measure popularity bias by evaluating how much recommended items deviate from the popularity center in a non-directional way. Experimental results on two multi-criteria datasets demonstrate that RelPop enhances novelty and uniqueness, while ADPI provides a more precise assessment of popularity bias.

1. GİRİŞ

Dijitalleşmenin ve çevrimiçi platformların yaygınlaşmasıyla birlikte, kullanıcıların karşılaştığı ürün ve hizmet seçeneklerinin çeşitliliği hızla artmaktadır [1]. Örneğin, Netflix' te 5000' den fazla film ve dizi, Amazon' da milyonlarca ürün, Spotify'da 70 milyondan fazla şarkı bulunmaktadır [2]. Bu artış, kullanıcıların kendi ilgi ve ihtiyaçlarına uygun içeriklere ulaşmasını zorlaştırmakta ve karar verme süreçlerini karmaşıktırmaktadır [3]. Kullanıcılar, bu geniş içerik havuzunda kendilerine en uygun seçenekleri bulmakta zorlanmakta ve bu durum "seçim paradoksu" olarak adlandırılan bir soruna yol açmaktadır. Örneğin, bir kullanıcı Netflix' te film seçerken, binlerce seçenek arasından kendisine en uygun olanı bulmakta zorlanabilir ve bu durum karar verme sürecini olumsuz etkileyebilmektedir [4]. Dolayısıyla e-ticaret, dijital medya ve çevrimiçi hizmet platformlarında, kullanıcıların ihtiyaçlarına ve zevklerine uygun içeriklere ulaşabilmesi için öneri sistemleri vazgeçilmez bir araç haline gelmiştir [5]. Bu sistemler, kullanıcıların geçmiş davranışlarını, tercihlerini ve etkileşimlerini analiz ederek kişiselleştirilmiş öneriler sunmaktadır [6]. Ancak, geleneksel öneri sistemlerinin çoğu, kullanıcıların değerlendirmelerini tek bir genel puan üzerinden modellemekte ve ürünlerin çok boyutlu doğasını göz ardı etmektedir [7,8]. Örneğin, bir film değerlendirmesinde kullanıcılar oyunculuk, senaryo, yönetmenlik, görsel efektler gibi farklı kriterleri dikkate alırken, sistem bu değerlendirmeleri tek bir genel puan üzerinden ele almaktadır [7,9]. Bu yaklaşım, kullanıcıların her bir kriter için farklı beklentileri ve öncelikleri olduğu gerçeğini göz ardı etmektedir [10,11]. Bu çok yönlü değerlendirme ihtiyacını karşılamak üzere geliştirilen çok ölçütlü öneri sistemleri, her bir ürün veya hizmetin çeşitli yönlerini ayrı ayrı analiz ederek daha rafine ve kullanıcı odaklı öneriler sunmaktadır [11,12]. Bu sistemler, kullanıcıların farklı kriterlere verdikleri önem derecelerini dikkate alarak, daha doğru ve kişiselleştirilmiş öneriler oluşturabilmektedir [13]. Örneğin, bir film öneri sisteminde, bir kullanıcı oyunculuk performansına daha fazla önem verirken, başka bir kullanıcı senaryo kalitesine daha fazla önem verebilir. Çok ölçütlü öneri sistemleri olarak adlandırılan bu yapılar, bu farklılıkları dikkate alarak her kullanıcı için daha uygun öneriler sunabilmektedir. Ancak, her ne kadar oldukça başarılı tavsiyeler üretip kullanıcı memnuniyetini artırsa da bu sistemler popülerlik yanlılığı olarak ifade edilen problemle karşı karşıya kalabilmektedir. Bu problem, sistemlerin çoğunlukla yüksek oylama almış ya da geniş kitlelerce bilinen içerikleri öne çıkarması; buna karşılık daha az bilinen, ancak kullanıcıya potansiyel olarak daha uygun içerikleri arka planda bırakması şeklinde ortaya çıkmaktadır [14,15]. Literatürde bu sorunun etkilerini azaltmaya yönelik çeşitli algoritmalar ve kullanıcı eğilimi modelleme yaklaşımları geliştirilmiş olmakla birlikte, çalışmaların büyük çoğunluğu ya tek ölçütlü sistemlere odaklanmakta ya da kullanıcıların kriterler arası tercih farklılıklarını yeterince dikkate almamaktadır [16], [17]. Örneğin, Klimashevskaya ve arkadaşları [18], işbirlikçi filtreleme tabanlı algoritmaların popüler ürünleri aşırı öne çıkardığını ve bunun kişiselleştirme kalitesini olumsuz etkilediğini göstermiştir. Kowald ve Lacić [19] ise popülerlik yanlılığının hem ürün hem de kullanıcı düzeyinde farklı yansımalarını ele alarak, popüler içeriklere ilgi duymayan kullanıcılar için öneri doğruluğunun anlamlı şekilde düştüğünü ortaya koymuştur. Ayrıca, Abdollahpouri ve arkadaşları [20], bu yanlılığın adalet ve kalibrasyon üzerindeki etkilerini incelemiş, popüler olmayan ürünlerin bu ürünlere ilgi duyan kullanıcılara dahi yeterince önerilmediğini vurgulamıştır. Bu kapsamda, öneri sistemlerinde popülerlik etkisini daha doğru şekilde değerlendirmek üzere Steck [21] "popularity-stratified recall" metriğini önermiş, Wei ve arkadaşları [22] ise nedensel çıkarım ve karşıt-gerçeksel analiz teknikleriyle popülerlik etkisini ayırtmaya çalışmıştır. Dolayısıyla çalışmaların her birinde çok ölçütlü yapılar üzerinde bu problemin etkisinin göz ardı edildiği görülmektedir.

Bu çalışmada, öneri sistemlerinde sıklıkla karşılaşılan popülerlik yanlılığının çok ölçütlü sistemlerde etkisini azaltmaya yönelik olarak, kullanıcıların geçmiş puanlama davranışlarını dikkate alan yenilikçi bir kullanıcı popülerlik eğilimi hesaplama yöntemi olan Göreceli Popülerlik Skoru (RelPop) önerilmektedir. RelPop yöntemi, her bir kullanıcının puanlama ölçeğini bağlamsal olarak analiz ederek, benzer puanların farklı kullanıcılar nezdinde farklı anlamlar taşıyabileceği gerçeğini temel almaktadır. Bu bağlamda yöntem, mutlak puan değerlerinden ziyade, puanların kullanıcının kendi puanlama eğilimleri içindeki göreceli konumunu dikkate alarak, bireysel memnuniyet düzeyini daha hassas biçimde modellemektedir. Örneğin, genel olarak düşük puanlar veren bir kullanıcının bir ürüne verdiği 4 ya da 5 gibi yüksek bir puan, bu ürünün söz konusu kullanıcı için gerçekten öne çıkan ve tatmin edici bir deneyim sunduğunu işaret ederken; sıkça yüksek puanlar veren bir kullanıcı için benzer bir puan daha sıradan bir değerlendirme olarak yorumlanabilir. RelPop yöntemi bu farklılıkları sistematik biçimde analiz ederek, kullanıcıların göreceli memnuniyet düzeylerini nesnel şekilde hesaplamayı ve bu sayede öneri listelerinde daha çeşitli, daha az popüler ama potansiyel olarak daha ilgi çekici içeriklerin yer almasını sağlamayı amaçlamaktadır. Bu

yaklaşım, öneri sistemlerinin bireysel bağlamları daha etkin biçimde dikkate almasını sağlayarak hem içerik çeşitliliğini hem de kullanıcı adaletini artırma potansiyeline sahiptir.

Çalışmanın başlıca katkıları aşağıda özetlenmektedir:

- Kullanıcıların göreceli puanlama eğilimlerine dayalı olarak popülerlik eğilimini daha hassas biçimde ölçebilen yeni bir hesaplama yöntemi geliştirilmiştir. Bu yöntem, kullanıcıların genel değerlendirme davranışlarını dikkate alarak sistematik bir popülerlik analizi yapma olanağı sunmaktadır.
- Bu yöntem, çok ölçütlü öneri sistemlerine uyarlanarak farklı algoritmalarla kapsamlı karşılaştırmalı deneyler yapılmıştır.
- Önerilen yöntem, çok ölçütlü öneri sistemlerine uyarlanarak farklı algoritmalarla kapsamlı karşılaştırmalı deneyler gerçekleştirilmiştir. Bu kapsamda, öneri performansı ve popülerlik eğilimi açısından çeşitli sistemlerin duyarlılıkları karşılaştırmalı olarak değerlendirilmiştir.
- Popülerlikten sapma düzeyini yönsüz biçimde ölçen yeni bir değerlendirme metriği olan Mutlak Popülerlik Eğilimindeki Sapma (Absolute Deviation in Popularity Inclination-ADPI) önerilmiştir. Bu metrik, mevcut literatürdeki yönlü ölçütlerin tamamlayıcısı olarak sistemlerin popülerlikten sapma derecelerini daha bütüncül biçimde analiz etme olanağı sunmaktadır.
- Ayrıca, çok ölçütlü öneri sistemleri alanında popülerlik yanlılığına yönelik çalışmaların sınırlı sayıda olması göz önünde bulundurularak, bu çalışma söz konusu boşluğu doldurmayı ve alana yeni bir bakış açısı kazandırmayı hedeflemektedir.

Bu çalışma altı bölümden oluşmaktadır. Çalışma boyunca kullanılan yöntemler ve metotlar Bölüm 2’ de yer alırken; Bölüm 4 ve 5’ te sırasıyla Deney metodolojisi ve Deneysel sonuçlar yer almaktadır. Son bölüm ise sonuç ve önerilerden oluşmaktadır.

2. ÖN ÇALIŞMALAR

Bu bölümde çalışma boyunca kullanılan yöntemlere ve metotlara yer verilmiştir.

2.1. Kullanıcı Popülerlik Eğilimi Belirleme Yöntemleri

Öneri sistemlerinde popülerlik yanlılığını azamaya yönelik geliştirilen pek çok yöntem, genellikle öneri listelerini sonradan işleme teknikleriyle yeniden sıralamaya odaklanır; bu süreçte popüler ürünler geri plana itilirken, daha az bilinen ürünler öne çıkarılır. Ancak bu yaklaşımlar, çoğu zaman kullanıcıların ürünlerin popülerliğiyle ilgili gerçek tercihlerini dikkate almaz. Oysa pratikte, bazı kullanıcılar daha çok popüler içerikleri tercih ederken, bazıları ise daha az bilinen seçeneklere yönelmektedir. Bu nedenle, kullanıcılara daha kişiselleştirilmiş ve tatmin edici öneriler sunabilmek için, onların popülerlik eğilimlerinin doğru şekilde tespit edilmesi ve öneri algoritmalarına entegre edilmesi gerekmektedir. Bu bölümde, literatürde kullanıcıların popülerlik eğilimini ölçmek amacıyla yaygın olarak kullanılan yöntemler özetlenmektedir.

2.1.1. Popüler Ürünlerin Oranı Yaklaşımı

Popüler ürünlerin oranı yaklaşımı (RPI), kullanıcıların öneri sistemlerinde yaygın olarak tercih edilen içeriklere ne kadar ilgi gösterdiğini anlamak için kullanılan bir metriktir. Bu yaklaşımda, sistemdeki tüm öğeler, aldıkları toplam oy veya puan sayısına göre sıralanır. En çok oy alan ve sistemdeki toplam oyların yaklaşık beşte birini oluşturan öğeler “popüler” olarak tanımlanırken, geri kalanlar “az popüler” olarak kabul edilir [20]. Bu doğrultuda bu yöntem, bir kullanıcının etkileşimde bulunduğu içeriklerin ne kadarının bu popüler gruba girdiğini hesaplar. Örneğin, bir film öneri platformunda, kullanıcı A’nın izlediği 10 filmin 7’si en çok izlenenler listesindeyse, bu kullanıcının RPI değeri 0,7 olur. Böylece, öneri sistemi kullanıcıların popüler içeriklere mi yoksa daha niş seçeneklere mi yöneldiğini daha iyi analiz edebilir.

$$RPI_u = \frac{|\{i \in I_u \mid i \in H\}|}{|I_u|} \quad (1)$$

Burada, I_u kullanıcının değerlendirdiği tüm öğeleri, H ise sistemde popüler olarak tanımlanan öğeleri temsil etmektedir.

2.1.2. Derecelendirilen Ürünlerin Ortalama Popülerliği

Derecelendirilen ürünlerin ortalama popülerliği (APRI), kullanıcıların popüler içeriklere olan eğilimini ölçmek için geliştirilmiş bir metriktir. Bu metrik, her bir öğenin popülerliğini, o öğeyi değerlendiren kullanıcı sayısının sistemdeki toplam kullanıcı sayısına oranı olarak tanımlamaktadır [23].

$$APRI_u = \frac{1}{|I_u|} \sum_{i \in I_u} \frac{|\{u' \in U \mid i \in I_{u'}\}|}{|U|} \quad (2)$$

Burada, U tüm kullanıcı kümesini, I_u herhangi bir kullanıcının daha önce değerlendirme yaptığı ürünler kümesini ifade etmektedir.

2.1.3. Ortalamadan Daha İyi Yaklaşımı

Bu yöntemde kullanıcıların derecelendirdiği tüm ürünler yerine yalnızca beğendikleri ürünler dikkate alınmaktadır. Önceki bölümlerde açıklanan RPI ve APRI yöntemlerinin bu mantığa (BTA_{RPI} , BTA_{APRI}) uyarlanmasıdır [24]. BTA_{RPI} değeri, beğenilen öğelerin popüler kategorideki oranını ölçerek, kullanıcının popüler öğelere olan eğilimini ortaya koyarken; BTA_{APRI} metriği, kullanıcının beğendiği öğelerin popülerlik değerlerinin ağırlıklı ortalamasını alarak popüler içerik eğilimini ölçmektedir. Her iki yaklaşımda da yalnızca beğeni eşiğini aşan öğeler dikkate alınarak, kullanıcının gerçek tercihleri üzerinden popülerlik eğilimi analiz edilir.

2.1.4. Kullanıcı Popülerlik Eğilimini Tespit Etmek İçin Önerilen Yeni Yaklaşım: Göreceli Popülerlik Skoru

Kullanıcıların puanlama davranışlarının daha derinlemesine ve niteliksel olarak analiz edilebilmesi için, Lee ve arkadaşları [25] geliştirdiği tercih modeli bu çalışmada referans alınmıştır. Söz konusu model, kullanıcının bir ürüne verdiği puanı, o kullanıcının genel puanlama alışkanlıklarıyla birlikte ele alarak değerlendirir. Bu sayede, aynı puan değeri, farklı kullanıcılar için farklı anlamlar taşıyabilmektedir. Örneğin, çoğunlukla yüksek puanlar veren bir kullanıcı ile nadiren yüksek puan veren bir kullanıcının aynı ürüne verdiği yüksek puan, tercih düzeyi açısından aynı şekilde yorumlanmamalıdır. Tercih Skoru (Preference Score) olarak adlandırılan bu yöntemde, her bir kullanıcının değerlendirdiği ürünler arasında, belirli bir ürüne verdiği puanın göreceli konumu dikkate alınarak bir tercih skoru bulunmaktadır (Eşitlik 3).

$$Pref_{(r_{u,i})} = \alpha \frac{pref > (r_{u,i})}{R_u^+} + \beta \frac{pref = (r_{u,i})}{R_u^+} \quad (3)$$

Eşitlik 3'te:

- $pref > (r_{u,i})$: Kullanıcının i ürününe verdiği puandan daha düşük puan verdiği ürünlerin sayısı,
- $pref = (r_{u,i})$: Kullanıcının i ürününe verdiği puanla aynı puan verdiği ürünlerin sayısı,
- R_u^+ : Kullanıcının değerlendirdiği toplam ürün sayısı,
- α ve β : Ağırlık parametreleri olup, literatürde genellikle sırasıyla 1 ve 0,5 olarak alınmaktadır

Önerilen bu yaklaşımda (Göreceli Popülerlik Skoru, RelPop), yukarıda tanımlanan tercih puanları, her kullanıcı için değerlendirdiği tüm ürünler üzerinde ortalama alınarak, kullanıcının genel popülerlik eğilimi skoru aşağıdaki denklem kullanılarak elde edilmiştir.

$$RelPop_u = \frac{1}{|I_u|} \sum_{i \in I_u} Pref_{(r_{u,i})} \quad (4)$$

Eşitlik 4'de I_u , u kullanıcısının değerlendirdiği ürünlerin kümesini ifade etmektedir. Bu yöntem sayesinde, kullanıcıların puanlama davranışlarındaki bireysel farklılıklar dikkate alınarak, daha adil ve kişiselleştirilmiş bir popülerlik eğilimi ölçümü sağlanmıştır.

RelPop yöntemi ile geleneksel işbirlikçi filtreleme algoritmalarının sıklıkla göz ardı ettiği önemli bir soruna, yani kullanıcıların puanlama davranışlarındaki bağlamsal ve görelî farklılıkların dikkate alınmaması problemine doğrudan çözüm sunması amaçlanmıştır. Klasik yöntemler, tüm kullanıcıların aynı puanı benzer anlamlarla verdiği varsayımıyla hareket ederek, bireysel değerlendirme ölçeklerindeki farklılıkları göz ardı etmekte ve bu durum önerilerin doğruluğunu, çeşitliliğini ve kullanıcıya olan uygunluğunu sınırlamaktadır. Bu çalışmada önerilen bu yaklaşım, kullanıcıların geçmiş puanlama dağılımlarını analiz ederek her bir puanın o kullanıcı özelindeki anlamını modellemekte ve böylece daha gerçekçi bir tercih temsili sunmaktadır. Bu yöntem sayesinde, önerilen içeriklerin yalnızca popülerliğe dayalı olarak değil, kullanıcıların görelî memnuniyet düzeyleri doğrultusunda kişiselleştirilmiş biçimde sunulması mümkün hale getirmesi amaçlanmıştır. Sonuç olarak, sistem hem popülerlik yanlılığını azaltacak hem de kullanıcı memnuniyetini ve öneri kalitesini arttıracaktır. Özellikle çok ölçütlü karar ortamlarında bu bağlamsal analiz, öneri sistemlerinin adil, dengeli ve kullanıcı odaklı hale gelmesine önemli katkılar sağlayacaktır.

2.2. Popülerlik Yanlılığını Azaltmaya Yönelik Algoritmalar

Öneri sistemlerinde karşılaşılan temel sorunlardan biri olan popülerlik önyargısı, sistemin etkinliğini ve kullanıcı memnuniyetini önemli ölçüde düşürmektedir [20]. Bu sorunun çözümüne yönelik olarak, literatürde: artırılmış fayda üretimi (Aug), çarpımsal fayda öğrenme (Mul) ve popülerlik duyarlı öge ağırlıklandırma (Paiv) gibi yaklaşımlar sunulmuş ve her biri bu bölümde incelenmiştir.

2.2.1. Artırılmış Fayda Üretimi Algoritması

Artırılmış fayda üretimi algoritması (Aug), öneri sistemlerinde içerik dağılımını dengelemek için geliştirilmiş yenilikçi bir yaklaşımdır [26]. Bu yöntem, makine öğrenmesi modelinin tahmin ettiği kullanıcı ilgisini, içeriğin popülerlik seviyesine göre belirlenen bir dengeleme katsayısıyla birleştirerek daha dengeli öneriler sunmayı amaçlar. Bu sayede, çok popüler içeriklerin ağırlığı azaltılırken, daha az bilinen içerikler öne çıkarılır.

2.2.2. Çarpımsal Fayda Öğrenme Algoritması

Çarpımsal fayda öğrenme algoritması (Mul), öneri sistemlerinde popüler içeriklerin etkisini azaltmak için geliştirilmiş yenilikçi bir sıralama tekniğidir. Her içeriğin tahmini puanı, popülerlik seviyesiyle ters orantılı bir katsayıyla çarpılarak, popüler içeriklerin puanları belirgin şekilde düşürülür. Bu sayede, ana akım içeriklerin öneri listelerinde baskın olması engellenir ve kullanıcıya daha çeşitli içerikler sunulur [27].

2.2.3. Popülerlik Duyarlı Öge Ağırlıklandırma Algoritması

Popülerlik duyarlı öge ağırlıklandırma (Paiv), öneri sistemlerinde içerik çeşitliliğini artırmak için geliştirilmiş bir yaklaşımdır. Yöntem, içeriklerin popülerlik seviyelerine göre logaritmik bir dönüşümle ağırlıklar hesaplar; popüler içerikler düşük, az bilinenler ise yüksek ağırlık alır. Bu ağırlıklar, tahmini öneri puanlarıyla λ parametresiyle birleştirilerek, hem kullanıcı tercihlerine uygun hem de çeşitli öneriler sunulur. Paiv, öneri kalitesini koruyarak çeşitliliği artırır [28].

3. DENEY METODOLOJİSİ

Bu çalışmada, önerilen RelPop yönteminin etkinliğini değerlendirmek için kapsamlı bir deneysel çerçeve tasarlanmıştır. Değerlendirme süreci, yaygın olarak kullanılan tek-çıkışlı çapraz (leave-one out) doğrulama stratejisi temel alınarak yapılandırılmıştır. Her bir değerlendirme iterasyonunda, veri setinden rastgele bir kullanıcı test kullanıcısı olarak seçilirken, kalan kullanıcılar eğitim setini oluşturmak için kullanılmıştır. Eğitim verileri kullanılarak, test kullanıcısının ürünleri için tahmini puanları oluşturmak üzere bir işbirlikçi filtreleme algoritması uygulanmış ve bu tahmini değerlerden en yüksek 10 ürün kullanıcıya öneri olarak verilmiştir. Bu kapsamda, iki farklı analiz senaryosu tasarlanmıştır. İlk senaryoda, kullanıcıların film ve otellere verdikleri genel puanlar dikkate alınarak analiz gerçekleştirilmiştir. Bu yaklaşımda her bir kullanıcının tek bir bütünsel değerlendirmesi üzerinden tercih eğilimleri modellenmiş ve öneri sistemi

performansı, bu genel puanlar üzerinden test edilmiştir. Böylece, sistemin genel beğeni düzeyini ne ölçüde yakalayabildiği ortaya konulmuştur. İkinci senaryoda ise, kullanıcıların her bir kriter için (örneğin filmler için hikâye, oyunculuk, yönetmenlik, görsel efektler; oteller için konum, temizlik, hizmet, değer ve odalar) ayrı ayrı verdikleri puanlar kullanılarak kriter bazlı analiz uygulanmıştır. Bu yöntemde, kullanıcıların tercihlerinin farklı boyutlarını yansıtan çok ölçütlü verilerden yararlanılmış ve öneri sistemi performansının bu ayrıntılı tercihler altında nasıl değiştiği incelenmiştir. Bu iki senaryonun karşılaştırmalı değerlendirilmesi sayesinde, önerilen yöntemin hem genel değerlendirmeler temelinde hem de ayrıntılı kriter bazında kullanıcı tercihlerini ne ölçüde yansıttığı ortaya konulmuştur. Böylece, öneri sisteminin farklı veri senaryolarında gösterdiği başarı ve esneklik hakkında daha kapsamlı bir analiz yapılmıştır.

Çalışmamızın ilk aşaması olan kullanıcı popülerlik analizinde, test setindeki her bir kullanıcı geleneksel yöntemlere ($APRI$, RPI , BTA_{APRI} , BTA_{RPI}) ek olarak önerilen RelPop yöntemine göre eğilimler tespit edilip, bunların karşılaştırmalı analiz sonuçları verilirken; ikinci aşamada önerilen bu yöntemin performansı, Mul, Paiw ve Aug gibi popüler yeniden sıralama yöntemleriyle karşılaştırılmıştır.

3.1. Veri Seti

Bu çalışmada, öneri sistemleri alanında yaygın olarak kullanılan çok ölçütlü Yahoo! Movies (YM) ve TripAdvisor (TA) veri setleri kullanılmıştır. YM20 veri seti, 429 kullanıcı ve 491 filmde oluşmakta, toplam 18.504 değerlendirme içermektedir. Her film, hikâye, oyunculuk, yönetmenlik ve görsel efektler olmak üzere dört kriter üzerinden değerlendirilmiştir [29]. TA veri seti ise 1.354 kullanıcı ve 1.291 otelden oluşmakta, toplam 38.512 değerlendirme içermektedir [30]. %99,72 seyreklik oranı ile oldukça seyrek bir yapıdadır ve kullanıcı başına ortalama 28,44 değerlendirme ile düşük etkileşim göstermektedir. TA, konum, temizlik, hizmet, değer ve odalar olmak üzere beş kriter içermektedir. Her iki veri seti de 1-5 arası değerlendirme ölçeği kullanmakta ve seyreklik oranları ile kullanıcı başına değerlendirme sayıları bakımından farklılık göstermektedir. Bu farklılıklar, öneri sistemleri için çeşitli zorluk seviyeleri oluşturmakta ve kullanıcı davranış profillerinin analizine olanak sağlamaktadır. Veri setlerine ilişkin özet bilgiler Çizelge 1’de sunulmaktadır.

Çizelge 1. YM ve TA veri seti ile ilgili bilgiler

Veri seti	#Kullanıcı	#Ürün	#Oylama
YM20	429	491	18.504
TA	1346	1289	38.512

3.2. Değerlendirme Metrikleri

Öneri sistemlerinde popülerlik önyargısının değerlendirilmesinde çeşitli performans ölçütleri kullanılmaktadır. Bu çalışmada, ilk olarak sistemin kullanıcıya özel öneriler sunma başarısını ve popüler içeriklere olan eğilimini analiz etmek amacıyla kesinlik metriği temel alınmıştır.

3.2.1. Kesinlik

Kesinlik metriği, öneri sistemlerinde hem alaka düzeyinin hem de yanlılık etkisinin ölçülmesinde literatürde yaygın olarak tercih edilmekte ve bu metrikte aktif kullanıcı için ilginç ürünlerin önerilme oranı belirlenmektedir [31]. Aşağıdaki eşitlik kullanılarak hesaplama yapılabilmektedir:

$$Kesinlik@K(u) = \frac{|Rel_u \cap Rec_u^K|}{|Rec_u^K|} \quad (5)$$

Eşitlik 5’te, Rel_u kullanıcı u için ilgili öğeler kümesini, Rec_u^K ise kullanıcı u için önerilen en iyi K öğe kümesini ifade etmektedir.

3.2.2. Geri Çağırma

Öneri sistemlerinin başarısını değerlendirmek için yaygın olarak kullanılan temel ölçütlerden biridir. Bu metrik, sistemin kullanıcıya ait tüm ilgili öğeler arasından ne kadarını öneri listesine dahil edebildiğini gösterir [32] ve matematiksel formülasyonu aşağıda yer almaktadır:

$$\text{Geri çağırma}@K(u) = \frac{|Rel_u \cap Rec_u^K|}{|Rec_u|} \quad (6)$$

Kullanıcı u için ilgili öğelerin tamamı Rel_u kümesinde yer alırken, sistemin kullanıcıya sunduğu en iyi K adet öneri ise Rec_u^K kümesinde bulunmaktadır (Eşitlik 6).

3.2.3. F1-Ölçütü

Kesinlik ve geri çağırma değerlerinin harmonik ortalaması olan F1-ölçütü, bu iki performans metriğini dengeli bir şekilde birleştirerek sistemin genel başarısını değerlendirmeye olanak tanır. Bu ölçütün matematiksel ifadesi şu şekildedir (Eşitlik 7):

$$F1@K(u) = 2 \times \frac{\text{Precision}@K(u) \times \text{Recall}@K(u)}{\text{Precision}@K(u) + \text{Recall}@K(u)} \quad (7)$$

3.2.4. Yenilik

Yenilik metriği, öneri sistemlerinin kullanıcılara tanıdık olmayan ancak ilgi alanlarına uygun yeni içerikleri sunma yeteneğini ölçer. Bu metrik, sistemin kullanıcıların daha önce karşılaşmadığı fakat beğenme olasılığı yüksek olan öğeleri önerme başarısını değerlendirir. Yüksek bir yenilik değeri, sistemin kullanıcıları yeni ve çeşitli içeriklerle tanıştırdığını, böylece kullanıcı deneyimini zenginleştirdiğini gösterir [33].

$$\text{Yenilik} = \frac{1}{|R|} \sum_{i \in R} -\log_2(p(i)) \quad (8)$$

Öneri listesindeki toplam öğe sayısı R ile gösterilirken, $p(i)$ değişkeni, i . öğenin herhangi bir kullanıcı tarafından değerlendirilme olasılığını ifade eder (Eşitlik 8).

3.2.5. Ortalama Uzun Kuyruk Popülerliği

Ortalama uzun kuyruk popülerliği (APLT), öneri sistemlerinin kullanıcılara sunduğu öneri listesinde yer alan az popüler veya niş içeriklerin dağılımını değerlendiren bir metriktir [33]. Herhangi bir K kullanıcı için, sistemin oluşturduğu N_K öneri listesindeki uzun kuyruk öğelerinin oranı, aşağıdaki denklem kullanılarak hesaplanabilir

$$APLT_K = \frac{|T \cap N_K|}{N} \quad (9)$$

Burada, uzun kuyruk öğelerinin kümesi T ile gösterilirken, APLT değerinin düşük olması, sistemin öneri listesinde yeterli çeşitlilik sağlayamadığını ve genellikle popüler içeriklere ağırlık verdiğini ifade etmektedir. Buna karşılık, yüksek APLT değerleri, sistemin uzun kuyruk içeriklerini başarıyla öneri listesine dahil edebildiğini işaret eder. Ancak, yüksek APLT değeri tek başına öneri kalitesinin iyi olduğunu garanti etmez, çünkü sistem az popüler içerikleri önerirken kullanıcı tercihlerini göz ardı edebilir [34].

3.2.6. Uzun Kuyruk Kapsamı

Uzun Kuyruk Kapsamı (LTC) metriği, öneri sisteminin tüm kullanıcılar için oluşturduğu öneri listelerinde yer alan az popüler içeriklerin genel dağılımını değerlendirilmektedir (Eşitlik 10):

$$LTC = \frac{|I_N \cap T|}{|T|} \quad (10)$$

LTC metriğinin yüksek değerler göstermesi, öneri sisteminin az popüler veya niş içerikleri daha kapsamlı bir şekilde öneri listelerine dahil edebildiğini işaret eder. Bu durum, sistemin sadece popüler içeriklere odaklanmak yerine, daha geniş bir içerik yelpazesini kullanıcılara sunabildiğini gösterir. Yüksek LTC

değerleri, öneri sisteminin çeşitlilik açısından daha başarılı olduğunu ve popülerlik yanlılığından daha az etkilendiğini ortaya koyar. Bu da kullanıcıların daha çeşitli ve kişiselleştirilmiş öneriler almasını sağlar [35].

3.2.7. Gini İndeksi Ölçütü

Gini İndeksi ölçütü, veri setindeki sınıfların dağılımını analiz ederek, bir düğümdeki verilerin ne kadar heterojen olduğunu ortaya koymaktadır.

$$Gini(D) = 1 - \sum p_i^2 \quad (11)$$

Gini değeri hesaplanırken Eşitlik 11’de, D ile gösterilen veri setindeki her bir sınıfın (i) dağılım oranı p_i değişkeni ile ifade edilir. Gini değerinin sıfır olması, incelenen düğümdeki tüm verilerin aynı sınıfa ait olduğu, yani tam bir homojenlik sağlandığını açıklayabilir [36].

3.2.8. Popülerlik Eğilimindeki Sapma

Popülerlik Eğilimindeki Sapma (Deviation in Popularity Inclination-DPI): Kullanıcıların mevcut yöntemlerle hesaplanan popülerlik eğilimleriyle yeni önerilen teknikler sonucunda elde edilen popülerlik eğilimleri arasındaki farkı normalize ederek ölçmektedir:

$$DPI_u = \frac{P_u - \varepsilon_u}{\varepsilon_u} \quad (12)$$

Eşitlik 12’de, ε_u mevcut yöntemlerle hesaplanan kullanıcı popülerlik eğilimi iken, P_u yeni yöntemle hesaplanan kullanıcı popülerlik eğilimini temsil etmektedir. DPI değerinin negatif çıkması, önerilen yöntemin kullanıcılara daha az popüler, dolayısıyla daha özelleştirilmiş ya da farklı içerikler sunduğunu göstermektedir [24].

3.2.9. Mutlak Popülerlik Eğilimindeki Sapma

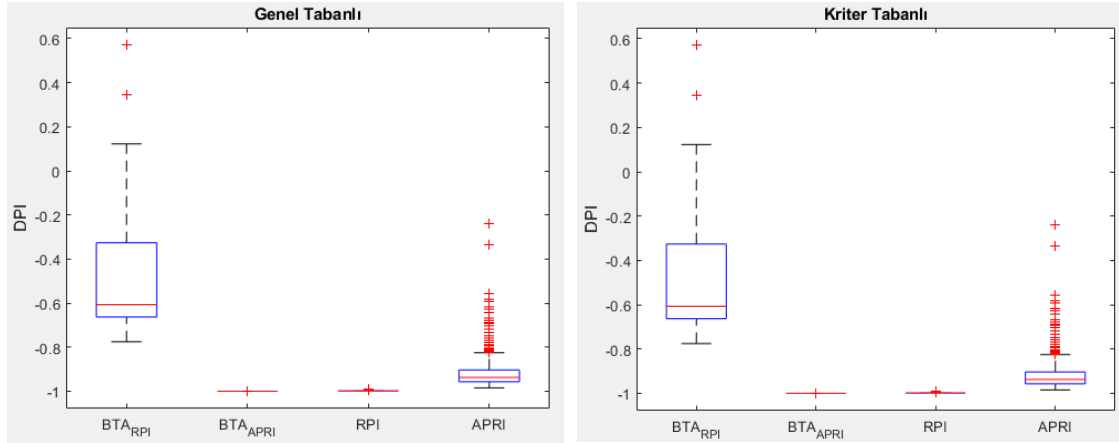
Mutlak Popülerlik Eğilimindeki Sapma (Absolute Deviation in Popularity Inclination-ADPI): Kullanıcıların mevcut yöntemlerle hesaplanan popülerlik eğilimleriyle yeni önerilen teknikler sonucunda elde edilen popülerlik eğilimleri arasındaki mutlak farkı normalize ederek ölçmektedir. Önerilen bu yöntemde popülerlik eğilim farkının artık pozitif/negatif yönü değil, sadece büyüklüğü dikkate alınır.

$$ADPI_u = \left| \frac{P_u - \varepsilon_u}{\varepsilon_u} \right| \quad (13)$$

Mutlak DPI değeri, klasik yöntemle önerilen içeriklerle yeni yöntemin önerileri arasındaki popülerlik eğilimi farkının büyüklüğünü göstermektedir. Bu ölçüt sayesinde sistemin öneri davranışındaki sapmaların yönünden bağımsız olarak, öneri çeşitliliği ve popülerlikten uzaklaşma düzeyi değerlendirilebilmektedir (Eşitlik 13).

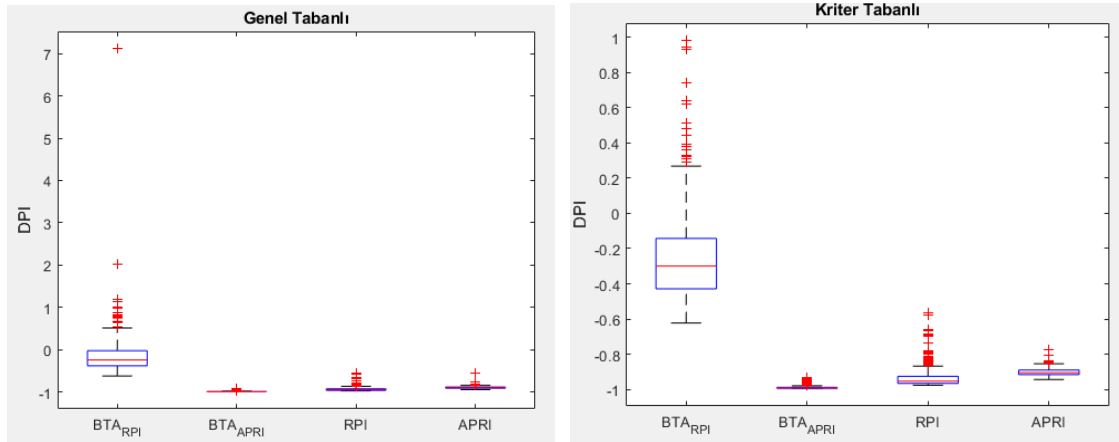
4. DENEYSEL SONUÇLAR

Çalışmanın temel amacı, öneri sistemlerinde kullanıcıların değerlendirme davranışlarındaki bağlamsal farklılıkları dikkate alarak popülerlik yanlılığını azaltmaya yönelik daha etkili ve kullanıcı odaklı bir yaklaşım geliştirmektir. Bu kapsamda, önerilen *RelPop* yöntemi ile literatürde yaygın olarak kullanılan *BTA_{APRI}*, *BTA_{RPI}*, *APRI* ve *RPI* yöntemleri, iki farklı çok ölçütlü veri kümesi (YM20 ve TA) üzerinde karşılaştırmalı olarak incelenmiştir. Analizlerde, geleneksel performans metriklerine ek olarak popülerlik yanlılığını daha hassas ölçen yeni bir ADPI metriği de kullanılmıştır. Kullanıcı popülerlik eğilimlerinin tespitine yönelik karşılaştırmalı analiz sonuçları Şekil 1-4’te sunulmuştur.



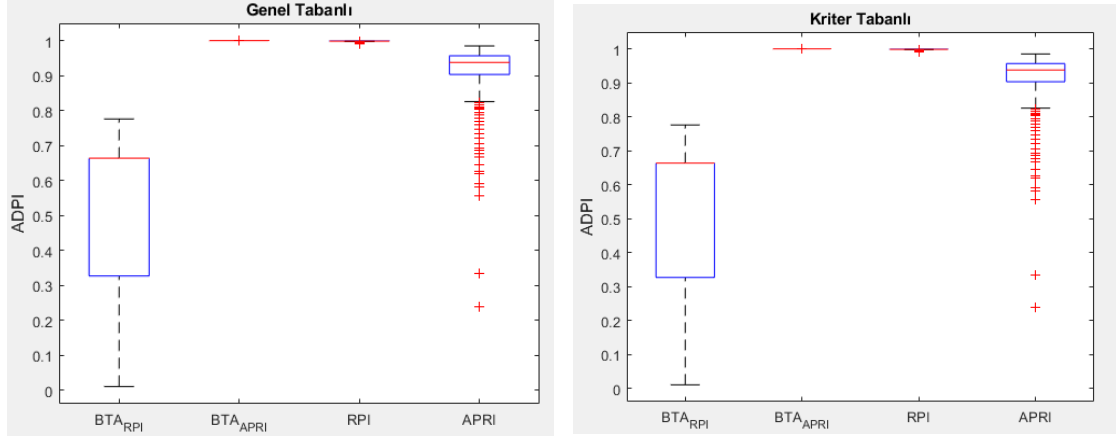
Şekil 1. TA veri seti üzerinde DPI analiz sonuçları

Şekil 1’de TA veri kümesi üzerinde gerçekleştirilen DPI analizleri sunulmaktadır. Genel tabanlı değerlendirmede, BTARPI yöntemi, negatif medyan değeri ve geniş dağılım aralığıyla dikkat çekmekte; bu durum yöntemin, kullanıcıya sunulan içerikleri daha az popüler hale getirme potansiyeline sahip olduğunu göstermektedir. Diğer taraftan, APRI yöntemi de negatif değerlere sahip olmakla birlikte, daha dar bir dağılım sergilemekte ve önerilerin popülerlikten sınırlı biçimde uzaklaştığını ortaya koymaktadır. RPI ve BTA_{APRI} yöntemleri ise DPI açısından neredeyse sıfıra yakın değerlere sahip olup, kullanıcılara sunulan içeriklerin mevcut popüler eğilimlerden anlamlı biçimde sapmadığını göstermektedir. Kriter tabanlı değerlendirmede de benzer bir eğilim gözlemlenmiştir; BTA_{RPI} yöntemi en düşük medyan değeri ile öne çıkarken, APRI yöntemi negatif değerlere sahip olmakla birlikte daha sınırlı sapma göstermiştir. RPI ve BTA_{APRI} yöntemlerinin ise çeşitliliği sınırlı tutan daha sabit sonuçlar ürettiği görülmektedir.



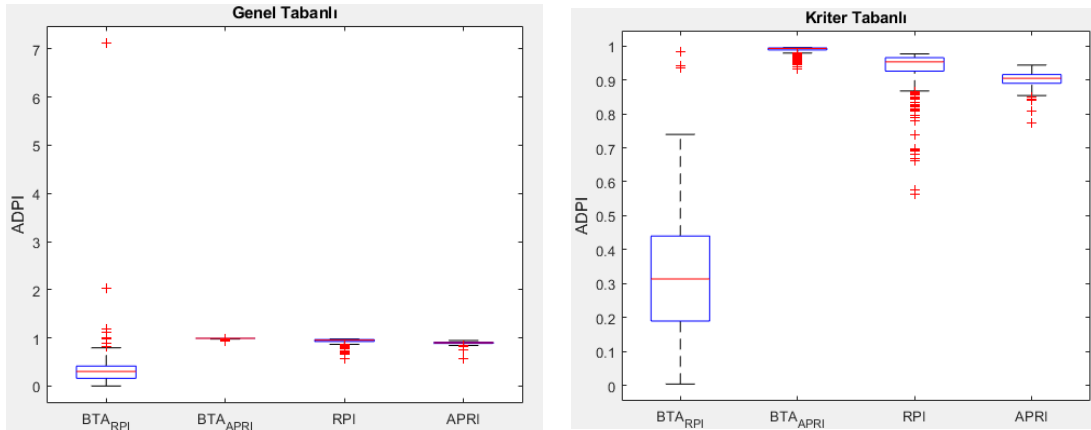
Şekil 2. YM20 veri seti üzerinde DPI analiz sonuçları

Şekil 2’ de ise bir diğer veri seti olan YM20 üzerindeki sonuçlara yer verilmektedir. Genel tabanlı analizde, BTA_{RPI} yöntemi uç değerlerin etkisiyle geniş bir varyans aralığına sahiptir; bu durum, yöntemin bazı kullanıcılar için önerilen içeriklerin popülerlikten güçlü biçimde sapabildiğini göstermektedir. Buna karşın, RPI ve BTA_{APRI} yöntemleri, düşük varyans ve sıfıra yakın DPI değerleriyle daha tutarlı ancak popülerliğe yakın önerilerde bulunmaktadır. APRI yöntemi ise negatif değerler üretmesine rağmen öneri çeşitliliği açısından oldukça sınırlı bir etki göstermektedir. Kriter tabanlı değerlendirmede ise BTA_{RPI} yöntemi, yine popülerlik yanlılığından uzak öneriler sunarken, diğer yöntemler –özellikle BTA_{APRI} ve RPI– istikrarlı ancak oldukça benzer dağılımlar göstermiştir. Bu durum, klasik yöntemlerin kullanıcıya önerdiği içeriklerin popülerliğe olan bağımlılığının sürdüğünü ortaya koymaktadır.



Şekil 3. TA veri seti üzerinde ADPI analiz sonuçları

Şekil 3 incelendiğinde, mutlak sapma bağlamında BTA_{RPI} yöntemi, medyan açısından en düşük ADPI değerine sahip olup önerilerin, kullanıcıların genel eğilimlerinden sapma derecesinin kontrollü olduğunu göstermektedir. RPI ve BTA_{APRI} yöntemleri neredeyse sabit ve yüksek ADPI değerleriyle dikkat çekmektedir; bu durum, önerilerin kullanıcıların mevcut popülerlik eğilimlerinden anlamlı düzeyde sapmadığını göstermektedir. APRI yöntemi ise genel tabanlı analizde en yüksek ADPI değerine sahiptir; bu durum, önerilerin farklılaşma açısından göreceli olarak daha uzak olduğunu ancak bu farklılaşmanın rastlantısal biçimde oluşabileceğini düşündürmektedir. Kriter tabanlı sonuçlarda da benzer örüntü korunmuş, BTA_{RPI} göreceli olarak daha düşük ADPI değerleri üretirken, diğer yöntemler sabit ve yüksek değerlere sahip olmuştur.

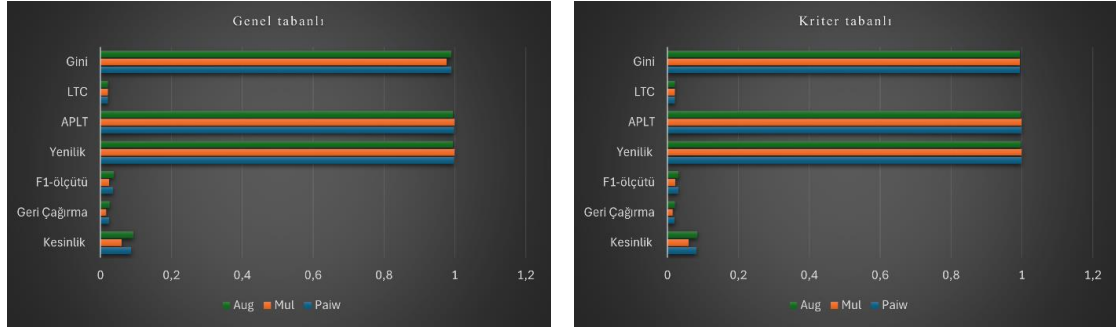


Şekil 4. YM20 veri seti üzerinde ADPI analiz sonuçları

Aynı performans metriğinin diğer veri setindeki sonuçları Şekil 4'te gösterilmektedir. Genel tabanlı analizde BTA_{RPI} yöntemi, ADPI değerlerinde düşük medyan ve geniş dağılım aralığı ile bazı kullanıcılar için önerileri ciddi biçimde özelleştirirken, bazıları için daha dengeli içerikler sunmaktadır. RPI ve BTA_{APRI} yöntemleri ise sabit ve yüksek ADPI değerleriyle, sistemin popülerlik merkezli çalıştığını göstermektedir. APRI ise yüksek medyan ve dar varyansla, önerilerin kullanıcıların popülerlik eğilimlerinden çok sapmadığını ortaya koymaktadır. Kriter tabanlı analizde de benzer bir görünüm sergilenmiş, BTA_{RPI} görece düşük sapma ile öne çıkarken, diğer yöntemler yüksek ve homojen sapma düzeylerinde kalmıştır.

Genel olarak, DPI ve ADPI analizleri öneri sistemlerinin popülerlik yanlılığı açısından farklı performanslar sergilediğini göstermektedir. BTA_{RPI} , her iki veri kümesinde de popülerlik eğiliminden saparak daha özelleştirilmiş öneriler sunma potansiyeli taşımaktadır. APRI, önerilerin popülerlikten uzaklaştığını gösterse de bu sapma genellikle sınırlı kalmıştır. RPI ve BTA_{APRI} ise popülerlik odaklı öneriler sunmaya devam etmekte, bu da çeşitliliğin artırılmasında yetersiz kaldığını göstermektedir. Sonuçlar, öneri sistemlerinde adil ve çeşitli öneriler için DPI ve ADPI metriklerinin birlikte değerlendirilmesinin önemini vurgulamaktadır. Özellikle BTA_{RPI} yöntemi, farklı kullanıcı grupları için popülerlik yanlılığını azaltmada etkili bir yaklaşım sunmaktadır.

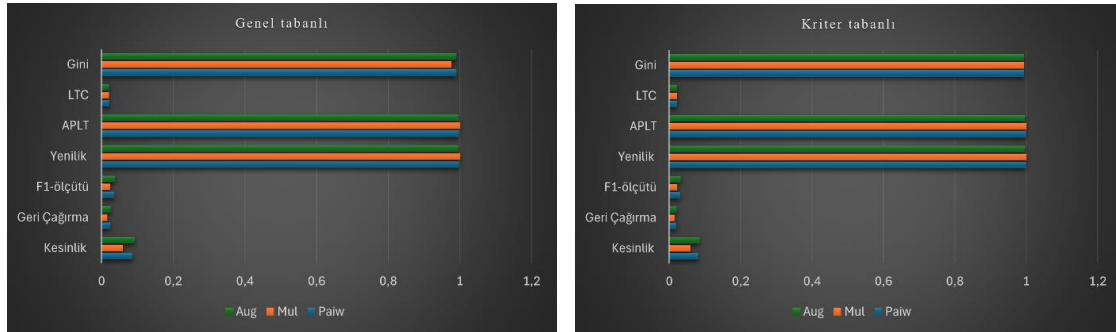
Çalışmanın ikinci aşaması olan, kullanıcıların popülerlik eğilimlerinde meydana gelen değişimlerin kapsamlı analizi Şekil 5-10 arasında gösterilmektedir.



Şekil 5. YM20 veri seti üzerinde ADPI analiz sonuçları

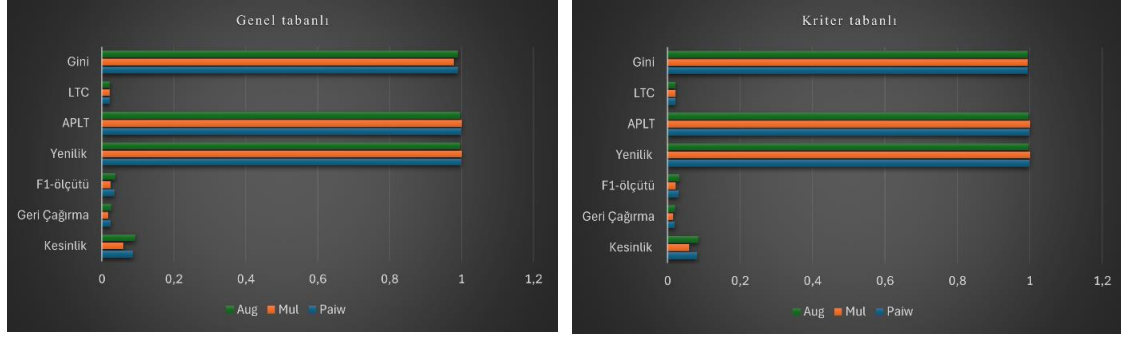
Şekil 5'te, YM20 veri seti üzerinde BTA_{RPI} yöntemiyle yapılan genel ve kriter tabanlı analizlerin sonuçları sunulmaktadır. Aug, Mul ve Paiw algoritmalarının performansları yedi farklı metrik üzerinden karşılaştırılmıştır. Her iki analizde de Gini, APLT ve yenilik metriklerinde tüm algoritmalar yüksek ve benzer değerlere ulaşmıştır; bu da öneri sistemlerinin çeşitli ve yenilikçi içerikler sunmada başarılı olduğunu göstermektedir. Yüksek Gini indeksi, öneri listelerinin kullanıcılar arasında dengeli ve adil dağıldığını, popülerlik yalnlığının azaldığını göstermektedir. Buna karşılık, LTC metriği tüm algoritmalar için düşük kalmış, sistemlerin uzun kuyruktaki niş içerikleri önerme oranının sınırlı olduğunu ortaya koymuştur.

Doğruluk metrikleri olan kesinlik, geri çağırma ve F1-ölçütü ise tüm algoritmalarda düşük çıkmıştır; bu da sistemlerin isabetli ve ilgili öneriler sunmada yetersiz kaldığını göstermektedir. Genel ve kriter tabanlı analizler karşılaştırıldığında, metriklerin dağılımı ve algoritmaların sıralaması büyük ölçüde benzerdir. Sonuç olarak, Aug, Mul ve Paiw algoritmaları çeşitlilik ve yenilik açısından başarılı olsa da, doğruluk ve uzun kuyruk kapsamı açısından geliştirmeye ihtiyaç duymaktadır. Bu analizler, öneri sistemlerinin yalnızca doğruluk değil, aynı zamanda yeni ve çeşitli içerik sunma kapasitesiyle de değerlendirilmesi gerektiğini göstermektedir.



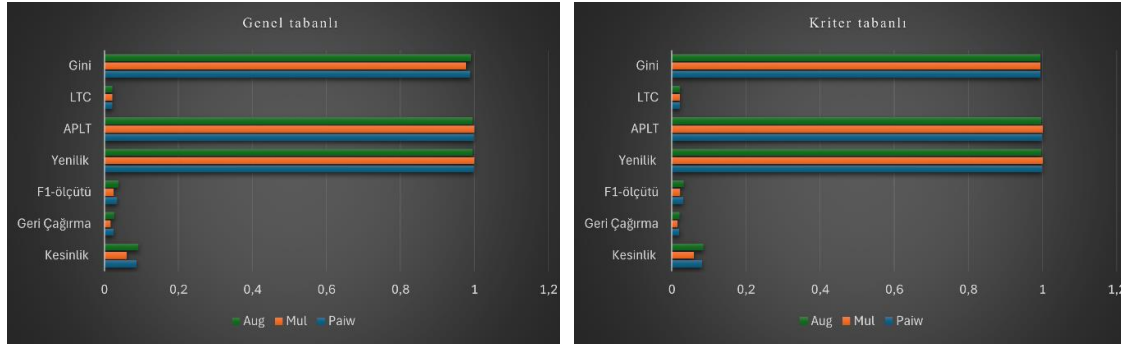
Şekil 6. YM20 veri seti üzerinde BTA_{APRI} analiz sonuçları

Şekil 6'da, aynı veri setinin BTA_{APRI} yöntemiyle yapılan analiz sonuçları sunulmaktadır. Grafikler, Aug, Mul ve Paiw algoritmalarının Gini, APLT ve yenilik metriklerinde 0,95'in üzerinde, yüksek değerlere ulaştığını göstermektedir. Özellikle yenilik değerinin 0,98'e yaklaşması, öneri listelerinin büyük ölçüde yeni ve kullanıcı için keşfedilmemiş ürünlerden oluştuğunu ortaya koymaktadır. Buna karşılık, LTC metriğinin 0,03'ün altında kalması, öneri listelerinde niş içeriklerin sınırlı olduğunu göstermektedir. Doğruluk metriklerinde ise kesinlik yaklaşık 0,04, geri çağırma ve F1-ölçütü ise 0,01'in üzerindedir; bu da sistemlerin isabetli ve kişiselleştirilmiş öneriler sunmada yetersiz kaldığını göstermektedir. Genel ve kriter tabanlı analizler benzer sonuçlar vermiş, yalnızca kriter tabanlı analizde APLT ve yenilikte küçük artışlar gözlenmiştir. Sonuç olarak, Aug, Mul ve Paiw algoritmaları çeşitlilik ve yenilikte başarılı olsa da, doğruluk ve uzun kuyruk kapsamı açısından geliştirmeye ihtiyaç duymaktadır. Özellikle LTC ve doğruluk metriklerindeki düşük değerler, sistemlerin niş içeriklere ve kullanıcıya özel önerilere daha fazla odaklanması gerektiğini göstermektedir.



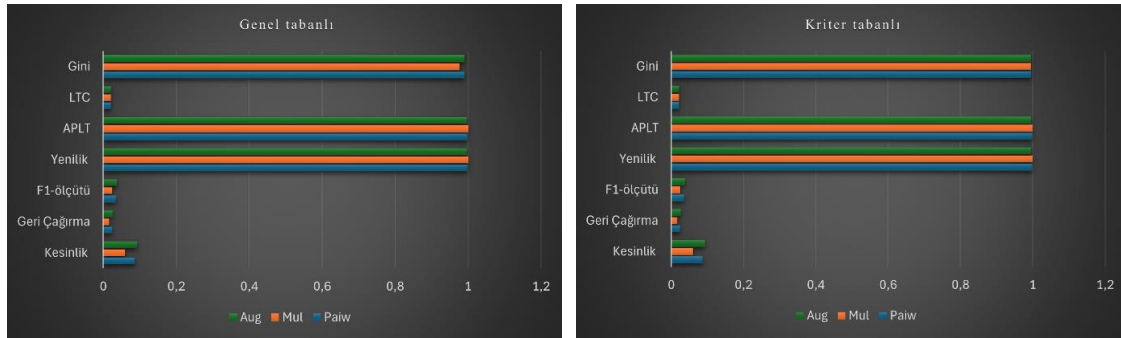
Şekil 7. YM20 veri seti üzerinde RPI analiz sonuçları

Şekil 7’de, Aug, Mul ve Paiw algoritmalarının Gini, APLT ve yenilik metriklerinde 0,95’in üzerinde ve birbirine yakın değerlere ulaştığı görülmektedir; bu da sistemlerin yenilikçi ve çeşitli içerikler sunmada başarılı olduğunu göstermektedir. LTC metriğinde ise Mul algoritması, diğerlerine göre daha yüksek bir değere ulaşarak uzun kuyrukta yer alan içerikleri önerme konusunda kısmi bir avantaj sağlamıştır. Doğruluk metrikleri genel olarak düşük kalırken, Aug algoritması ise kesinlik ve F1-ölçütlerinde diğerlerine göre daha iyi performans göstermiştir. Kriter tabanlı analizde ise APLT ve yenilikte küçük artışlar gözlenmiş, bu da kriter bazlı yaklaşımın çeşitlilik ve yenilik üretimini bir miktar artırdığını göstermektedir. Sonuç olarak, RPI yöntemi altında algoritmalar yenilik ve çeşitlilikte güçlü performans sergilerken, Mul uzun kuyruk kapsamı, Aug ise doğruluk metriklerinde öne çıkmıştır.



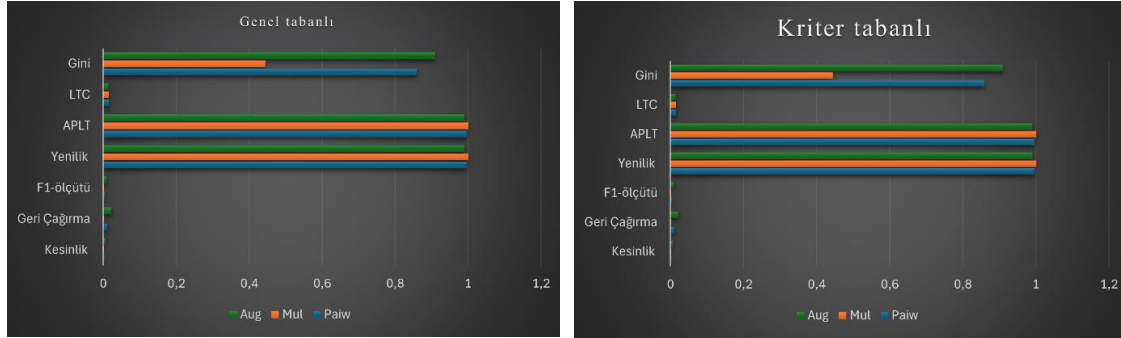
Şekil 8. YM20 veri seti üzerinde APRI analiz sonuçları

Şekil 8’de, YM20 veri setinde APRI yöntemiyle yapılan analizlerde Aug, Mul ve Paiw algoritmaları Gini, APLT ve yenilik metriklerinde yüksek değerler elde ederek çeşitlilik ve yenilikte başarılı olmuştur. Ancak, LTC ve doğruluk metrikleri tüm algoritmalarda düşük kalmış, özellikle kesinlik değerlerinin 0,1’in altında olması sistemlerin isabetli önerilerde yetersiz olduğunu göstermiştir. Kriter tabanlı ve genel analizler arasında anlamlı bir fark görülmemiş, APRI yöntemi tutarlı bir performans sergilemiştir. Sonuç olarak, algoritmalar yenilik ve çeşitlilikte güçlü, doğruluk ve uzun kuyruk kapsamı açısından ise geliştirmeye açıktır.



Şekil 9. YM20 veri seti üzerinde RelPop analiz sonuçları

Şekil 9’da, RelPop yöntemiyle YM20 veri setinde algoritmaların genel ve kriter tabanlı performansları karşılaştırılmıştır. Tüm algoritmalar Gini, APLT ve yenilik metriklerinde üst sınıra yakın değerlere ulaşarak, RelPop’un hem çeşitliliği hem de öneri özgünlüğünü koruduğunu göstermiştir. Gini değerlerinin yüksek ve birbirine yakın olması, önerilerin kullanıcılar arasında dengeli dağıldığını işaret etmektedir. LTC metriği ise orta düzeyde kalarak, RelPop’un niş içerikleri dengeli şekilde önerdiğini göstermektedir. Doğruluk metriklerinde Aug algoritması, kesinlik ve F1’de diğerlerine göre daha tutarlı sonuçlar vermiştir. Kriter tabanlı analizde ise yenilik ve APLT’deki küçük artışlar, kullanıcı tercih ağırlıklarının öneri kalitesine olumlu katkı sağladığını göstermektedir. Sonuç olarak, RelPop yöntemiyle algoritmalar, çeşitlilik ve kullanıcı odaklılıkta dengeli ve tatmin edici bir performans sergilemiş, önerilerde denge odaklı bir yaklaşım sunmuştur.



Şekil 10. TA veri seti üzerinde $BTAPRI$, $BTAAPRI$, RPI , $APRI$ ve $RelPop$ analiz sonuçları

Şekil 10’da, TA veri seti üzerinde farklı yöntemlerle yapılan genel ve kriter tabanlı analizlerin sonuçları sunulmuştur. Tüm algoritmalar Gini, APLT ve yenilik metriklerinde yüksek skorlar elde ederek, kullanıcıya çeşitli ve özgün içerikler sunmada başarılı olmuştur. Ancak, LTC metriği tüm yöntemlerde düşük kalmış, sistemlerin uzun kuyruktaki niş içerikleri önermede yetersiz olduğunu göstermiştir. Doğruluk metrikleri de genel olarak düşük çıkmış, özellikle kesinlik ve geri çağırma oranlarının zayıf olması, sistemlerin isabetli öneriler sunmakta zorlandığını ortaya koymuştur.

Genel ve kriter tabanlı analizlerde benzer sonuçlar alınması ise, TA veri setinin seyrek yapısından kaynaklanmaktadır. TA veri seti, çalışmadaki en seyrek veri seti olduğundan, kullanıcıya kişisel öneriler sunabilmek için kullanılan oylama değerleri oldukça sınırlıdır. Bu durum, yöntemlerin tüm metriklerde aynı değerleri üretmesine yol açmış ve algoritmalar arası farkların belirginleşmesini engellemiştir. Sonuç olarak, yöntemler çeşitlilik ve yenilikte başarılı, ancak doğruluk ve uzun kuyruk kapsamı açısından sınırlı kalmıştır; veri setinin seyrekliği ise performans farklarının ortaya çıkmasını zorlaştırmıştır.

5. SONUÇ VE ÖNERİLER

Bu çalışmada, öneri sistemlerinde popülerlik yanlılığını azaltarak kullanıcıya daha çeşitli, adil ve bağlama duyarlı içerikler sunmayı amaçlayan RelPop yöntemi ile bu amaca yönelik geliştirilen yeni bir değerlendirme metriği olan ADPI önerilmiştir. RelPop, kullanıcıların mutlak puanlamalarından ziyade kendi puanlama geçmişleri içindeki göreceli eğilimlerine odaklanarak, kişisel anlamı daha etkili bir şekilde modellemektedir. Böylece öneri listelerinde hem çeşitliliğin hem de kişiselleştirme kalitesinin artırılması hedeflenmiştir. YM10 ve TA veri setleri üzerinde yürütülen deneysel çalışmalar sonucunda, RelPop yönteminin özellikle yenilik, APLT ve Gini endeksi gibi çeşitlilik ve adalet odaklı metriklerde yüksek performans sergilediği gözlemlenmiştir. Ayrıca, RelPop’un Aug algoritmasıyla birlikte kullanıldığında, az bilinen ve potansiyel olarak kullanıcıya uygun içeriklerin öneri listelerine daha dengeli bir şekilde dahil edildiği görülmüştür. Bu durum, RelPop’un popüler içeriklerin baskınlığını azaltarak daha kapsayıcı ve kullanıcı odaklı öneriler sunduğunu göstermektedir. Özellikle düşük etkileşimli kullanıcılar veya farklı kriterlere göre değerlendirme yapan kullanıcılar açısından da sistemin etkili sonuçlar ürettiği belirlenmiştir. Ancak, kesinlik, geri çağırma ve F1-ölçütü gibi klasik doğruluk metriklerinde RelPop’un geleneksel yöntemlere benzer sınırlılıkları bulunmaktadır. Ayrıca, TA veri setinin seyrek yapısı ve YM20 veri setindeki kullanıcı sayısının azlığı, yöntemin potansiyelinin tam olarak ortaya konmasını kısıtlamıştır.

Performans değerlendirmesi kapsamında kullanılan DPI ve önerilen ADPI metrikleri, öneri sistemlerinin popülerlik yanlılığından ne ölçüde saptığını değerlendirmede birbirini tamamlayıcı niteliktedir. DPI,

sapmanın yönünü ve büyüklüğünü gösterirken; ADPI, sapmanın yönünden bağımsız olarak büyüklüğünü ölçmektedir. DPI metriklerinin kullanıcı odaklı çeşitlilik analizlerinde; ADPI'nın ise farklı yöntemler arasındaki genel performans farklarını karşılaştırmada daha uygun olduğu sonucuna ulaşılmıştır.

Bu kapsamda elde edilen bulgular doğrultusunda aşağıdaki öneriler sunulabilir:

- Özellikle seyrek veri setlerinde kullanıcı-öge etkileşimlerinin artırılması, RelPop'un bağlamsal analiz gücünü daha da geliştirecektir.
- RelPop'un çeşitlilik ve adalet açısından sunduğu avantajlar, doğruluk odaklı algoritmalarla birleştirilerek hibrit öneri sistemleri tasarlanabilir.
- Gelecek çalışmalarda farklı veri setleriyle benzer analizler yapılarak yöntemin genellenebilirliği ve sınırlılıkları daha kapsamlı biçimde değerlendirilebilir.

Bununla birlikte, yöntemin genellenebilirliğini artırmak için yalnızca film ve otel gibi belirli alanlarla sınırlı kalınmaması, farklı türde veri setlerinde (ör. müzik, e-ticaret, sosyal medya) test edilmesi önemlidir. Veri seti sınırlılıklarının azaltılması amacıyla sentetik veri üretme veya transfer learning tabanlı yaklaşımlar gelecekte değerlendirilebilir. Ayrıca RelPop'un, çok ölçütlü sistemlerde her bir kriter için ayrı ayrı veya kriterlerin ağırlıklandırılması yoluyla hesaplanması, daha rafine kişiselleştirme imkânı sağlayarak yöntemin çok ölçütlü yapılara özgü avantajını ortaya koyacaktır. Bu tür genişletmeler, RelPop'un yalnızca popülerlik yanlılığını azaltmakla kalmayıp, aynı zamanda kullanıcıların bireysel kriter önceliklerini de dikkate alan daha esnek ve kapsamlı öneri sistemlerinin geliştirilmesine katkı sunacaktır.

6. KAYNAKLAR

1. Zhang, S., Yao, L., Sun, A. & Tay, Y. (2019). Deep learning based recommender system: A survey and new perspectives. *ACM Computing Surveys*, 52(1), 1-38.
2. Gomez-Urbe, C. A. & Hunt, N. (2016). The Netflix recommender system: Algorithms, business value, and innovation. *ACM Transactions on Management Information Systems*, 6(4), 1-19.
3. Ricci, F., Rokach, L. & Shapira, B. (2022). Recommender systems: Techniques, applications, and challenges. In F. Ricci, L. Rokach & B. Shapira (Eds.), *Recommender Systems Handbook* (1-35). Cham: Springer.
4. Schwartz, B. (2016). *The paradox of choice: Why more is less* (Rev. ed.). New York: Harper Perennial.
5. Jannach, D. & Adomavicius, G. (2017). Price and profit awareness in recommender systems. In *Proceedings of the 11th ACM Conference on Recommender Systems* (192-200).
6. Koren, Y. & Bell, R. (2015). Advances in collaborative filtering. In F. Ricci, L. Rokach & B. Shapira (Eds.), *Recommender Systems Handbook* (77-118). Cham: Springer.
7. Adomavicius, G. & Kwon, Y. (2015). Multi-criteria recommender systems. In F. Ricci, L. Rokach & B. Shapira (Eds.), *Recommender Systems Handbook* (847-880). Cham: Springer.
8. Baltrunas, L. & Ricci, F. (2014). Experimental evaluation of context-dependent collaborative filtering using item splitting. *User Modeling and User-Adapted Interaction*, 24(1-2), 7-34.
9. Kalkan, İ.E. ve Şahin, C. (2023). Çapraz satışı destekleyebilecek transformer ile geliştirilmiş bir öneri sistemi. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 38(2), 571-584.
10. Liu, H., Wang, Y. & Chen, X. (2021). Multi-criteria recommendation with user preference learning. In *Proceedings of the 15th ACM Conference on Recommender Systems* (456-465).
11. Jannach, D. & Adomavicius, G. (2016). Recommendations with a purpose. In *Proceedings of the 10th ACM Conference on Recommender Systems* (7-10).
12. Koşar, O. ve Atak, M. (2023). Bulut bilişim sanal sunucu ürün seçiminde çok kriterli bir karar destek modeli. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 38(4), 939-953.
13. Baltrunas, L., Ludwig, B. & Ricci, F. (2011). Matrix factorization techniques for context-aware recommendation. In *Proceedings of the 5th ACM Conference on Recommender Systems* (301-304).
14. Wang, X., He, X., Wang, M., Feng, F. & Chua, T. S. (2019). Neural graph collaborative filtering. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval* (165-174).
15. Steck, H. (2018). Calibrated recommendations. In *Proceedings of the 12th ACM Conference on Recommender Systems* (154-162).

16. Ekstrand, M.D., Azpiazu, I.M., Anuyah, O., McNeill, D., Ekstrand, J.D. & Pera, M.S. (2018). All the cool kids, how do they fit in?: Popularity and demographic biases in recommender evaluation and effectiveness. *In Proceedings of the Conference on Fairness, Accountability and Transparency* (172-186).
17. Abdollahpouri, H., Mansoury, M., Burke, R. & Mobasher, B. (2020). The connection between popularity bias, calibration, and fairness in recommendation. *In Proceedings of the 14th ACM Conference on Recommender Systems* (726-731).
18. Klimashevskaya, A., Abdollahpouri, H. & Mobasher, B. (2022). Popularity bias in collaborative filtering: A survey. *In Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization* (1-11).
19. Kowald, D. & Lacić, E. (2022). Popularity bias in different domains: A multi-domain analysis. *In Proceedings of the 16th ACM Conference on Recommender Systems* (1-11).
20. Abdollahpouri, H., Burke, R. & Mobasher, B. (2019). The unfairness of popularity bias in recommendation. *In Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* (148-154).
21. Steck, H. (2011). Item popularity and recommendation accuracy. *In Proceedings of the 5th ACM Conference on Recommender Systems* (125-132).
22. Wei, D., Wang, X., Li, Y. & He, X. (2021). Causal inference for recommender systems. *In Proceedings of the 30th ACM International Conference on Information and Knowledge Management* (3043-3049).
23. Mansoury, M., Abdollahpouri, H., Smith, J., Dehpanah, A., Pechenizkiy, M., Mobasher, B. (2019). Investigating potential factors associated with gender discrimination in collaborative recommender systems. *Information Processing & Management*, 57(2), 102371.
24. Taçlı, Y., Yalçın, E. & Bilge, A. (2022). Novel approaches to measuring the popularity inclination of users for the popularity bias problem. *In Proceedings of the 6th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)* (555-560).
25. Lee, J., Lee, D., Lee, Y., Hwang, W. & Kim, S. (2016). Improving the accuracy of top-n recommendation using a preference model. *Information Sciences*, 348, 290-304.
26. Yalçın, E. & Bilge, A. (2021). Investigating and counteracting popularity bias in group recommendations. *Information Processing & Management*, 58(5), 102608.
27. Zhang, Y., Chen, X., Ai, Q., Croft, W. B. & Guo, J. (2017). Towards conversational search and recommendation: System ask, user respond. *In Proceedings of the 27th ACM International Conference on Information and Knowledge Management* (177-186).
28. Abdollahpouri, H., Burke, R. & Mobasher, B. (2017). Controlling popularity bias in learning-to-rank recommendation. *In Proceedings of the 11th ACM Conference on Recommender Systems (RecSys)* (42-46).
29. Lakiotaki, K., Matsatsinis, N.F. & Tsoukias, A. (2011). Multicriteria user modeling in recommender systems. *IEEE Intelligent Systems*, 26(2), 64-76.
30. Wang, H., Lu, Y. & Zhai, C. (2011). Latent aspect rating analysis without aspect keyword supervision. *In Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (618-626).
31. Herlocker, J.L., Konstan, J.A., Terveen, L.G. & Riedl, J.T. (2004). Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS)*, 22(1), 5-53.
32. Bellogín, A., Castells, P. & Cantador, I. (2011). Precision-oriented evaluation of recommender systems: An algorithmic comparison. *In Proceedings of the 5th ACM Conference on Recommender Systems* (333-336).
33. Vargas, S. & Castells, P. (2011). Rank and relevance in novelty and diversity metrics for recommender systems. *In Proceedings of the 5th ACM Conference on Recommender Systems* (109-116).
34. Park, Y.J. & Tuzhilin, A. (2008). The long tail of recommender systems and how to leverage it. *In Proceedings of the 2008 ACM Conference on Recommender Systems* (11-18).
35. Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, 12(4), 331-370.
36. Breiman, L., Friedman, J., Stone, C.J. & Olshen, R.A. (1984). *Classification and regression trees*. Chapman & Hall/CRC.



Bağlayıcı Olarak Bazaltik Pomza Tozu Kullanımının Geopolimer Harçların Mekanik Dayanımına Etkisi

Taha Salah Wahhab AL-ANTAKİ^{1,a}, Anıl NİŞ^{1,b}

¹*İstanbul Gelişim Üniversitesi, Mühendislik ve Mimarlık Fakültesi, İnşaat Mühendisliği Bölümü, İstanbul, Türkiye*

^a*ORCID: 0000-0001-8050-4062; ^bORCID: 0000-0001-9092-8088*

Makale Bilgileri

Geliş : 10.09.2024

Kabul : 24.09.2025

DOI: 10.21605/cukurovaumfd.1790372

Sorumlu Yazar

Anıl NİŞ

anis@gelisim.edu.tr

Anahtar Kelimeler

Bazaltik pomza

Geopolimer

Basınç dayanımı

Eğilme dayanımı

Ultrasonik ses geçiş hızı (UPV)

Atf şekli: AL-ANTAKİ, T.S.W., NİŞ, A., (2025). Bağlayıcı Olarak Bazaltik Pomza Tozu Kullanımının Geopolimer Harçların Mekanik Dayanımına Etkisi. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 687-697.

ÖZ

Bu çalışmada, %20 ve %40 oranlarında volkanik kökenli bazaltik pomza kullanımı ile üretilen geopolimer harçların mekanik özelliklerine olan etkisi incelenmiştir. Geopolimer numuneler üretilirken bağlayıcı olarak F-tipi uçucu kül ile yüksek fırın cürufu, alkali aktivatör olarak ise sodyum silikat ile sodyum hidroksit kullanılmıştır. Çalışmada %20 ve %40 oranlarında bağlayıcı malzeme yerine kullanılan pomza (P) ile üretilmiş numunelerin 7., 28., ve 56. günlerde basınç ve eğilme dayanımları, ultrasonik ses geçiş hızları irdelenmiştir. Ayrıca farklı sodyum hidroksit molaritelerinin (8M, 12M ve 16M) mekanik dayanımına olan etkisi araştırılmıştır. Sonuçlara göre, pomza oranı arttıkça basınç dayanımının azaldığı, sodyum hidroksit molaritesi arttıkça da basınç dayanımının arttığı bulunmuştur. 56. günde en yüksek dayanım değerleri P0-16M numunelerinde 61.62 MPa bulunurken, en düşük dayanımsa P40-8M numunelerinde 44.14 MPa bulunmuştur. Bağlayıcı olarak cüruf ve uçucu kül yerine %40 oranlarına kadar bazaltik pomza kullanılarak üretilen geopolimer numuneler, mevcut standartlarda belirtilen yapıların inşasında kullanılabilecek minimum C25/30 beton basınç dayanımı sınıfını sağlamaktadır.

The Influence of Using Basaltic Pumice Powder as a Binder on the Mechanical Strength of Geopolymer Mortars

Article Info

Received : 10.09.2024

Accepted : 24.09.2025

DOI: 10.21605/cukurovaumfd.1790372

Corresponding Author

Anıl NİŞ

anis@gelisim.edu.tr

Keywords

Basaltic pumice

Geopolymer

Compressive strength

Flexural strength

Ultrasonic pulse velocity (UPV)

How to cite: AL-ANTAKİ, T.S.W., NİŞ, A., (2025). The Influence of Using Basaltic Pumice Powder as a Binder on the Mechanical Strength of Geopolymer Mortars. Çukurova University, Journal of the Faculty of Engineering, 40(3), 687-697.

ABSTRACT

In this study, the influence of 20% and 40% volcanic origin basaltic pumice on the mechanical properties of geopolymer mortars was examined. F-type fly ash and ground granulated blast furnace slag were used as binders, and sodium silicate and sodium hydroxide were used as alkali activators. Compressive and bending strengths, and ultrasonic pulse velocities of geopolymers produced with pumice (P) at 20% and 40% ratios were examined on the 7., 28., and 56. days. Also, effects of different sodium hydroxide molarities (8M, 12M and 16M) on mechanical strengths were investigated. It was found that compressive strength decreases as pumice ratio increases, and compressive strength enhances as sodium hydroxide molarity increases. The highest compressive strength was obtained as 61.62 MPa in the P0-16M sample, while the lowest compressive strength was obtained as 44.14 MPa in the P40-8M sample at 56 days. Geopolymers up to 40% pumice provide the minimum C25/30 concrete compressive strength class that can be used in the construction of structures specified in current standards.

1. GİRİŞ

Dünyada 2014 yılından sonra geleneksel Portland çimentosu üretimi 4,3 milyar ton değerine ulaştığı belirtilmektedir ve 2050 yılında geleneksel Portland çimentosu üretiminin 5,5 milyar ton değerine ulaşacağı tahmin edilmektedir [1]. Çimento üretiminin çok önemli çevresel etkileri bulunmaktadır ve üretimi sırasında 3 milyar ton hammadde tüketirken 4700 MJ/ton enerjiye ihtiyaç duymaktadır [2]. Genel olarak 1 ton çimento üretimi sırasında atmosfere yaklaşık 900 kg karbondioksit (CO_2) salınımı gerçekleşmektedir. Bu da şu an geleneksel Portland çimento üretimiyle beraber çevreye çok yüksek miktarda CO_2 salındığı ve bu miktarın 2050 yılına kadar daha da artması beklenmektedir. Bu çevresel etkiyi önlemek adına son zamanlarda araştırmacılar çimentosuz beton olarak adlandırdıkları yeni nesil çevreci bir yapı malzemesi üzerine çalışmalarını yoğunlaştırdılar. Geopolimer olarak isimlendirilen çimentosuz yapı malzemesi, bağlayıcı olarak çimento yerine düşük kalsiyum içerikli uçucu kül ve öğütülmüş granüle yüksek fırın cürufu kullanılmaktadır. Bu malzemelerden F-tipi uçucu kül termik santral atığı, öğütülmüş granüle yüksek fırın cürufu ise demir-çelik endüstrisi atığı olmasından ötürü bu malzemelerle üretilecek yeni nesil sürdürülebilir yapı malzemeleri atık malzemelerin yeniden değerlendirilmesi açısından da çevreye çok büyük katkı sağlamaktadır. Ayrıca çimento yerine uçucu kül ve cüruf gibi malzemelerin kullanılması diğer atık malzemelerin depolanma veya gömülme problemlerini de ortadan kaldırmakta ve doğal kaynakların daha az kullanımı açısından da son derece önem arz etmektedir [3].

Diğer taraftan pomza, volkanik patlamalar sonunda ortaya çıkmış, boşluklu ve süngerimsi yapıya sahip, kimyasal ve fiziksel etmenlere karşı dirençli bir volkanik kayadır. Pomzanın oluşması sırasında bünyesinde bulunan gazların yapıyı hızlı bir şekilde terk etmesi ve sonrasında soğuması sonucunda bünyesinde çok sayıda gözenek içermektedir. Dolayısıyla pomzanın birim hacim ağırlığı düşerek yaklaşık 500 kg/m³ olmaktadır. Boşluklu yapısı sayesinde ses ve ısı yalıtımına uygun bir malzemedir. Pomza kışın soğuğa karşı, yazın ise sığağa karşı yüksek yalıtım görevi görmektedir. Pomzanın kimyasal içeriğinde ise yaklaşık %75 civarlarında silis bulunmaktadır. Pomzada bulunan silisyum oksit (SiO_2) miktarı pomza malzemesine abrasif özelliği kazandırmaktadır. Bu abrasif özellikten dolayı pomzanın çeliği bile rahatça aşındıracak bir kimyasal yapısı vardır. Pomzanın kimyasal içeriğindeki %15 alüminyum oksit (Al_2O_3) bileşeni ise ateş ve ısıya karşı yüksek yalıtım sağlar [4-6].

Pomza asidik ve bazaltik olmak üzere ikiye ayrılmaktadır. Bazaltik pomzanın özgül ağırlığı 1-2 g/cm³ civarında, rengi ise siyah, kahverengi ve koyu olmaktadır. Asidik pomzanın ise özgül ağırlığı 0,5-1 g/cm³ civarında, rengi ise genellikle beyaz ve beyaza yakın olmaktadır. İçeriğinde alüminyum, potasyum, sodyum ve silis içerdiği için açık renge sahiptir [7]. Asidik pomza genelde inşaat endüstrisinde kullanılmakta ve ülkemizdeki pomzanın %80'i hafif beton agregası olarak kullanılmaktadır. Pomza kullanılarak üretilen betonların daha hafif olmasıyla hem işçilikten hem de zamandan tasarruf edilmektedir [8,9]. Ayrıca pomza ile üretilen betonlar hafif beton olacağı için deprem bölgelerinde yapılara daha az deprem yükünün gelmesini sağlayabilirler. Pomzanın beton agregası olarak kullanıldığında, kırılmadan büyük deformasyon yapabilme yetenekleri sayesinde, deprem bölgelerinde önemli şekil değiştirme kapasitesi sağlayacağı düşünülmektedir. Şu an pomza kullanımı bims ve briket üretimi ile duvarların inşasıyla sınırlı kalmaktadır. Pomza kullanımı ile ısı yalıtımı da sağlanması enerji tasarrufu sağlaması bakımından önem arz etmektedir. Ek olarak pomza zemin stabilizasyonunda da kullanılabilir [10].

Pomza hafif agrega olarak kullanılabilmesi gibi, kimyasal bileşimindeki yüksek oranda silis içermesinden dolayı puzolanik özelliği gösterip bağlayıcı olarak kullanılabilmesi de mümkündür. Bunun yanı sıra pomza bir dolgu malzemesi olarak kullanılabilir. Yapılan çalışmalarda %15-%25 oranlarında pomzanın dolgu olarak kullanımı ASTM standardına göre uygun olabileceği belirtilmiş olup, daha yüksek oranlarda pomza kullanımının dayanım değerlerini düşürdüğü belirtilmiştir [10].

Son yıllarda yapılan araştırmalar çimento üretimi sebebiyle neden olan çevre kirliliğini ve enerji talebini azaltmak amacıyla pomza ve diğer malzemelerin kullanılması üzerinedir. Yapılan bir çalışmada hafif betonda iri agrega ve çimento yerine pomza kullanımı araştırılmıştır. Çalışmada çimento yerine %0-25 oranlarında öğütülmüş pomza tozunun kullanılarak, pomzanın taze ve sertleşmiş hal özelliklerine olan

etkileri araştırılmıştır. Sonuç olarak çimento yerine %15 oranlarına kadar kullanılan öğütülmüş pomzanın betonların fiziksel ve kimyasal özelliklerini olumlu etkilediği belirtilmiştir [11]. Başka bir çalışmada %20 uçucu kül ve bazaltik pomza agregası kullanılarak üretilen betonlarda basınç dayanımının 29 MPa olduğunu bunun da yapılarda kullanılması gereken minimum basınç dayanımı olan 25 MPa değerinden büyük olduğunu belirtmişlerdir [12]. Yapılan diğer bir çalışmada bazaltik pomzanın hafif yapı malzemesi olarak kullanılabilirliği irdelenmiş olup, bazaltik pomza kullanılarak 28. günde 30 MPa basınç dayanımına sahip hafif beton üretiminin mümkün olduğu belirtilmiştir [13]. Başka bir çalışmada ise pomza agregası, silis dumanı ve uçucu kül kullanılarak hafif betonlar üretilmiş olup, üretilen numuneler farklı dayanım testlerine tabii tutulmuştur. Elde ettikleri sonuçlara göre, pomza kullanılarak üretilen hafif beton dayanımlarının ACI 213R standardında belirtilen yük taşıyan hafif beton dayanımını aştığını belirtmişlerdir [14].

Geopolimer beton ise çimentosuz olarak üretilen yeni nesil bir beton tipidir. Cüruf bazlı geopolimer ısı kürüne ihtiyaç duymaksızın erken basınç dayanımına sahip olurken, uçucu kül bazlı geopolimerlerde ise yüksek erken dayanım için ısı kürüne ihtiyaç vardır. Cüruf bazlı geopolimerlerde büzülme çatlakları ve erken priz alma problemleri olurken, uçucu kül kullanılarak üretilen geopolimerlerde ise bu sorun büyük oranda ortadan kalkmaktadır. Son zamanlarda yapılan çalışmalarda %50 uçucu kül ve %50 cüruf kullanılarak üretilen geopolimer numunelerde büzülme çatlakları ve ani priz alma gibi problemlerin ortadan kalktığı rapor edilmiştir. Ayrıca üretilen bu betonlar yüksek basınç ve eğilme dayanımına ve farklı kimyasal etkilere karşı da yüksek dirence sahiptir [3].

Literatürde bazaltik pomza tozunun bağlayıcı olarak kullanımının geopolimer harçların üzerindeki etkisini inceleyen çok az çalışma mevcuttur. Zeyad ve arkadaşları yaptığı bir çalışmada pomzanın boşlukları doldurma etkisi ve pomzanın puzolanik reaktif olması sebebiyle geleneksel Portland çimentolu numunelerde basınç dayanımının arttığını belirtmiştir [15]. Geopolimer numunelerde yapılan çalışmalarda ise benzer sonuçlar göstermektedir. Geopolimer bağlayıcının bir kısmının volkanik pomza tozu ile yer değiştirmesinin daha yoğun bir mikro yapı ile sonuçlandığı, gözenekliliğin azaldığı ve basınç dayanımının arttığı rapor edilmiştir [16-18]. Yapılan başka bir çalışmada [19], farklı oranlarda pomza tozu ikamesinin (%25, %50 ve %75) geopolimer numunelerin farklı günlerdeki (7, 28 ve 56 gün) basınç dayanımına olan etkisi araştırılmıştır. Elde edilen sonuçlara göre, geopolimer numunelerin basınç dayanımının zamanla arttığı fakat 28 gün sonrasındaki dayanım artışın çok düşük olduğu hatta bazı numunelerde basınç dayanımında düşüşler gözlemlenmiştir. Bunun sebebine ise değinilmemiştir. %25 pomza eklenmesinin geopolimer numunelerinin 7. gün basınç dayanımını %8.79, 28. gün basınç dayanımını ise %8.68 oranında düşürdüğü belirtilmiştir. Daha yüksek pomza ikamesinde (%50 ve %75) ise bu dayanımdaki azalmanın çok fazla olduğu rapor edilmiştir. %50 oranında pomza ilavesinin basınç dayanımını 42.6 MPa değerinden 34.1 MPa değerine, %75 oranında pomza ilavesinin ise basınç dayanımını 16.69 MPa değerine düşürdüğü gözlemlenmiştir [19].

Dolayısıyla yapılan bu çalışma kapsamında %20 ve %40 oranlarında bazaltik pomza tozu, F-tipi uçucu kül ve öğütülmüş granüle yüksek fırın cürufu yerine ikame edilerek geopolimer harçlardaki mekanik özelliklere olan etkisinin araştırılması amaçlanmıştır. Ayrıca, bazaltik pomza tozunun farklı sodyum hidroksit molaritesinin (8M, 12M ve 16M) değişimi ile hangi oranda değiştiği 7., 28. ve 56. günlerde detaylıca irdelenecektir.

2. DENEYSEL ÇALIŞMA

2.1. Malzemeler ve Karışım Oranları

Geopolimer harçların üretiminde bazaltik pomza tozu (P), F-tipi uçucu kül ve öğütülmüş granüle yüksek fırın cürufu bağlayıcı olarak kullanılmıştır. Bazaltik pomza tozu %0, %20 ve %40 oranlarında düşük kalsiyum içerikli F-tipi uçucu kül ve cüruf yerine eşit oranda kullanılmıştır. Pomzasız serilerde uçucu kül ve cüruf oranları sırasıyla %50 ve %50, %20 oranında pomza içeren numunelerde uçucu kül ve cüruf oranları %40 ve %40, %40 oranında pomza içeren numunelerde uçucu kül ve cüruf oranları %30 ve %30 olarak kullanılmıştır. Kullanılan bazaltik pomza, F-tipi uçucu kül ve yüksek fırın cürufu Sika şirketi

tarafından temin edilmiştir. Kullanılan ince agrega, 4 mm'nin altında dane boyutuna sahip doğal kumdur. Toz halindeki pomzanın parçacık boyutu 200 μm 'den küçüktür; pomzanın özgül ağırlığı 2,5 (g/cm^3) ve Blaine inceliği 3700 (cm^2/g) olarak belirlenmiştir. Agrega dane dağılımı eğrileri, diğer çalışmalara benzer bulunmuştur [20, 21]. Çizelge 1'de pomza, uçucu kül ve cürufun fiziksel ve kimyasal özellikleri verilmiştir. Alkali aktivatör olarak sodyum hidroksit (NaOH) ve sodyum silikat (Na_2SiO_3) kullanılmıştır. Çalışmada ayrıca 3 farklı sodyum hidroksit molaritesi 8 M, 12 M ve 16 M kullanılmıştır. Sodyum silikat çözeltisi (Na_2O : %11,4, SiO_2 : %30,1, su: %57,5 kütlece) sıvı olarak, NaOH ise %97-98 saflıkta peletler halinde temin edilmiştir. NaOH çözeltisi kullanılmadan en az bir gün önce hazırlanıp, soğuması sağlanmıştır.

Çizelge 1. Pomza, cüruf ve uçucu külün fiziksel özellikleri ve kimyasal bileşimi

Kimyasal bileşim	SiO_2 (%)	Al_2O_3 (%)	CaO (%)	Fe_2O_3 (%)	MgO (%)	SO_3 (%)	Na_2O (%)	K_2O (%)	Cl (%)	Yanma kayıp (%)	Özgül ağırlık (g/cm^3)
Bazaltik pomza	58.12	21	11.47	7	7	0.41	5.16	1.86	-	1.68	2.5
Cüruf	37.97	13.27	37.92	1.16	5.64	0.23	0.84	0.56	0.01	0.01	2.95
F-tipi uçucu kül	56.15	26.37	1.79	6.44	2.35	0.056	1.1	3.8	0.09	2.2	2.05

2.2. Numunelerin Hazırlanması ve Kür Koşulları

Pomza bazlı geopolimer harç numuneleri beton üretimine benzer şekilde laboratuvar ortamında üretilmiştir. Bir gün öncesinde alkali çözeltisi hazırlanmış ve laboratuvar ortamında 24 saat soğutulmuştur. İlk önce kuru malzemeler (kum, cüruf, kül, varsa pomza) homojenlik sağlamak için mikserde 2 dakika karıştırılmıştır. Daha sonra, önceden hazırlanan alkali aktivatör çözeltisi mikserin içine eklenmiş ve 2 dakika daha karıştırılmıştır. Ardından, süper akışkanlaştırıcı karışıma dâhil edilmiş ve ek olarak 2 dakika daha karıştırma işlemi gerçekleştirilmiştir. Pomza bazlı geopolimer harçların karıştırma işlemi toplamda 6 dakika sürmüştür. İşlenebilirliği artırmak için yüksek oranda su azaltıcı katkı maddesi (Sika Plastiment-15) kullanılmış olup, dayanım için ise sodyum silikat/sodyum hidroksit oranı 2.5, aktivatör/bağlayıcı oranı ise 0.45 olarak belirlenmiştir [22, 23]. Numunelerde bazaltik pomza olmayan numuneler P0, %20 pomza içeren numuneler P20 ve %40 pomza içeren numuneler ise P40 olarak isimlendirilmiştir. İkinci terim NaOH molaritesini göstermektedir. Örneğin, P40-16M numunesi %40 pomza içermekte olup, sodyum hidroksit molaritesi 16 Molar olan geopolimer numuneleri göstermektedir.

Basınç dayanım testi için 50×50×50 mm küp numuneleri üretilip ASTM C39 standardına göre test edilmiştir. Eğilme dayanım testi için ise 40×40×160 mm prizmatik üretilerek eğilme dayanımları EN1015-11 standardı kullanılarak belirlenmiştir. Numuneler üretildikten hemen sonra üzeri alkali çözeltinin buharlaşmasını önlemek amacıyla plastik torbalarla kaplanmış ve geopolimerizasyon aktivasyonu için 65°C'de 48 saat süreyle ısı kürü uygulanmıştır. Isı kürü sonrası geopolimer numuneler test zamanı gelene kadar oda sıcaklığında (23±2°C) tutulmuştur. Şekil 1'de bazaltik pomza tozu, üretilen numuneler, uygulanan ısı kürü ve eğilme testi gösterilmiştir. Çizelge 2'de ise geopolimer numunelerin üretilmesinde kullanılan malzeme miktarları verilmiştir. 8 M, 12 M, 16 M ifadelerinde ise sodyum hidroksit molaritesi değişmektedir.



Şekil 1. Bazaltik pomza katkılı geopolimer numuneler

Çizelge 2. Geopolimer harçlarda kullanılan malzeme miktarları (g/l)

Malzeme	P0 (50C+50U)	P20 (40C+40U+20P)	P40 (30C+30U+40P)
Kum	1608	1598	1587
Bazaltik pomza	0	107	212
Cüruf	268	213	159
Uçucu kül	268	213	159
Sodyum silikat	172	172	172
Sodyum hidroksit	69	69	69
Akışkanlaştırıcı	2	2	2

3. BULGULAR VE TARTIŞMA

3.1. Birim Hacim Ağırlık

Çizelge 3’de 50x50x50 mm ebatlarındaki küp geopolimer harçların birim hacim ağırlıkları verilmiştir. Elde edilen sonuçlara göre, pomza oranının artması numunenin birim hacim ağırlığını azaltmaktadır. Birim hacim ağırlıklarındaki azalmanın sebebi ise pomzanın özgül ağırlığı ile mineral katkıların özgül ağırlığı arasındaki farkla ilişkilidir [24]. Başka bir çalışmada da pomza tozunun kademeli olarak artması birim hacim ağırlığını düşürmüştür [25]. Ayrıca, birim hacim ağırlıkları sodyum hidroksit molaritesi arttıkça az da olsa artmaktadır. Sodyum hidroksit molaritenin artması katı sodyum hidroksit parçacıklarının miktarının artırdığı için numunelerdeki birim hacim ağırlığı artmıştır.

Çizelge 3. Geopolimer numunelerin birim hacim ağırlıkları

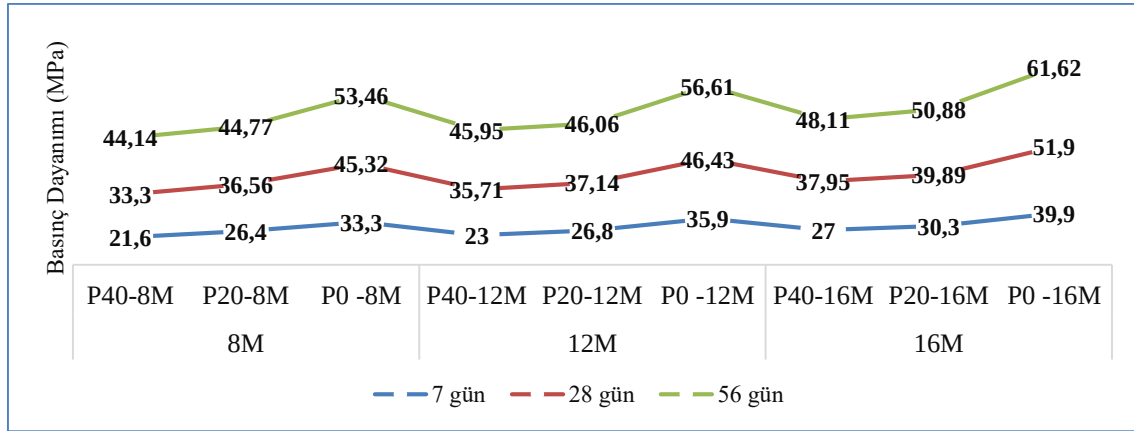
Numune adı	(g/cm ³)	Numune adı	(g/cm ³)	Numune adı	(g/cm ³)
P0 -8M	2.36	P0 -12M	2.38	P0 -16M	2.42
8M P20-8M	2.34	12M P20-12M	2.35	16M P20-16M	2.38
P40-8M	2.28	P40-12M	2.30	P40-16M	2.32

3.2. Basınç Dayanımı Sonuçları

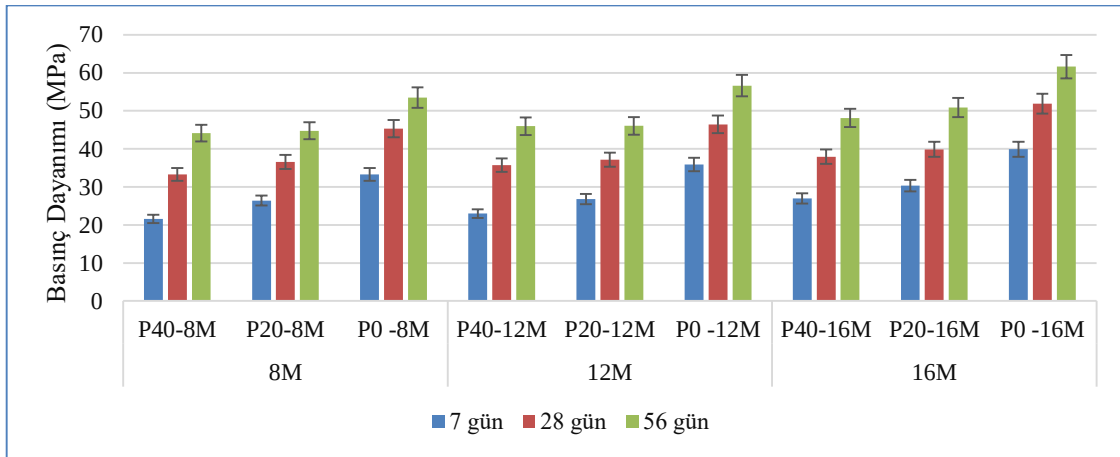
Çizelge 4’de ise farklı sodyum hidroksit molaritesine sahip ve pomza katkılı geopolimer numunelerin 7., 28. ve 56. günlerdeki basınç dayanımı sonuçları verilmiştir. Her bir seri için 3 numunenin ortalama basınç dayanım değeri alınarak numunelerin basınç dayanımı elde edilmiştir. Elde edilen sonuçlara göre, en yüksek basınç dayanımı P0-16M numunelerinde elde edilirken, en düşük basınç dayanımı ise P40-8M geopolimer harç numunelerinde elde edilmiştir. Sonuçlara göre, geopolimer numunelere 65 °C 48 saat ısı kürü uygulanmasına rağmen, basınç dayanımlarındaki artışın zamanla devam ettiği görülmektedir (Şekil 2). Bu da geopolimerizasyon reaksiyonlarının tamamlanmadığını göstermektedir. Isı kürü uygulandıktan sonra artan basınç dayanımları literatürde de rapor edilmiştir [26,27]. Kullanılan pomza oranının artması geopolimer harç numunelerinde basınç dayanımını düşürmüştür. Fakat %20 pomza kullanımı ile %40 pomza kullanımı arasındaki dayanım farkı özellikle çok azdır.

%40 pomza kullanımında dahi 28. günde elde edilen basınç dayanım değerleri birçok standartta yapıların inşasında kullanılması gereken beton numunelerde minimum basınç dayanımı olan 25 MPa (silindir) ve 30 MPa (küpe) değerlerinden daha fazladır. Şekil 3’de ise sodyum hidroksit molaritesi değişiminin geopolimer harç numunelerde basınç dayanımına olan etkisi gösterilmektedir. Sonuçlara göre, sodyum hidroksit molaritesi arttıkça pomza katkılı geopolimer numunelerin basınç dayanımı da artmaktadır. En düşük basınç dayanımı 8M numunelerinde, sonra 12M numunelerde, en yüksek dayanım ise 16M geopolimer numunelerde elde edilmiştir. Molaritenin artması sonucu elde edilen daha alkali bir ortam, pomzadan silika çözünmesini kolaylaştırarak, dayanım artışına katkı sağlayan sodyum alüminyum silikat-hidratlı jel (N-A-S-H), C-S-H (kalsiyum silikat-hidratlı jel) ve C-A-S-H (kalsiyum alüminyum silikat-hidratlı jel) geopolimerizasyon ürünlerinin artmasını sağlayarak dayanım katkıda bulunmaktadır [3]. Özellikle yüksek kalsiyum içeriğine sahip cürufun alkali aktivasyonu sonucunda C-S-H ile C-A-S-H jelleri oluşurken, uçucu külün aktivasyonu sırasında ise ana jel N-A-S-H jelleri oluşarak, geopolimer numunelerde basınç dayanımının artmasını sağlamaktadır [28,29]. Sonuçlara göre, geopolimer harç numunelerde

öğütülmüş granüle yüksek fırın cürufu ve F-tipi uçucu kül yerine %40 oranlarına kadar bazaltik pomza ikamesi yapılmasının uygun olabileceği söylenebilir.



Şekil 2. Numunelerinin farklı günlerdeki ortalama basınç dayanımı (MPa)



Şekil 3. Numunelerinin farklı günlerdeki ortalama basınç dayanımı (MPa)

Çizelge 4. Numunelerin farklı günlerdeki ortalama basınç dayanımları

		Ortalama basınç dayanımı (MPa)		
Numune adı		7 gün	28 gün	56 gün
8 M	P40-8M	21,6	33,3	44,14
	P20-8M	26,4	36,56	44,77
	P0-8M	33,3	45,32	53,46
12 M	P40-12M	23	35,71	45,95
	P20-12M	26,8	37,14	46,06
	P0-12M	35,9	46,43	56,61
16 M	P40-16M	27	37,95	48,11
	P20-16M	30,3	39,89	50,88
	P0-16M	39,9	51,90	61,62

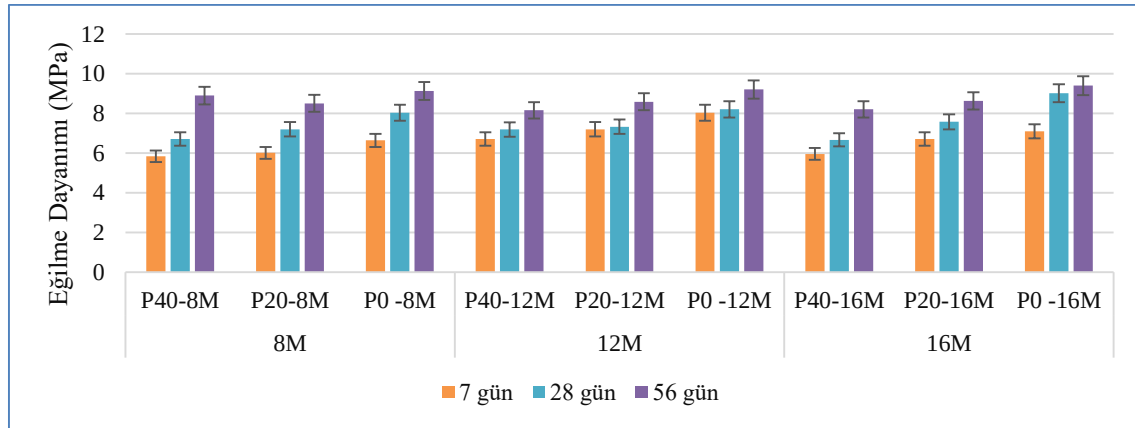
3.3. Eğilme Dayanımı Sonuçları

Çizelge 5'te farklı sodyum hidroksit molaritesine sahip ve pomza katkılı geopolimer numunelerin 7., 28. ve 56. günlerdeki üç noktalı eğilme dayanımı sonuçları verilmiştir. Her bir seri için 3 numunenin ortalama eğilme dayanım değeri alınarak numunelerin ortalama eğilme dayanımı elde edilmiştir. Basınç dayanımı sonuçlarına benzer şekilde, en yüksek eğilme dayanımı P0-16M numunelerinde 9,40 MPa olarak elde edilirken, en düşük eğilme dayanımı ise P40-8M numunelerinde 5,84 MPa olarak elde edilmiştir. Pomza oranı arttıkça eğilme dayanımı değerleri düşmektedir. Şekil 4'de bazaltik pomza katkılı geopolimer harç

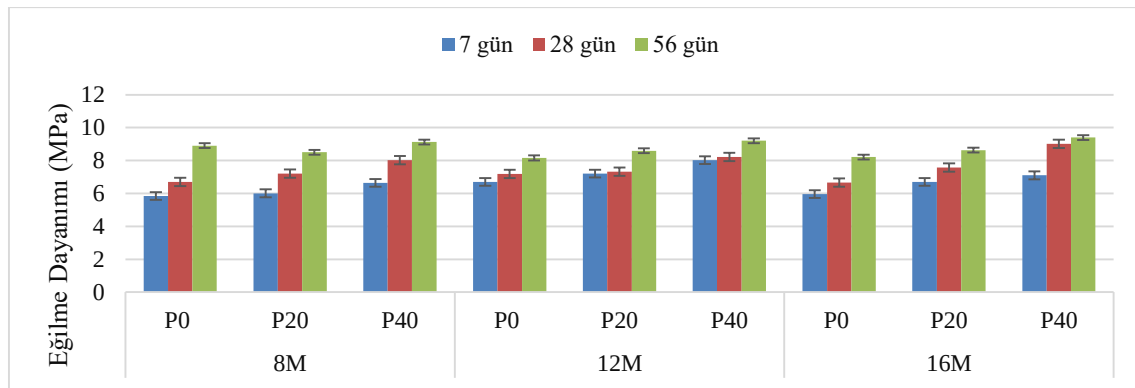
numunelerin farklı günlerdeki ortalama eğilme dayanım sonuçları gösterilmiştir. Geopolimer numunelere ısı kürü uygulanmasına rağmen, zaman arttıkça numunelerin eğilme dayanımları artmıştır. Bu da geopolimerizasyon sürecinin ilk günlerde tamamlanmadığını, devam eden geopolimerizasyon reaksiyonlarından ötürü oluşan C-S-H, C-A-S-H ve N-A-S-H jel oluşumunun eğilme dayanımını arttırdığı söylenebilir. Şekil 5’de ise sodyum hidroksit molaritesi değişiminin bazaltik pomza katkıli geopolimer harç numunelerin eğilme dayanıma olan etkisi gösterilmektedir. Sonuçlara göre, sodyum hidroksit molaritesi arttıkça pomza katkıli geopolimer numunelerin eğilme dayanımı az da olsa artmaktadır. Özellikle 7. gün eğilme dayanımları sonuçlarına göre en yüksek eğilme dayanımına sahip numuneler sodyum hidroksit molaritesi 12M olan numuneler olmuştur. Sonraki günlerde ise en yüksek eğilme dayanımları 16M geopolimer harç numunelerde olmuştur. Bunun sebebi olarak alkali ortamın kısa zamanda pomzadaki silisi çözmemesi olarak gösterilebilir. Elde edilen bu bulgular, pomza oranı ve sodyum hidroksit molaritesinin eğilme dayanımı üzerinde belirleyici bir etkiye sahip olduğunu göstermektedir.

Çizelge 5. Bazaltik pomza katkıli numunelerinin farklı günlerdeki ortalama eğilme dayanımı (MPa)

		Ortalama eğilme dayanımı (MPa)		
Numune adı		7 gün	28 gün	56 gün
8M	P40-8M	5,84	6,71	8,90
	P20-8M	6,01	7,20	8,51
	P0-8M	6,64	8,03	9,13
12M	P40-12M	6,71	7,19	8,16
	P20-12M	7,20	7,33	8,59
	P0 -12M	8,03	8,21	9,21
16M	P40-16M	5,96	6,67	8,21
	P20-16M	6,71	7,58	8,63
	P0-16M	7,10	9,02	9,40



Şekil 4. Bazaltik pomzalı geopolimer numunelerinin farklı günlerdeki eğilme dayanımı (MPa)



Şekil 5. Farklı molariteye sahip geopolimer numunelerinin ortalama eğilme dayanımı (MPa)

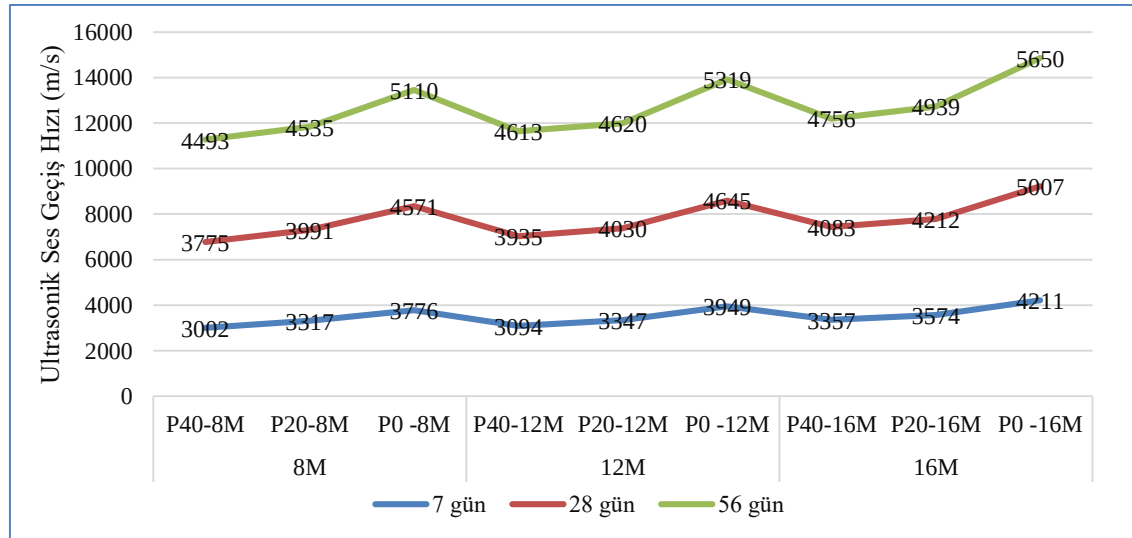
3.4. Ultrasonik Geçiş Hızı Sonuçları

Yapılan çalışma kapsamında bazaltik pomza katkılı geopolimer harç numunelerinin ultrasonik geçiş hızı (UPV) değerleri de elde edilmiştir. Deneyde TS EN 12504-4'te belirtilen ultrases geçiş hızının tayini deneyi uygulanmıştır. Deneyde, 0,1 μ s duyarlıklı ultrases aleti ile t, μ s ses geçiş süreleri ve ultrasonik geçiş hızı (m/s) değerleri ölçülmüştür. Çizelge 6'da bazaltik pomza katkılı geopolimer harç numunelerin UPV sonuçları gösterilmiştir. Neville [30] çalışmasında UPV değerlerine göre betonun mükemmel, iyi, orta ve kötü olarak sıralamıştır. UPV değerleri < 3000 m/s ise kötü, 3000-3500 m/s arası ise orta, 3500-4500 m/s arası ise iyi, >4500 m/s ise mükemmel olarak sıralamıştır.

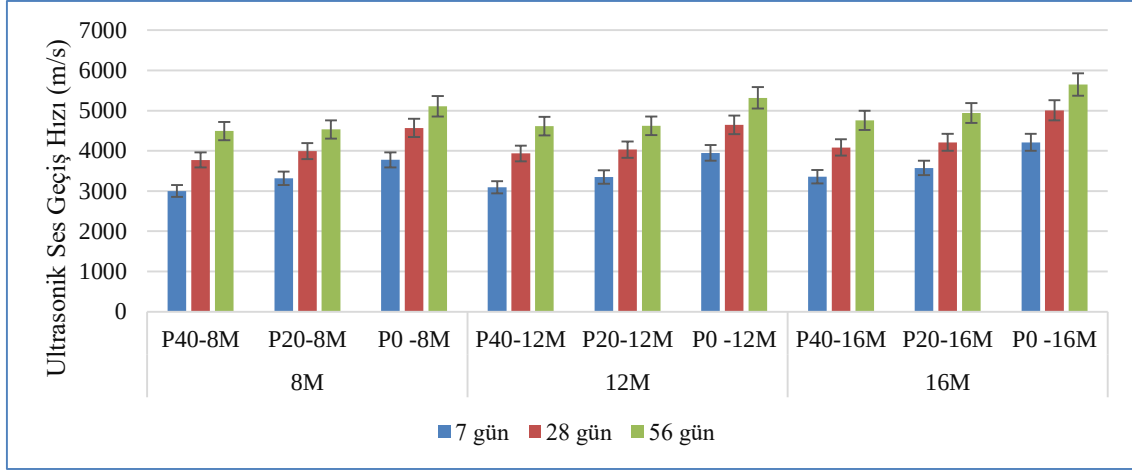
Çizelge 6. Pomza katkılı numunelerinin farklı günlerdeki ultrasonik ses geçiş hızı (UPV) (m/s)

		Ultrasonik ses geçiş hızı (UPV) (μ s)		
Numune adı		7 gün	28 gün	56 gün
8M	P40-8M	3002	3775	4493
	P20-8M	3317	3991	4535
	P0 -8M	3776	4571	5110
12M	P40-12M	3094	3935	4613
	P20-12M	3347	4030	4620
	P0 -12M	3949	4645	5319
16M	P40-16M	3357	4083	4756
	P20-16M	3574	4212	4939
	P0 -16M	4211	5007	5650

Şekil 6'da 7., 28. ve 56. günlerde UPV sonuçları gösterilmektedir. 7 günlük UPV sonuçları irdelendiğinde geopolimer harç numuneleri orta, 28 günlük UPV verilerine göre iyi, 56 günlük UPV sonuçlarına göre mükemmel olarak yorumlanabilir. UPV verilerinin düşük olması boşluklu bir iç yapıya ve düşük yoğunluğa işaret ederken, UPV verilerinin yüksek olması için boşluksuz ve yoğun bir iç yapıyı ve yüksek yoğunluğu göstermektedir. Şekil 7'de ise farklı molariteye sahip bazaltik pomza katkılı geopolimer harç numunelerin UPV değerleri gösterilmektedir. Pomza kullanım oranının artması geopolimer harç numunelerde UPV değerlerini düşürmüştür. UPV sonuçları basınç dayanımı sonuçları ile kıyaslandığında benzer sonuçlar elde edilmiştir. Dolayısıyla UPV sonuçları ile basınç dayanımı sonuçları arasında bir ilişki vardır. UPV değerlerinin yüksek olması geopolimer harç numunelerinde basınç dayanımının yüksek olacağına işaret etmektedir. Literatürde de UPV değerleri ile basınç dayanım değerleri arasında doğru orantılı bir ilişki olduğu rapor edilmiştir [31].



Şekil 6. Bazaltik pomzalı geopolimer numunelerinin ortalama ultrasonik ses geçiş hızı (m/s)



Şekil 7. Farklı molariteye sahip geopolimer numunelerinin ultrasonik ses geçiş hızı (m/s)

4. SONUÇLAR

Yapılan bu çalışmada %20 ve %40 oranlarında F-tipi uçucu kül ve öğütülmüş granüle yüksek fırın cürufu yerine kullanılan bazaltik pomza ile üretilmiş geopolimer numunelerin 7., 28., ve 56. günlerde basınç ve eğilme dayanımları ve ultrasonik ses geçiş hızı (UPV) sonuçları irdelenmiştir. Ayrıca farklı sodyum hidroksit (NaOH) molaritelerinin (8 M, 12 M ve 16 M) bazık pomza ikameli geopolimer harçların mekanik dayanımına olan etkisi de araştırılmıştır. Elde edilen sonuçlara göre;

Üretilen bazaltik pomzalı geopolimer harçların taze birim hacim ağırlığı, pomzasız karışımlara göre azalmıştır. Özellikle %40 ve %20 pomza içeren karışımlar, %0 pomza içeren karışımlara kıyasla daha düşük birim hacim ağırlığı değerlerine sahiptir. Bu durum, pomza kullanım oranı arttıkça geopolimer harçların ağırlığında azalma olduğunu ortaya koymaktadır. Sonuç olarak, pomza kullanımı geopolimer harçlarla daha hafif yapıların elde edilmesi sayesinde inşaat uygulamalarında önemli avantajlar sağlayabilir.

Basınç ve eğilme dayanımı sonuçlarına göre, pomza kullanım oranındaki artışın dayanım sonuçlarında bir düşüşe neden olduğu gözlemlenmiştir. Elde edilen dayanım sonuçlarına göre, en yüksek basınç ve eğilme dayanımları P0-16M numunelerinde elde edilirken, en düşük basınç ve eğilme dayanımları ise P40-8M geopolimer harç numunelerinde elde edilmiştir. Sonuçlara göre, geopolimer numunelere 65 °C 48 saat ısı kürü uygulanmasına rağmen, basınç ve eğilme dayanımlarındaki artışın zamanla devam ettiği görülmektedir. Bu da geopolimer numunelerde geopolimerizasyon reaksiyonlarının uzun süre devam ettiğini kanıtlamaktadır.

Basınç ve eğilme dayanım değerleri ultrasonik ses geçiş hızı değerleri ile ilişkili olup, elde edilen UPV değerlerine göre 7 günlük pomza katkılı numuneler orta, 28 günlük numuneler iyi, 56 günlük numuneler ise mükemmel olarak sınıflandırılmaktadır. Hem dayanım hem de UPV sonuçlarına göre yüksek fırın cürufu ve uçucu kül yerine %40 oranlarında bazaltik pomza bağlayıcı olarak geopolimer numunelerin üretiminde kullanılabilir.

8M, 12M ve 16M sodyum hidroksit molaritesinin dayanım sonuçlarına etkisi incelendiğinde, artan sodyum hidroksit molaritesi pomza katkılı geopolimer numunelerin dayanım değerlerini artırmıştır. Molaritenin artmasıyla beraber pomzadaki silisin çözünmesini daha kolay hale gelmekte, bu da dayanım artışına katkı sağlayan N-A-S-H, C-A-S-H ve C-S-H geopolimerizasyon ürünlerinin oluşumunu sağlayarak dayanıma katkıda bulunmaktadır.

5. KAYNAKLAR

1. Munir, Q., Abdulkareem, M., Horttanainen, M. & Kärki, T. (2023). A comparative cradle-to-gate life cycle assessment of geopolimer concrete produced from industrial side streams in comparison with traditional concrete. *Science of the Total Environment*, 865, 161230.

2. Qureshi, L.A., Ali, B. & Ali, A. (2020). Combined effects of supplementary cementitious materials (silica fume, GGBS, fly ash and rice husk ash) and steel fiber on the hardened properties of recycled aggregate concrete. *Construction and Building Materials*, 263, 120636.
3. Çevik, A. & Niş, A. (2023). Introduction to fiber-reinforced alkali-activated composites. In *Advanced Fiber-Reinforced Alkali-Activated Composites*, 1-21, Elsevier.
4. Yazıcıoğlu, S., Arıcı, E. ve Gönen, T. (2003). Pomza taşının kullanım alanları ve ekonomiye etkisi, *Fırat Üniversitesi DAUM Dergisi*, 1, 118-123.
5. Öz, E. (2007). Nevşehir dolaylarında yüzeylenen asidik pomzanın hafif beton agregası olarak kullanılabilirliği. *Yüksek Lisans Tezi*, Çukurova Üniversitesi Fen Bilimleri Enstitüsü, Adana, 71.
6. Dolgun, O. (2010). Kendiliğinden yerleşen betonlarda öğütülmüş pomza kullanılabilirliğinin araştırılması. *Yüksek Lisans Tezi*, Niğde Üniversitesi Fen Bilimleri Enstitüsü, Niğde, 88.
7. Pınarcı, İ. & Kocak, Y. (2022). Hydration mechanisms and mechanical properties of pumice substituted cementitious binder. *Construction and Building Materials*, 335, 127528.
8. Elmastaş, N. (2012). Türkiye ekonomisi için önemi giderek artan bir maden: Pomza (sünger taşı). *Journal of International Social Research*, 5(23), 197-206.
9. Aksay, E.K. (2005). İzmir-Menderes yöresi pomza cevherinin kullanımına yönelik teknolojik özelliklerinin araştırılması. *Doktora Tezi*, Dokuz Eylül Üniversitesi Fen Bilimleri Enstitüsü, İzmir, 289.
10. Yıldırım, A.N. (2007). Pomza ve uçucu kül kullanılarak imal edilen hafif betonların agresif su ortamlarında mekanik özelliklerinin araştırılması. *Yüksek Lisans Tezi*, Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya, 118.
11. Khandaker, M. & Hossain, A. (2002). Properties of volcanic pumice-based cement and lightweight concrete. *Cement and Concrete Research*, 2478.
12. Yasar, E., Atis, C.D., Kilic, A. & Gulsen, H. (2003). Strength properties of lightweight concrete made with basaltic pumice and fly ash. *Materials Letters*, 57(15), 2267-2270.
13. Gülşen, H. (2004). Bazaltik pomzanın hafif yapı malzemesi olarak kullanılabilirliği. *Yüksek Lisans Tezi*, Çukurova Üniversitesi, Fen Bilimleri Enstitüsü, Adana, 73.
14. Türkel, S. & Kadiroğlu, B. (2007). Pomza agregali taşıyıcı hafif betonun mekanik özelliklerinin incelenmesi. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 13(3), 353-359.
15. Zeyad, A.M., Tayeh, B.A. & Yusuf, M.O. (2019). Strength and transport characteristics of volcanic pumice powder based high strength concrete. *Construction and Building Materials*, 216, 314-324.
16. Tahwia, A.M., Aldulaimi, D.S., Abdellatief, M. & Youssf, O. (2024). Physical, mechanical and durability properties of eco-friendly engineered geopolimer composites. *Infrastructures*, 9(11), 191.
17. Yadollahi, M.M., Benli, A. & Demirboğa, R. (2015). The effects of silica modulus and aging on compressive strength of pumice-based geopolimer composites. *Construction and Building Materials*, 94, 767-774.
18. Safari, Z., Kurda, R., Al-Hadad, B., Mahmood, F. & Tapan, M. (2020). Mechanical characteristics of pumice-based geopolimer paste. *Resources, Conservation and Recycling*, 162, 105055.
19. Tahwia, A.M., Abdellatief, M., Salah, A. & Youssf, O. (2025). Valorization of recycled concrete powder, clay brick powder, and volcanic pumice powder in sustainable geopolimer concrete. *Scientific Reports*, 15(1), 11049.
20. Vafaei, M. & Allahverdi, A. (2016). Influence of calcium aluminate cement on geopolimerization of natural pozzolan. *Construction and Building Materials*, 114, 290-296.
21. Kurtoglu, A.E., Alzebaree, R., Aljumaili, O., Nis, A., Gulsan, M.E., Humur, G. & Cevik, A. (2018). Mechanical and durability properties of fly ash and slag based geopolimer concrete. *Advances in Concrete Construction*, 6(4), 345.
22. Erol, F., Al-mashhadani, M.M., Aygörmmez, Y. & Niş, A. (2023). Effect of ceramic waste powder content and sodium hydroxide molarity on the residual mechanical strength of alkali-activated mortars. *Materials Chemistry and Physics*, 309, 128403.
23. Olivia, M. & Nikraz, H. (2012). Properties of fly ash geopolimer concrete designed by Taguchi method. *Materials & Design*, 36, 191-198.
24. Balun, B. & Karataş, M. (2021). Influence of curing conditions on pumice-based alkali activated composites incorporating Portland cement. *Journal of Building Engineering*, 43, 102605.
25. Kabay, N., Mert, M., Miyan, N. & Omur, T. (2021). Pumice as precursor in geopolimer paste and mortar. *Journal of Civil Engineering and Construction*, 10(4), 225-236.
26. Bakharev, T. (2005). Geopolimeric materials prepared using Class F fly ash and elevated temperature curing. *Cement and Concrete Research*, 35(6), 1224-1232.
27. Palomo, A., Grutzeck, M.W. & Blanco, M.T. (1999). Alkali-activated fly ashes: A cement for the future. *Cement and Concrete Research*, 29(8), 1323-1329.

28. Thunuguntla, C.S. & Rao, T.G. (2018). Effect of mix design parameters on mechanical and durability properties of alkali activated slag concrete. *Construction and Building Materials*, 193, 173-188.
29. Al-Antaki, T.S.W. & Niş, A. (2025). The influences of sulfuric acid and magnesium sulfate attacks on pumice powder incorporated alkali-activated mortars with different sodium hydroxide molarities. *Construction and Building Materials*, 493, 143136.
30. Neville, A.M. (1995). *Properties of concrete*. 4th Edition, Addison Wesley Longman Ltd., Essex.
31. Mehmetoğlu, M. (2023). Strength investigation of slag-based geopolymer composites incorporating different amounts of colemanite waste and silica fume under different exposure conditions. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 38(3), 841-849.
32. Gündüz, L. & Kalkan, Ş.O. (2022). Anhidrit III katkısının çimento esaslı harcın performansına etkisi. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 37(4), 959-972.



Structural Design of the Electric Vehicle Components Using Runge-Kutta Optimization Algorithm

Dildar GÜRSES^{1,a}

¹Bursa Uludağ University, Hybrid and Electric Vehicle Technology Division, Department of Electric and Energy, Bursa, Türkiye

^aORCID: 0000-0002-1517-1692

Article Info

Received : 26.08.2025

Accepted : 23.09.2025

DOI: 10.21605/cukurovaumfd.1770296

Corresponding Author

Dildar GÜRSES

dildargurses@uludag.edu.tr

Keywords

Shape optimization

Structural optimization

Battery pack housing profile

Artificial neural network

Runge-Kutta optimization algorithm

How to cite: GÜRSES, D., (2025).

Structural Design of The Electric Vehicle Components Using Runge-Kutta Optimization Algorithm. Çukurova University, Journal of the Faculty of Engineering, 40(3), 699-705.

ABSTRACT

This research focuses on the optimization of a battery pack housing profile of an electric vehicle using advanced structural optimization techniques, specifically topology and shape optimization. The newly developed Runge-Kutta optimization algorithm(RKOA) was applied to minimize the component's weight without violating allowable stress constraints for the shape optimization phase. The artificial neural network (ANN) surrogate modeling method was employed to model the objective and constraint functions accurately. The findings indicate that RKOA delivers improved optimization results compared to the Harris Hawks algorithm, demonstrating its effectiveness in achieving an electric vehicle design's lightweight and structurally sound battery pack housing profile.

Runge-Kutta Optimizasyon Algoritması Kullanılarak Elektrikli Araç Bileşenlerinin Yapısal Tasarımı

Makale Bilgileri

Geliş : 26.08.2025

Kabul : 23.09.2025

DOI: 10.21605/cukurovaumfd.1770296

Sorumlu Yazar

Dildar GÜRSES

dildargurses@uludag.edu.tr

Anahtar Kelimeler

Şekil optimizasyonu

Yapısal optimizasyon

Batarya paketi muhafaza profili

Artificial neural network

Runge-Kutta optimizasyon algoritması

Atıf şekli: GÜRSES, D., (2025). Runge-

Kutta Optimizasyon Algoritması Kullanılarak Elektrikli Araç Bileşenlerinin Yapısal Tasarımı. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 699-705.

ÖZ

Bu araştırma, gelişmiş yapısal optimizasyon teknikleri, özellikle topoloji ve şekil optimizasyonu kullanılarak bir elektrikli otomobil batarya paketi muhafaza profili optimizasyonuna odaklanmaktadır. Şekil optimizasyonu aşamasında, izin verilen gerilme kısıtlamalarını ihlal etmeden bileşenin ağırlığını en aza indirmek için yeni geliştirilen Runge-Kutta optimizasyon algoritması (RKOA) uygulanmıştır. Hedef ve kısıt fonksiyonlarını doğru bir şekilde modellemek için yapay sinir ağı (ANN) vekil modelleme yöntemi kullanılmıştır. Bulgular, RKOA'nın Harris Hawks algoritmasına kıyasla daha iyi optimizasyon sonuçları verdiğini ve hafif ve yapısal olarak sağlam bir elektrikli araç batarya paketi muhafaza kutusu tasarımı elde etmede etkinliğini gösterdiğini ortaya koymaktadır.

1. INTRODUCTION

The automobile has been widely used in daily life for over a century. The demand for low-cost and lightweight vehicle components is increasing in the automotive industry to meet cost-effective solutions. Automotive components are often subjected to unpredictable loads that can cause localized deformations. Therefore, defining proper design criteria is crucial to ensure the desired performance characteristics. Among these components, the electric vehicle's battery pack housing profile plays a vital role in vehicle safety.

Recently, software tools based on structural optimization techniques have gained considerable attention for enabling rapid and efficient design of automotive components [1-9]. For instance, [2] aimed to reduce weight in electric vehicle body structures by employing carbon twill weave fabric composites. In [7], an automotive cross-car beam's topology, size, and shape optimization were investigated to enhance its structural performance. Using advanced finite element analysis and optimization techniques, they explored design variations that minimize weight while maintaining stiffness and crashworthiness. The study confirmed that optimized cross-car beam designs can significantly reduce mass without sacrificing safety or functionality. Various design scenarios were analyzed numerically to evaluate structural performance under lightweight conditions. The results demonstrated that the composite body structure can provide significant mass savings without compromising safety and stiffness. As the service life of a vehicle is closely linked to its weight, weight reduction efforts are paramount in automotive design. Topology optimization is a widely adopted structural optimization approach used to generate optimal geometries and material distributions for components. This method offers an efficient material layout that assists designers in initiating the product development process with a strong foundation.

Swarm intelligence-based optimization techniques have been successfully implemented to solve real-world problems across many disciplines, including the automotive and healthcare sectors. These methods offer effective solutions for almost all design problems and outperform conventional techniques by not requiring derivative information [10-20].

With the increasing interest in automobiles, the selection of appropriate optimization methods has become crucial to achieving optimal designs and contributing to efficient product development. In this study, the Runge-Kutta optimization algorithm (RKO) was employed to design the optimal battery pack housing profile of an electric vehicle. Introduced in [10], the Runge-Kutta optimization algorithm (RKO) has been successfully applied to problems such as gear design, three-bar truss structures, and marine propeller optimization. Moreover, [16] demonstrated its superiority in solving complex engineering design optimization problems, including brake disc design, welded beam problems, DC motor design, and motor speed reducers, outperforming other algorithms in these cases.

Given its proven success in engineering design applications, the Runge-Kutta optimization algorithm (RKO) was selected for use in this study. An electric vehicle's battery pack housing profile was optimized using the Runge-Kutta optimization algorithm and the Harris Hawks algorithm. Comparative results demonstrate that the Runge-Kutta optimization algorithm (RKO) performs better. These findings indicate that the Runge-Kutta optimization algorithm (RKO) can be effectively employed to solve challenging design optimization problems.

2. MATERIAL AND METHOD

This paper uses the Runge-Kutta optimization algorithm (RKO) to optimize battery pack housing profiles of electric vehicles. The Runge-Kutta optimization algorithm (RKO) [16] is a metaheuristic technique inspired by the classical Runge-Kutta method used in numerical analysis. It applies a step-by-step prediction and correction strategy to improve the precision of solution updates throughout the optimization process. The Runge-Kutta optimization algorithm (RKO) effectively balances global exploration and local exploitation by modeling solution trajectories in a multi-dimensional space. It is especially suited for tackling challenging nonlinear optimization problems where conventional algorithms often fall short. Additionally, RKO's adaptive nature helps preserve diversity within the population and reduces the risk of getting trapped in local optima.

3. STRUCTURAL OPTIMIZATION OF THE ELECTRIC VEHICLE BATTERY PACK HOUSING PROFILE

The battery pack housing profile of an electric vehicle has a vital role in vehicle design, as shown in Figure 1. This research optimized the structural design using the Runge-Kutta optimization algorithm (RKOA). ST37 steel was selected as the construction material due to its suitable mechanical properties. The optimization aimed for performance while maintaining structural integrity.



Figure 1. Battery pack housing profile of electric vehicle three-dimensional representation

This design problem can be formulated as follows:

$$\text{Min } F(x) = \text{Weight}(x) \quad (1)$$

Constraints:

$$\text{Stress}(x) \leq \text{Maximum Stress} \quad (2)$$

$$x_i^l \leq x_i \leq x_i^u, \quad i = 1, \text{TSD}, \quad (3)$$

The goal of the problem is to reduce the structural mass by adjusting the design variables within the specified lower (x_l) and upper (x_u) bounds. A stress constraint is included to ensure the design remains safe under operational loading conditions. The maximum stress is limited to 200 MPa, which corresponds to the yield strength of the chosen material. This constraint is represented by the peak stress experienced by the component under the defined boundary conditions.

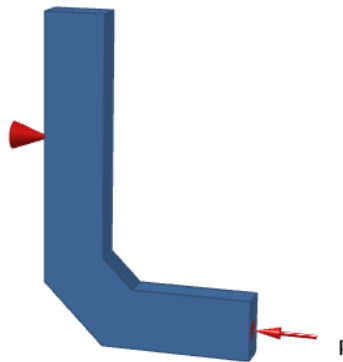


Figure 2. Boundary conditions of the battery pack housing profile of an electric vehicle

The initial geometry and boundary conditions applied to the battery pack housing profile of the electric vehicle are presented in Figures 1 and 2, respectively. The finite element model developed for the analysis comprises 13.026 nodes and 11.415 elements. In the conducted topology optimization, the objective function was set as compliance minimization, while volume reduction served as the constraint. This configuration established the foundation for the battery pack housing profile of the electric vehicle's topology optimization process.

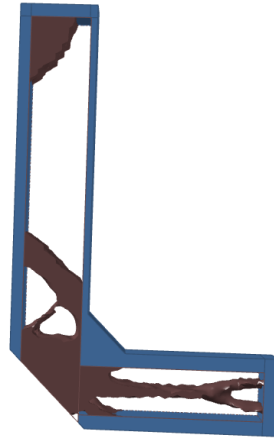


Figure 3. Optimized material layout from topology optimization

Using the topology optimization results in Figure 3, a new battery pack housing profile for electric vehicle design was developed. The white areas in Figure 3 represent regions where material can be eliminated. Guided by these results, a redesigned 3D model was created, as shown in Figure 4, reflecting the optimized geometry suggested by the finite element analysis. Stress distribution for the redesigned battery pack housing profile of the electric vehicle is shown in Figure 5.

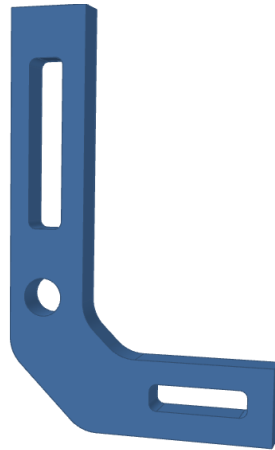


Figure 4. Redesigned the battery pack housing profile of an electric vehicle

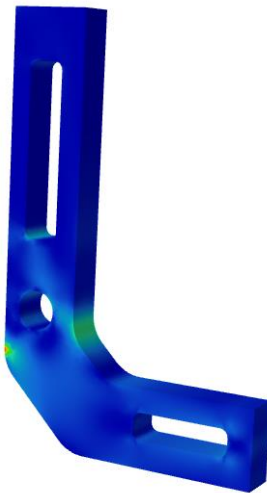


Figure 5. Stress distribution for the redesigned battery pack housing profile of an electric vehicle

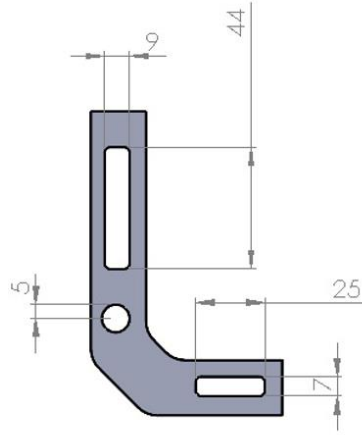


Figure 6. Design variables

In the battery pack housing profile of electric vehicle's shape optimization, three key design parameters were selected, as illustrated in Figure 6, with their values restricted between specific upper and lower bounds: $5 < X_1 < 20$, $6 < X_2 < 22$, $8 < X_3 < 26$, $30 < X_4 < 50$, and $5 < X_5 < 15$. The optimization problem, governed by Equations (1–3), was addressed using the Runge-Kutta optimization algorithm (RKOA) applied in this study. The artificial neural network (ANN) surrogate modeling technique was employed to obtain the objective and constraint functions accurately.

The objective was to achieve the minimum possible weight for an electric vehicle's battery pack housing profile while maintaining stress levels within safe limits. Both the goal and the constraint functions were explicitly formulated to reflect these conditions.



Figure 7. The optimum battery pack housing profile of electric vehicle design obtained using the Runge-Kutta optimization algorithm

The final optimized design obtained through the shape optimization process is depicted in Figure 7. In Figure 8, the stress distribution for the optimum battery pack housing profile of an electric vehicle is shown.

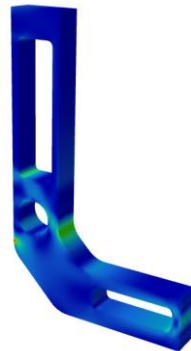


Figure 8. Stress distribution for optimum battery pack housing profile of an electric vehicle

Table 1 summarizes the outcomes of the battery pack housing profile of electric vehicle optimization using the RKO algorithm. Following the optimization, the weight of the battery pack housing profile of the electric vehicle decreased from 1252.6 grams to 1068.3 grams. This change represents a weight reduction of approximately 14.7% compared to the initial design.

Table 1. Optimization outcomes for the electric vehicle battery box side profile

Method	Weight (grams)	Stress (MPa)
Initial design	1252.6	118
Redesigned model of the side profile	1194.1	136
Harris Hawks algorithm	1134.1	178
Runge-Kutta optimization algorithm	1068.3	190

Table 2 displays the comparison results obtained from 30 separate optimization runs. Each run was conducted with a population size of 30 and iterated over 100 generations. The Runge-Kutta optimization algorithm (RKO) outperformed the Harris Hawks algorithm (HHA) across all key metrics, including minimum, maximum, mean, and standard deviation. These findings highlight the effectiveness and consistency of RKO in addressing the shape optimization of the battery pack housing profile of an electric vehicle.

Table 2. Optimization statistics for the battery pack housing profile of an electric vehicle

Method	Minimum weight (grams)	Mean weight (gram)	Maximum weight (gram)
Initial design	1252.6	1252.6	1252.6
Redesigned model of the side profile	1194.1	1194.1	1194.1
Harris Hawks algorithm	1134.1	1150.4	1180.6
Runge-Kutta optimization algorithm	1068.3	1085.7	1098.6

4. RESULTS AND DISCUSSION

This research explores the application of the Runge-Kutta optimization algorithm (RKO) for the structural optimization of the battery pack housing profile of electric vehicle components. The study aims to support the development of lightweight yet structurally sound designs, addressing the increasing demands of optimization in the automotive sector. By applying the Runge-Kutta optimization algorithm (RKO), the weight of the battery pack housing profile of the electric vehicle was reduced by 14.7% from 1252.6 g to 1068.3 g without exceeding allowable stress thresholds. A comparative evaluation in Tables 1 and 2 confirms that the Runge-Kutta optimization algorithm (RKO) outperforms the Harris Hawks algorithm in terms of efficiency, reliability, and consistency. The results demonstrate that the Runge-Kutta optimization algorithm (RKO) is a promising tool for tackling various engineering optimization tasks.

5. REFERENCES

1. Belingardi, G. & Scattina, A. (2023). Battery pack and underbody: integration in the structure design for battery electric vehicles-challenges and solutions. *Vehicles*, 5(2), 498-514.
2. Liu, Q., Lin, Y., Zong, Z., Sun, G. & Li, Q. (2013). Lightweight design of carbon twill weave fabric composite body structure for electric vehicle. *Composite Structures*, 97, 231-238.
3. Zhang, J., Ning, L., Hao, Y. & Sang, T. (2021). Topology optimization for crashworthiness and structural design of a battery electric vehicle. *International Journal of Crashworthiness*, 26(6), 651-660.
4. Acar, E., Jain, N., Ramu, P., Hwang, C. & Lee, I. (2024). A survey on design optimization of battery electric vehicle components, systems, and management. *Structural and Multidisciplinary Optimization*, 67(3), 27.
5. Roper, S.W.K. & Kim, I.Y. (2023). Integrated topology and packaging optimization for conceptual-level electric vehicle chassis design via the component-existence method. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 237(9), 2118-2131.
6. Yu, L., Gu, X., Qian, L., Jiang, P., Wang, W. & Yu, M. (2021). Application of tailor rolled blanks in optimum design of pure electric vehicle crashworthiness and lightweight. *Thin-Walled Structures*, 161, 107410.

7. Li, C. & Kim, I.Y. (2015). Topology, size and shape optimization of an automotive cross car beam, *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 229(10), 1361-1378.
8. Kim, J., Kim, J.J. & Jang, I.G. (2022). Integrated topology and shape optimization of the five-spoke steel wheel to improve the natural frequency. *Structural and Multidisciplinary Optimization*, 65(3), 78.
9. Grujicic, M., Arakere, G., Pisu, P., Ayalew, B., Seyr, N., Erdmann, M. & Holzleitner, J. (2008). Application of topology, size and shape optimization methods in polymer metal hybrid structural lightweight engineering. *Multidiscipline modeling in Materials and Structures*, 4(4), 305-330.
10. Mirjalili, S. (2015). The ant lion optimizer. *Advances in engineering software*, 83, 80-98.
11. Miao, Y., Fadel, G.M. & Gantovnik, V.B. (2008). Vehicle configuration design with a packing genetic algorithm. *International Journal of Heavy Vehicle Systems*, 15(2-4), 433-448.
12. Montazeri-Gh, M. & Poursamad, A. (2006). Application of genetic algorithm for simultaneous optimisation of HEV component sizing and control strategy. *International Journal of Alternative Propulsion*, 1(1), 63-78.
13. Li, Z., Pourmehrab, M., Elefteriadou, L. & Ranka, S. (2018). Intersection control optimization for automated vehicles using genetic algorithm. *Journal of Transportation Engineering, Part A: Systems*, 144(12), 04018074.
14. Hui, S. (2010). Multi-objective optimization for hydraulic hybrid vehicle based on adaptive simulated annealing genetic algorithm. *Engineering Applications of Artificial Intelligence*, 23(1), 27-33.
15. Lu, X., Wu, Y., Lian, J., Zhang, Y., Chen, C., Wang, P. & Meng, L. (2020). Energy management of hybrid electric vehicles: A review of energy optimization of fuel cell hybrid power system based on genetic algorithm. *Energy Conversion and Management*, 205, 112474.
16. Ahmadianfar, I., Heidari, A.A., Gandomi, A.H., Chu, X. & Chen, H. (2021). RUN beyond the metaphor: An efficient optimization algorithm based on Runge Kutta method. *Expert Systems with Applications*, 181, 115079.
17. Pınarbaşı, A., Külekçi, M.K., Boğa, C. & Eşme, U. (2020). Optimization of the effect of processing parameters on surface roughness and cutting energy in CNC Milling of Al-7075 material. *Çukurova Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 35(2), 345-356.
18. Alhasan, H.A. & Güneş, M. (2017). Yeni bir adaptif parçacık sürü optimizasyon algoritması kullanarak DC motor için öz ayarlamalı PID kontrolör tasarımı. *Çukurova Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 32(3), 243-250.
19. Ulukök, M.K. (2023). Çift-girişim tabanlı iyileştirme algoritmasının sayısal iyileştirme fonksiyonları üzerinde performans analizi. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 38(2), 545-552.
20. Keleş, M.K. & Keleş, A.E. (2017). Veri madenciliği uygulamalarının ve sezgisel optimizasyon algoritmalarının yapım yönetimindeki yeri. *Çukurova Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 32(1), 235-242.



Sağlık Alanında Açıklanabilir Yapay Zekâ Kullanımı Üzerine Bir Uygulama

Gülhan TOĞA^{1,a}

¹Erciyes Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Kayseri, Türkiye

^aORCID: 0000-0001-8835-1769

Makale Bilgileri

Geliş : 30.06.2025

Kabul : 25.09.2025

DOI: 10.21605/cukurovaumfd.1729986

Sorumlu Yazar

Gülhan TOĞA

gpala@erciyes.edu.tr

Anahtar Kelimeler

Açıklanabilir yapay zeka

SHAP

Veri madenciliği

Meme kanseri

Atıf şekli: TOĞA, G., (2025). Sağlık Alanında Açıklanabilir Yapay Zekâ Kullanımı Üzerine Bir Uygulama. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 707-723.

ÖZ

Büyük verilerin hayatımıza girmesi ile bu verilerle elde edilebilecek bilgiler de artmıştır. Özellikle derin öğrenmeye yönelik olmak üzere makine öğreniminin yaygın kullanımı sayesinde büyük veri kümelerinden anlamlı çıkarımlar yapabilmek mümkün hale gelmiştir. Ancak son kullanıcılar, oluşturulan modellerin şeffaflık, yorumlanabilirlik ve güvenilirlik gibi bazı eksiklikleri ile karşı karşıya kalmaktadır. Bu eksikliklerin giderilmesi için geliştirilen “Açıklanabilir Yapay Zekâ (AYZ)”, sistematik bir yaklaşım ile oluşturulan modelin ve alınan kararların kullanıcılar tarafından anlaşılmasını ve modellerin güvenilirliğinin nasıl değerlendirilmesi gerektiğini açıklayabilmektedir.

Bu çalışmada sağlık alanında AYZ kullanımı ve faydaları hakkında bilgi verilmektedir. Ayrıca çalışma, sağlık alanında AYZ kullanımına örnek olması açısından, meme kanseri hastalarının çeşitli parametrelere bağlı olarak sağ/ölü durumlarının öngörülmesine olanak sağlayan bir uygulama içermektedir. Öncelikle, makine öğrenmesi algoritmaları kullanılarak sınıflandırma çalışması yapılmıştır. Ardından en başarılı sınıflandırma modelinin sonuçlarının açıklanabilirliği için bir AYZ yöntemi olan SHAP tekniğine başvurulmuştur. Sonuçlar, AYZ kullanımının modelin karar verme sürecinin nedenlerini daha iyi anlamamıza yardımcı olduğunu göstermektedir.

An Application on the Use of Explainable Artificial Intelligence in the Healthcare

Article Info

Received : 30.06.2025

Accepted : 25.09.2025

DOI: 10.21605/cukurovaumfd.1729986

Corresponding Author

Gülhan TOĞA

gpala@erciyes.edu.tr

Keywords

Explainable artificial intelligence

SHAP

Data mining

Breast cancer

How to cite: TOĞA, G., (2025). An Application on the Use of Explainable Artificial Intelligence in the Healthcare. Çukurova University, Journal of the Faculty of Engineering, 40(3), 707-723.

ABSTRACT

With the advent of big data, the potential to derive valuable insights has significantly increased. The widespread use of machine learning, especially deep learning, has made it possible to extract meaningful patterns from large datasets. However, end-users often encounter issues related to the transparency, interpretability, and reliability of these models. To address such concerns, Explainable Artificial Intelligence (XAI) has been developed to systematically clarify how models work and how their decisions can be understood and trusted.

This study explores the use and benefits of XAI in the healthcare domain. As an example, it presents an application that predicts the survival status (alive or deceased) of breast cancer patients based on various clinical parameters. Initially, classification was performed using machine learning algorithms. To interpret the results of the most successful model, the SHAP (SHapley Additive exPlanations) technique, a prominent XAI method, was applied. Findings indicate that XAI enhances understanding of model decisions and supports trustworthy AI applications in healthcare.

1. GİRİŞ

Yapay zeka (YZ) günümüzde birçok endüstride önemli bir rol oynamaktadır. YZ, bilgisayar sistemlerine insan benzeri zekaya sahip olma yeteneği kazandıran bir bilim dalıdır. Bu teknoloji, büyük veri setlerinden anlamlı bilgiler çıkarma, karmaşık problemleri çözme, öngörü yapma ve kararlar alma gibi bir dizi avantaj sağlamaktadır. YZ, birçok farklı alanda çeşitli uygulamalara sahip çok yönlü bir teknolojidir. YZ, finansal hizmetlerde; risk değerlendirmesi, sahtecilik tespiti, portföy yönetimi ve müşteri hizmetleri gibi alanlarda kullanılmaktadır. Perakende ve e-ticarette, büyük veri analizi ve öneri sistemleri gibi YZ teknikleri, müşteri davranışlarını anlama, stok yönetimi, fiyatlandırma ve kişiselleştirilmiş pazarlama gibi alanlarda kullanılmaktadır. YZ Ulaşım ve lojistik alanında; trafik yönetimi, rotalama, teslimat planlaması ve sürücüsüz araç teknolojileri gibi alanlarda önemli bir rol oynamaktadır. Enerji ve çevre sektörlerinde, enerji verimliliği, güneş paneli optimizasyonu, akıllı şebeke yönetimi gibi alanlarda enerji tüketimini optimize etmeye yardımcı olurken, çevre koruma alanında da su kalitesi izleme, hava kirliliği izleme gibi uygulamalarda kullanılmaktadır. Eğitim sektöründe ise YZ, öğrenci performansını izleme, kişiselleştirilmiş öğrenme deneyimleri sunma ve eğitim materyallerini geliştirme gibi uygulama alanlarına sahiptir. YZ, sağlık sektöründe, hastalıkların erken teşhisi, tedavi planlaması, tıbbi görüntü işleme ve ilaç keşfi gibi konularda kullanılabilmektedir [1].

YZ görüldüğü üzere, veri yoğun ve karar verme gerektiren hemen hemen her alanda kullanılabilecek geniş bir potansiyele sahiptir. Ancak, YZ'nin kullanımı bazı dezavantajları da beraberinde getirmektedir. Özellikle, derin öğrenme gibi karmaşık YZ modelleri, genellikle karar alma süreçlerinin nedenlerini açıklamakta zorlanmaktadır ve bu da güvenilirlik, şeffaflık ve yorumlanabilirlik eksikliklerine yol açabilmektedir. [2]. YZ'nin karar alma süreçlerinde nihai karara ulaşmasına olanak sağlayan yapı kara kutu modeller olarak nitelendirildiği için, alınan kararların gerekçelerinin son kullanıcı tarafından anlaşılabilirliği bir sorun olarak karşımıza çıkmakta, bu da güvenilirlik noktasında eksikliklere neden olmaktadır. Bu eksiklikleri gidermek için Açıklanabilir Yapay Zeka geliştirilmiştir [3]. AYZ, yapay zeka modellerinin karar verme süreçlerini şeffaf hale getirerek kullanıcıların modelin işleyişini anlamasına ve alınan kararların güvenilir biçimde değerlendirilmesine olanak sağlamaktadır. AYZ için kullanılan teknikler arasında LIME (Local Interpretable Model-agnostic Explanations), SHAP (SHapley Additive exPlanations) ve ICE (Individual Conditional Expectation) gibi yöntemler bulunmaktadır [4].

Sağlık alanında, YZ teknolojisinin geniş bir kullanım alanı mevcuttur. Hastalara daha iyi bakım sağlamak ve sağlık hizmetlerini optimizasyonu, hastalıkların erken teşhisi, tedavi planlaması, hasta monitörizasyonu, tıbbi görüntüleme analizi ve ilaç keşfi gibi birçok alanda sıklıkla kullanılmaktadır [5]. Ancak, bu alanlarda kullanılan karmaşık YZ modellerinin karar alma süreçlerini açıklamakta zorlanması, sağlık profesyonelleri ve hastalar arasında güven sorunlarına neden olabilir. Bu nedenle AYZ'nin sağlık alanında kullanımı eksikliklerin giderilmesini sağlar. AYZ, karmaşık yapay zeka modellerinin karar alma süreçlerini açıklayarak, sağlık profesyonellerine ve hastalara kararların nedenlerini anlama ve güvenle değerlendirme olanağı sunar. Örneğin, bir hastalığın teşhis edilmesinde kullanılan AYZ modelleri, hangi belirtilerin teşhis için belirleyici olduğunu ve teşhisin nasıl yapıldığını anlatarak, doktorların hastalık durumunu daha iyi anlamalarına yardımcı olabilir [6]. Bu da tedavi planlaması ve hasta bakımında daha iyi kararlar alınmasına olanak tanımaktadır. Sağlık alanında, AYZ'nin kullanımı sadece hastaların daha iyi bakım almasını sağlamakla kalmaz, aynı zamanda sağlık hizmetlerinin etkinliğini artırır, maliyetleri azaltır ve sağlık profesyonellerinin karar alma süreçlerinde daha güvenli hissetmelerini sağlar. Bu nedenle, sağlık alanında AYZ'nin kullanımı giderek daha büyük bir önem kazanmaktadır [7].

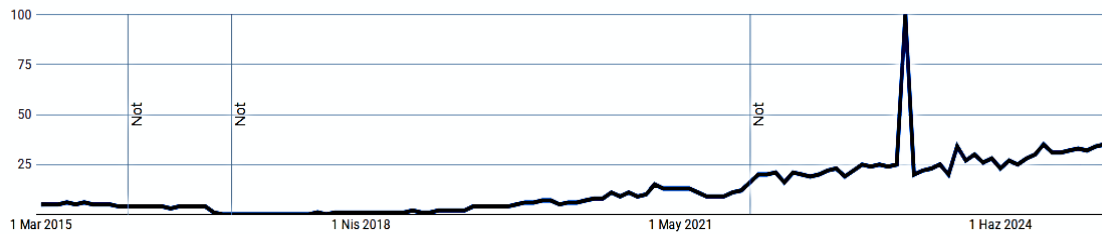
AYZ'nin sağlık alanında kullanımı da hızla yaygınlaşmaktadır. Hastalık teşhisi için literatürde çok farklı örnekleri bulunmaktadır. Maymun Çiçeği deri lezyonlarının sınıflandırılmasında [8], tüberküloz teşhisinde [9], beyin tümörü tespitinde [10], Alzheimer hastalığının sınıflandırılmasında [11], akciğer kanseri tespiti [12] vb. çalışmalarda çoğunlukla da görüntü işleme ile entegre AYZ teknikleri kullanılmaktadır. Ayrıca ameliyat yöntemlerinin değerlendirilmesi aşamasında da literatür hızla genişlemekte ve ameliyat yöntemlerinin hangisinin tercih edileceği noktasında hekimlere açıklanabilir, şeffaf kararlar sunulmaktadır [13]. Sağlık alanında AYZ uygulamaları ile ilgili detaylı literatür taramaları için Yang [14], Tjoa ve Guan [15] ve Sun ve arkadaşları [16]'nın çalışmaları irdelenebilir.

Mevcut literatürde sağlık alanında AYZ uygulamaları ağırlıklı olarak görüntü işleme temelli tanı sistemlerine odaklanmaktadır. Bununla birlikte, klinik ve demografik parametrelerin açıklanabilir modellerle değerlendirilmesine yönelik çalışmalar sınırlıdır. Bu boşluğu gidermek amacıyla, bu çalışma meme kanseri hastalarının hayatta kalma durumunu makine öğrenmesi algoritmaları ile sınıflandırmakta ve en başarılı modelin sonuçlarını SHAP yöntemi ile yorumlamaktadır. Çalışmanın özgün katkısı, yalnızca doğru tahminler yapmakla kalmayıp, aynı zamanda tahminlerin hangi faktörlerden kaynaklandığını şeffaf biçimde ortaya koymasındır. Böylece sağlık alanında AYZ tabanlı karar destek sistemlerine duyulacak güvenin artırılması hedeflenmektedir.

Bu çalışmada AYZ'nin sağlık alanında ne amaçla ve nasıl kullanılabileceğini açıklamak için detaylı bir örnek verilecektir. Çalışmanın 2. Bölümünde kullanılacak olan yöntemlerden bahsedilecek olup, 3. Bölümde sağlık alanında örnek bir AYZ uygulaması yapılacaktır. 4. Bölümde sonuçlar irdelenecektir. 5. Bölümde gelecek çalışmalardan bahsedilecektir.

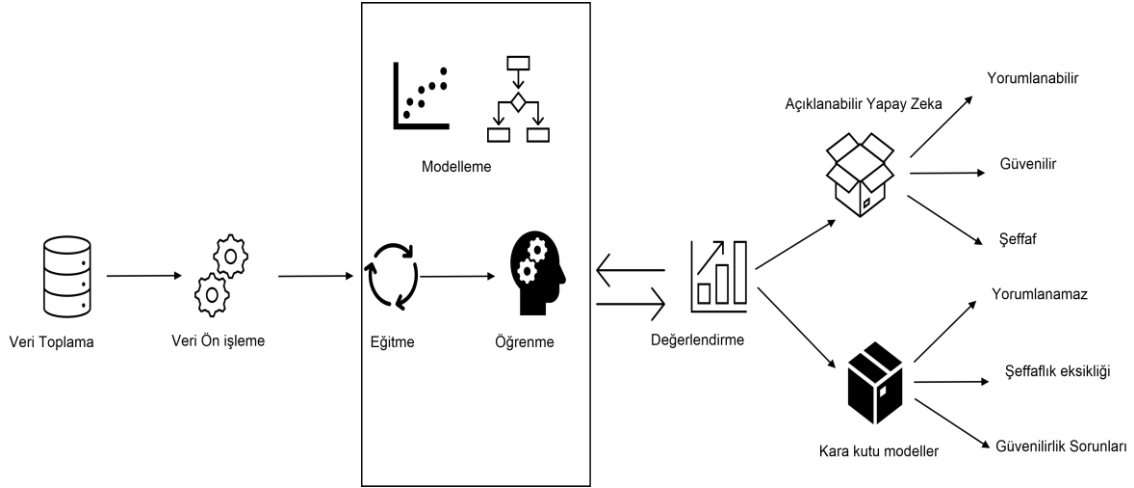
2. YÖNTEM

Şeffaflık, yorumlanabilirlik ve açıklanabilirlik AYZ'nin kara kutu modellere nazaran sağladığı en önemli avantajlardır. Kara kutu modeller, son kullanıcıların modelin davranışını anlamasını zorlaştırmakla beraber, potansiyel önyargıları veya hataları tespit etmelerini veya modeli kararlarından sorumlu tutmalarını engeller. Bu eksikliklerin giderilmesi amacı ile kullanımı gittikçe yaygınlaşan AYZ, akademik çalışmalar ve uygulayıcılar tarafından yaygın şekilde kullanılmaya başlanmıştır. Google Trend verilerinden elde edilen bilgilere göre son 10 yıl için dünya genelinde “Açıklanabilir Yapay Zeka (Explainable Artificial Intelligence)” terimine gösterilen ilgi Şekil 1’de görüldüğü üzere 2018 sonrasında keskin bir şekilde artış göstermiştir. Bu artışın arkasında Avrupa Birliği'nin GDPR (Genel Veri Koruma Tüzüğü) gibi düzenlemelerle otomatik karar alma mekanizmalarında şeffaflığı zorunlu kılması, yapay zekanın sağlık ve finans gibi yüksek riskli alanlarda yaygınlaşmasıyla hesap verilebilirliğe olan ihtiyacın artması ve etkili açıklanabilirlik yöntemlerinin literatüre kazandırılması gibi önemli etkilerin olduğu düşünülmektedir.



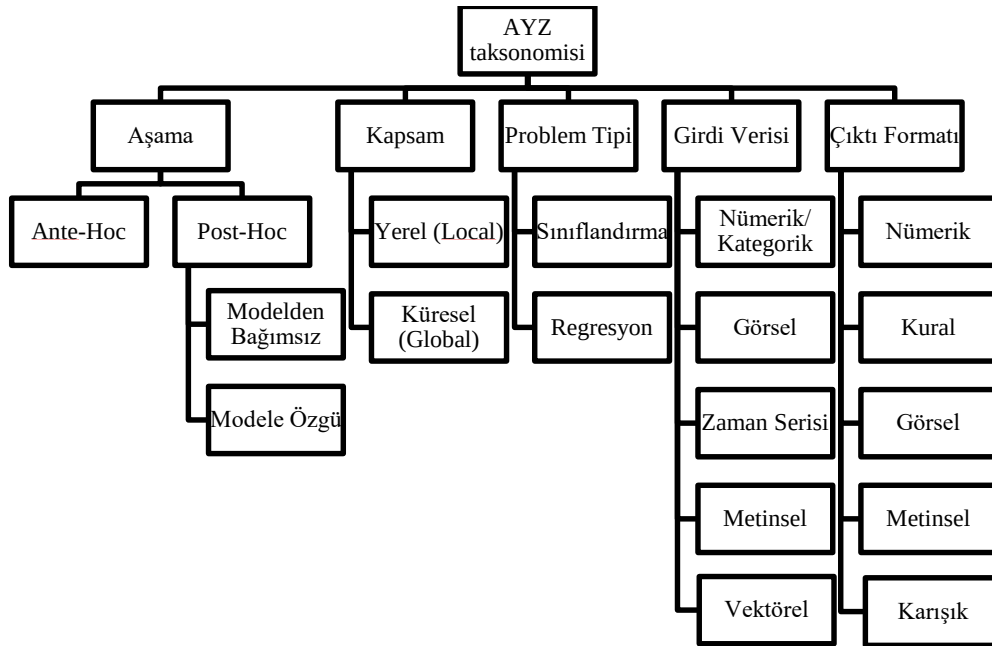
Şekil 1. AYZ Google Trend verileri

AYZ terimi literatürde ilk olarak simülasyon oyunları uygulamalarında yapay zeka tabanlı entitilerin (nesnelerin) davranışlarını açıklamak için Van Lent vd.'nin çalışmasında tanımlanmıştır [17]. Genel olarak AYZ Şekil 2’de gösterilen bir akışa sahiptir. İlk olarak, veri toplama ve ön işleme adımları gerçekleştirilmektedir. Bu aşamada veriler temizlenir, eksik değerler düzenlenir ve modellemeye uygun hale getirilir. İkinci aşamada, eğitim, test ve doğrulama süreçleri ile makine öğrenmesi veya derin öğrenme temelli kara kutu modelleri oluşturulur. Ancak bu modellerin iç işleyişi kullanıcı tarafından doğrudan görülemez veya kolayca yorumlanamaz. Bu modeller yüksek doğruluk oranları sağlayabilse de, karar süreçlerinin neden ve nasıl oluştuğu çoğunlukla anlaşılamamaktadır. Üçüncü aşamada, modelin performansı çeşitli istatistiksel ölçütler (doğruluk, hassasiyet, duyarlılık vb.) ile değerlendirilmektedir. Ancak kara kutu yapının temel sınırlılığı, yalnızca bu performans ölçümleri ile yetinilmesi ve modelin karar mantığının kullanıcıya açıklanamamasıdır. Oluşturulan YZ modelinin başarısının değerlendirilmesinin ardından AYZ için uygun model (LIME, SHAP, ICE vb.) yöntemleri seçilir ve uygulanır [18] [19]. Son aşamada devreye giren AYZ teknikleri, oluşturulan modellerin karar mekanizmasını şeffaflaştırmayı amaçlamaktadır. Bu sayede, kullanıcılar modelin hangi girdilerden nasıl etkilendiğini anlayabilmekte ve sonuçları daha güvenilir biçimde yorumlayabilmektedir [20].



Şekil 2. Kara kutu modeller ve AYZ

Son yıllarda hızla büyüyen bir literatüre sahip olan AYZ'ye ait taksonomi Şekil 3'te verilmiştir [21]. Bu çalışmada modellerin açıklanabilirliği için kullanılan SHAP yöntemi tercih edilmiştir. SHAP yönteminin tercih edilmesinin başlıca nedenlerinden biri sağlık alanındaki modellerde yalnızca lokal değil, küresel açıklamaların da yapılabilmesinin büyük önem taşıyor olmasıdır [22]. SHAP, hem tüm hasta grubu için genel eğilimleri hem de tek bir hasta için bireysel tahminleri yorumlama imkânı sunmaktadır. Ayrıca SHAP, özellikler arasındaki doğrusal olmayan ilişkileri daha iyi yakalayabilmekte [23] ve bu durum karmaşık yapıli klinik verilerin açıklanmasında ek avantaj olarak karşımıza çıkmaktadır. Bu nedenlerle, çalışmada açıklanabilirlik yöntemi olarak SHAP kullanılması uygun bulunmuştur. Kullanılacak olan SHAP AYZ tekniğinin detayları Bölüm 2.1'de verilmiştir.



Şekil 3. AYZ taksonomisi [21]

2.1. SHAP

Makine öğrenimi modellerinin karar verme süreçlerini şeffaf hale getirmek için geliştirilen SHAP yöntemi, oyun teorisine dayalı olarak her bir özelliğın tahmine katkısı adil biçimde hesaplamayı amaçlar [24]. Bu yaklaşım, Shapley [25] tarafından ortaya konulan Shapley değerlerini temel alır ve her bir özelliğın (özellikle i) model çıktısına yaptığı marjinal katkıyı, tüm olası özellik kombinasyonları üzerinden ortalama olarak tanımlar. SHAP değeriğın matematiksel formülasyonu Eşitlik 1'deki gibidir:

$$\phi_i(f) = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} [f(S \cup \{i\}) - f(S)] \quad (1)$$

Bu denklemde F özellikler kümesini, S ise bu kümenin i 'yi içermeyen bir alt kümesini temsil eder. Ayrıca $\phi_i(f)$, özellik i 'nin katkısı; $f(S)$ ise S altkümesinin modele sağladığı katkıyı ifade etmektedir. Eşitlik, i özelliğinin modele ne kadar katkı sağladığını hesaplamak için, i 'nin dahil edilip edilmemesi durumundaki model çıktısı farklarının ortalamasını alır [26]. Bu yaklaşım, modelin açıklanabilirliğini artırırken karar destek sistemlerinde güvenilirlik açısından önemli avantajlar sunmaktadır.

3. UYGULAMA ÇALIŞMASI

Bu uygulama çalışmasında, meme kanseri hastalarının teşhis aldıktan sonra ölü/ hayatta kalma durumlarının veri madenciliği teknikleri ile tahmini ve ardından modelde ele alınan hasta özelliklerinin model tahmini üzerinden etkisi ve modelin sonuçlarının AYZ teknikleri ile açıklanabilirliği üzerine odaklanılmıştır.

Kullanılan veri seti açık bir kaynaktır ve meme kanseri hasta veri seti, Ulusal Kanser Enstitüsü'nün (National Cancer Institute) nüfus tabanlı kanser istatistikleri sağlayan SEER Programı'ndan elde edilmiştir. Ulusal Kanser Enstitüsü'nün veri tabanında bilimsel çalışmalarda kullanılmak üzere paylaşılmış olan veri seti, 2006-2010 yılları arasında infiltratif duktal ve lobüler karsinom meme kanseri tanısı almış hastaların kayıtlarını içermektedir [27]. Veri seti, meme kanseri teşhisi sonrası 3408 sağ kalım, 616 ölüm olmak üzere 4024 hastaya ait 14 farklı parametrenin bilgilerini içermektedir.

Veri setine ilişkin özellikler aşağıda verilmiştir:

A. Girdi değişkenleri:

1. Demografik Özellikler:

- Yaş, Irk ve Medeni Durum

2. Tümör Özellikleri:

- Tümör Boyutu: Tümörün boyutunu temsil etmektedir.
- T Evresi: Tümörün T evresini (tümör büyüklüğü ve yayılım ölçüsü) temsil etmektedir.
- N Evresi: Tümörün N evresini (lenf düğümü tutulumu ölçüsü) temsil etmektedir.
- 6. Evre: Kanserın genel evresini (T, N ve M evrelerinin kombinasyonu) temsil etmektedir.
- Derece (Grade): Tümör derecesini (kansere hücrelerinin ne kadar anormal görüldüğünü) temsil etmektedir.
- Farklılaşma: Tümör hücrelerinin ne kadar farklılaştığını temsil etmektedir.

3. Hormonal Durum:

- Östrojen Durumu: Tümörün östrojen reseptör durumunu temsil eder, sayısal olarak kodlanmıştır.
- Progesteron Durumu: Tümörün progesteron reseptör durumunu temsil eder, sayısal olarak kodlanmıştır.

4. Bölgesel Lenf Düğümleri:

- Bölgesel Lenf Düğümü İncelendi: İncelenen bölgesel lenf düğümlerinin sayısını temsil etmektedir.
- Bölgesel Lenf Pozitif: Kanser açısından pozitif olan bölgesel lenf düğümlerinin sayısını temsil etmektedir.

5. A Evresi:

- A Evresi: Kanserın A evresini (metastaz ölçüsü) temsil eder, sayısal olarak kodlanmıştır.

B. Çıktı Değişkeni:

Durum: Hastanın hayatta kalma durumunu temsil etmektedir (0: Ölü, 1: Hayatta).

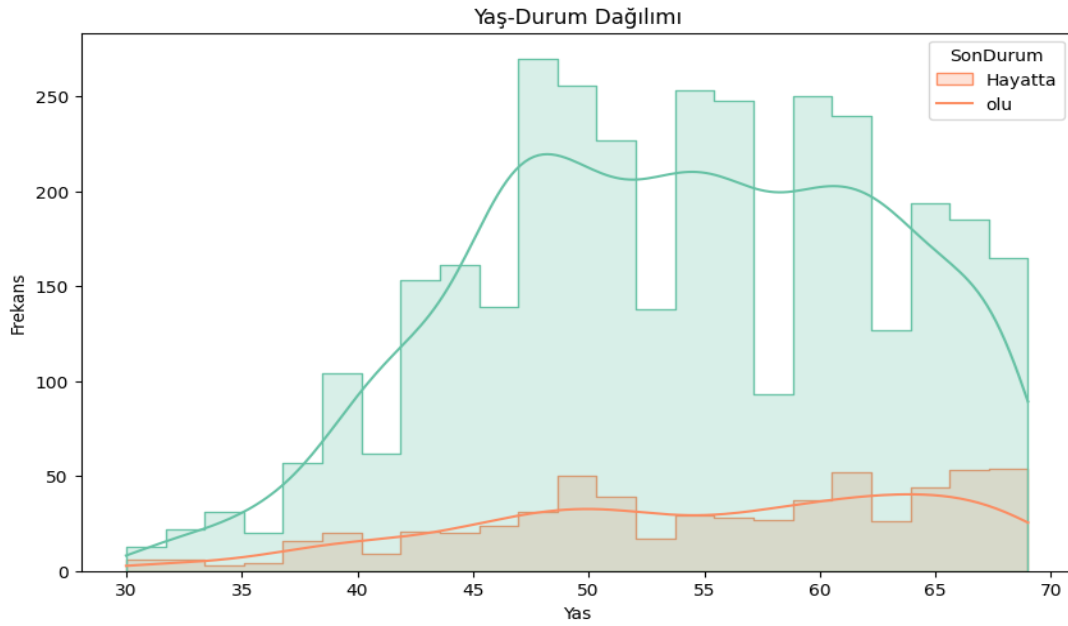
3.1. Veri Ön İşleme

Çalışmanın tüm kodlamaları Python programlama dili aracılığı ile yapılmıştır. Öncelikle veri seti veri ön işleme aşamasına tabii tutulmuştur. Ön işlemede girdi ve çıktı sütunlarında yer alan verilerin yapısı tespit edilmiştir. Verilerin tamsayı, kategorik vs. gibi sınıflandırmaları yapılmıştır. Boş hücre olup olmadığı da eş zamanlı olarak kontrol edilmiştir. Çizelge 1'de verilerin özellikleri görülmektedir.

Çizelge 1. Veri yapıları

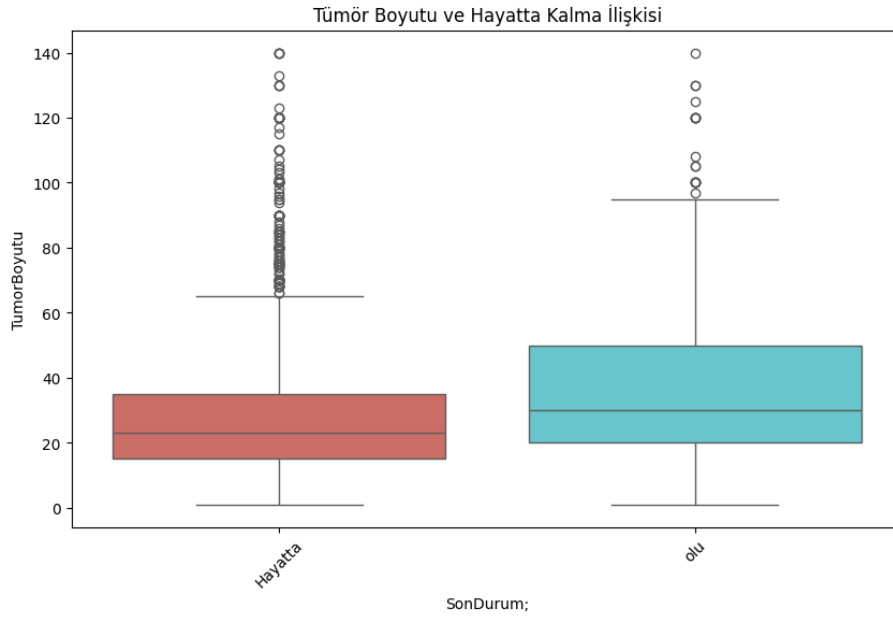
Değişkenler	Dolu hücre	Veri tipi
Yaş	4024	Nümerik
İrk	4024	Kategorik
Medeni durum	4024	Kategorik
T_evresi	4024	Kategorik
N_evresi	4024	Kategorik
6.Evre	4024	Kategorik
Farklılaşma	4024	Kategorik
Derece	4024	Kategorik
A_evre	4024	Kategorik
Tümör boyutu	4024	Nümerik
Östrojen durumu	4024	Kategorik
Progesteron durumu	4024	Kategorik
Bölgesel düğüm incelendi	4024	Nümerik
Bölgesel düğüm pozitif	4024	Nümerik
Son durum	4024	Kategorik

Bazı girdi değişkenlerinin çıktı değişkeni ile ilişkisi görselleştirilmiş olup bunlar Şekil 4 ve Şekil 5'te verilmiştir. Yaş ve meme kanseri sonucu hayatta kalma durumu arasındaki ilişki için Şekil 4'e ve tümör boyutu ve hayatta kalma arasındaki ilişki için ise Şekil 5'e bakılabilir. Şekil 4'te yaşa bağlı ölüm artışı eğilimi olduğu görülmektedir. Ayrıca Şekil 5'te tümör boyutunun artmasının hayatta kalma üzerindeki olumsuz etkisi de gözlemlenebilmektedir.



Şekil 4. Yaş ve son durum değişkenleri arasındaki ilişki grafiği

Bu grafikler ham veriler yardımı ile her girdi değişkeni için çizilebilir. Ancak bu grafiklerde hastanın son durumu ile seçilen girdi değişkeni arasında bir korelasyon görsel olarak algılansa da hastanın sağ kalma/ölme durumu tek bir parametreden kaynaklı olmayıp birçok parametrenin etkileşimi ile ortaya çıkabilecek karmaşık bir olaydır. Farklı parametre etkileşimleri ile YZ yardımı ile herhangi bir vakanın durumunun ne olacağının öngörülmesi ve AYZ yardımı ile kurulan modellerin hangi parametrelerinin ne seviyede olmasının sağ kalma/ölme durumu üzerinde etkili olduğunun irdelenmesi gerekmektedir.



Şekil 5. Tümör boyutu ve son durum değişkenleri arası ilişki grafiği

Veri ön işlemede önemli safhalardan birisi verilerin kodlanmasıdır. Veriler kategorik ya da nümerik gibi farklı yapılarda olabilir. Kategorik değişkenler, birçok makine öğrenimi modelinde olduğu gibi doğrudan analiz için kullanılamadığından, her model için uygun bir formata dönüştürülmelidir. Bir değişkenin bir metin (string) olarak temsil edilmesine ek olarak, eğer değerlerin büyüklüğü veya sıralaması sayısal olarak anlam ifade etmiyorsa, bu değişkenin kategorik bir tür olarak ele alınması gerekir. Kategorik değişkenleri kodlamak için One-Hot Encoding (Tek Sıcaklık Kodlaması), Dummy Encoding (Sahte Değişken Kodlaması), Ordinal Encoding (Sıralı Kodlama), Binary Encoding (İkili Kodlama), Count Encoding (Frekans Kodlama) ve Target Encoding (Hedef Kodlama) gibi çeşitli yöntemler kullanılmaktadır [28]. Verinin yapısına göre farklı kodlama teknikleri seçilebilir. Bu çalışma özelinde değişkenlerin yapısı dikkate alınarak sıralı, ikili ve tek sıcaklık kodlaması yöntemleri uygulanmıştır.

- **Sıralı Kodlama:**

A evresi: Bölgesel, metastaz ve uzak metastaz olmak üzere üç kategoriden oluşmaktadır.

T evresi: T1, T2, T3, T4 olmak üzere dört kategoriden oluşmaktadır.

N evresi: N1, N2, N3 olmak üzere üç kategoriden oluşmaktadır.

6. Evre: IIB, IIIA, IIIB, IIIC olmak üzere dört kategoriden oluşmaktadır.

Farklılaşma: Farklılaşma yok, Çok az farklılaşma, Orta dereceli farklılaşma, Çok fazla farklılaşma olmak üzere dört kategoriden oluşmaktadır.

Kodlama aşamada bu kategoriler sayısal değerler ile değiştirilmiştir. Grup sayısı kadar kod numarası verilmiştir. Örneğin 6. Evre: IIB, IIIA, IIIB, IIIC gruplarından oluşmaktadır. Sırası ile gruplar 1,2,3,4 kod numaraları ile değiştirilmiştir.

- **İkili kodlama ile:**

Östrojen ve Progesteron durumu Pozitif ve Negatif olmak üzere iki kategoriden ve Son durum değişkeni ise Hayatta ve Ölü olmak üzere iki kategorilerinden oluşmaktadır. Bu değişkenler 0-1 şeklinde kodlanmıştır.

- **Tek sıcaklık kodlama ile** Irk ve Medeni durum değişkenleri kodlanmıştır.

Irk ve medeni durum için aşağıda verilen kategoriler yeni birer öznitelik gibi ele alınmış ve sadece hangi satırda bu kategoriler bulunuyorsa o satıra "1" değeri verilmiş, diğerleri "0" şeklinde kodlanmıştır. Irk değişkeni: Irk Beyaz, Irk Diğer, Irk Siyah; Medeni Durum değişkeni: Medeni Durum Ayrı, Medeni Durum Bekar, Medeni Durum Dul, Medeni Durum Evli, Medeni Durum Boşanmış sütunlarına dönüştürülmüştür. Nümerik veriler için de ön işleme gerekmektedir. Amacı, özellikleri benzer bir ölçeğe getirmek olan normalizasyon, sayısal verileri diğer sayısal verilerle aynı ölçeğe dönüştürmek anlamına gelmektedir. Bu, modelin performansını ve eğitim kararlılığını artırmaktadır [29]. Nümerik değişkenleri normalize için L1-

L2 normalizasyonu, bir aralığa ölçeklendirme (Min-Max), kırpma (Clipping), Logaritmik ölçeklendirme (Log scaling) ve Standardizasyon/Z-skoru (Z-score) gibi yöntemler bulunmaktadır. Bu çalışmada sayısal veriler için yaş, tümör boyutu, bölgesel düğüm gibi özellikler min-max normalizasyon yöntemi ile normalize edilmiştir. Kullanılan min-max normalizasyon formülü Eşitlik 2'deki gibidir:

$$norm_x = \frac{x - \min}{\max - \min} \quad (2)$$

Veri ön işleme aşamasının tamamlanmasının ardından tüm veriler kategorik ve normalize edilmiş sürekli değerlere dönüştürülmüştür. Bu aşama sayesinde, daha sağlıklı YZ modellemesi yapılabilecektir. Çizelge 2'de, verilerin bir kısmına ait ön işleme öncesi ve sonrası değerler görülmektedir.

Çizelge 2. Veri ön işleme öncesi (a) ve sonrası (b)

a) Veri ön işleme öncesi										
Yaş	İrk	Medeni durum	T evresi	N evresi	6.Evre	Farklılaşma	Derece	A evre	Östrojen durumu	
68	Beyaz	Evli	T1	N1	IIA	Çok az farklılaşma	3	Bölgesel	Pozitif	
50	Beyaz	Evli	T2	N2	IIIA	Orta dereceli farklılaşma	2	Bölgesel	Pozitif	
58	Beyaz	Boşanmış	T3	N3	IIIC	Orta dereceli farklılaşma	2	Metastaz	Negatif	
58	Siyah	Evli	T1	N1	IIA	Çok az farklılaşma	3	Bölgesel	Pozitif	
47	Beyaz	Boşanmış	T2	N1	IIB	Çok az farklılaşma	3	Bölgesel	Pozitif	

b) Veri ön işleme sonrası										
Yaş	İrk Beyaz	Medeni durum Boşanmış	Medeni durum Evli	T evresi	N evresi	6.Evre	Farklılaşma	Derece	A Evre	Östrojen durumu
0,71	1	0	1	1	1	1	2	3	1	1
0,45	1	0	1	2	2	3	3	2	1	1
0,51	1	1	0	3	3	5	3	2	2	0
0,51	0	0	1	1	1	1	2	3	1	1
0,41	1	1	0	2	1	2	2	3	1	1

3.2. Makine Öğrenmesi

Çeşitli değişkenlere göre çıktı değerine ait bir sınıfın tahmin edildiği modeller sınıflandırma modelleri olarak adlandırılmaktadır. Bu çalışmada makine öğrenmesinde sınıflandırma amacı için geliştirilmiş olan Rassal Orman Sınıflayıcı (ROS), Gradient Boosting Sınıflayıcı (GBS), Lojistik Regresyon (LR), Karar Ağaçları (KA), Destek Vektörü Sınıflayıcı (DVS) ve K-En yakın Komşu (KNN) sınıflayıcı modelleri kullanılmış ve performans yönünden modeller birbiri ile kıyaslanmıştır. Kullanılan yöntemlerin genel çalışma mekanizmaları şu şekilde tanımlanabilir: Lojistik Regresyon, ikili sınıflandırma problemlerinde sık kullanılan, doğrusal ilişkilere dayalı olasılıksal bir modeldir [30]. Karar Ağaçları, veriyi dallara ayırarak sınıflandıran, yorumlanabilirliği yüksek, hiyerarşik yapıda bir yöntemdir [31]. Rassal Orman Sınıflayıcı, birden fazla karar ağacının topluluk (ensemble) yöntemiyle birleştirilmesiyle oluşur; aşırı öğrenmeyi azaltarak daha kararlı sonuçlar üretilmesine olanak tanır [32]. Gradient Boosting Sınıflayıcı, hataları ardışık olarak düzelten zayıf öğrenicilerin (weak learners) birleşiminden oluşan güçlü bir topluluk yöntemidir [33]. Destek Vektörü Sınıflayıcı, sınıflar arasındaki ayrımı en iyi sağlayacak hiper düzlemi bulmaya çalışır; özellikle karmaşık ve yüksek boyutlu verilerde etkili olmaktadır [34]. K-En Yakın Komşu algoritması ise, yeni bir örneği, eğitim verisindeki en yakın komşularının sınıfına göre sınıflandıran, basit ve örnek tabanlı bir algoritmadır [34].

Çalışmada kullanılan algoritmalar farklı yapısal özelliklere sahiptir ve birbirleriyle kıyaslanarak en iyi performans gösteren modelin belirlenmesi amaçlanmıştır. Örneğin, Lojistik Regresyon doğrusal ilişkilerin anlaşılmasına imkân tanırken, Rassal Orman ve Gradient Boosting algoritmaları değişkenler arası karmaşık ilişkileri yakalayabilmektedir. Karar Ağaçları yorumlanabilirlik açısından güçlüdür, KNN ise örnekler arasındaki benzerliklere dayalı tahminler yapabilmektedir. Destek Vektör Makineleri ise sınıflar arasındaki

sınırları optimize etmeye odaklanmaktadır. Bu çeşitlilik sayesinde, farklı bakış açıları sunan modellerin karşılaştırılması ve en uygun modelin seçilmesi hedeflenmiştir.

Modeller için eğitim ve test verisi bölümlendirmesi, %80 eğitime ve %20 test olacak şekilde yapılmıştır. Makine öğrenmesi aşamasında kullanılan sınıflandırma modellerine ait parametreler Python scikit-learn kütüphanesinde tanımlanan varsayılan değerler olarak kullanılmıştır. Veri seti, sınıf sayılarının 3408 sağ vaka ve 616 ölü vaka olması nedeni ile dengesiz bir yapıdadır. Dengesiz veri setlerinde model başarısı düşebilmekte, makine öğrenmesi algoritmalarında performansı artırmak için SMOTE (Synthetic Minority Over-sampling Technique) [35], Sınıf Ağırlığı (SA)[36] gibi farklı tekniklerin algoritmalara entegre edilmesi gerekmez. Bu çalışmada, dengesiz veri setinin tahmin performansına olan etkisini azaltmak için Sınıf Ağırlığı ve SMOTE yöntemlerine yer verilmiştir. Sınıf Ağırlığı yöntemi, modelin eğitim sırasında farklı sınıfların örneklerine farklı ağırlıklar vermesini sağlar. Azınlık sınıfının örneklerine daha yüksek ağırlık verilirken, çoğunluk sınıfının örneklerine daha düşük ağırlık verilir. Modelin kayıp fonksiyonu (loss function) hesaplanırken, her örneğin katkısı o örneğin ait olduğu sınıfın ağırlığı ile çarpılır. Daha yüksek ağırlık verilen azınlık sınıfı örneklerindeki yanlış tahminler, modelin toplam kaybına daha fazla katkıda bulunur. Bu da modelin eğitim sırasında azınlık sınıfı örneklerini doğru tahmin etmeye daha fazla odaklanmasını sağlar. SMOTE yöntemi ise dengesiz veri kümelerinde azınlık sınıfının örneklerini sentetik olarak artırmak için kullanılan popüler bir üst örnekleme (oversampling) tekniğidir [35]. Gradient Boosting ve KNN algoritmaları yapıları gereği sınıf ağırlığı modellerinin kullanımına uygun değildir. İki yaklaşım dikkate alınarak makine öğrenmesi modelleri uygulanmış, elde edilen sonuçlar Çizelge 3'te verilmiştir.

Çizelge 3. Makine öğrenmesi modelleri sonuç değerleri

Makine öğrenmesi tekniği	Yaklaşım	Doğruluk	Kesinlik (Ölü)	Duyarlılık (Ölü)	F1 skoru (Ölü)	Kesinlik (Hayatta)	Duyarlılık (Hayatta)	F1 skoru (Hayatta)
ROS	SA	0,85	0,54	0,13	0,21	0,86	0,98	0,92
LR	SA	0,70	0,27	0,62	0,38	0,91	0,72	0,80
DVS	SA	0,71	0,27	0,61	0,37	0,91	0,72	0,81
KA	SA	0,78	0,28	0,30	0,29	0,88	0,87	0,87
GBS	SMOTE	0,79	0,33	0,39	0,35	0,89	0,86	0,87
ROS	SMOTE	0,79	0,30	0,32	0,31	0,88	0,87	0,87
KNN	SMOTE	0,69	0,23	0,48	0,31	0,89	0,72	0,80
DVS	SMOTE	0,73	0,27	0,50	0,35	0,90	0,77	0,83
KA	SMOTE	0,75	0,25	0,34	0,28	0,88	0,82	0,85

Modellerin performanslarını değerlendirmek için doğruluk (accuracy), kesinlik (precision), duyarlılık (recall) ve F1 skoru değerleri hesaplanmıştır. Bu metriklerin hesaplanmasında kullanılan temel kavramlar: Gerçekte pozitif olan (örneğin, “ölü” sınıfı) ve doğru tahmin edilen örnekleri ifade eden “TP (True Positive)”, Gerçekte negatif olan (örneğin, “hayatta” sınıfı) ve doğru tahmin edilen örnekleri ifade eden TN (True Negative), Gerçekte negatif olduğu halde pozitif tahmin edilen örnekleri ifade eden FP (False Positive) ve Gerçekte pozitif olduğu halde negatif tahmin edilen örnekleri ifade eden FN (False Negative) parametreleridir. Ardından bu parametrelere bağlı olarak modelleri birbiri ile kıyaslayabilmek adına aşağıdaki formüller yardımı ile doğruluk, kesinlik, duyarlılık ve F1 skoru değerleri hesaplanmıştır.

Tüm doğru sınıflandırmaların toplam örnek sayısına oranını ifade eden Doğruluk değeri Eşitlik 3'teki gibi hesaplanmaktadır.

$$\text{Doğruluk} = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (3)$$

Pozitif olarak tahmin edilenlerin içinde gerçekten pozitif olanların oranını ifade eden kesinlik değeri Eşitlik 4'teki gibi hesaplanmaktadır:

$$\text{Kesinlik} = \frac{(TP)}{(TP+FP)} \quad (4)$$

Gerçekte pozitif olanların içinde doğru tahmin edilenlerin oranını ifade eden Duyarlılık değeri Eşitlik 5 kullanılarak hesaplanmaktadır:

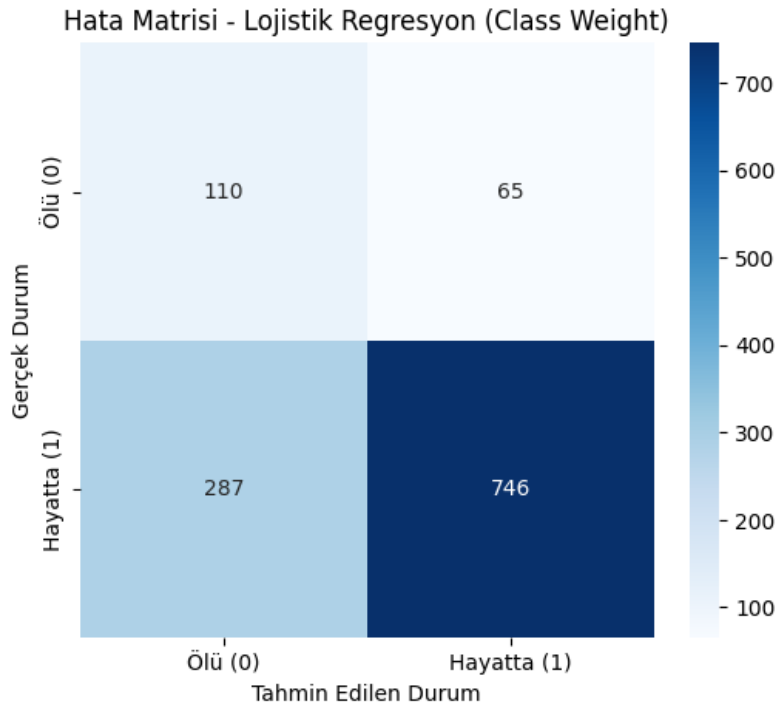
$$\text{Duyarlılık} = \frac{(TP)}{(TP+FN)} \quad (5)$$

Kesinlik ve duyarlılığın harmonik ortalamasını ifade eden F1 skoru değeri ise Eşitlik 6 ile hesaplanmaktadır:

$$F1 = 2X \frac{\text{Kesinlik} \times \text{Duyarlılık}}{(\text{Kesinlik} + \text{Duyarlılık})} \quad (6)$$

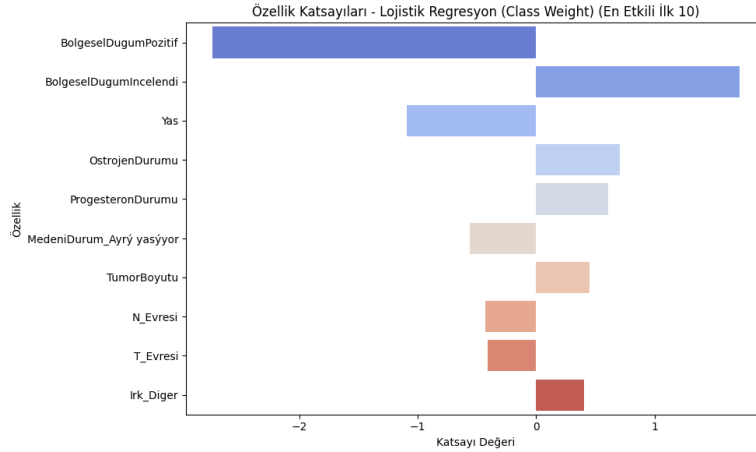
Bu ölçütler sayesinde modeller yalnızca doğruluk açısından değil, aynı zamanda azınlık sınıfı olan “ölü” grubunu doğru tahmin edebilme başarısı açısından da daha kapsamlı bir şekilde değerlendirilmiştir. Ayrıca özellikle dengesiz sınıf dağılımına sahip veri kümelerinde Doğruluk değerine göre model performansına karar vermek yanıltıcı olabilmektedir. Yüksek doğruluk, modelin sadece çoğunluk sınıfını doğru tahmin etmesinden kaynaklanabilir. F1 Skoru ise hem Duyarlılık hem de Kesinliği dikkate aldığı için dengesiz veri kümelerinde daha bilgilendiricidir. Yüksek bir F1 Skoru, modelin hem yanlış pozitifleri hem de yanlış negatifleri düşük tutarak iyi bir denge kurduğunu göstermektedir.

Çizelge 3’te görüldüğü üzere genel doğru tahmin açısından ve hayatta kalanların doğru sınıflandırılması açısından Rassal Orman Sınıflayıcı model daha başarılıdır. Dengesiz veri seti üzerinde performans artırıcı tüm çalışmalara rağmen sayıca az olan sınıfın tahminleme başarısı % 62’yi geçememektedir. Ancak bu değer SMOTE ve Sınıf Ağırlığı yöntemleri entegre edilmeden önce % 15 civarında olduğu tespit edilmiştir. Özellikle ele alınan veri setinde “Ölü” sınıf sayısının azlığı dikkate alındığında ve ayrıca ölü bir hastayı yanlış sınıfa atamanın maliyeti daha yüksek olacağı için ölü sınıfına ait duyarlılık değeri baz alınarak bir tercih yapılmalıdır. Şekil 6’da F1 skoru açısından en başarılı olan Lojistik Regresyon sınıflayıcı modeline ait hata matrisleri verilmiştir. LR hata matrisinde de görüldüğü üzere 746 sağ hasta sağ olarak tahmin edilirken, 110 hasta ölü hasta ölü sınıfında tahmin edilmiştir.



Şekil 6. Lojistik Regresyona ait hata matrisi

Makine öğrenmesi aşamasında LR modeli için özellik önemine ait bilgiler elde edebiliriz. Şekil 7’de ve Çizelge 4’te bu iki yönetime ait özellik önem grafiği ve değerleri görülmektedir.



Şekil 7. Özellik önem sıralaması grafiği

Şekil 7’de ve Çizelge 4’te LR sınıflandırma modelinin tahmininde en etkili parametrenin Pozitif Bölgesel Düğüm olduğu görülmektedir. Ardından sırası ile incelenen bölgesel düğüm sayısı, yaş, östrojen durumu, progesteron durumu gibi parametreler gelmektedir. Şekil 7’de LR modelinde model performansı üzerinde en etkili parametrelerin önem sırası ve ne derece önemli olduğu çubuk grafiklerle görselleştirilmiştir. Katsayı değeri negatif bölgede olan parametrelerin sağ kalım üzerinde olumsuz yönlü etkisi olduğu anlaşılmaktadır.

Çizelge 4. Özellik önem dereceleri

Lojistik regresyon		
	Özellik	Önem
1	Bölgesel düğüm pozitif	2,73
2	Bölgesel düğüm incelendi	1,71
3	Yaş	1,08
4	Östrojen durumu	0,70
5	Progesteron durumu	0,61
6	Medeni durum_ayrı	0,55
7	Tümör boyutu	0,45
8	N_evresi	0,43
9	T_evresi	0,40
10	Irk_diger	0,40

Makine öğrenmesi modellerinin performans ölçümleri, modelin ne kadar iyi tahmin ettiği ve nerede zorlandığı hakkında bilgi verirken; bu sonuçları modelin neden belirli bir tahmin yaptığını veya hangi özelliklerin belirli bir sonuç üzerinde en önemli etkiye sahip olduğunu açıklayamamaktadır. Burada devreye AYZ teknikleri girmektedir. SHAP gibi yöntemler, bireysel tahminler için "yerel" açıklamalar sunarak, hangi özelliklerin (örneğin yaş, tümör boyutu, evre vb.) belirli bir hastayı "Ölü" veya "Hayatta" olarak tahmin etmede en etkili olduğunu ve bu etkinin yönünü anlamamızı sağlamaktadır. Bu durum, yalnızca model performansını değerlendirmek için değil, aynı zamanda modele güven oluşturmak, çıktıları doğrulamak, olası önyargıları engellemek ve hatta alan uzmanlarının (doktorlar gibi) modelin tahminlerini yorumlamalarına ve kullanmalarına yardımcı olmak için de önemlidir.

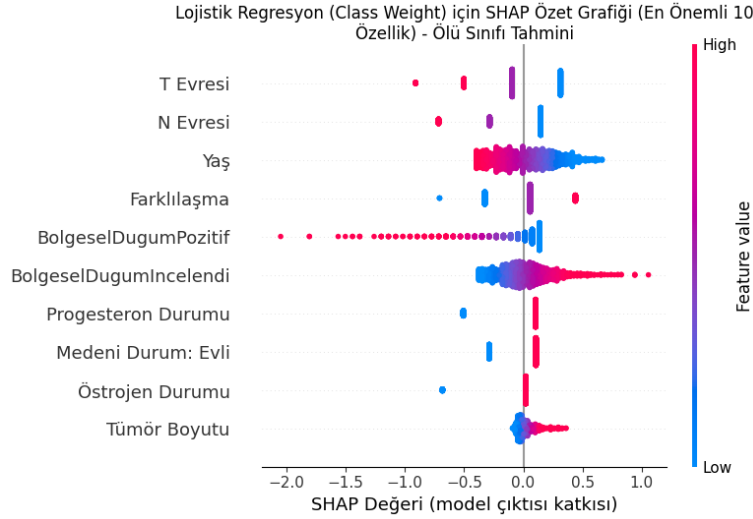
3.3. SHAP Tekniği ile Açıklanabilir Yapay Zeka Analizi

Bu aşamada, sadece “Ölü” sınıfını tahmin etmede başarı gösteren LR sonuçları AYZ tekniklerinden olan SHAP tekniği yardımı ile çıkarımlarda bulunulmuştur. LR katsayıları genel eğilimleri gösterirken, ROS gibi daha karmaşık modeller birer kara kutu gibi davranabilmektedir. Ancak SHAP, her bir tekil gözlem için, her bir özelliğin modelin nihai tahminine ne kadar ve hangi yönde katkıda bulunduğunu matematiksel olarak sağlam bir temelde açıklayarak yerel yorumlanabilirlik sağlamakta ve model performansının ötesinde, güven oluşturmaya ve alınan kararları anlamlandırmaya yardımcı olmaktadır.

Bu bağlamda, hangi özelliklerin çıktığı ne yönde etkilediğini anlamak için SHAP özet tablosundan faydalanmak mümkündür. Şekil 8’de “Ölü” sınıfının tahmininde kullanılan LR modeli çıktıları üzerinden

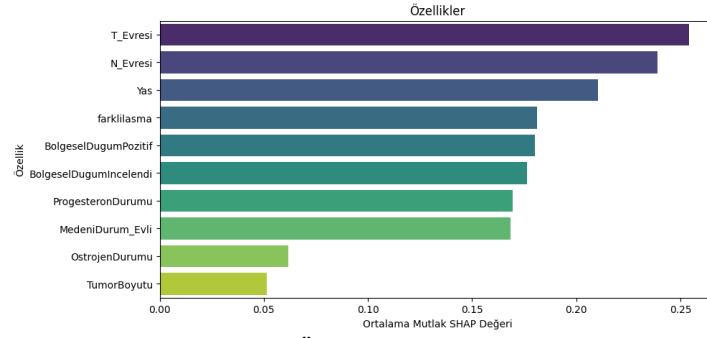
en önemli 10 özelliğe ait SHAP özet tablosu verilmektedir. Şekil 8’de eğer özellik nümerik bir girdi ise; kırmızı veriler ilgili özelliğin yüksek sayısal değerlerini ifade ederken, mavi veriler düşük sayısal değerleri ifade etmektedir. SHAP değerinde 0 değerinin sağ tarafı “Hayatta” sınıfının tahminini 1 olarak bulmamıza neden olan özellikleri ifade ederken sol tarafı “Hayatta” sınıfının değerini 0 olarak tahmin etmemize neden olan değerleri göstermektedir (“Ölü” olarak tahmin edilen sonuçlar). Ayrıca kategorik sınıflandırma yapılan özellikler için mavi renk, kod numarası düşük olan grubu ifade ederken, kırmızı renkli veriler kod değeri yüksek olan verileri ifade etmektedir. Bu sonuçlara göre Ölü sınıfı tahminlemede en başarılı olan LR modeline ait Şekil 8’de verilen SHAP özet tablosu yorumlandığında yapılan çıkarımlar aşağıdadır:

- T Evresi için T4, T3 ve T2 sırası ile (tümörün meme duvarına veya cilde yayılımının artmasını ifade eder) “Ölü” sınıfı tahminini artırmaktadır.
- N Evresi için N3, N2 sırası ile (kanserin daha uzak veya daha fazla sayıda lenf düğümüne yayıldığı durumu ifade eder) “Ölü” sınıfı tahminini artırmaktadır.
- Yaş değişkeni için ileri yaşlarda olmak “Ölü” sınıfı tahminini artırmaktadır.
- Farklılaşma değişkeni için “Farklılaşma yok” ve “Çok az farklılaşma” olması (kansere hücrelerinin normal meme hücrelerine çok az benzediği anlamına gelir ki bu agresif bir kanser türüne işaret eder [37]) “Ölü” sınıfı tahminini artırmaktadır.



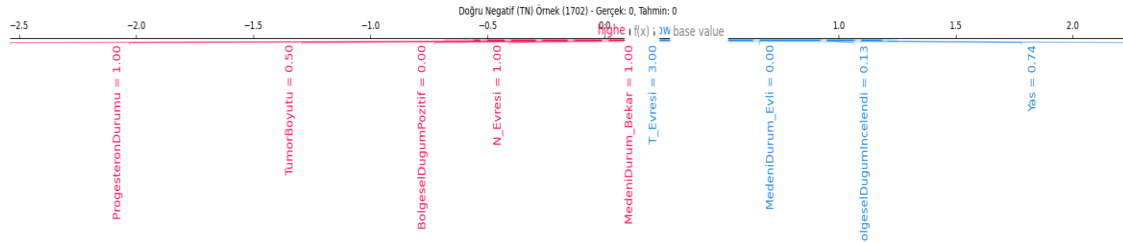
Şekil 8. Ölü sınıfının tahmini için kullanılan LR modeline ait en önemli 10 özellik için SHAP özet tablosu

- Bölgesel düğüm pozitif değişkeni için kanser hücrelerinin lenfatik sisteme yüksek yayılımını ifade eden değerler “Ölü” sınıfı tahminini artırmaktadır.
- Bölgesel düğüm incelendi değişkeni için daha az sayıda lenf düğümü incelemesi “Ölü” sınıfı tahminini artırmaktadır.
- Negatif progesteron ve östrojen değeri meme kanserinin hormonlara duyarsız olduğu ve bu nedenle hormonal tedavinin işe yaramadığı grup olarak “Ölü” sınıfı tahminini artırmaktadır.
- Tümör boyutu küçük değerler için “Ölü” grubuna ait tahmin olasılığı artıyor görülmektedir. Ancak bu durum tek başına değerlendirme için yeterli gözükmemektedir. Negatif progesteron ve östrojen değerli küçük bir tümör, daha büyük boyutlu ancak pozitif progesteron ve östrojen değerli bir tümörden daha agresif olabilmektedir.
- Evli olmayan grup için “Ölü” sınıfı tahmini artırmaktadır. Literatürde birçok çalışmada evli olmanın meme kanseri dahil olmak üzere birçok kanserde bağımsız bir sağkalım artırıcı faktör olduğu gösterilmiştir [38]. Ayrıca Şekil 9’da LR modeli için “Ölü” sınıfını tahmin etmede en etkili 10 özellik ve etki oranlarına ilişkin grafik görülmektedir. Buradan modelin “Ölü” değeri tahminlemede en çok etkilendiği özellikler önem sırasına göre görülmektedir. Elde edilen bulguların bilimsel gerçeklerle de uyumlu olduğu görülmektedir.



Şekil 9. SHAP analizine göre LR modelinin “Ölü” sınıfına ait tahmin yapmasında etkili olan 10 girdi ve ortalama mutlak SHAP değerleri

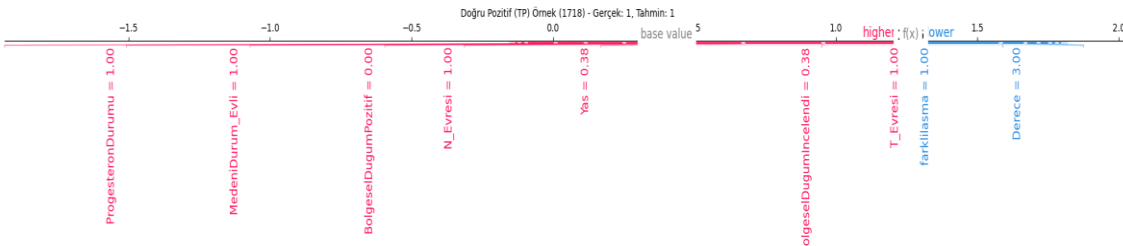
AYZ ayrıca, tekil bir örnek için modelin neden belirli bir tahmin ürettiğini derinlemesine anlamak için çok kullanışlıdır. Bu durumda karar grafiğinden faydalanabilir. Şekil 10 (a, b, c ve d)’da belirli hastaların çıktısının tahmin edilmesine etki eden parametreler ve ne yönde etki ettiği görülmektedir. Karar grafiği üzerinde tahmin hattının sağa doğru ilerlemesine neden olan özellik değerlerinin, bu özelliği “Ölü” sınıfı olarak tahmin etme yönünde katkı sağladığını göstermektedir.



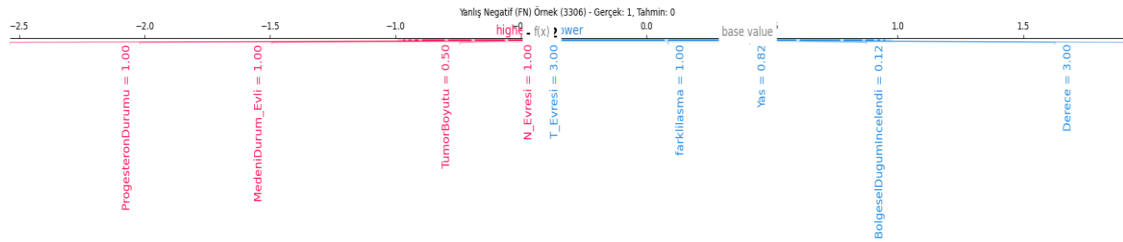
(a). Doğru negatif tahmin; örnek no: 1702



(b). Yanlış pozitif tahmin; örnek no: 1937



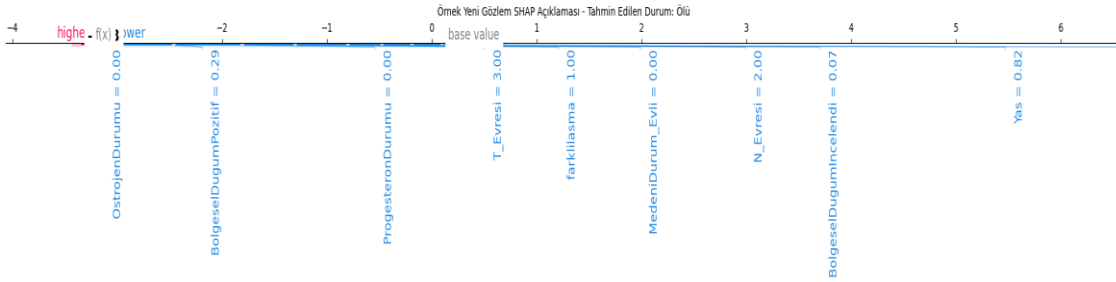
(c). Doğru pozitif tahmin; örnek no: 1718



(d). Yanlış negatif tahmin; örnek no: 3306

Şekil 10. Tahmin sonuçları üzerine girdi değişkenlerinin etkisini gösteren SHAP kuvvet grafiği

Ayrıca yeni gelen bir vakanın belirlenen en iyi makine öğrenmesi algoritmasına göre sonucunun ne olacağı ve ardından bu sonucun tahmin edilmesinde etkili olan girdi değişkenlerinin neler olduğu da SHAP yardımı ile öngörülebilmektedir. Bu bağlamda belirli girdi değerlerine sahip yeni bir hasta için modelin tahmini ve bunu etkileyen girdi değerleri incelenebilmektedir. Örnek hastaya ait özellikler şu şekildedir: Yaş: 62, Irk: Beyaz, Medeni Durum: Dul, Tümör Boyutu: 38, Derece: 3, T Evresi: T3, N Evresi: N2, 6.Evre: IIIC, A Evre: Bölgesel, Östrojen Durumu: Negatif, Progesteron Durumu: Negatif, Farklılaşma: Çok az farklılaşma, Bölgesel Düğüm İncelendi:5, Bölgesel Düğüm Pozitif:14. Belirtilen yeni özelliklere sahip vaka için LR-SA yöntemi sonucu “Ölü” sınıfı olarak tahmin edilmiştir. Şekil 11’de SHAP kuvvet grafiği analizi ile sonuca etki eden parametreler irdelenmiştir. Vakanın “Ölü” sınıfında tahmin edilmesinin en önemli etkenleri olarak Bölgesel Düğüm Pozitif değişkeninin değeri, Progesteron ve Östrojen varlığının negatif olması tespit edilmiştir.



Şekil 11. Belirli özelliklere sahip bir vakanın tahminine etki eden değişkenlere ait SHAP kuvvet grafiği

4. TARTIŞMA VE SONUÇ

Bu çalışmada elde edilen bulgular, literatürdeki benzer araştırmalarla karşılaştırıldığında genel olarak literatüre uyumluluk göstermektedir. Öncelikle, literatürde belirtildiği gibi [39] dengesiz veri setleri ile yapılan çalışmada SMOTE tekniğinin kullanımının, meme kanseri hastalarının sağkalım tahmininde kullanılan makine öğrenmesi algoritmalarının başarısını artırdığı yönünde bulgular elde edilmiştir. Ayrıca SHAP'ın meme kanseri hastalarında sağkalım tespitinde daha önceki çalışmalarda da belirtilen etkin kullanımını [40] sayesinde makine öğrenmesi modellerinden klinik olarak anlamlı sonuçlar çıkarılmıştır. SHAP yöntemi ile elde edilen çıkarımlar, Şekil 8 ve 9'da da verildiği üzere literatürde meme kanseri sağkalımına etki eden risk faktörleriyle benzerlik göstermektedir. Yaş, T evresi, N evresi gibi değişkenlerin model çıktıları üzerinde belirleyici etkilerinin bulunduğu bulgusu, mevcut literatürde rapor edilen sonuçlarla örtüşmektedir [41]. İlave olarak medeni durumun da bulguları destekler şekilde sağkalım üzerinde etkisi olduğu tespit edilmiştir [38]. Bu durum, AYZ'nin meme kanseri sağkalım tahminlemede ve nedenlerini açıklamada başarılı sonuçlar verdiğinin göstergesidir. Ayrıca genelleme yapılacak olursa, sağlık alanında AYZ kullanımı sayesinde hastalar, hastalıklar ve tedaviler ile ilgili özel ve genel çıkarımlar yapılabileceği anlaşılmaktadır.

AYZ, makine öğrenmesi ile oluşturulan modellerin çalışma mekanizmasını anlamamız, böylece model tahminlerinin neden oluştuğunu anlamlandırmamız açısından oldukça önemlidir. Bu modeller sayesinde daha şeffaf ve anlamlandırılabilir çıkarımlar yapmak mümkündür. Bu durum özellikle sağlık gibi yanlış karar vermenin maliyetinin yüksek olduğu alanlarda karar vericilere önemli bir katkı sağlamaktadır. Literatür örneklerinden de anlaşıldığı üzere çok farklı alanlarda kara kutu modelleri açıklama ve anlamlandırma yeteneği olan AYZ, sağlık alanında tanı, teşhis, tedavi, girişimsel ve girişimsel olmayan tedavilerin sonuçlarının tahmini için yaygın olarak kullanılmaya başlanmıştır.

Bu çalışmada öncelikle sağlık alanında AYZ'nin hangi amaçlarla kullanıldığına ilişkin literatür sunulmuş, ardından dengesiz sınıf kümelerinden oluşan meme kanseri verileri üzerinde çeşitli makine öğrenmesi algoritmaları uygulanmıştır. Veri setinin dengesiz sınıf dağılımının negatif etkilerini gidermek için SMOTE ve sınıf ağırlığı yöntemlerinden faydalanılmış bu yöntemler sayesinde model performansları önemli ölçüde iyileştirilmiştir. Modellerin başarısı özellikle F1 skoru baz alınarak değerlendirilmiş, genel doğruluk açısından en yüksek performans %85 ile Rassel Orman modelinde, ölü sınıfına ait en yüksek duyarlılık ise %62 ile Lojistik Regresyon modelinde elde edilmiştir. F1 skoru dikkate alındığında en başarılı modelin Lojistik Regresyon olduğu görülmüştür. Ardından bu modelin test sonuçları dikkate alınarak AYZ tekniklerinden SHAP ile sınıflandırma sonuçlarının nedenleri araştırılmıştır. Elde edilen çıkarımlar, meme

kanseri hastalarının hangi özellikler dolayısıyla riskli sınıfta yer alabileceği hakkında bilgi vermiş ve sonuçların literatürle uyumlu olduğu gözlemlenmiştir.

Bu çalışma, yalnızca doğru tahminler yapabilen bir model geliştirmekle sınırlı kalmayıp, aynı zamanda model kararlarının nedenlerini şeffaf bir biçimde açıklaması ve klinik karar süreçlerine açıklanabilirlik getirmesi açısından literatüre özgün bir katkı sunmaktadır. Literatürdeki birçok çalışmanın görüntü tabanlı uygulamalara odaklandığı dikkate alındığında, bu yaklaşımın sağlık verilerinin farklı boyutlarını değerlendirmede özgün bir katkı sunduğu düşünülmektedir.

5. GELECEK ÇALIŞMALAR

Bu çalışma, meme kanseri sağkalım tahmininde AYZ kullanımının önemini ortaya koysa da, bulguları daha da güçlendirmek ve eksikliklerini gidermek amacıyla çeşitli yönlerde gelecek araştırmalara açıktır. İlk olarak, makine öğrenmesi aşamasında parametre optimizasyonu ile performans artırımı sağlanabilir. Çalışmada uygulanan SHAP yönteminin yanı sıra, diğer AYZ tekniklerinin karşılaştırmalı bir analizi yapılabilir. Bu sayede, farklı açıklanabilirlik yaklaşımlarının uzmanlar tarafından anlaşılabilirlik, güvenilirlik ve kullanılabilirlik açısından nasıl algılandığı değerlendirilebilir. Veri setinin niteliği ve kapsamı üzerine de önemli iyileştirmeler yapılabilir. Mevcut çalışmadaki dengesiz veri yapısının model performansı üzerindeki sınırlayıcı etkisi göz önüne alındığında, daha büyük ve dengeli veri setleriyle çalışılması modelin genelleme yeteneğini artıracaktır. Ayrıca, genetik faktörler, tedavi yöntemleri veya yaşam tarzı faktörlerin de modele dahil edilmesi, tahminlerin klinik geçerliliğine önemli ölçüde katkı sağlayabilir. Bir diğer önemli çalışma alanı ise, bu modellerin gerçek dünyada klinik ortamlara entegrasyonudur. Geliştirilen tahminleme modellerinin ve AYZ yönteminin entegre bir klinik karar destek sistemi haline getirilmesi ve hekimlerle yapılacak kullanılabilirlik çalışmalarından geçirilmesi bu gibi çalışmaların değerini artıracaktır. Son olarak, bu çalışmada sunulan metodolojik yaklaşımın akciğer, prostat veya kolon kanseri gibi diğer kanser türlerine uyarlanarak test edilmesi, yaklaşımın genel geçerliliğini gösterecektir.

6. KAYNAKLAR

1. Kaplan, M., Çakar, F. & Bingöl, H. (2024). Sağlık alanında yapay zekanın kullanımı: Derleme. *Muş Alparslan Üniversitesi Sağlık Bilimleri Dergisi*, 4(3), 75-85.
2. Adadi, A. & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access*, 6, 52138–52160.
3. Terzi, R. (2021). Sağlık sektöründe açıklanabilir yapay zekâ. Yapay zekâ ve büyük veri çalışmaları, *Siber Güvenlik ve Mahremiyet*, 157-175
4. Kalasampath, K., Spoorthi, K.N., Sajeev, S., Kuppa, S.S., Ajay, K. & Maruthamuthu, A. (2025). A literature review on applications of explainable artificial intelligence (XAI). *IEEE Access*, 13, 41111-41140.
5. Yildirim, H. (2024). Hastalıkların teşhis ve tedavisi için tıbbi görüntülemelerden yapay zeka kullanımı ile daha hassas ve daha hızlı sonuç elde etme konusundaki gelişmeler: FDA (Amerikan Gıda ve İlaç Dairesi) tarafından onaylanmış teknolojilerle geliştirilebilecek yeni uygulamalar. *Academic Social Resources Journal*, 6(24), 558-568.
6. Alkhanbouli, R., Abdulla Almadhaani, H.M., Alhosani, F. & Simsekler, M.C.E. (2025). The role of explainable artificial intelligence in disease prediction: A systematic literature review and future research directions. *BMC Medical Informatics and Decision Making*, 25(1), 110.
7. Albahri, A.S., Duham, A.M., Fadhel, M.A., Alnoor, A., Bager, N.S., Alzubaidi, L., Albahri, O.S., Alamoodi, A.H., Bai, J., Salhi, A., Santamaria, J., Ouyang, C., Gupta, A., Gu, Y. & Deveci, M. (2023). A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion. *Information Fusion*, 96, 156-191.
8. Akin, K.D., Gurkan, C., Budak, A. & Karataş, H. (2022). Açıklanabilir yapay zeka destekli evrimsel sinir ağları kullanılarak maymun çiçeği deri lezyonunun sınıflandırılması. *European Journal of Science and Technology*, 40, 106-110.
9. Nafisah, S.I. & Muhammad, G. (2024). Tuberculosis detection in chest radiograph using convolutional neural network architecture and explainable artificial intelligence. *Neural Computing and Applications*, 36(1), 111-131.
10. Orman, A., Köse, U. & Yiğit, T. (2021). Brain tumor detection via explainable convolutional neural networks: Açıklanabilir evrimsel sinir ağları ile beyin tümörü tespiti. *El-Cezeri Journal of Science and Engineering*, 8(3), 1323-1337.

11. Viswan, V., Shaffi, N., Mahmud, M., Subramanian, K. & Hajamohideen, F. (2024). Explainable artificial intelligence in Alzheimer's disease classification: A systematic review. *Cognitive Computation*, 16(1), 1-44.
12. Wani, N.A., Kumar, R. & Bedi, J. (2024). DeepXplainer: An interpretable deep learning based approach for lung cancer detection using explainable artificial intelligence. *Computer Methods and Programs in Biomedicine*, 243, 107879.
13. Zhang, Y., Weng, Y. & Lund, J. (2022). Applications of explainable artificial intelligence in diagnosis and surgery. *Diagnostics*, 12(2), 237.
14. Yang, C.C. (2022). Explainable artificial intelligence for predictive modeling in healthcare. *Journal of Healthcare Informatics Research*, 6(2), 228-239.
15. Tjoa, E. & Guan, C. (2021). A survey on explainable artificial intelligence (XAI): Toward medical XAI. *IEEE Transactions on Neural Networks and Learning Systems*, 32(11), 4793-4813.
16. Sun, Q., Akman, A. & Schuller, B.W. (2025). Explainable artificial intelligence for medical applications: A review. *ACM Transactions on Computing for Healthcare*, 6(2), 1-31.
17. van Lent, M., Fisher, W. & Mancuso, M. (2004). An explainable artificial intelligence system for small-unit tactical behavior. In *Proceedings of the 16th Conference on Innovative Applications of Artificial Intelligence (IAAI'04)* (900-907). AAAI Press.
18. Barredo Arrieta, A., Diaz-Rodriguez, N., Del Ser, J., Benetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R. & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82-115.
19. Belle, V. & Papantonis, I. (2021). Principles and practice of explainable machine learning. *Frontiers in Big Data*, 4, 688969.
20. Kalasampath, K., Spoorthi, K.N., Sajeev, S., Kuppa, S.S., Ajay, K. & Maruthamuthu, A. (2025). A literature review on applications of explainable artificial intelligence (XAI). *IEEE Access*, 13, 41111-41140.
21. Rawal, A., McCoy, J., Rawat, D.B., Sadler, B.M. & Amant, R.S. (2022). Recent advances in trustworthy explainable artificial intelligence: Status, challenges, and perspectives. *IEEE Transactions on Artificial Intelligence*, 3(6), 852-866.
22. Nguyen, H.T.T. & Cao, H.Q. (n.d.). Evaluation of explainable artificial intelligence: SHAP, LIME, and CAM. [Manuscript in preparation].
23. Holzinger, A., Saranti, A., Molnar, C., Biecek, P. & Samek, W. (2022). Explainable AI methods-A brief overview. In A. Holzinger et al. (Eds.), *xxAI - Beyond Explainable AI* (13-38). Springer.
24. Salih, A.M., Raisi, Z., Galazzo, I.B., Radeva, P., Petersen, S.E., Lekadir, K. & Menegaz, G. (2025). A perspective on explainable artificial intelligence methods: SHAP and LIME. *Advanced Intelligent Systems*, 7(1), 2400304.
25. Shapley, L. (2020). A value for n-person games. In H. W. Kuhn (Ed.), *Classics in game theory* (69-79). Princeton University Press.
26. Meng, Y., Yang, N., Qian, Z. & Zhang, G. (2021). What makes an online review more helpful: An interpretation framework using XGBoost and SHAP values. *Journal of Theoretical and Applied Electronic Commerce Research*, 16(3), 3.
27. Teng, J. (2019). SEER breast cancer data. *IEEE DataPort*.
28. Potdar, K., Pardawala, T. & Pai, C. (2017). A comparative study of categorical variable encoding techniques for neural network classifiers. *International Journal of Computer Applications*, 175(4), 7-9.
29. Ahsan, M.M., Mahmud, M.A.P., Saha, P.K., Gupta, K.D. & Siddique, Z. (2021). Effect of data scaling methods on machine learning algorithms and model performance. *Technologies*, 9(3), 3.
30. Hua, Y., Stead, T.S., George, A. & Ganti, L. (2025). Clinical risk prediction with logistic regression: Best practices, validation techniques, and applications in medical research. *Academic Medicine & Surgery*, 1(1), 1-13.
31. Patel, H.H. & Prajapati, P. (2018). Study and analysis of decision tree based classification algorithms. *International Journal of Computer Sciences and Engineering*, 6(10), 74-78.
32. Masetic, Z. & Subasi, A. (2016). Congestive heart failure detection using random forest classifier. *Computer Methods and Programs in Biomedicine*, 130, 54-64.
33. Bentéjac, C., Csörgő, A. & Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, 54(3), 1937-1967.
34. Wasule, V. & Sonar, P. (2017). Classification of brain MRI using SVM and KNN classifier. In *2017 Third International Conference on Sensing, Signal Processing and Security (ICSSS)* (218-223). IEEE.

35. Blagus, R. & Lusa, L. (2013). SMOTE for high-dimensional class-imbalanced data. *BMC Bioinformatics*, 14(1), 106.
36. Anand, A., Pugalenth, G., Fogel, G.B. & Suganthan, P.N. (2010). An approach for classification of highly imbalanced data using weighting and undersampling. *Amino Acids*, 39(5), 1385-1391.
37. Rakha, E., Toss, M. & Quinn, C. (2022). Specific cell differentiation in breast cancer: A basis for histological classification. *Journal of Clinical Pathology*, 75(2), 76-84.
38. Aizer, A.A., Chen, M.-H., McCarthy, E., Mendu, M.L., Koo, S., Wilhite, T.J., Graham, P.L., Choueiri, T.K., Hoffman, K.E., Martin, N.E., Hu, J.C. & Nguyen, P.L. (2013). Marital status and survival in patients with cancer. *Journal of Clinical Oncology*, 31(31), 3869-3876.
39. Ghavidel, A. & Pazos, P. (2025). Machine learning (ML) techniques to predict breast cancer in imbalanced datasets: A systematic review. *Journal of Cancer Survivorship*, 19(1), 270-294.
40. Moncada-Torres, A., van Maaren, M.C., Hendriks, M.P., Siesling, S. & Geleijnse, G. (2021). Explainable machine learning can outperform Cox regression predictions and provide insights in breast cancer survival. *Scientific Reports*, 11(1), 1-13.
41. Park, S.W., Park, Y.L., Lee, E.G., Chae, H., Park, P., Choi, D.W., Choi, Y.H., Hwang, J., Ahn, S., Kim, K., Kim, W.J., Kong, S.-Y., Jung, S.-Y. & Kim, H.J. (2024). Mortality prediction modeling for patients with breast cancer based on explainable machine learning. *Cancers*, 16(22), 3799.



Basamak Patlatma Tasarımı ve Analizi Yapabilen Bir Mobil Uygulama Yazılımının Geliştirilmesi

Barış KADEM^{1,a}, Abdulkadir KARADOĞAN^{1,b}

¹İstanbul Üniversitesi - Cerrahpaşa, Mühendislik Fakültesi, Maden Mühendisliği Bölümü, İstanbul, Türkiye

^aORCID: 0009-0004-2518-7681; ^bORCID: 0000-0001-7321-3320

Makale Bilgileri

Geliş : 28.05.2025

Kabul : 25.09.2025

DOI: 10.21605/cukurovaumfd.1708218

Sorumlu Yazar

Abdulkadir KARADOĞAN

akadir@iuc.edu.tr

Anahtar Kelimeler

Patlatma tasarımı

Çevresel etkiler

Parça boyut dağılımı

Mobil uygulama geliştirme

Atf şekli: KADEM, B., KARADOĞAN, A., (2025). Basamak Patlatma Tasarımı ve Analizi Yapabilen Bir Mobil Uygulama Yazılımının Geliştirilmesi. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 725-741.

ÖZ

Bu çalışmada, patlatma tasarımı ve analizleri süreçlerini bütünleşik bir şekilde gerçekleştirilebilen mobil bir uygulama yazılımı geliştirilmiştir. iRockBlast adı verilen bu uygulama, kullanıcıdan alınan saha verileri ile birincil patlatma tasarımı yapabilmekte; yer titreşimi, hava şoku, kaya savrulması gibi çevresel etkileri ve parça boyut dağılım oranlarını tahmin edebilmektedir. Uygulama, Flutter uygulama geliştirme aracı ve Dart programlama dili kullanılarak iOS ve Android mobil işletim sistemlerinde çalışabilecek biçimde geliştirilmiştir. Literatür bilgileri ile oluşturulan algoritmalar üzerine inşa edilmiştir. Bir vaka çalışması ile uygulamanın doğruluğu test edilmiş ve saha sonuçları ile yüksek düzeyde uyum sağladığı gösterilmiştir. Geliştirilen yazılım, mevcut diğer uygulamalardan farklı olarak tasarım ve analiz aşamalarında yenilikler sunmaktadır. Bu uygulama, patlatma operasyonlarının planlanması aşamasında zaman verimliliği ve karar desteği sağlamaktadır.

Development of a Mobile Application Software Capable of Bench Blast Design and Analysis

Article Info

Received : 28.05.2025

Accepted : 25.09.2025

DOI: 10.21605/cukurovaumfd.1708218

Corresponding Author

Abdulkadir KARADOĞAN

akadir@iuc.edu.tr

Keywords

Blast design

Environmental effects

Particle size distribution

Mobile application development

How to cite: KADEM, B., KARADOĞAN, A., (2025). Development of a Mobile Application Software Capable of Bench Blast Design and Analysis. Çukurova University, Journal of the Faculty of Engineering, 40(3), 725-741.

ABSTRACT

In this study, a mobile application software capable of performing integrated blast design and analysis processes has been developed. Named iRockBlast, the application enables users to carry out primary bench blast designs using field data provided on-site. It also allows for the prediction of environmental impacts such as ground vibration, air overpressure, and fly rock, as well as fragment size distribution. The application was developed using the Flutter framework and Dart programming language, making it compatible with both iOS and Android operating systems. It is built upon algorithms derived from well-established literature models. The accuracy of the application was verified through a case study, demonstrating a high degree of consistency between the predicted and actual field results. Unlike existing solutions, the developed software introduces innovative features in both the design and analysis phases. This tool offers significant time efficiency and decision support during the planning of blasting operations.

1. GİRİŞ

İnsanlık tarihinin çeşitli dönemlerinde karşımıza çıkan patlayıcı maddeler, zamanla kaya kazısında da kullanılmaya başlanmıştır. Patlatmalı kazılar, madencilik, taş ocağı işletmeciliği, inşaat sektörleri ve kazı gerektiren diğer altyapı işlerinde yaygın olarak kullanılmaktadır ve bu işlemin kendisinden sonra gelen tüm işlemlere etkisi bulunmaktadır. Bundan dolayı patlatma tasarımı kritik bir önem taşımaktadır. Patlatma tasarımı başarıya veya başarısızlığa bağlı olarak değişkenlik gösterirken, planlayıcılar kontrol edilemeyen parametreleri göz önüne alarak kontrol edilebilen parametreleri düzenlemektedirler. Başarılı ve çevresel etkileri en az olan kontrollü patlatmaları gerçekleştirmek, atım yapılacak alandaki kayacın özelliklerinin iyi bilinmesine ve bu özelliklere uygun patlatma tasarımının yapılmasına bağlıdır. İdeal bir patlatma tasarımı; en düşük delme, patlatma, yükleme, taşıma ve kırma maliyetleriyle istenen parçalanma oranını sağlamayı hedefler. Artan şehirleşmeyle birlikte madencilik, taş ocaklığı, metro, tünel, baraj, yol ve temel kazıları gibi birçok alanda patlatmadan kaynaklanan çevresel etkilerle yoğun bir şekilde karşılaşmaktadır. Bu çevresel etkiler titreşim, aşırı hava basıncı ve kaya savrulmaları gibi sorunları içerir ve bu etkiler belli şartlar altında çevreye zarar verebilmektedir. Patlatma tasarımında çevresel etkilerin kontrolü hem etkin hem de ekonomik çözümler sunarak madencilik faaliyetlerinin sürdürülebilirliğini sağlamaktadır [1-7].

Literatürde basamak patlatması tasarımı yapabilmek için geliştirilmiş pek çok tasarım modeli mevcuttur. Kayaç özellikleri, işletme koşulları, istenen parçalanma oranları gibi değişkenlere bağlı olarak patlatma tasarım parametreleri belirlenmektedir. Birincil tasarımların sahada uygulanması ve sonuçlara göre tasarımın geliştirilmesi ile nihai tasarım elde edilmektedir. Olofsson [8] tarafından basamak patlatma tasarımı ve bir delikte bir ara sıkılama uygulaması ile çift yemleme kullanılmasına yönelik geliştirilmiş bir model mevcuttur. Patlatma tasarımında delik gecikme süreleri Konya ve Walter'ın [9] önerdiği eşitlikler ile belirlenebilmektedir. Patlatma kaynaklı çevresel etkilerin önceden tahmin edilebilmesi için araştırmacılar tarafından geliştirilmiş çeşitli ampirik eşitlikler bulunmaktadır. Yer titreşimi tahmininde Duvall ve Fogelson [10] tarafından geliştirilen ölçekli mesafe – en yüksek parçacık hızı eşitliği açık ocak patlatmalarında sıkça kullanılmaktadır. Patlatma kaynaklı kaya savrulma mesafeleri için farklı senaryolara göre geliştirilmiş çeşitli eşitlikler mevcuttur. Lundborg ve arkadaşları [11] tarafından en kötü durum senaryosu eşitliği, Richards ve Moore [12] tarafından basamak aynası (yatay) ve sıkılama (düşey) yönlerde kaya savrulma mesafeleri eşitlikleri, McKenzie [13] tarafından ise ölçekli gömme derinliği eşitliği önerilmiştir. Patlatma kaynaklı hava şoku miktarlarının tahmininde Siskind ve arkadaşları [14] tarafından geliştirilmiş bir eşitlik bulunmaktadır. Patlatma işlemi sonucu ana kayadan kopup parçalanan kayaçların oluşturacağı yığın tane boyut dağılımının önceden tahmin edilebilmesi amacıyla Cunningham'ın [15] önerdiği iyileştirilmiş Kuz-Ram modeli gerçek sonuçlar ile uyumluluğu nedeniyle tercih edilmektedir.

Son yıllarda yapay zekâ (AI) ve nesnelerin interneti (IoT) gibi teknolojiler, patlatma mühendisliği alanında önem kazanmaya başlamıştır. AI, makine öğrenmesi (ML) ve derin öğrenme (DL) gibi alt dallara ayrılmaktadır. Bu algoritmalar, basamak patlatma tasarımı, çevresel etkilerin tahmini ve parça boyut dağılımının tahmini amacıyla birçok çalışmada kullanılmıştır. Literatürde, patlatma kaynaklı yer titreşiminin tahmini için [16-20]; aşırı hava basıncı tahmini için [21-26]; kayaç savrulması tahmini için [27-31]; parça boyut dağılımı tahmini için ise [32-36] gibi çalışmalar mevcuttur. Gradyan inişi gibi yöntemler kullanılarak, AI algoritmalarından alınan çıktıların belirli bir limitin altına düşürülmesi için girdiler optimize edilebilmektedir. Bu durum, tahmin sonuçlarına göre patlatma tasarımının revize edilmesini sağlamaktadır. Literatürde yer alan pek çok çalışmanın sonuçlarına göre AI tabanlı yöntemler, ampirik yöntemlere oranla daha yüksek tahmin başarıları sunmaktadır. Ancak doğası gereği kara-kutu olarak tanımlanan bu yöntemler, tahmin başarısını artırırken ampirik yöntemlerin sağladığı açıklanabilirliği azaltmaktadır. Ayrıca, yapılan çalışmaların birçoğu bilimsel ölçekli olup, saha uygulamalarında kullanımı son derece kısıtlıdır. IoT tabanlı sensörler ile saha koşullarında verilerin gerçek zamanlı olarak toplanması ve AI algoritmalarıyla analiz edilmesi, patlatma mühendisliğinde yeni bir yaklaşım olarak ortaya çıkmaktadır. Bu yaklaşım sayesinde, titreşim, aşırı hava basıncı ve kayaç savrulması gibi çevresel etkiler anlık olarak takip edilebilecek ve elde edilen sonuçlara göre patlatma tasarımları hızla revize edilebilecektir. Aruna ve diğ. [37] tarafından yapılan çalışmada, IoT teknolojisinin AI ile kullanıldığında, patlatma süreçlerinde çevresel etkilerin daha doğru tahmin edilmesine katkı sağladığı gösterilmiştir.

Patlatmanın tasarım aşaması dinamik bir yapıdadır. Patlatmanın tasarımı ve analizi fazla sayıda değişkenin eşzamanlı olarak değerlendirilmesini gerektiren karmaşık bir süreçtir. Patlatma planlayıcıları saha üzerinde her zaman anlık olarak yeni bir patlatma tasarımı geliştirebilecek zamanı bulamayabilmektedir. Patlatma tasarımı, parçalanma tahmini ve çevresel etki tahminleri amacıyla geliştirilmiş pek çok masaüstü yazılım

mevcuttur ancak saha koşullarında bilgisayarlara erişim her zaman mümkün olamamaktadır. Bu durum, patlatma planlayıcılarının ani değişikliklere hızlıca tepki vermesini zorlaştırmakta, tasarım-analiz-yenileme döngüsünü yavaşlatmaktadır. Gelişen teknolojiye paralel olarak gelişen akıllı telefon teknolojisi, akıllı telefonları kişisel birer cep bilgisayarı haline getirmiştir. Son dönemde akıllı telefonlar, patlatma tasarımı ve analizi için uygun bir donanım olarak karşımıza çıkmaktadır. Mobil cihazlara entegre çözümler, karar süreçlerinde zaman avantajı sağlayabilmektedir [38].

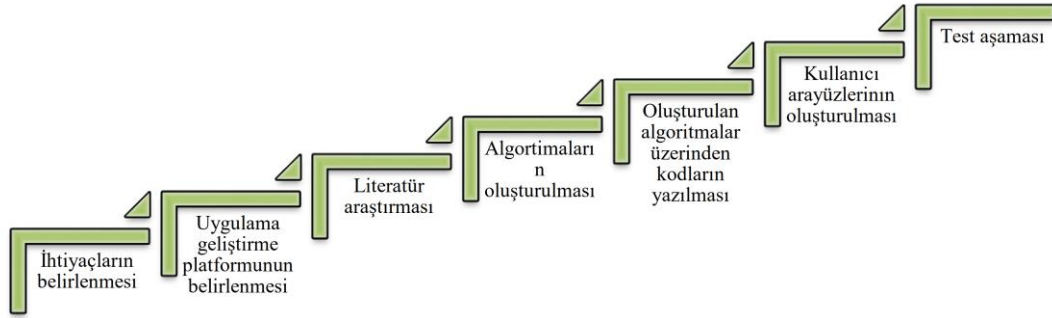
Daha önce benzer amaçlarla patlayıcı firmaları ve/veya yazılım firmaları tarafından geliştirilen uygulamalar [39-45] incelendiğinde hiçbir uygulamanın patlatma tasarımı, çevresel etki analizi ve parça boyut dağılımı analizi konularının tamamını bütüncül bir şekilde ele alamadığı, çeşitli konularda eksiklikleri olduğu ve kullanıcı odaklı çözümler sunamadıkları görülmüştür.

Bu çalışmanın amacı, anlık olarak kullanıcı tarafından uygulamaya girilecek olan saha, patlayıcı madde ve patlatma geometrisi ile ilgili veriler üzerinden birincil patlatma tasarımı yapabilen, bu tasarımdan yola çıkarak yer titreşimi, kaya savrulması, hava şoku gibi çevresel etkileri ve parça boyut dağılımlarını patlatma öncesi tahmin edebilen ve tüm sonuçları raporlayarak bu anlamda patlatma planlayıcılarına karar desteği sağlayabilecek bir mobil uygulama yazılımı geliştirmektir. Bu çözüm, sahada gerçek zamanlı değerlendirme yapılabilmesini mümkün kılarak hem güvenlik hem de ekonomik verimlilik açısından önemli avantajlar sağlayacaktır.

2. YÖNTEM

Patlatma tasarımı ve analizleri süreçlerini bütünlük bir şekilde gerçekleştirilebilen mobil bir uygulama yazılımı geliştirilmesi kapsamında, aşağıdaki aşamalar sistematik biçimde gerçekleştirilmiştir. Çalışmada izlenen yöneme ait metodolojik akış Şekil 1’de verilmiştir.

1. Kullanıcı ihtiyaçlarının belirlenmesi,
2. Uygun yazılım dili ve mobil uygulama geliştirme platformunun belirlenmesi,
3. Uygulama algoritmasında kullanılacak literatür bilgilerin tespiti,
4. İlgili algoritmaların geliştirilmesi ve yazılım kodlarına dönüştürülmesi,
5. Kullanıcı arayüzlerinin tasarlanması,
6. Gerçek zamanlı testlerin gerçekleştirilmesi.



Şekil 1. Metodolojik akış

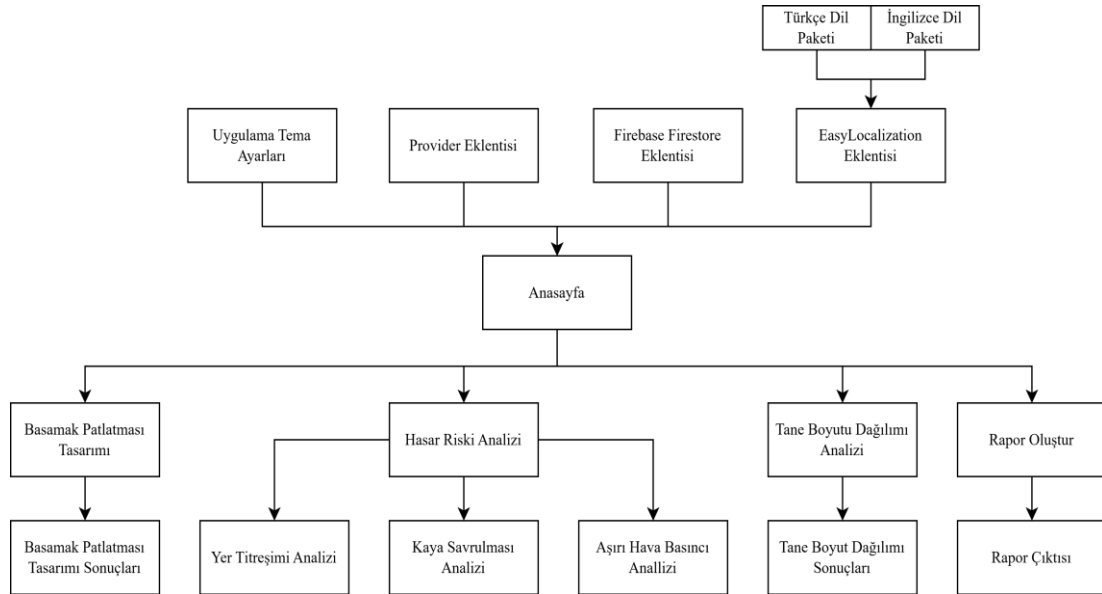
Çalışmanın birinci aşamasında mevcut olan benzer mobil uygulamalar analiz edilmiştir. Bu uygulamaların sahip olduğu ve eksik kalan yönleri belirlenerek geliştirilecek olan yazılımın fonksiyonel kapsamı tanımlanmıştır. Uygulamanın hem iOS hem de Android işletim sistemlerine sahip cihazlarda çalışabilir olması hedeflenmiştir. Bu nedenle, tek bir kod tabanı ile birden fazla platformda çalışabilen, zaman ve maliyet açısından daha verimli olan Flutter çapraz platform geliştirme platformu tercih edilmiştir. Programlama dili olarak ise Flutter ile entegre şekilde çalışan Dart dili seçilmiştir. Uygulamanın ihtiyaç duyacağı hesaplamalar ve bu hesaplamalarda kullanılacak bilgiler için literatür taraması yapılmıştır. Elde edilen bilgiler ile algoritmalar geliştirilmiş, bu algoritmalar kodlara dönüştürülmüştür. Kullanıcı arayüzleri tasarlanmış ve sistemin çalışabilirliği gerçek zamanlı olarak test edilmiştir. Nihai testler ve uygulamadan elde edilen sonuçların saha verileri ile karşılaştırılması ile geliştirme aşaması sonlandırılmıştır. Uygulamaya iRockBlast adı verilmiştir. Uygulama; birincil patlatma tasarımının yapılmasını, yer titreşimi, kaya savrulması ve aşırı hava basıncı gibi çevresel etkilerin önceden tahmin edilmesini, parçalanma

oranlarının analiz edilmesini ve tüm bu adımların sonuçlarının bir rapor halinde kullanıcıya sunulmasını sağlamaktadır. Uygulama toplamda 6209 satır kod içermektedir. Uygulama algoritmasında, kullanıcıya kolaylık sağlamak amacıyla üç ayrı veri tabanı oluşturulmuştur. Bu veri tabanları, uygulamaya entegre biçimde çalışmaktadır. Veri tabanları kullanıcı tarafından yapılan seçimlere göre sorgulanmakta ve bu sorgulardan gelen değerler hesaplamalar için gerekli bazı parametreleri sağlamaktadır. Bu veri tabanları şu şekildedir;

- Yer Titreşimi Veri Tabanı:** Bu veri tabanı, titreşim tahminlerinde kullanılmak üzere tasarlanmıştır. Saha iletim (k) ve jeolojik (β) sabitlerinin ilgili sahaya özgü olarak belirlenebilmesi için istatistikî güvenilirlik düzeyi (≥ 30 atım ölçümü) gerekmektedir. Ön veri eksikliği bulunan sahalarda kullanılabilmesi için ulusal literatürden alınan denklem sabitleri derlenerek bir veri tabanı oluşturulmuştur.
- Hava Şoku Veri Tabanı:** Hava şoku tahminleri için de yer titreşiminde olduğu gibi benzer bir yaklaşım benimsenmiştir. Literatürdeki güvenilir hava şoku yayılım formülleri ile oluşturulan veri tabanı, kullanıcıların sahadan yeterli sayıda ölçüm verisine sahip olmadan öncesinde güvenli tasarım yapabilmesine imkân tanımaktadır.
- Patlayıcı Malzeme Bilgi Veri Tabanı:** Türkiye’de ve dünyada yaygın olarak kullanılan ticari patlayıcıların marka, model ve yoğunluk bilgilerini içermektedir. Kullanıcı, tasarım aşamasında patlayıcı seçimi yaparken veri tabanındaki teknik özelliklere doğrudan erişmekte; böylece hesaplama sürecinde ek bir araştırma yapmasına gerek olmamaktadır.

3. UYGULAMA ÖZELLİKLERİ VE KULLANICI ARAYÜZÜ

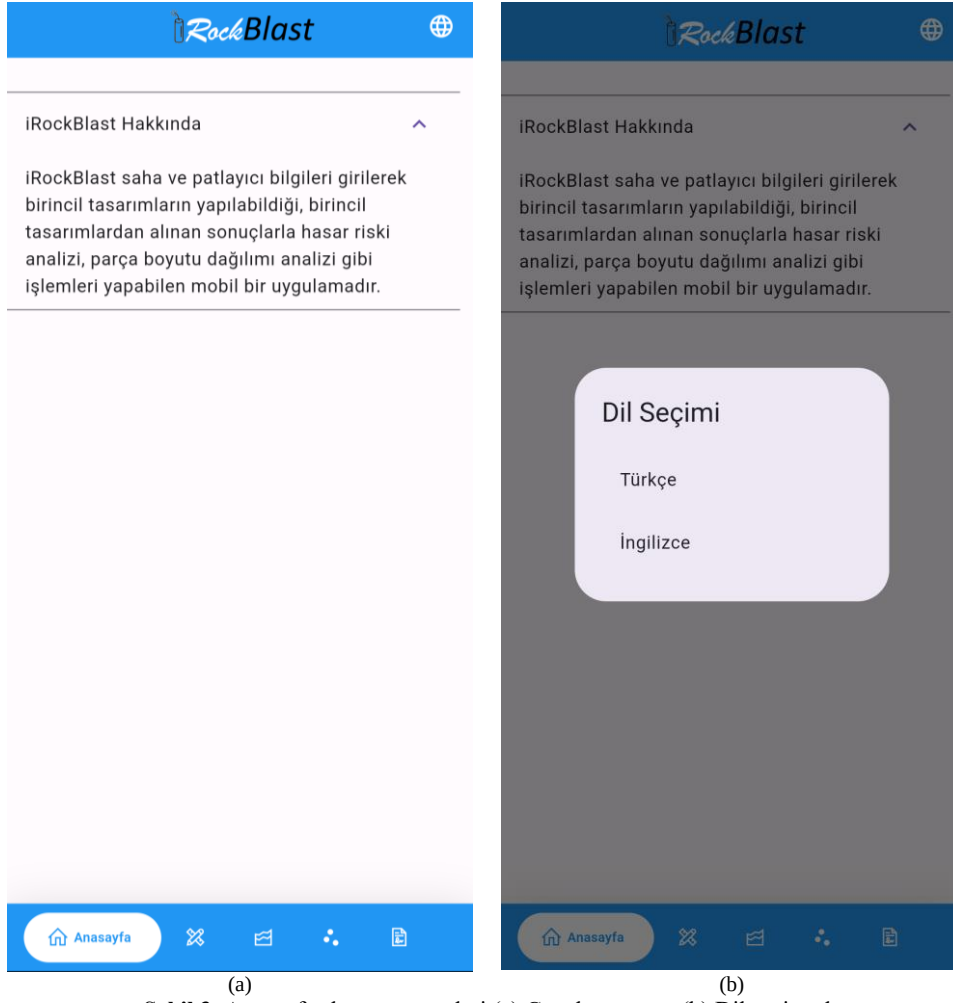
Bu bölümde geliştirilen uygulamanın çalışma mantığı açıklanmış ve kullanıcı arayüzü görselleri verilmiştir. Uygulamanın yapısı beş ana modül üzerine kurulmuştur. Bunlar sırası ile anasayfa, basamak patlatması tasarımı, hasar riski analizi, parça boyut dağılımı analizi ve rapor oluşturma modülleridir. Uygulamadaki sayfaların alt-üst ilişkileri Şekil 2’de sunulmuştur.



Şekil 2. Uygulama sayfaları arasındaki ilişkiler

3.1. Anasayfa

Geliştirilen uygulamanın anasayfa algoritması, uygulamanın temel kodlarına yer vermektedir. Bu kodlar uygulamanın başlatılması, uygulamanın ihtiyaç duyacağı eklentilerin tanımlanması, kullanıcı arayüzü tasarımları, sayfalar arası geçişler, uygulama dil seçenekleri gibi içeriklere sahiptir. Uygulamaya entegre olarak çalışan veri tabanları ve uygulamada hesaplanan değerlerin bellekte saklanması ve gerekli sayfalardaki değişkenlere atanması için gerek duyulan durum yönetimi sistemi ve uygulamanın çalışacağı dil dosyaları bu sayfa içerisinde tanımlanmıştır. Tasarım ve analiz işlemlerine geçişler anasayfa üzerinden yapılmaktadır. Uygulamanın anasayfa ekran görüntüleri Şekil 3’te verilmiştir.



(a) Genel görünüm (b) Dil seçim ekranı
Şekil 3. Anasayfa ekran görüntüleri (a) Genel görünüm (b) Dil seçim ekranı

3.2. Basamak Patlatma Tasarımı Sayfası

Bu sayfada kullanıcıdan istenen veriler ile Olofsson [8] tarafından geliştirilen tek ve çift kademeli patlatma tasarım modeli eşitlikleri kullanılarak tasarım parametreleri hesaplanmaktadır. Olofsson modeli, tasarım için daha az sayıda veri girişine gerek duyması ve daha pratik olması nedeni ile tercih edilmiştir. Ayrıca, önerilen çift kademeli hesaplamaların uygulamaya dahil edilmesi ile çevresel etki değerlerinin izin verilen limitleri aşması durumunda bir ara sıkılama uygulaması ile delik içerisindeki patlayıcı madde miktarının bölünüp farklı zamanlarda ateşlenmesi sonucunda çevresel etki değerleri düşürülebilmektedir. Delikler ve sıralar arası gecikme hesaplamaları Konya ve Walter [9] tarafından önerilen yaklaşım ile yapılmaktadır. Bu sayfanın ekran görüntüleri Şekil 4'te verilmiştir. Arayüzde kullanıcıdan girişi istenen parametreler şu şekildedir;

- Atım Bilgileri: Delik çapı, delik eğimi, basamak yüksekliği, atım uzunluğu, atım genişliği ve kaya sabiti.
- Patlayıcı Bilgileri: Kademe sayısı (tek/çift), patlayıcı şarj türü (kolon/kolon ve dip), yemleyici patlayıcı türü, yemleyici patlayıcı miktarı, delik içi kapsül gecikme süresi, ara kapsül gecikme süresi, ara sıkılama uzunluğu, dip şarj patlayıcı modeli, dip şarj patlayıcı türü, dip şarj patlayıcı yoğunluğu, kolon şarj patlayıcı modeli, kolon şarj patlayıcı türü, kolon şarj patlayıcı yoğunluğu.
- Gecikme Bilgileri: Atım sonrası istenen sonuçlar, kayaç faktörü.

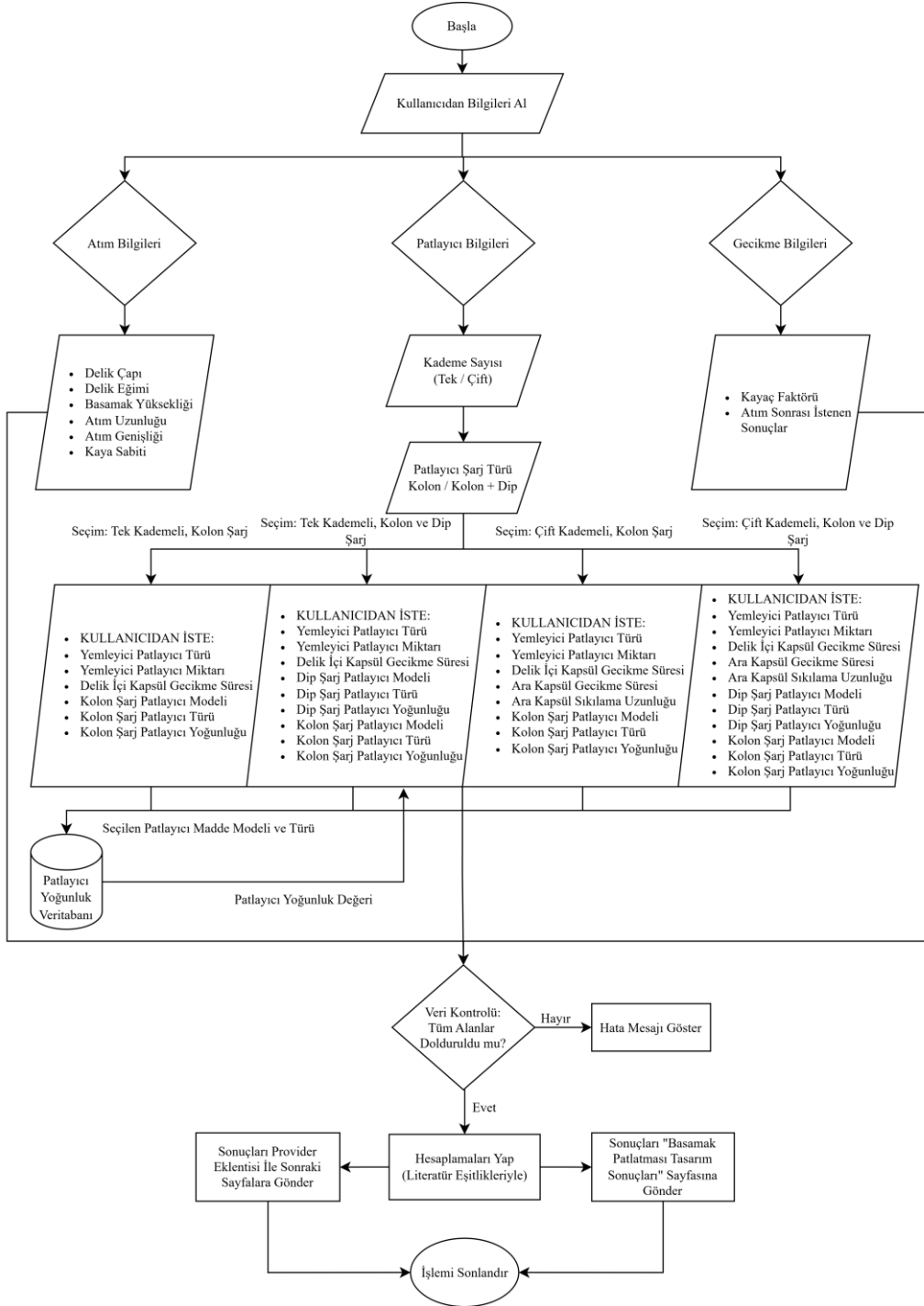
Yapılan parametre girişlerinin uygunluğu arayüz tarafından kontrol edilmekte, uygun değil ise hata mesajı gösterilmektedir. Uygun ise literatür eşitliklerine göre hesaplamalar yapılmakta ve aşağıdaki parametreler hesaplanmaktadır;

- Kolon şarj konsantrasyonu, dip şarj konsantrasyonu, maksimum dilim kalınlığı, delik taban payı, delik boyu, delik hata payı, pratik dilim kalınlığı, sıralar arası mesafe, atım genişliğine uyarlama, pratik sıralar arası mesafe, delik sayısı, sıra sayısı, toplam delik sayısı, özgül delme, sıkılama boyu, seçildi ise dip şarj boyu, dip şarj miktarı, kolon şarj boyu, kolon şarj miktarı, seçildi ise çift kademeli şarj hesaplamaları, toplam şarj miktarı, gecikme başı en yüksek şarj miktarı, özgül şarj, delikler arası gecikme süresi, sıralar arası gecikme süresi.

Parametre	Sonuç	Birim
Delik Çapı	89	mm
Delik Eğimi	90	°
Basamak Yüksekliği	8.000	m
Atım Uzunluğu	30.000	m
Atım Genişliği	6.000	m
Maks. Dilim Kalınlığı	3.315	m
Delik Taban Payı	0.994	m
Delik Boyu	8.994	m
Delik Hata Payı	0.359	m
Pratik Dilim Kalınlığı	2.956	m
Sıralar Arası Mesafe	3.695	m
Delik Sayısı	10	adet
Sıra Sayısı	4	adet
Toplam Delik Sayısı	40	adet
Patlatma Kademe Türü	Tek kademeli	
Patlayıcı Şarj Tipi	Kolon	
Sıkılama Boyu	2.956	m
Ara Sıkılama Boyu	0.000	m
Dip Şarj Konsantrasyonu	0.000	kg/m
Dip Şarj Boyu	0.000	m
Dip Şarj Miktarı	0.000	kg
Kolon Şarj Konsantrasyonu	4.977	kg/m
Kolon Şarj Boyu	6.039	m
Kolon Şarj Miktarı	30.053	kg
Yemleyici Patlayıcı Miktarı	1.000	kg
Toplam Şarj Boyu	6.039	m
Toplam Şarj Miktarı	31.053	kg
Gecikme Başı Maksimum Şarj	31.053	kg
Özgül Delme	0.103	m/m³
Özgül Şarj	0.355	kg/m³
Delik İçi Kapsül Gecikme	500	ms
Ara Kapsül Gecikme	0	ms
Delikler Arası Gecikme Süresi	17	ms
Sıralar Arası Gecikme Süresi	50	ms

Şekil 4. Basamak patlatması tasarımı sayfası ekran görüntüleri (a) Veri giriş kutuları (b) Tasarım hesap sonuçları

Uygulama, bu tasarım işlemleri sırasında deliklerdeki su durumuna göre esnek patlatma tasarımı yapılabilmesi adına kullanıcıya kolon ve dip şarj seçeneklerini ve bu şarjlarda kullanılacak patlayıcı maddelerin seçimine imkân tanımaktadır. Patlayıcı madde seçimi, uygulamanın patlayıcı madde veri tabanı ile bağlantılıdır. Bu veri tabanı Türkiye’de ve dünyada yaygın olarak kullanılan birçok patlayıcı üreticisinin ürün portföyünde yer alan patlayıcıların yoğunluk değerlerini içermektedir. Yapılan seçime göre patlayıcı yoğunluk değeri veri tabanından alınmaktadır. Patlayıcı yoğunluk değerleri el ile de uygulamaya girilebilmektedir. Yapılan hesaplama sonuçları “Basamak Patlatma Tasarımı Sonuçları” sayfasında bir tablo halinde kullanıcıya gösterilmekte ve durum yönetimi sistemi ile kendisinden sonra gelen analiz ve rapor sayfalarına iletilmektedir. Bu modülde kullanılan algoritmanın akım şeması Şekil 5’te verilmiştir.



Şekil 5. Basamak patlatması tasarımı sayfası algoritması

3.3. Hasar Riski Analizi Sayfası

Uygulamanın bu modülünde patlatma kaynaklı çevresel etkilerin atım öncesi tahmin edilebilmesi amacı ile oluşturulmuş olan yer titreşimi analizi, kaya savrulması analizi ve aşırı hava basıncı analizi alt sayfalarına yer verilmiştir. Her bir tahmin sayfasının açıklaması aşağıda sunulmuştur.

3.3.1. Yer Titreşimi Analizi Sayfası

Yer titreşimi analizi sayfasında açık ocak patlatmalarında titreşimin en yüksek parçacık hızını tahmin etmedeki başarısı nedeniyle Duvall ve Fogelson [10] tarafından önerilen titreşim yayılım denklemi eşitliği

kullanılarak en yüksek parçacık hızı değeri girildi ise, gecikme başı en yüksek şarj miktarı; gecikme başı en yüksek şarj miktarı girildi ise en yüksek parçacık hızı değerleri hesaplanmaktadır. Bu sayfada kullanıcıdan istenen parametreler şu şekildedir;

- En yüksek parçacık hızı, gecikme başı en yüksek şarj miktarı, mesafe, il, ilçe, kayaç türü, arazi iletim sabiti, saha jeolojik sabiti.

Bu sayfa ile bağlantılı olarak çalışan yer titreşimi veri tabanı içerisinde ulusal yayınlardan elde edilen titreşim yayılım denklemi sabitleri yer almaktadır. Sayfa içerisinde kullanıcının yapacağı il-ilçe ve kayaç türü seçimlerine göre uygun titreşim yayılım denklemi sabitleri veri tabanından alınmakta ve hesaplamalarda kullanılmaktadır. Bu değerler el ile de girilebilmektedir. Yer titreşimi analizi sayfasının ekran görüntüleri Şekil 6'da verilmiştir.

3.3.2. Kaya Savrulması Analizi Sayfası

Kaya savrulması analizi sayfası, kullanıcının ihtiyaç duyabileceği farklı yönlerdeki kaya savrulma mesafelerini tahmin edebilmek amacıyla atım sonucu savrulabilecek kayaçların ortalama tane boyutunun ve en kötü durum senaryosu [11], basamak aynası yönünde ve sıkılama yönünde [12] ve ölçekli gömme derinliği [13] eşitliklerine göre kaya savrulma mesafelerinin hesaplandığı sayfadır. Bu sayfanın ekran görüntüleri Şekil 7'de verilmiştir. Kullanıcıdan istenen girdi parametreleri şu şekildedir;

- Delik çapı, delik eğimi, dilim kalınlığı, kaya sabiti, sıkılama boyu, toplam şarj boyu, toplam şarj miktarı, ortalama patlayıcı yoğunluğu.

3.3.3. Hava Şoku Analizi Sayfası

Hava şoku analizi sayfası, atım sonrası oluşacak olan aşırı hava basıncı değerlerinin önceden tahmin edilebilmesi amacıyla oluşturulmuştur. Hesaplamalar Siskind ve arkadaşları [14] tarafından önerilen eşitlikler ile yapılmaktadır. Bu sayfada kullanıcıdan istenen girdi parametreleri şu şekildedir;

- Mesafe, gecikme başı en yüksek şarj miktarı, patlatma türü, saha iletim sabiti, saha jeolojik sabiti.

Bu sayfa ile bağlantılı olarak çalışan hava şoku veri tabanı içerisinde patlatma türlerine göre güvenilir olarak kabul edilen hava şoku yayılım denklemi sabitleri yer almaktadır. Uygulama içerisinde kullanıcı tarafından yapılan patlatma türü seçimine denk gelen yayılım denklemi sabitleri hesaplamalarda kullanılmaktadır. Bu değerler el ile de girilebilmektedir. Hava şoku analizi sayfasının ekran görüntüleri Şekil 8'de verilmiştir.

3.4. Parça Boyut Dağılımı Analizi Sayfası

Parça boyut dağılımı analizi sayfası, kullanıcıdan alınan veriler ve/veya yapılan birincil patlatma tasarımından alınan sonuçlar üzerinden tahmin başarısının gerçek değerlere yüksek uygunluğu ve kullanım kolaylığıyla bilinen iyileştirilmiş Kuz-Ram parçalanma modeli [15] kullanılarak, atım öncesinde parça boyut dağılım oranlarının belirlenmesi amacıyla kullanılmaktadır. Bu sayfada kullanıcıdan istenen parametre girişleri şu şekildedir;

- Kayaç Bilgileri: Özgül ağırlık, elastisite modülü, tek eksenli basınç dayanımı, kaya kütle sınıflaması, eklem düzlem aralığı, eklem düzlem yönelimi.
- Atım Bilgileri: Delik çapı, dilim kalınlığı, sıralar arası mesafe, basamak yüksekliği, dip şarj boyu, kolon şarj boyu, delik düzeni.
- Patlayıcı Bilgileri: Patlayıcının ağırlıkça kuvveti, gecikme başı en yüksek şarj miktarı.
- Parçalanma Dağılım Hedefleri: Elek üstü boyutu, optimum boyut, elek altı boyutu.

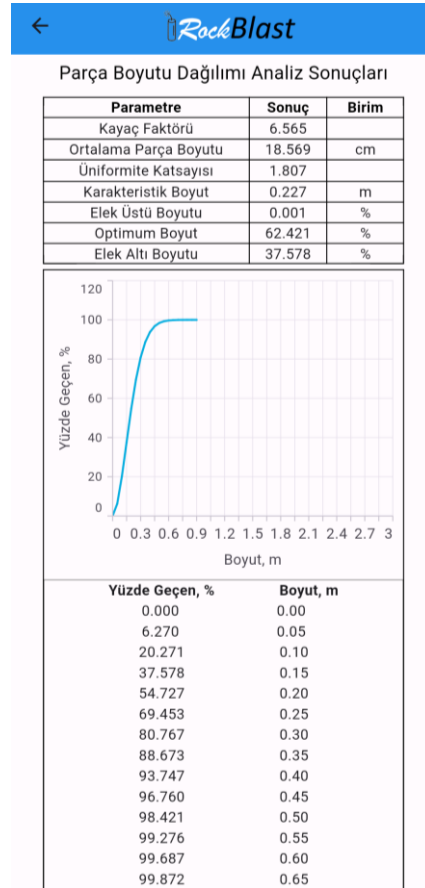
Kayaç faktörü, ortalama parça boyutu, üniformite indeksi, karakteristik boyut, patlatılabilirlik indeksi gibi parametreler hesaplandıktan sonra % geçen – boyut eğrisi ve tablosu oluşturulmakta ve tüm hesaplama sonuçları arayüzde kullanıcıya gösterilmektedir. Bu sayfanın ekran görüntüleri Şekil 9'da verilmiştir.

Şekil 6. Yer titreşimi tahmini sayfası ekran görüntüsü

Şekil 7. Kaya savrulması tahmini sayfası ekran görüntüsü

Şekil 8. Hava şoku tahmini sayfası ekran görüntüsü

(a)

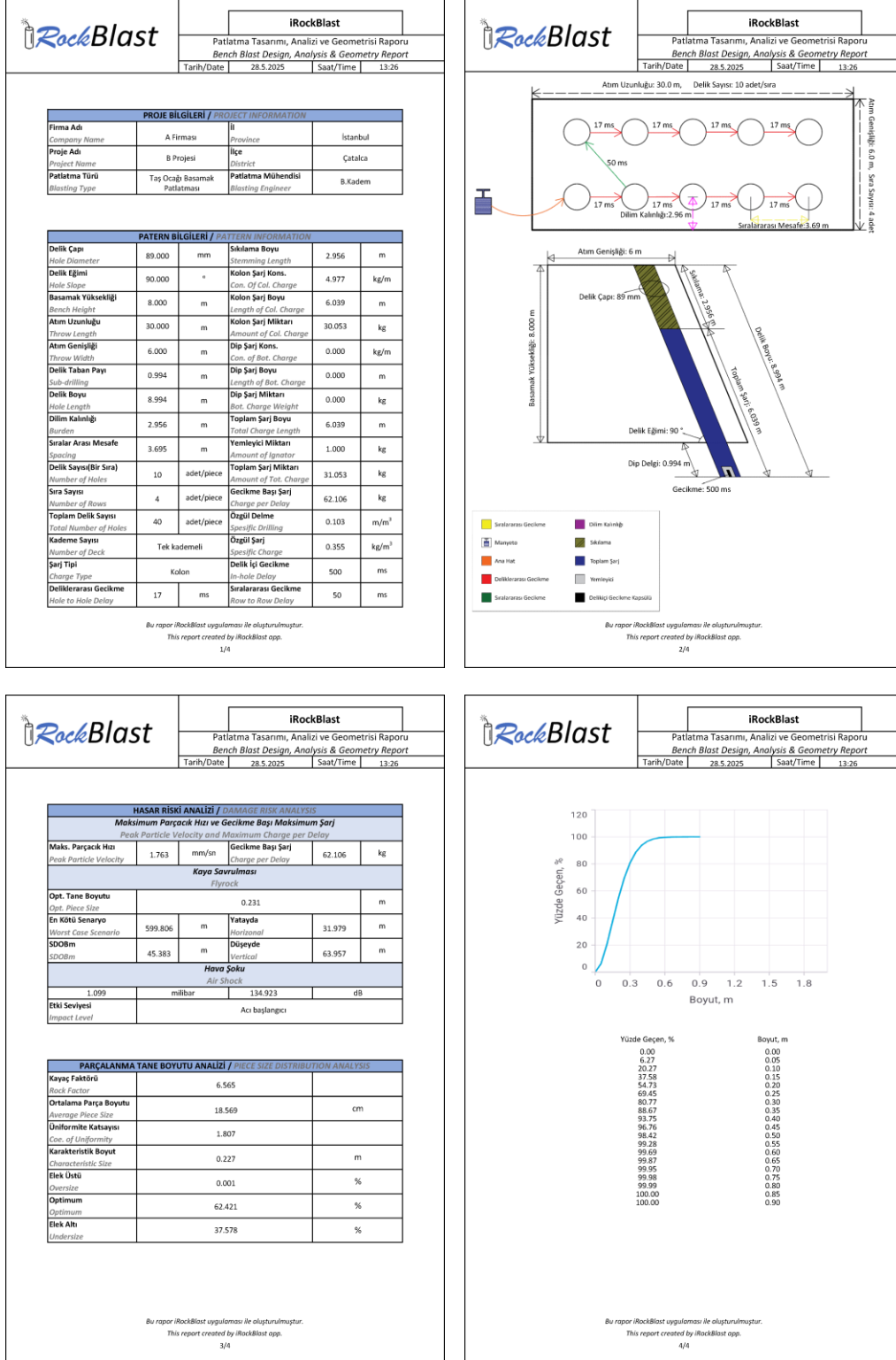


(b)

Şekil 9. Parça boyut dağılımı analizi sayfası ekran görüntüleri (a) Veri giriş kutuları (c) Hesaplama sonuçları

3.5. Rapor Oluştur Sayfası

Rapor oluştur sayfası, uygulama tarafından üretilen patlatma tasarımı, çevresel etki analizi ve parça boyut dağılımı analizi sonuçlarının PDF formatında yazılı ve görsel olarak bir rapor halinde üretilbildiği sayfadır. Bu sayede oluşturulan tasarım ve analiz sonuçları kaydedilebilmekte ve paylaşılabilir. Bu modül tarafından üretilen raporun ekran görüntüleri Şekil 10'da sunulmuştur.



Şekil 10. Uygulama tarafından üretilen raporun sayfaları

4. VAKA ÇALIŞMASI

Bu bölümde, Özyurt ve arkadaşları [39] tarafından Trakya bölgesinde gerçekleştirilen örnek bir atıma ait bilgiler ve sahadan elde edilen ölçüm sonuçları kullanılarak geliştirilen uygulamanın tasarım, çevresel etki ve parçalanma tahminleri sonuçları karşılaştırılmaktadır. Araştırmacılar tarafından sahada uygulanan patlatma tasarımına ait bilgiler ve sahadan ölçülen sonuçlar Çizelge 1’de sunulmuştur.

Çizelge 1. Örnek saha çalışmasına ait bilgiler [39]

Uygulanan patlatma tasarımı			Ölçülen çevresel etki sonuçları		
Parametre	Değer	Birim	Parametre	Değer	Birim
Patern bilgileri			Ölçüm cihazı: Instanfel BE10709		
Basamak yüksekliği	8	m	Maksimum Parçacık Hızı	1,397	mm/s
Delik çapı	89	mm	Frekans	12,80	Hertz
Delik boyu	10	m	Hava Şoku	126,60	dB
Dip delgi	2	m	Mesafe	149,00	m
Dilim kalınlığı	2,70	m	Kaya savrulma mesafesi		
Sıralar arası mesafe	2,70	m			
Özgül delme	0,17	m/m ³			
Delik adeti	40	adet			
Patlayıcı yoğunluğu	0,80	g/cm ³			
Sıkılama boyu	2,50	m			
Şarj boyu	7,50	m			
Şarj konsantrasyonu	4,97	kg/m			
Bir delikteki yemleyici miktarı	1	kg			
Bir delikteki ANFO miktarı	37,275	kg			
Bir delikteki toplam patlayıcı miktarı	38,275	kg			
Toplam patlayıcı miktarı	1491	kg			
Çakışan delik sayısı	2	adet			
Gecikme başı şarj miktarı	76,55	kg			
Özgül şarj	0,47	kg/m ³			
Delik içi gecikme	500	ms			
Yüzey gecikmesi	25	ms			
Sıralar arası gecikme	42	ms	Parça boyut dağılımı analizi sonuçları		
Saha - kayaç bilgileri			P ₂₀	118,33	mm
Tek eksenli basınç dayanımı	47,63	MPa	P ₅₀	227,01	mm
Süreksizlik aralığı	1,70	m	P ₈₀	354,01	mm
Yoğunluk	2,497	g/cm ³	En Büyük Parça Boyutu	657,59	mm

4.1. Basamak Patlatması Tasarımı Sonuçlarının Karşılaştırılması

Örnek saha çalışmasından alınan tasarım parametreleri (Çizelge 1) uygulamaya girilerek hesaplamalar yapılmıştır. Uygulama tarafından üretilen hesaplama çıktıları Çizelge 2’de sunulmuştur. Uygulama çıktıları ile saha uygulamasında kullanılan tasarım parametreleri arasında genel olarak yüksek benzerlik söz konusudur. Farklılıkların bir kısmı uygulamanın literatür bilgilere göre hesaplama yaparken, planlamacıların bu bilgilere ek olarak deneyimsel tecrübelerini de hesaplamalarda kullanmasından kaynaklanmıştır. Elde edilen sonuçlar, geliştirilen uygulamanın literatür temelli tasarım algoritmasının, saha değerleri ile yüksek düzeyde örtüşüğünü göstermektedir.

Araştırmacılar, örnek çalışmada aynı gecikme sürelerinde iki delik ateşlemiş, bu nedenle gecikme başı en yüksek şarj miktarı, delik içi toplam şarj miktarının iki katı olmuştur. Uygulamada bundan sonraki hesaplamalarda gecikme başı en yüksek şarj değeri için saha şartlarına uygun biçimde delik içi toplam şarj miktarının iki katı kabul alınmıştır.

Çizelge 2. Basamak patlatması tasarımı için uygulamaya girilen değerler ve alınan sonuçlar

Girdiler			Çıktılar		
Parametre	Değer	Birim	Parametre	Sonuç	Birim
<i>Atım bilgileri</i>			Delik taban payı	0,994	m
Delik çapı	89	mm	Delik boyu	8,994	m
Delik eğimi	90	°	Delik hata payı	0,359	M
Basamak yüksekliği	8	m	Dilim kalınlığı	2,956	m
Atım uzunluğu	30	m	Sıralar arası mesafe	3,695	m
Atım genişliği	6	m	Delik sayısı	10	adet
Kayaç türü	Zayıf kayaç		Sıra sayısı	4	adet
<i>Patlayıcı bilgileri</i>			Toplam delik sayısı	40	adet
Patlatma kademe sayısı	Tek kademeli		Sıkılama boyu	2,956	m
Patlayıcı şarj tipi	Kolon şarj		Şarj konsantrasyonu	4,997	kg/m
Yemleyici patlayıcı türü	Jelatinit dinamit		Şarj boyu	6,039	m
Yemleyici patlayıcı miktarı	1	kg	Kolon şarj miktarı	30,053	kg
Delik içi kapsül gecikmesi	500	ms	Delik şarj miktarı	31,053	kg
Kolon şarj modeli	ANFO		Gecikme başı şarj miktarı	31,053	kg
Kolon şarj türü			Toplam şarj miktarı	1242,12	kg
Kolon şarj patlayıcı yoğunluğu	0.8	g/cm ³	Özgül delme	0,103	m/m ³
<i>Gecikme bilgileri</i>			Özgül şarj	0,355	kg/m ³
Kayaç faktörü	Bazı kireçtaşları, şeyl vb.		Delikler arası gecikme	25	ms
Atım sonrası istenen sonuçlar	Minimum geri kırılma ve dağılık yığın		Sıralar arası gecikme	50	ms

4.2. Patlatma Kaynaklı Çevresel Etki Sonuçlarının Karşılaştırılması

Çevresel etki tahminleri, uygulamanın yer titreşimi, kaya savrulması ve hava şoku modülleri kullanılarak gerçekleştirilmiş ve Çizelge 1’deki saha ölçümleriyle karşılaştırılmıştır. Hasar riski analizi için kullanılan girdi parametreleri ve alınan sonuçlar Çizelge 3’te verilmiştir. Aşağıda her bir etki türü için kullanılan yöntem ve elde edilen sonuçlar aşağıda özetlenmiştir:

a. Yer Titreşimi: Uygulamanın titreşim veri tabanından, aynı sahada daha önce ölçülmüş ve veri tabanında kayıtlı İstanbul-Çatalca-Sarıkayatepe Taş Ocağı titreşim yayılım denklemi sabiti verileri alınmıştır ve hesaplamalarda kullanılmıştır. Uygulama tahmini ile saha ölçümlerinin en yüksek parçacık hızı değerleri arasındaki farklılığın, jeolojik heterojenite ve ölçüm anındaki atmosferik koşullar gibi kontrol edilemeyen değişkenlerden kaynaklandığı düşünülmektedir. Önemli olan husus, bu sapmanın kabul edilebilir düzeyde kalması ve tahmin edilen en yüksek parçacık hızı değerlerinin çevresel sınırlar içerisinde kalmasıdır.

b. Kaya Savrulması: Sahada gerçekleştirilen atıma ait video görüntüsü Kinovea uygulaması ile basamak yüksekliği (8 m) baz alınarak analiz edilmiştir. Video görüntülerinin görünen kısmında kaya parçalarının yatay yönde yaklaşık 20 metre savrulduğu görülmüştür. Bu parçacıklar bir miktar daha serbest hareketine devam edecek, sonrasında ise yerde sürüklenecektir. Sahadaki kaya savrulma mesafelerinin yaklaşık 30 metre olacağı ön görülmüştür. Geliştirilen uygulama ile yapılan kaya savrulma mesafelerinin tahmin sonuçları ile sahada uygulanan atımın yatay yöndeki sonuçları örtüşmektedir. Atım videosundan elde edilen yatay yöndeki savrulma mesafesi, uygulamada hesaplanan kayaç savrulma mesafesi sınırlarının içerisinde kalmaktadır.

c. Hava Şoku: Bu hesaplama için uygulamada patlatma türü olarak “taş ocağı basamak patlatması” seçimi yapılmış, bu seçime karşılık gelen değer uygulamanın veri tabanından alınarak hesaplamalarda kullanılmıştır. Hesaplanan hava şoku değeri ile örnek saha çalışmasında yapılan ölçüm sonuçları arasında yaklaşık 8 dB’lik bir fark söz konusudur. Elde edilen sonuçlar aynı etki seviyesi içerisinde kalmaktadır. Cihaz hassasiyeti ve yerel atmosferik koşullar göz önünde bulundurulduğunda bu sonuç kabul edilebilir sapma seviyesindedir.

4.3. Parça Boyut Dağılımı Sonuçlarının Karşılaştırılması

Uygulamaya girilen parametre değerlerinden alınan parça boyut dağılım analizi sonuçları ile örnek saha çalışmasında araştırmacılar tarafından görüntü işleme teknikleri kullanılarak elde edilen yığın parça boyutu dağılımı sonuçları büyük oranda benzerlik göstermektedir. Sahadan elde edilen verilerde, sadece yığının fotoğrafta görünen yüzeysel kısımları analiz edilebildiğinden, uygulamadan alınan sonuçlar ile sahada

yapılan ölçümler arasında bazı farklılıkların olması normal bir durumdur. Basamak patlatması sonuçlarının karşılaştırılması amacı ile uygulamaya girilen değerler ve alınan sonuçlar Çizelge 4'te verilmiştir.

Bu vaka çalışması, geliştirilen uygulamanın sahada uygulanan gerçek tasarım ve ölçümlerle yüksek uyum sağladığını, aynı zamanda literatür tabanlı hesaplamaların pratikte geçerli sonuçlar verdiğini ortaya koymaktadır.

Çizelge 3. Hasar riski analizi için uygulamaya girilen değerler ve alınan sonuçlar

Girdiler			Çıktılar		
Parametre	Değer	Birim	Durum	Sonuç	Birim
<i>Yer titreşimi analizi</i>					
Gecikme başı maksimum şarj	62,106	kg			
Mesafe	149	m		1,763	mm/s
Saha iletim sabiti	340				
Saha jeolojik sabiti	1,79				
<i>Kaya savrulması analizi</i>					
Delik çapı	89	mm	Ortalama parça boyutu	0,231	m
Delik eğimi	90	°	En kötü durum senaryosu	599,806	m
Dilim kalınlığı	2,956	m	Ölçekli gömme derinliği	45,383	m
Kaya sabiti	25		Basamak aynası yönünde	31,979	m
Sıkılama boyu	2,956	m	Sıkılama yönünde	63,957	m
Şarj boyu	6,036	m			
Şarj miktarı	31,053	kg			
Patlayıcı yoğunluğu	0,8	g/cm ³			
<i>Hava şoku analizi</i>					
Mesafe	149	m	Hava basıncı	1,099	mBar
Gecikme başı maksimum şarj	62,106	kg	Hava basıncı	134,923	dB
Saha iletim sabiti	37,1		Etki seviyesi	Acı başlangıcı	
Saha jeolojik sabiti	0,97				

Çizelge 4. Parça boyut dağılımı analizi için uygulamaya girilen değerler ve alınan sonuçlar

Girdiler			Çıktılar		
Parametre	Değer	Birim	Parametre	Sonuç	Birim
<i>Kayaç bilgileri</i>			Kayaç faktörü	6,717	
Özgül ağırlık	2.497	t/m ³	Ort. parça boyutu	19,191	cm
Elastisite modülü	40,00	GPa	Üniformite katsayısı	1,807	
Tek eksenli basınç dayanımı	47,63	MPa	Karakteristik boyut	0,235	m
Kaya kütle sınıflaması	Bloklı		Elek üstü boyutu	0,001	%
Eklem düzlem aralığı	>1.0	m	Optimum boyut	64,143	%
Eklem düzlem yönelimi	Aynaya paralel		Elek altı boyutu	35,856	%
<i>Atım bilgileri</i>			Yüzde geçen, %	Boyut, m	
Delik çapı	89	mm	0,000	0,00	
Dilim kalınlığı	2,956	m	5,919	0,05	
Sıralar arası mesafe	3,695	m	19,220	0,10	
Basamak yüksekliği	8,000	m	35,856	0,15	
Dip şarj boyu	0,000	m	52,606	0,20	
Kolon şarj boyu	6,039	m	67,287	0,25	
Delik düzeni	Şaşırtmalı/şeşbeş		78,845	0,30	
<i>Patlayıcı bilgileri</i>			87,154	0,35	
Patlayıcı ağırlıkça kuvveti	100		92,661	0,40	
Gecikme başı maksimum şarj	62,106	kg	97,994	0,50	
<i>Parçalanma dağılım hedefleri</i>			99,563	0,60	
Elek üstü boyutu	0,90	m	99,971	0,75	
Optimum boyut	0,45	m	99,996	0,90	
Elek altı boyutu	0,15	m	100,000	0,95	

5. KARŞILAŞTIRMA VE SONUÇLAR

5.1. Mevcut Uygulamaların Karşılaştırılması

Bu çalışmada ele alınan konu ile ilgili daha önce çeşitli patlayıcı ve/veya yazılım firmaları tarafından geliştirilmiş mobil uygulamalar bulunmaktadır. Bu uygulamaların listesi ve özellikleri Çizelge 5'te

sunulmuştur. Bu bölümde, bu çalışmada bahsedilen uygulama ile daha önce geliştirilmiş olan uygulamaların özellikleri karşılaştırılmıştır.

Çizelge 5: Benzer amaçla daha önce geliştirilmiş olan mobil uygulamalar

Uygulama adı	Uygulama özellikleri	Yayın durumu	Kaynak
O-Pit Blast Guide	Patlatma tasarım parametrelerinin her biri farklı bir sayfada hesaplanmaktadır. Yer titreşimi ve aşırı hava basıncı tahminleri yapılabilmektedir. Kayaç savrulması ve parça boyut dağılım analizine yer vermemektedir.	AP + GPS	[40]
Pocket Blast Guide Mobile App	Delik başı şarj miktarı, özgül şarj, patlatılacak toplam kayaç hacmi, yer titreşimi hesaplamaları gibi birkaç temel özelliğe sahiptir.	AP + GPS	[41]
Explosives Engineers' Mobile App	Birkaç temel patlatma tasarım parametresi hesabı yapılabilmektedir. Yer titreşimi ve aşırı hava basıncı tahmini yapılabilmektedir.	AP + GPS	[42]
Nelson Brothers App	Planlanan basamak patlatma tasarım parametreleri uygulamaya girildiğinde toplam delgi miktarı, patlatılacak kayaç hacmi gibi sonuçlar üretmekte, planlanan tasarımda düzeltmeler önermekte ve en yüksek parçacık hızı, ölçekli mesafe gibi sonuçları üretmektedir.	AP + GPS	[43]
BME Guide	Birden fazla patlatma tasarım parametresinin tek sayfada hesaplanabildiği bir modüle ve bu parametrelerin ayrı ayrı hesaplanabildiği ek modüllere sahiptir. Yer titreşimi tahmini yapılabilmektedir.	AP + GPS	[44]

*AP: AppStore, GPS: Google Play Store

Benzer amaçla daha önce geliştirilen Smart Blasting App [45] adlı uygulama yayından kaldırılmıştır. Sakcalı ve arkadaşları [46] tarafından geliştirilen uygulama ise hiçbir uygulama marketinde yayınlanmamıştır. Bu uygulamalar karşılaştırmaya dahil edilememiştir. Çalışma kapsamındaki uygulama ile, benzer amaçla daha önce geliştirilmiş diğer uygulamalar karşılaştırıldığında elde edilen sonuçlar şu şekildedir;

- Geliştirilen uygulama diğer uygulamaların aksine patlatma tasarımı, çevresel etki tahminleri ve parçalanma oranı analizlerini tek bir uygulamada birleştirmiştir. Bu hesaplamaların her birini farklı modüller içinde bütünleşik olarak yapmaktadır. Diğer uygulamaların birçoğunda bu özelliklerin biri/birkaçı mevcut değildir ya da her bir modüldeki parametreler ayrı ayrı hesaplanmaktadır. Uygulama, diğer uygulamalardan farklı olarak kaya savrulması tahminlerinde birden fazla senaryoya göre hesaplamalar yapılabilmektedir. Parça boyut dağılımı analizinde ise sonuçları tablo ve grafik üzerinde sunabilmektedir.
- Geliştirilen uygulamada diğer uygulamaların aksine patlatma tasarımı aşaması sırasında patlatma deliklerindeki su varlığı durumuna göre kolon ya da kolon ve dip şarj yöntemleri uygulanarak patlatma tasarımı yapılabilmektedir. Ayrıca şarj miktarlarının çevresel etkiler açısından izin verilen limitleri aşması durumunda bir delikte bir ara sıkılama uygulaması ile şarj miktarını bölerek iki ayrı patlatma gerçekleştirilmesine imkân tanıyan çift kademeli patlatma tasarımlarının yapılabilmesini sağlamaktadır.
- Uygulama, çeşitli veri tabanları ile bağlantılıdır. Yer titreşimi ve hava şoku veri tabanları sayesinde arazi iletim ve jeolojik sabitleri istatistiksel olarak belirlenene kadar daha güvenli patlatma tasarımlarının yapılabilmesi amaçlanmıştır. Uygulamanın patlayıcı bilgi veri tabanında ise, diğer uygulamaların aksine Türkiye’de ve dünyada faaliyet gösteren birçok patlayıcı firmasının ürün bilgilerine yer vermektedir. Patlayıcı firmaları tarafından geliştirilen uygulamalarda sadece kendi ürünlerine yer vermişlerdir.
- Geliştirilen uygulama, kullanıcıya tüm sonuçların bir rapor halinde sunulmasını sağlamaktadır. Görsel ve metinsel olarak tüm hesaplamalar tek bir rapor üzerinde kaydedilebilmekte ve paylaşılabilmektedir. Bu özellik diğer uygulamalar tarafından desteklenmemektedir.

Diğer uygulamalarda yer alan metrik/İngiliz ölçü sistemi seçenekleri, patlatma literatür dokümantasyonu, bazı patlatma tasarım parametreleri ve hesaplama sonuçları (sıkılama malzemesinin optimum tane boyutları gibi) özelliklerine sahip değildir.

5.2. Sonuçlar ve Gelecek Çalışmalar

Bu çalışmada, kullanıcıdan alınan saha bilgileri, patlayıcı bilgileri ve patlatma geometrisi ile ilgili parametreleri kullanarak patlatma tasarımı yapabilen, parçalanma oranları ve çevresel etki (yer titreşimi, hava şoku, kaya savrulması) analizleri sunabilen, elde edilen tüm verileri rapor halinde PDF formatında

kullanıcıya sunabilen bir mobil uygulama geliştirilmiştir. Dart programlama dili ve Flutter mimarisi kullanılarak inşa edilen iRockBlast adlı bu uygulama hem iOS hem de Android işletim sistemlerine sahip cihazlarda çalışmaktadır.

Uygulamanın güvenilirliğini tespit etmek amacıyla, örnek bir saha çalışması üzerinden geliştirilen uygulamanın tasarım ve analiz hesaplamalarından elde edilen sonuçlar hem literatürdeki bilgilerle hem de gerçek saha verileriyle karşılaştırılmıştır. Bu kapsamda, uygulama tarafından üretilen çıktılar ile sahada yapılan ölçümler karşılaştırmalı olarak analiz edilmiştir. Saha çalışmasında, sahada fiilen uygulanan tasarım parametreleri temel alınarak, uygulamanın ürettiği tasarım, çevresel etki ve yığın parça boyut dağılımı sonuçları karşılaştırılmıştır. Tasarım aşamasında elde edilen bulgular, tasarım parametrelerine ilişkin uygulama çıktılarının saha verileri ile büyük ölçüde örtüştüğünü göstermektedir. Gözlemlenen küçük farklılıkların ise, planlamacıların kişisel deneyimlerini tasarım sürecine dâhil etmelerinden kaynaklandığı değerlendirilmiştir. Çevresel etki analizlerine ilişkin sonuçlar da genel anlamda saha verileriyle yüksek uyum göstermiştir. Gözlemlenen küçük miktardaki farklılıkların, sahadaki jeolojik heterojenite, patlatma anındaki atmosferik koşullar gibi değişkenlerden etkilendiği düşünülmektedir. Parçalanma analizine ilişkin karşılaştırmalarda ise, uygulamadan elde edilen sonuçlar ile araştırmacılar tarafından görüntü işleme teknikleri kullanılarak yapılan analiz arasında genel bir uyum gözlenmiştir. Bununla birlikte, görüntü işleme teknikleri ile yapılan analizlerde yalnızca yığının yüzey kısmı analiz edildiğinden, küçük farklılıkların ortaya çıkması doğal kabul edilmektedir. Genel bir değerlendirme yapıldığında, uygulamanın ürettiği çıktıların hem saha verileri hem de literatürdeki bilgilerle yüksek derecede tutarlılık gösterdiği ve saha koşullarına uyum sağladığı sonucuna varılmıştır.

Mevcut durumda geliştirilen uygulama, benzer mobil çözümlerden farklı olarak tek bir platformda bütünleşik patlatma tasarımı, çevresel etki analizi ve parçalanma analizi modüllerini daha gelişmiş bir biçimde sunmaktadır. Getirdiği yenilikçi çözümler ile patlatmanın tasarım ve analiz aşamalarında kullanıcı dostu bir yapı sergilemektedir. Patlatma operasyonları; yüksek maliyetleri, çevresel etkileri ve parçalanma derecelerinin sonraki proselere olan etkisi nedeniyle titizlikle planlanmalıdır. Basamak patlatma tasarımlarının hazırlanması, çevresel etki analizlerinin yapılması ve parça boyut dağılımının öngörülmesi zaman alıcı ve karmaşık işlemler gerektirmektedir. Güçlü bir literatür altyapısı üzerine inşa edilen bu uygulama, planlamacıların bu süreçlerdeki zaman kayıplarını azaltmakta ve karar verme süreçlerini hızlandırmaktadır. Tasarım parametrelerinin anlık olarak değiştirilebilmesi ve sonuçların eşzamanlı hesaplanabilmesi, olası tasarım hatalarının minimize edilmesini sağlayarak operasyonların daha verimli yürütülmesine katkı sunmaktadır. İyi planlanmış bir patlatma tasarımı yalnızca patlatma aşamasını değil, aynı zamanda yükleme, nakliye, kırma ve öğütme gibi sonrasında gelen tüm madencilik işlemlerini doğrudan etkilemekte; işlem maliyetlerini azaltarak toplam işletme verimliliğini artırmaktadır. Dolayısıyla geliştirilen bu ve benzeri mobil uygulamaların yaygınlaşması, patlatma planlamacılarının sahada tasarım ve analiz süreçlerini anlık olarak gerçekleştirebilmesine olanak tanıyacak; böylece yalnızca teknik doğruluk değil, aynı zamanda maliyet, zaman ve güvenlik açısından da somut faydalar sağlayacaktır.

İlerleyen dönemlerde, geliştirilen bu mobil uygulamanın kapsamının genişletilmesi planlanmaktadır. Özellikle yeraltı ve tünel patlatmaları için tasarım modüllerinin eklenmesi kullanılabilirliği artıracaktır. Ayrıca, patlatma maliyet analizini gerçekleştirebilecek yeni modüllerin entegrasyonu ile ekonomik değerlendirme süreçlerinin desteklenmesi hedeflenmektedir. Çevresel etkilerin daha doğru modellenebilmesi için kullanılan yer titreşimi ve hava şoku yayılım denklemlerine ait sabitlerin bulunduğu veri tabanlarının veri çeşitliliğinin artırılması da planlanan geliştirmeler arasındadır. Bunun yanında, IoT sensör entegrasyonu ile sahadan gerçek zamanlı verilerin toplanması, uygulamanın yalnızca tahmine dayalı değil, ölçüm verilerine dayalı bir karar destek sistemi olarak kullanılabilmesine olanak sağlayacaktır. Yapay zekâ ve optimizasyon algoritmalarının sisteme entegre edilmesiyle birlikte tasarım parametrelerinin otomatik optimize edilmesi hem güvenlik hem de maliyet açısından daha etkin çözümler sunabilecektir. Mobil platformlara bu yeni teknolojilerin entegrasyonunun sağlanması ile uygulamanın, patlatma mühendisliği alanında hem yüzey hem de yeraltı koşullarına kullanılabilen, kapsamlı ve akıllı bir platform haline getirilmesi öngörülmektedir.

6. KAYNAKLAR

1. Ceylanoğlu, A., Kahrıman, A. ve Demirci, A. (1993). Delme patlatmanın önemi, kullanıldığı alanlar ve maden mühendisliği ile ilgisi. *1. Delme ve Patlatma Sempozyumu*, Ankara, Türkiye.

2. MacKenzie, A.S. (1966). Cost of Explosives – Do You Evaluate it Properly? *Mining Congress Journal*, 52(5), 32-41.
3. ISEE (International Society of Explosives Engineers), (2015). Field practice guidelines for blasting seismographs. *International Society of Explosives Engineers*, Cleveland, USA.
4. Jimeno, C.L., Jimeno, E.L. & Carcedo, F.J.A. (1995). *Drilling and Blasting of Rocks*. Ramiro, Y.V. (Çev.), A.A. Balkema Publishers, Brookfield.
5. Karadoğan, A. (2008). Patlatmadan kaynaklanan titreşimler için ulusal yapı hasar kriterlerinin oluşturulabilirliğinin araştırılması. *Doktora tezi*, İstanbul Üniversitesi Fen Bilimleri Enstitüsü, 187.
6. Kalaycı Şahinoğlu, Ü. (2024). Patlatma sonucu meydana gelen hava şoku ve partikül madde yayılımı arasındaki ilişkinin araştırılması. *Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi*, 39(3), 831-837.
7. Ercins, S. ve Tosun, A. (2022). Bir taş ocağındaki patlatma uygulamalarına ait en yüksek parçacık hızı ile frekans ilişkisi. *Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi*, 37(3), 827-834.
8. Olofsson, S.O. (1990). *Applied Explosives Technology for Construction and Mining*. 2nd Ed., Applex, Arla, Sweden.
9. Konya, C.J. & Walter, E.J. (1990). *Surface Blast Design*. Prentice Hall, New Jersey.
10. Duvall, W.I. & Fogelson, D.E. (1962). *Review of Criteria for Estimating Damage to Residences from Blasting Vibrations*. U.S. Department of the Interior, Bureau of Mines, 19.
11. Lundborg, N., Persson, A., Ladegaard-Pedersen, A. & Holmberg, R. (1975). Keeping the lid on flyrock in open-pit blasting. *Engineering & Mining Journal*, 176, 95-100.
12. Richards, A. & Moore, A. (2004). Flyrock control-by chance or design. *The Annual Conference on Explosives and Blasting Technique*, 1, ISEE.
13. McKenzie, C.K. (2009). Flyrock range and fragment size prediction. *35th Annual Conference on Explosives and Blasting Technique*, 2, ISEE, Denver, CO.
14. Siskind, D.E., Stagg, M.S., Kopp, J.W. & Dowding, C.H. (1980). Structure response and damage produced by ground vibration from surface mine blasting. *Rapor*, RI 8507, Bureau of Mines, 118.
15. Cunningham, C.V.B. (1983). The Kuz-Ram model for prediction of fragmentation from blasting. *1st International Symposium on Rock Fragmentation by Blasting*, Lulea, Sweden.
16. Ullah, S., Ren, G., Ge, Y., Fisssha, Y., Kinyua, E.M. & Zhang, L. (2025). Machine learning-based prediction of blast-induced ground vibration in open-pit mining. *Journal of Vibration Engineering & Technologies*, 13, 293.
17. Fisssha, Y., Khatti, J., Ikeda, H., Grover, K.S., Owada, N., Toriya, H., Adachi, T. & Kawamura, Y. (2024). Predicting ground vibration during rock blasting using relevance vector machine improved with dual kernels and metaheuristic algorithms. *Scientific Reports*, 14, 20026.
18. Chen, L., Fisssha, Y., Hasanipanah, M., Ghodhbani, R., Dehghani, H. & Khatti, J. (2025). Accurate prediction of blast-induced ground vibration intensity using optimized machine learning models. *Defence Technology*, 1-15.
19. Fattahi, H. ve Hasanipanah, M. (2021). Prediction of blast-induced ground vibration in a mine using relevance vector regression optimized by metaheuristic algorithms. *Natural Resources Research*, 30, 1849-1863.
20. Mohamed, M.T. (2009). Artificial neural network for prediction and control of blasting vibrations in Assiut (Egypt) limestone quarry. *International Journal of Rock Mechanics and Mining Sciences*, 46(2), 426-431.
21. Temeng, V.A., Ziggah, Y.Y. & Arthur, C.K. (2021). Blast-induced noise level prediction model based on brain inspired emotional neural network. *Journal of Sustainable Mining*, 20.
22. He, Z., Jahed Armaghani, D., Masoumnezhad, M., Khandelwal, M., Zhou, J. & Murlidhar, B.R. (2021). A combination of expert-based system and advanced decision-tree algorithms to predict air-overpressure resulting from quarry blasting. *Natural Resources Research*, 30, 1889-1903.
23. Ramesh Murlidhar, B., Yazdani Bejarbaneh, B., Armaghani, D.J., Mohammed, A.S. & Mohamad, E.T. (2021). Application of tree-based predictive models to forecast air overpressure induced by mine blasting. *Natural Resources Research*, 30(2), 1865-1887.
24. Gao, W., Alqahtani, A.S., Mubarakali, A., Mavaluru, D. & Khalafi, S. (2020). Developing an innovative soft computing scheme for prediction of air overpressure resulting from mine blasting using GMDH optimized by GA. *Engineering with Computers*, 36, 647-654.
25. Nguyen, H., Bui, X.N., Bui, H.B. & Mai, N.L. (2020). A comparative study of artificial neural networks in predicting blast-induced air-blast overpressure at Deo Nai open-pit coal mine, Vietnam. *Neural Computing and Applications*, 32(8), 3939-3955.

26. Hasanipanah, M., Jahed Armaghani, D., Khamesi, H., Bakhshandeh Amnieh, H. & Ghoraba, S. (2016). Several non-linear models in estimating air-overpressure resulting from mine blasting. *Engineering with Computers*, 32, 441-455.
27. Trivedi, R., Singh, T.N. & Gupta, N. (2015). Prediction of blast-induced flyrock in opencast mines using ANN and ANFIS. *Geotechnical and Geological Engineering*, 33(4), 875-891.
28. Monjezi, M., Bahrami, A., Varjani, A.Y. & Sayadi, A.R. (2011). Prediction and controlling of flyrock in blasting operation using artificial neural network. *Arabian Journal of Geosciences*, 4(3), 421-425.
29. Han, H., Armaghani, D.J., Tarinejad, R., Zhou, J. & Tahir, M.M. (2020). Random forest and bayesian network techniques for probabilistic prediction of flyrock induced by blasting in quarry sites. *Natural Resources Research*, 29(2), 655-667.
30. Barkhordari, M.S., Armaghani, D.J. & Fakharian, P. (2022). Ensemble machine learning models for prediction of flyrock due to quarry blasting. *International Journal of Environmental Science and Technology*, 19(9), 8661-8676.
31. Hosseini, S., Poormirzaee, R., Hajihassani, M. & Kalatehjari, R. (2022). An ANN-fuzzy cognitive map-based Z-number theory to predict flyrock induced by blasting in open-pit mines. *Rock Mechanics and Rock Engineering*, 55(7), 4373-4390.
32. Kulatilake, P.H.S.W., Hudaverdi, T. & Wu, Q. (2012). New prediction models for mean particle size in rock blast fragmentation. *Geotechnical and Geological Engineering*, 30(3), 665-684.
33. Bahrami, A., Monjezi, M., Goshtasbi, K. & Ghazvinian A. (2011). Prediction of rock fragmentation due to blasting using artificial neural network. *Engineering with Computers*, 27, 177-181.
34. Miao, Y., Zhang, Y., Wu, D., Li, K., Yan, X. & Lin, J. (2021). Rock fragmentation size distribution prediction and blasting parameter optimization based on the muck-pile model. *Mining, Metallurgy & Exploration*, 38(2), 1071-1080.
35. Xie, C., Nguyen, H., Bui, X.N., Choi, Y., Zhou, J. & Nguyen-Trang, T. (2021). Predicting rock size distribution in mine blasting using various novel soft computing models based on meta-heuristics and machine learning algorithms. *Geoscience Frontiers*, 12(3), 101108.
36. Li, E., Yang, F., Ren, M., Zhang, X., Zhou, J. & Khandelwal, M. (2021). Prediction of blasting mean fragment size using support vector regression combined with five optimization algorithms. *Journal of Rock Mechanics and Geotechnical Engineering*, 13(6), 1380-1397.
37. Aruna, M., Vardhan, H., Tripathi, A. K., Parida, S., Reddy, N.V.R.S., Sivalingam, K.M., Yingqiu, L. & Elumalai, P.V. (2025). Enhancing safety in surface mine blasting operations with IoT based ground vibration monitoring and prediction system integrated with machine learning. *Scientific Reports*, 15, 3999.
38. Kadem, B. (2024). Bölgesel patlatma tasarımı ve analizi yapabilen bir mobil uygulama yazılımının geliştirilmesi. *Yüksek Lisans Tezi*, İstanbul Üniversitesi-Cerrahpaşa Lisansüstü Eğitim Enstitüsü, 117.
39. Özyurt, M.C., Arslan, A., Mutlu, H.E. ve Odabaşı, T.C. (2024). *Trakya bölgesinde gerçekleştirilen patlatmalı kazı çalışmalarının verimlilik açısından değerlendirilmesi*. Sonuç Raporu, TÜBİTAK 2209/A Üniversite Öğrencileri Araştırma Projeleri Destek Programı.
40. O-Pit Blast, (2024). *O-Pit Blast Guide uygulaması tanıtım sayfası*. <https://www.o-pitblast.com/products/o-pitblasting-guide> Erişim tarihi: 11.05.2024.
41. Orica-Nitro, (2024). *Pocket Blast Guide uygulaması tanıtım sayfası*. https://www.oricaminin gservices.com/au/en/page/products_and_services/pocket_blast_guide_app/pocket_blast_guide_mobile_app, Erişim tarihi: 12.05.2024.
42. Dyno-Nobel, (2024). *Explosives Engineers' Mobile App uygulaması tanıtım sayfası*. <https://dynonobel.com/practical-innovations/recent-innovation/mobile-app>, Erişim tarihi: 11.05.2024.
43. Nelson Brothers, (2024). *Nelson Brothers uygulaması Google Play store mağaza sayfası*. https://play.google.com/store/apps/details?id=com.nelsonbrothers.mobile&hl=en_US&pli=1 Erişim tarihi: 12.05.2024.
44. Bulk Mining Explosive, 2025. *BME Guide uygulama tanıtım sayfası*. <http://bme.co.za/blast-alliance/blasting-guide-app/>, Erişim tarihi: 27.04.2025.
45. MineExcellence, (2024). *Smart Blasting App uygulaması tanıtım sayfası*. <https://www.mineexcellence.com/blasting-mobile-app-for-mining/>, Erişim tarihi: 11.05.2024.
46. Sakcalı, A., Yavuz, H. ve Cevizci, H. (2016). Basamak patlatmasında kullanılmak üzere geliştirilen bir android uygulaması. 8. *Uluslararası Kırmataş Sempozyumu*, 13-14 Ekim 2016, Kütahya, Türkiye.



Thermophysiological Insights Into Madaline®-Based Multilayer Textiles For Medical Protection

Pelin ALTAY^{1,a}

¹*Istanbul Technical University, Faculty of Textile Technologies and Design, Textile Engineering Department, İstanbul, Türkiye*

^a**ORCID:** 0000-0000-0001-7888-9477

Article Info

Received : 03.09.2025

Accepted : 26.09.2025

DOI: 10.21605/cukurovaumfd.1777344

Corresponding Author

Pelin ALTAY

pelinaltay@itu.edu.tr

Keywords

Madaline® fabric

Multilayer textiles

Medical protective clothing

Thermophysiological comfort

Water vapour resistance (RET)

How to cite: ALTAY, P., (2025). Thermophysiological Insights Into Madaline®-Based Multilayer Textiles For Medical Protection. Çukurova University, Journal of the Faculty of Engineering, 40(3), 743-754.

ABSTRACT

Healthcare workers often experience thermal discomfort when using multilayer protective clothing, as barrier protection is often prioritized over comfort. This study evaluated a Madaline®-based multilayer textile system developed for medical applications, consisting of an outer Madaline® layer, two polyurethane membranes, and a quilted Nomex® Comfort liner. Results showed that adding membranes increased thermal resistance, with Membrane B performing slightly better than Membrane A. The inclusion of the quilted liner further enhanced insulation (65.43 mK·m²/W) but reduced vapour permeability (≈21–23%) and increased RET (14.2–15.8 Pa·m²/W), placing both three-layer systems in the moderate comfort range. The Membrane B + liner assembly is more suitable for high-risk environments requiring maximum protection, while the Membrane A + liner offers a better balance of comfort and safety for moderate-risk or extended-wear conditions.

Tıbbi Koruyucu Giysilerde Kullanılan Madaline® Tabanlı Çok Katmanlı Tekstillerin Termofizyolojik Analizi

Makale Bilgileri

Geliş : 03.09.2025

Kabul : 26.09.2025

DOI: 10.21605/cukurovaumfd.1777344

Sorumlu Yazar

Pelin ALTAY

pelinaltay@itu.edu.tr

Anahtar Kelimeler

Madaline® kumaş

Çok katmanlı tekstiller

Tıbbi koruyucu giysi

Termofizyolojik konfor

Su buharı direnci (RET)

Atf şekli: ALTAY, P., (2025). Tıbbi Koruyucu Giysilerde Kullanılan Madaline® Tabanlı Çok Katmanlı Tekstillerin Termofizyolojik Analizi. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 40(3), 743-754.

ÖZ

Sağlık çalışanları, çok katmanlı koruyucu giysiler kullanırken, genellikle koruyucu bariyer performansının konforun önünde tutulması nedeniyle termal rahatsızlık yaşamaktadır. Bu çalışma, tıbbi uygulamalar için geliştirilen, dış Madaline® katmanı, iki poliüretan membran ve kapitone Nomex® Comfort astardan oluşan çok katmanlı bir tekstil sistemini değerlendirmiştir. Bulgular, membran eklemenin termal direnci artırdığını ve Membran B'nin Membran A'ya göre biraz daha iyi performans gösterdiğini ortaya koymuştur. Kapitone astarın dahil edilmesi, yalıtımı daha da artırmış (65.43 mK·m²/W), ancak su buharı geçirgenliğini azaltmış (≈%21–23) ve RET değerini yükseltmiştir (14.2–15.8 Pa·m²/W). Bu durum, her iki üç katmanlı sistemi de orta konfor kategorisine yerleştirmiştir. Membran B + astar kombinasyonu, maksimum koruma gerektiren yüksek riskli ortamlar için daha uygunken, Membran A + astar kombinasyonu orta riskli veya uzun süreli kullanım koşullarında konfor ve güvenlik arasında daha iyi bir denge sağlamaktadır.

1. INTRODUCTION

Healthcare professionals frequently operate in thermally challenging environments where physiological comfort, moisture management, and adequate protection against biological hazards are essential for sustaining performance, safety, and well-being. Work in healthcare settings, such as wards, operating rooms, or emergency environments, often requires wearing multilayered clothing systems that prioritize barrier protection over thermal comfort, leading to increased thermal strain and discomfort [1-3]. Such discomfort has been shown to reduce perceived work performance, impair concentration, and increase the risk of heat-related illnesses, particularly under conditions of high physical activity or limited environmental control [4-8].

Protective medical clothing is designed primarily to act as a barrier against pathogens and hazardous fluids; however, these protective properties frequently come at the expense of breathability and moisture vapor permeability [9]. Previous research has shown that surgical gowns, scrub suits, and cleanroom apparel differ considerably in their thermal insulation and evaporative resistance values, directly affecting thermophysiological responses and subjective comfort of the wearer [10-13]. Studies have emphasized the necessity of balancing thermal insulation with vapor transmission to prevent excessive heat and moisture accumulation within the clothing microclimate. Inadequate moisture management can increase skin wetness and perceived clamminess, which in turn accelerates heat strain and impairs comfort [9-10].

A growing body of empirical work has evaluated thermal comfort of healthcare workers across diverse environments and garment systems. For example, Wang et al. demonstrated that medical protective clothing in hot and humid conditions significantly elevates skin temperature, sweat accumulation, and cardiovascular strain, particularly during high-intensity tasks [14]. Abreu, et al. compared the thermophysiological behavior of single-use scrub suits and demonstrated significant differences in thermal insulation, highlighting how ensemble design directly affects comfort [15]. Derks et al. showed that staff in hospital wards often report “slightly warm” sensations that lower work performance [1], while Khodakarami & Knight found that measured thermal comfort conditions in Iranian hospitals were often unacceptable compared with international standards [16].

Recent research highlights the growing importance of advanced textile solutions for managing thermal comfort in protective and functional clothing. Studies on personal thermal management garments show that localized cooling strategies and adaptive garment designs can effectively address body-specific comfort needs, improving both physiological stability and wearer perception [17]. Similarly, investigations into cold intolerance emphasize how innovative textile engineering and multilayer constructions enhance heat retention and user comfort in challenging conditions [18]. Research focused on high-thermal comfort wearables for sports underscores the relevance of material selection, coatings, and system-level textile design in enhancing breathability and heat dissipation [19]. Within medical contexts, pilot studies assessing disposable protective clothing confirm that safe working durations are limited by physiological heat strain, highlighting the need for improved thermal management in such garments [20]. Experimental investigations have also demonstrated that personal protective equipment substantially elevates heat stress in hot environments, affecting both physiological responses and subjective comfort [21]. To address these challenges, targeted interventions such as water-absorbing resins and condensation–dehumidification technologies have been proposed, showing measurable reductions in heat burden and discomfort [22]. Finally, recent reviews synthesize advances in thermoregulatory clothing, noting that dual-mode and responsive textile architectures offer promising pathways toward integrating protection with enhanced comfort [23]. These studies demonstrate that innovation in textile engineering from sportswear applications to medical protective clothing provides essential strategies for achieving both thermal comfort and protective functionality in healthcare and other high-intensity settings.

Recent reviews emphasize that advances in nanomaterials, phase-change composites, and multilayer textile architectures can significantly enhance thermoregulation without compromising protective performance [24]. Similarly, research on smart clothing systems has demonstrated how responsive fabric technologies can actively support personal thermal management, offering potential applications in healthcare apparel

[25]. Hygroscopic and adaptive fibers can play a key role in buffering rapid changes in humidity and temperature, thereby improving thermophysiological comfort during and after activity. For instance, Peng et al. introduced an integrated cooling (i-Cool) textile that combines heat conduction and sweat transportation, significantly enhancing perspiration management and thermal comfort [26]. Similarly, Chai et al. demonstrated that thermoregulatory clothing incorporating temperature-adaptive, multimodal body heat regulation can dynamically adjust to thermal loads, reducing heat stress while maintaining comfort [27]. Together, these studies confirm that advanced fiber engineering and textile architectures provide promising strategies to balance moisture management, thermal regulation, and comfort in protective clothing systems.

Phase-change inserts and active cooling systems, such as ice vests and forced-air devices, have also demonstrated effectiveness in reducing heat strain for healthcare personnel [28]. Experimental results demonstrate that integrating phase-change cold storage into protective apparel leads to marked improvements in both physiological and perceptual comfort [29]. Beyond garment engineering, several field investigations have highlighted how contextual factors, such as hospital ventilation systems, ward zoning, and local climate, affect comfort. For instance, Udom found that healthcare workers in remote hot-arid clinics often faced conditions outside comfort limits, showing the need for climate-specific comfort standards [30]. Likewise, indoor studies in Iran [31] revealed that standard PMV (Predicted Mean Vote)-based indices often underestimate discomfort in hospital environments due to clothing and activity-specific effects. These finding highlights the importance of integrating both environmental and clothing parameters into comfort evaluation.

Given these complexities, standardized and quantitative evaluation methods are essential. Tools such as the sweating guarded hot plate [ISO 11092] and thermal manikins [ASTM F1291, F2370] have been widely applied to measure thermal resistance (R_{ct}), evaporative resistance (R_{et}), and permeability index across single and multilayer systems [32-33]. Studies using thermal manikins to assess surgical gowns, scrub suits, and protective ensembles [28,32] provide objective metrics that, when combined with wearer trials, yield valuable insights into how fabric density, membrane type, and air layer thickness influence thermal and moisture transfer.

In light of these considerations, the present study focuses on the development and evaluation of a Madaline®-based multilayer textile system specifically designed for medical protective applications. The system integrates an outer Madaline® fabric with inherent liquid repellency, two types of polyurethane membranes with different densities, and a quilted Nomex® thermal liner. This configuration was selected to optimize the trade-off between thermal comfort and protective performance by enhancing moisture vapor transmission, regulating heat flow, and ensuring effective liquid barrier protection. By systematically assessing thermophysiological parameters such as thermal conductivity, resistance, and water vapor permeability, this study aims to provide new insights into how Madaline®-based multilayer assemblies can inform future standards for healthcare garment design and improve occupational comfort for medical personnel working under demanding conditions. While previous studies have explored thermal comfort and barrier properties of protective clothing in general, this work uniquely investigates the combined performance of Madaline® fabric with polyurethane membranes and a Nomex® Comfort liner under standardized ISO 11092 testing. This integrated approach offers a novel perspective on optimizing both comfort and protection in medical apparel.

2. MATERIALS AND METHODS

2.1. Materials

In this study, a multilayer textile system was developed to ensure user comfort, moisture regulation, and protection against biological fluids. The system consists of three distinct layers: a functional outer protective layer, a middle membrane layer for moisture management, and an inner quilted layer for thermal insulation.

Table 1 summarizes the detailed characteristics of the multilayer textile system components used in this study, including key parameters such as material type, basis weight, thickness, and the functional role of each layer. Figure 1 provides a visual representation of the multilayer textile system components, showing the arrangement of the layers within the system.

Outer Layer (Functional Protective Layer):

The outermost fabric used was Madaline®, a nonwoven microfiber textile provided by Mogul Nonwoven, composed of 70% polyester (PET) and 30% polyamide 6 (PA6), with a basis weight of 130 g/m². Produced using bi-component spunbond technology, this fabric combines softness, strength, and high dimensional stability. It was treated to be blood- and alcohol-repellent, making it well-suited for use in protective clothing intended for medical or hazardous settings.

Madaline® has a dense microfilament structure, which provides effective barrier and filtration properties while remaining breathable and quick-drying. It also shows good moisture management, thermal insulation, and wind resistance. In addition, its fabric-like handling characteristics allow for standard textile processes such as dyeing, cutting, and sewing, making it versatile for garment manufacturing.

Middle Layer (Moisture Management):

To promote moisture vapor transfer and improve overall physiological comfort, two types of polyurethane (PU) membranes were utilized in separate configurations. Membrane A is polyurethane membrane with a basis weight of 145 g/m², offering enhanced flexibility and improved moisture permeability. Membrane B is polyurethane membrane with a basis weight of 859 g/m², providing a more robust barrier effect and structural support.

These membranes were placed between the outer and inner layers in various combinations to assess their influence on comfort, moisture transport, and garment breathability.

Inner Layer (Thermal Insulation):

The innermost layer consisted of a two-layer nonwoven structure (55 g/m² + 55 g/m²) quilted to a Nomex® Comfort fabric, forming the inner lining of the system. This multilayer configuration was selected to provide thermal insulation, flame resistance, and wearer comfort. Nomex® Comfort offers inherent heat and flame protection while maintaining breathability and a soft textile feel, which enhances user comfort during extended wear. The quilted construction improves air permeability, contributes to thermal regulation, and ensures mechanical stability, making the layer suitable for high-performance protective garments.

Table 1. Characteristics of the multilayer textile system components

	Fabric type	Basis weight (g/m²)	Thickness (mm)	Function
Outer layer	Madaline® fabric (70% PET/ 30% PA6, bi-component spunbond nonwoven)	130	0.51 (± 0.02)	Blood and alcohol repellent barrier; primary defense against biological fluids and contaminants
Membrane A	Polyurethane (PU) lightweight membrane	145	0.31 (± 0.004)	Enhances moisture vapor permeability
Membrane B	Polyurethane (PU) dense membrane	859	0.48 (± 0.004)	Enhances moisture vapor permeability
Thermal liner	Two-layer nonwoven structure quilted to Nomex® Comfort	55+ 55	1.73 (± 0.05)	Provides thermal insulation, flame resistance, and wearer comfort

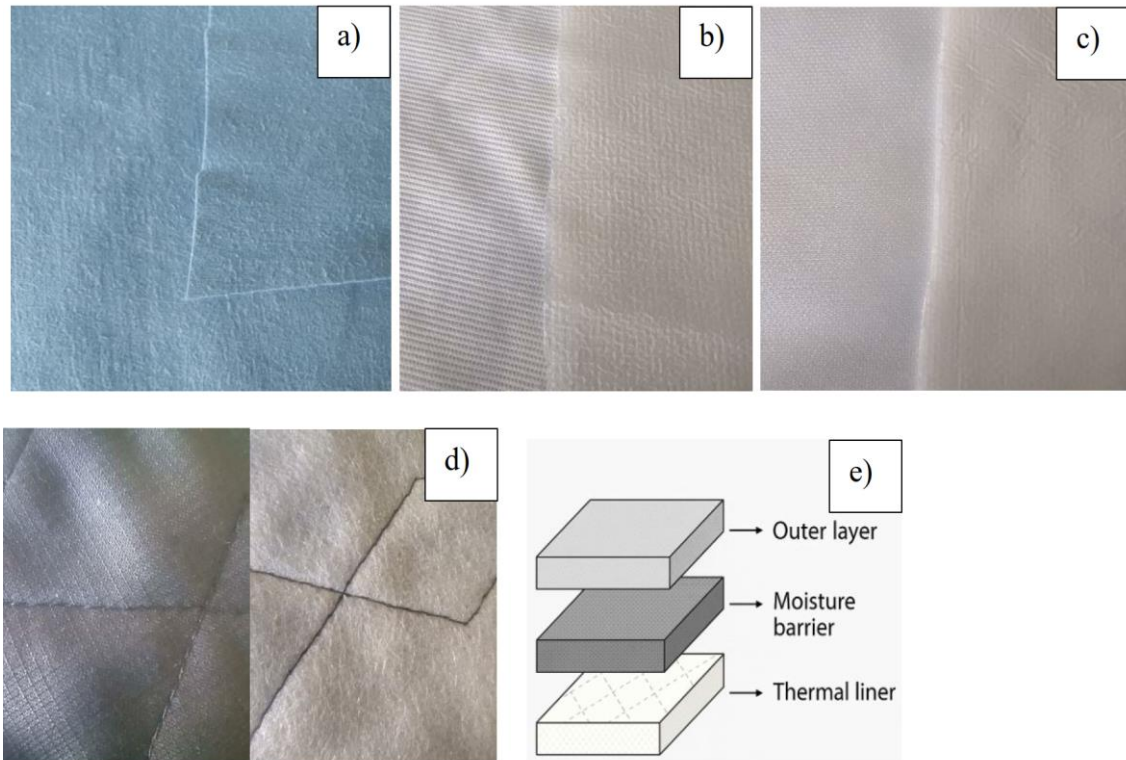


Figure 1. Visual representation of the multilayer textile components used in the study: Face and back surfaces of a) Outer layer (b) Membrane A (c) Membrane B (d) Thermal liner, and e) Placement/stacking sequence of layers

2.2. Method

The thermal properties of the fabric samples and their multilayer assemblies were evaluated in accordance with the ISO 11092 standard [34], which specifies the determination of thermal conductivity (λ), thermal absorptivity (α), thermal diffusivity (a), and thermal resistance (r).

The Alambeta instrument (Sensora, Czech Republic) was employed to precisely assess thermal characteristics, while the Permetest device (Sensora, Czech Republic) was used to measure water vapour permeability and evaporative resistance (RET), simulating skin perspiration under steady-state conditions. These evaluations provided key thermophysiological parameters essential for understanding the performance of different textile configurations. The calculations were based on standardized formulas and widely reported methodologies in the literature [35], ensuring reliable and consistent results across all fabric systems tested.

3. RESULTS AND DISCUSSION

3.1. Evaluation of Thermal Properties

The results of the thermal performance measurements are presented in Table 2, which summarizes the thermal conductivity, diffusivity, absorption, resistance, thickness, and maximum heat flux values for each fabric configuration. All measurements were conducted in accordance with related standard and average values are accompanied by standard deviations to reflect measurement repeatability and material uniformity. In addition, Figure 2 specifically illustrates the thermal conductivity and thermal resistance values for the same fabric configurations, offering a complementary graphical representation of these two key parameters.

Table 2. Thermal performance parameters of the fabric combinations, Including thermal conductivity (La), Thermal diffusion (a), Thermal absorption (b), Thermal resistance (r), Thickness (h), and Maximum heat flux (q_{max}) values measured for each sample configuration

Sample group	La (mW/m.K)	a (mm ² /s)	b (W.s ^{1/2} /m ² .K)	r (mK.m ² /W)	h (mm)	q _{max} (W/m ²)
Outer layer	34.4 ± 1.21	0.05 ± 0.02	154.50 ± 22.16	14.77 ± 0.21	0.51 ± 0.02	674.53 ± 34.39
Membrane A	33.63 ± 1.88	0.053 ± 0.02	150.46 ± 19.8	9.4 ± 0.43	0.31 ± 0.004	810.1 ± 4.05
Membrane B	40.53 ± 1.16	0.048 ± 0.02	191.4 ± 33.4	11.9 ± 0.25	0.48 ± 0.004	790.6 ± 52.3
Thermal layer	34.7 ± 2.76	0.128 ± 0.008	97.1 ± 5.83	50.03 ± 5.6	1.73 ± 0.05	489.4 ± 8.07
Outer layer + Membrane A	41.03 ± 1.06	0.08 ± 0.06	131.67 ± 21.97	21.40 ± 0.70	0.87 ± 0.02	663.93 ± 50.86
Outer layer + Membrane A + Thermal layer	41.47 ± 0.98	0.12 ± 0.01	122.33 ± 4.60	61.33 ± 1.29	2.54 ± 0.05	618.07 ± 15.45
Outer layer + Membrane B	44.93 ± 0.91	0.10 ± 0.02	140.77 ± 9.73	23.07 ± 0.50	1.04 ± 0.004	689.77 ± 22.34
Outer layer + Membrane B + Thermal layer	42.6 ± 0.52	0.11 ± 0.01	129.63 ± 4.45	65.43 ± 0.81	2.79 ± 0.02	647.80 ± 52.48

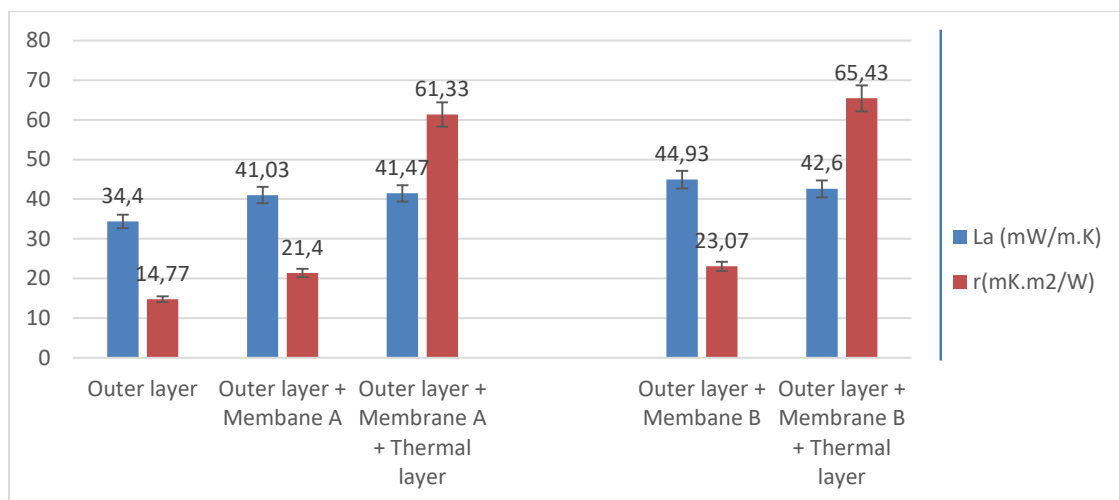


Figure 2. Thermal conductivity and thermal resistance of the fabric combinations

The outer layer alone, with the lowest basis weight of 130 g/m², exhibited the lowest thermal conductivity (34.4 ± 1.21 mW/m·K), which is expected for a single textile layer without any additional insulating or structural components. This result reflects the relatively simple construction of the fabric, which limits its ability to conduct heat and provides only a basic barrier to thermal transfer. When Membrane A (145 g/m²) was added, the thermal conductivity increased to 41.03 ± 1.06 mW/m·K. Despite being only slightly heavier than the outer layer, its denser polymer matrix offers more continuous pathways for heat transfer, increasing conduction. However, when the thermal liner (110 g/m²) was incorporated alongside Membrane A, the conductivity value (41.47 ± 0.98 mW/m·K) remained almost the same as the membrane-only configuration. The liner's quilted structure traps air within its layers, reducing heat transfer and compensating for the increase caused by the membrane [36]. This balance shows how lower-density, air-trapping components like the liner counteract the conductive effects of slightly denser materials, maintaining stability in overall conductivity [37,38].

The highest thermal conductivity was measured in the outer + Membrane B configuration (44.93 ± 0.91 mW/m·K) which can be attributed to the high density (859 g/m²) and compact microstructure of Membrane

B, facilitating more efficient heat conduction compared to other membrane type [39]. In contrast, the outer + Membrane B + thermal liner assembly exhibited a slightly lower conductivity ($42.6 \pm 0.52 \text{ mW/m}\cdot\text{K}$). This reduction is likely due to the insulating effect of the quilted Nomex® thermal liner, which introduces additional air gaps and reduces the direct conductive pathways present in the membrane-only configuration [40]. The layered construction of the liner, combined with its inherent low thermal conductivity, counteracts some of the conduction-enhancing effects of Membrane B, resulting in an overall decrease in measured conductivity [41].

Thermal diffusivity, a key measure of how quickly a material responds to temperature changes [42], was found to be $0.05 \pm 0.02 \text{ mm}^2/\text{s}$ for the outer layer and ranged between 0.08 and $0.12 \text{ mm}^2/\text{s}$ in the layered assemblies, exhibiting relatively low variation across samples. These values are consistent with typical textile diffusivity ranges and indicate that the materials respond at a stable and controlled rate when exposed to temperature fluctuations. The small standard deviations further suggest excellent repeatability and uniformity in sample structure. The diffusivity trend aligns with material density: denser layers like Membrane B react slightly faster to temperature changes due to their continuous, compact structure, while lighter layers provide more buffering effects.

Thermal absorption (b), also known as thermal absorptivity, represents a material's ability to absorb heat, especially in the initial contact phase, and is often associated with the feeling of immediate warmth or coolness upon touch. In our measurements, the outer layer alone exhibited the highest thermal absorption ($154.50 \pm 22.16 \text{ W}\cdot\text{s}^{1/2}/\text{m}^2\cdot\text{K}$), which then decreased with the addition of layers. Incorporating Membrane A reduced absorption to $131.67 \text{ W}\cdot\text{s}^{1/2}/\text{m}^2\cdot\text{K}$, and the further addition of the thermal liner brought it down to $122.33 \text{ W}\cdot\text{s}^{1/2}/\text{m}^2\cdot\text{K}$. A similar trend was observed with Membrane B combinations, where absorption dropped from $140.77 \text{ W}\cdot\text{s}^{1/2}/\text{m}^2\cdot\text{K}$ (outer + Membrane B) to $129.63 \text{ W}\cdot\text{s}^{1/2}/\text{m}^2\cdot\text{K}$ when the liner was added. This reduction is directly linked to the increasing total thickness and air entrapment in the lower-density quilted liner, which slows heat uptake and creates a warmer tactile sensation. This behavior is supported by the study reported by Sampath [43], which has demonstrated that fabrics with lower thermal absorptivity impart a warmer sensation upon contact due to slower heat uptake from the skin, typically seen in materials with thicker or more insulating structures.

For thermal absorption (b), the highest standard deviation was observed in the outer layer configuration ($\pm 22.16 \text{ W}\cdot\text{s}^{1/2}/\text{m}^2\cdot\text{K}$), indicating greater variability in heat uptake due to the fabric's unprotected, single-layer structure. The addition of Membrane A slightly reduced the standard deviation to $\pm 21.97 \text{ W}\cdot\text{s}^{1/2}/\text{m}^2\cdot\text{K}$, while the inclusion of the thermal liner alongside Membrane A significantly lowered it to $\pm 4.60 \text{ W}\cdot\text{s}^{1/2}/\text{m}^2\cdot\text{K}$. On the other hand, adding Membrane B to the outer layer reduced the standard deviation to $\pm 9.73 \text{ W}\cdot\text{s}^{1/2}/\text{m}^2\cdot\text{K}$, and combining it with the thermal liner yielded the lowest deviation of all configurations ($\pm 4.45 \text{ W}\cdot\text{s}^{1/2}/\text{m}^2\cdot\text{K}$). The substantial reduction in standard deviation when the thermal liner is incorporated can be attributed to its quilted, multi-layer structure, which increases measurement consistency by minimizing localized variations in heat absorption and reducing the influence of surface irregularities during testing [44].

Thermal resistance (r) increased progressively with material density. The outer layer (lightest, 130 g/m^2) had the lowest resistance ($14.77 \text{ mK}\cdot\text{m}^2/\text{W}$). Adding Membrane A (145 g/m^2) raised resistance to $21.40 \text{ mK}\cdot\text{m}^2/\text{W}$, while incorporating the liner significantly increased it to $61.33 \text{ mK}\cdot\text{m}^2/\text{W}$. Membrane B's dense structure produced slightly higher resistance ($23.07 \text{ mK}\cdot\text{m}^2/\text{W}$) than Membrane A. When the liner was added, the value peaked at $65.43 \text{ mK}\cdot\text{m}^2/\text{W}$. This demonstrates the strong insulating synergy between dense materials like Membrane B and lightweight, air-trapping layers like the liner, which together maximize insulation [45].

The maximum heat flux (q_{max}) values ranged from 618.07 W/m^2 to 689.77 W/m^2 across all configurations. For the outer layer, q_{max} was 674.53 W/m^2 , decreasing to 663.93 W/m^2 with the addition of Membrane A, and further to 618.07 W/m^2 when the thermal liner was incorporated, reflecting a progressive reduction in heat transfer capacity as insulation layers were added. In contrast, combining the outer layer with Membrane B resulted in the highest q_{max} value of 689.77 W/m^2 , slightly exceeding that of the outer + Membrane A combination. When the thermal liner was added to this configuration, q_{max} decreased to 647.80 W/m^2 , but remained higher than the equivalent Membrane A+liner assembly. This trend aligns with previous findings showing that thermal resistance and maximum heat flux, while often inversely related, are not always directly proportional. Thermal resistance describes a material's overall ability to impede steady-state heat

transfer over time, which improves with additional layers, trapped air, and low-conductivity components. In contrast, $q_{\max}$ represents the peak instantaneous heat transfer rate and is strongly influenced by surface density, intrinsic material conductivity, and contact characteristics. This occurs because $q_{\max}$ reflects the transient surface heat transfer upon contact, which depends heavily on surface density and thermal absorptivity, not just insulation [46,47]. Membrane B's dense and compact microstructure enhances barrier performance and increases thermal resistance but also provides more direct conductive pathways at the surface, resulting in higher peak heat transfer. The addition of the quilted Nomex® thermal liner mitigates this effect by introducing insulating air gaps and reducing both conductive and convective heat flow; however, the influence of Membrane B's higher intrinsic conductivity remains evident in the heat flux measurements.

3.2. Evaluation of Water Vapour Permeability & Water Vapour Resistance Properties

Figure 3 summarises the water vapour permeability and resistance (RET) values for each fabric configuration.

The results of the water vapour permeability and resistance (RET) measurements provide important insights into the balance between barrier protection and physiological comfort in multilayer textile systems. The outer layer alone exhibited the highest water vapour permeability (72.9%) and the lowest resistance to water vapour transfer ($RET = 1.7 \text{ Pa} \cdot \text{m}^2/\text{W}$). According to the ISO 11092 classification of comfort ranges, values below $6 \text{ Pa} \cdot \text{m}^2/\text{W}$ are considered to reflect *very good to excellent breathability* [34]. This confirms that the outer layer, with its relatively simple structure and absence of additional membranes or liners, provides minimal restriction to vapour transfer and is highly suitable for maintaining comfort during prolonged use.

Incorporating Membrane A reduced permeability to 44.4% and increased RET to $5.4 \text{ Pa} \cdot \text{m}^2/\text{W}$, reflecting the membrane's barrier effect, which restricts moisture diffusion through the fabric system. Nevertheless, its ability to maintain RET below 6 suggests that Membrane A still supports adequate moisture management while enhancing fluid protection. When the thermal liner was added alongside Membrane A, permeability further decreased to 21.2% and RET increased substantially to $15.8 \text{ Pa} \cdot \text{m}^2/\text{W}$. This classifies the assembly in the *satisfactory but limited comfort category* (13–20) [34]. The increase reflects the influence of the quilted Nomex® liner, which adds bulk, increases thickness, and introduces more tortuous diffusion paths for vapour. Consequently, although this configuration offers improved thermal insulation and protective performance, it significantly compromises breathability, which may elevate wearer discomfort during extended periods of activity.

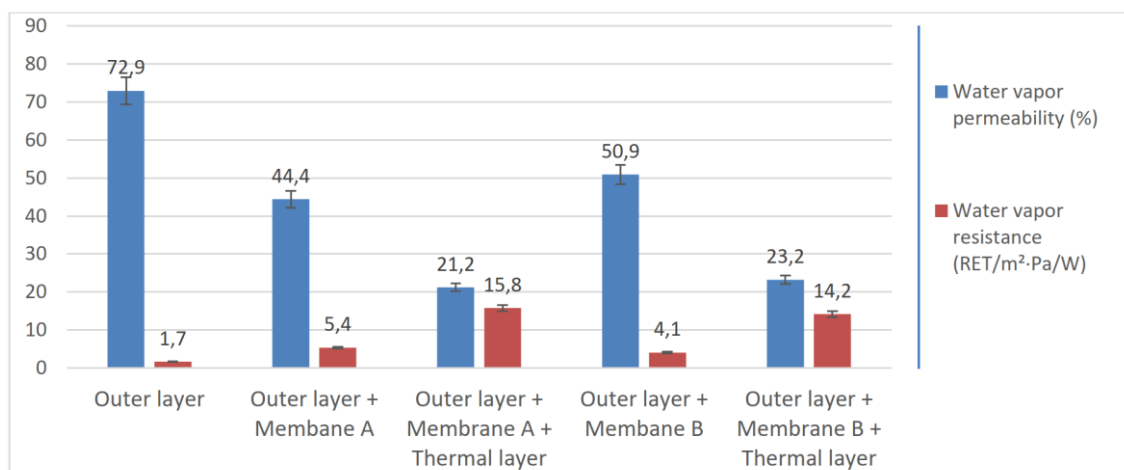


Figure 3. Water vapour permeability and resistance of the fabric combinations

The outer layer + Membrane B configuration demonstrated a water vapour permeability of 50.9% with a corresponding RET value of $4.1 \text{ Pa} \cdot \text{m}^2/\text{W}$, classifying it within the “very good breathability” category ($RET < 6$) according to ISO 11092 standards [34, 48]. The addition of the quilted thermal liner markedly reduced permeability to 23.2% and increased RET to $14.2 \text{ Pa} \cdot \text{m}^2/\text{W}$, which corresponds to “satisfactory but limited comfort” (Ret value of 13–20). A similar decline was observed in the Membrane A combinations, where

permeability dropped from 44.4% ($RET = 5.4 \text{ Pa} \cdot \text{m}^2/\text{W}$) to 21.2% ($RET = 15.8 \text{ Pa} \cdot \text{m}^2/\text{W}$) with the inclusion of the thermal liner. However, across both two- and three-layer configurations, Membrane B assemblies consistently showed slightly higher vapour permeability and lower RET values compared with Membrane A. This indicates that, although Membrane B is considerably heavier (859 g/m^2 vs. 145 g/m^2) and denser, its structural characteristics, likely related to its pore distribution and microstructural compactness, permit marginally greater vapour diffusion than Membrane A. Previous textile studies have reported that membrane morphology, including pore size and connectivity, exerts a dominant influence on moisture vapour transfer, sometimes overriding thickness effects [49-50]. The dense but potentially more interconnected microstructure of Membrane B may therefore facilitate limited vapour transmission while still ensuring a high barrier effect against liquids.

In both membrane types, the addition of the thermal liner significantly reduced permeability and increased RET. This effect is due to the liner's quilted nonwoven Nomex® structure, which increases thickness, introduces additional air gaps, and creates more tortuous diffusion paths, thereby impeding vapour transfer [48,51].

These results emphasize the inherent trade-off in multilayer protective textiles. While the integration of membranes and quilted liners improves thermal insulation and protective capacity, it also restricts moisture transmission and increases evaporative resistance. This dual effect highlights the importance of careful design optimization to achieve adequate protection while maintaining acceptable levels of wearer comfort in medical and other high-risk settings.

4. CONCLUSION

This study examined the thermophysiological performance of Madaline®-based multilayer textile systems for medical protective clothing, emphasizing the balance between thermal insulation, liquid barrier protection, and wearer comfort. The multilayer system consisted of a Madaline® outer layer, two polyurethane membranes with different densities, and a quilted Nomex® thermal liner. Results demonstrated that incorporating the membranes significantly improved thermal resistance and barrier performance compared to the outer layer alone. Notably, the outer + Membrane B + liner configuration provided the highest insulation ($65.43 \text{ mK} \cdot \text{m}^2/\text{W}$) and superior barrier protection. However, this improvement was accompanied by a marked decrease in water vapor permeability (23.2%) and an increase in evaporative resistance ($RET = 14.2 \text{ Pa} \cdot \text{m}^2/\text{W}$), placing it in the moderate comfort range according to ISO 11092 standards.

While the Membrane B + liner assembly offers maximum thermal protection and liquid barrier performance, making it highly suitable for high-risk medical applications such as surgical procedures or emergency response duties in contaminated environments, it compromises breathability during extended wear. In contrast, the outer + Membrane A + liner system demonstrated slightly lower insulation ($61.33 \text{ mK} \cdot \text{m}^2/\text{W}$) but provided a more favorable balance between comfort and protection, making it better suited for moderate-risk healthcare settings or long-duration activities, particularly in warmer climates.

These findings emphasize the inherent trade-off between protection and comfort in multilayer protective textiles. They also highlight the importance of selecting fabric configurations based on specific healthcare contexts, environmental conditions, and activity levels. Ultimately, this work provides practical guidance for optimizing Madaline®-based protective garments, supporting the development of future standards aimed at improving both safety and occupational comfort for healthcare professionals.

5. REFERENCES

1. Derks, M.T.H., Mishra, A.K., Loomans, M.G.L.C. & Kort, H.S.M. (2018). Understanding thermal comfort perception of nurses in a hospital ward work environment. *Building and Environment*, 140, 119-127.
2. Zhao, Y., Su, M., Meng, X., Liu, J. & Wang, F. (2023). Thermophysiological and perceptual responses of amateur healthcare workers: Impacts of ambient condition, inner-garment insulation and personal cooling strategy. *International Journal of Environmental Research and Public Health*, 20, 612.

3. Wibowo, R., Satow, M., Quartucci, C., Weinmann, T., Koller, D., Daanen, H.A.M., Nowak, D., Bose-O'Reilly, S. & Rakete, S. (2025). Impact of heat stress and protective clothing on healthcare workers: Health, performance, and well-being in hospital settings. *Annals of Work Exposures and Health*, 69(6), 665-675.
4. Eryürük, S.H. (2019). Effect of fabric layers on thermal comfort properties of multilayered thermal protective fabrics. *Autex Research Journal*, 19(3), 271-278.
5. Bröde, P., Fiala, D., Błażejczyk, K., Holmér, I., Jendritzky, G., Kampmann, B., Tinz, B. & Havenith, G. (2012). Deriving the operational procedure for the Universal Thermal Climate Index (UTCI). *International Journal of Biometeorology*, 56, 481-494.
6. McLellan, T.M. & Daanen, H.A.M. (2012). Heat strain in personal protective clothing: Challenges and intervention strategies. In: Kiekens, P., Jayaraman, S. (eds) *Intelligent Textiles and Clothing for Ballistic and NBC Protection*. NATO Science for Peace and Security Series B: Physics and Biophysics. Springer, Dordrecht.
7. Sarker, M.E. & Mezarcıöz, S. (2023). The effects of several washings on some of the comfort features of denim fabrics made of cotton and coolmax weft yarns with and without elastane. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 38(4), 1151-1159.
8. Şenkal Sezer, F. (2016). Sağlık ocaklarında konfor koşullarının değerlendirilmesi: Bursa/Nilüfer örneği. *Çukurova Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 30(1), 197-208.
9. Lou, L., Chen, K. & Fan, J. (2021). Advanced materials for personal thermal and moisture management of healthcare workers wearing PPE. *Materials Science and Engineering: R: Reports*, 146, 100639.
10. Konečná, K. & Adamovský, D. (2025). Investigation of cleanroom clothing thermal insulation by thermal manikin and impact on thermal comfort of cleanroom users. *Journal of Industrial Textiles*, 55, 1-19.
11. Roskotová, K. & Adamovský, D. (2021). Thermal insulation of clothing: Assessment of cleanroom clothing ensembles. In: *Healthy Buildings 2021 – Europe Proceedings*, 308-314.
12. Eryürük, S.H. (2021). Analyzing thermophysiological comfort and moisture management behavior of cotton denim fabrics. *Autex Research Journal*, 21(2), 248-254.
13. McCullough, E.A. (2005). The use of thermal manikins to evaluate clothing and environmental factors. In: Tochihara, Y., Ohnaka, T. (eds) *Environmental Ergonomics*. Elsevier Ergonomics Book Series, 3, 403-407.
14. Wang, F., Guo, W., Tian, Y., Liu, X., Pang, D., Lian, Z., Deng, C., Li, J. & Zhang, J. (2025). Thermal comfort of medical protective clothing under high temperature and high humidity. *Building and Environment*, 270, 112570.
15. Abreu, M.J., Abreu, I. & Ribeiro, P. (2014). Thermo-physiological behavior of single use scrub suits using a thermal manikin. *2nd International Congress on Healthcare and Medical Textiles*, September 25-26, Izmir, Turkey.
16. Khodakarami, J. & Knight, I. (2008). Required and current thermal conditions for occupants in Iranian hospitals. *HVAC&R Research*, 14(2), 175-193.
17. Tian, M., Qi, N., Jiang, Q., Su, Y. & Li, J. (2024). Addressing localized thermal comfort needs of the human body through advanced personal thermal management garments: Design and evaluation. *Textile Research Journal*, 95(3-4), 429-449.
18. Abuhay, A., Tadesse, M.G., Berhanu, B., Malengier, B. & Van Langenhove, L. (2025). Advancements in clothing thermal comfort for cold intolerance. *Fibers*, 13(2), 13.
19. Orjuela-Garzón, I.C., Rodríguez, C.F., Cruz, J.C. & Briceño, J.C. (2024). High-thermal comfort wearables for sports: Design, materials, and performance. *ACS Omega*, 9(50), 49143-49162.
20. Li, J., Cui, X., Huang, Q. & Li, J. (2024). Determining safe working hours of wearing medical disposable protective clothing from physiological thermal limits: A pilot study. *AATCC Journal of Research*, 11(3), 171-182.
21. Mao, Y., Zhu, Y., Guo, Z., Zheng, Z., Fang, Z. & Chen, X. (2022). Experimental investigation of the effects of personal protective equipment on thermal comfort in hot environments. *Building and Environment*, 222, 109352.
22. Jiang, H., Cao, B. & Zhu, Y. (2023). Improving thermal comfort of individual wearing medical protective clothing: Two personal cooling strategies integrated with the polymer water-absorbing resin material. *Building and Environment*, 243, 110730.
23. Lei, L., Shi, S., Wang, D., Meng, S., Dai, J. G., Fu, S. & Hu, J. (2023). Recent advances in thermoregulatory clothing: Materials, mechanisms, and perspectives. *ACS Nano*, 17(3), 1803-1830.
24. Tang, K.H.D. (2025). Advances in thermoregulating textiles: Materials, mechanisms, and applications. *Textiles*, 5(2), 22.

25. Zhang, Q., Cheng, H., Zhang, S., Li, Y., Li, Z., Ma, J. & Liu, X. (2024). Advancements and challenges in thermoregulating textiles: Smart clothing for enhanced personal thermal management. *Chemical Engineering Journal*, 488, 151040.
26. Peng, Y., Li, W., Liu, B., Jin, W., Schaadt, J., Tang, J., Zhou, G., Wang, G., Zhou, J., Zhang, C., Zhu, Y., Huang, W., Wu, T., Goodson, K.E., Dames, C., Prasher, R., Fan, S. & Cui, Y. (2021). Integrated cooling (i-Cool) textile of heat conduction and sweat transportation for personal perspiration management. *Nature Communications*, 12, 6122.
27. Chai, J., Kang, Z., Yan, Y., Lou, L., Zhou, Y. & Fan, J. (2022). Thermoregulatory clothing with temperature-adaptive multimodal body heat regulation. *Cell Reports Physical Science*, 3(7), 100958.
28. Zhou, Y., Lou, L. & Fan, J. (2023). Quantitative comparison of personal cooling garments in performance and design: A review. *Processes*, 11, 2976.
29. Cao, W., Zhang, X., Yao, W., Ruan, M., Zhang, Q., Cao, Z. & Shi, F. (2024). Experimental study on the cooling performance of medical protective clothing coupled with phase-change cold storage materials. *Applied Thermal Engineering*, 243, 122607.
30. Udom, S.U. (2019). Exploring thermal comfort band for healthcare workers in remote clinics in hot and arid climates: An approach for building performance improvement. In: *Proceedings of the 16th IBPSA Conference*, Rome, Italy, Sept. 2-4, 2019.
31. Khodakarami, J., Knight, I. (2007). Measured thermal comfort conditions in Iranian hospitals for patients and staff. In: *Proceedings of Clima 2007 WellBeing Indoors Conference*.
32. Ochoa, N., Giraldo, S.M., Angel, B.E. & Escobar, N.J. (2013). Development of a method for measuring thermal comfort in medical clothing through the hot plate apparatus. In: *2013 Pan American Health Care Exchanges (PAHCE)*, Medellin, Colombia, 1-6.
33. Kim, J.-H., Powell, J.B., Roberge, R.J., Shepherd, A. & Coca, A. (2014). Evaluation of protective ensemble thermal characteristics through sweating hot plate, sweating thermal manikin, and human tests. *Journal of Occupational and Environmental Hygiene*, 11(4), 259-267.
34. ISO 11092:2014. Textiles - Physiological effects - Measurement of thermal and water-vapour resistance under steady-state conditions (sweating guarded-hotplate test). *International Organization for Standardization*.
35. Romeli, D., Barigozzi, G., Esposito, S., Rosace, G. & Salesi, G. (2013). High sensitivity measurements of thermal properties of textile fabrics. *Polymer Testing*, 32(6), 1029-1036.
36. Li, R., Dolez, P.I., Lai, A., Gholamreza, F., Allen, S., Gathercole, R. & Li, R. (2025). Heat transfer through wavy clothing layers with varied permeability. *Building and Environment*, 280, 113114.
37. Firšt Rogale, S., Rogale, D., Knezić, Ž., Jukl, N. (2021). Measurement method for the simultaneous determination of thermal resistance and temperature gradients in the determination of thermal properties of textile material layers. *Materials (Basel)*, 14(22), 6853.
38. Rui, K., He, J., Xin, M., Chen, Z. & Guan, J. (2024). Effects of air gap and compression on the dual performance of multilayer thermal protective clothing under low radiant heat. *Journal of Industrial Textiles*, 54.
39. Mortazavinejad, S.M., Alakhdar, M., Vinches, L., Hallé, S. (2025). A novel methodology for calculating thermal conductivity of natural hollow fibers with validation in nonwoven fabric structures. *International Communications in Heat and Mass Transfer*, 167(A), 109269.
40. Wang, Y., Lu, Y., Li, J. & Pan, J. (2012). Effects of air gap entrapped in multilayer fabrics and moisture on thermal protective performance. *Fibers and Polymers*, 13, 647-652.
41. Das, T., Das, A. & Alagirusamy, R. (2022). Study on thermal protective performance of thermal liner in a multi-layer clothing under radiant heat exposure. *Journal of Industrial Textiles*, 51(5_suppl), 8208S-8226S.
42. Romeli, D., Barigozzi, G., Esposito, S., Rosace, G. & Salesi, G. (2013). High sensitivity measurements of thermal properties of textile fabrics. *Polymer Testing*, 32(6), 1029-1036.
43. Sampath, M., Aruputharaj, A., Senthilkumar, M. & Nalankilli, G. (2011). Analysis of thermal comfort characteristics of moisture management finished knitted fabrics made from different yarns. *Journal of Industrial Textiles*, 42(1), 19-33.
44. Jukl, N., Firšt Rogale, S. & Rogale, D. (2023). The influence of compressibility on the thermal contact conductivity of diamond-shaped quilted lining for special purpose clothing. *Tekstil*, 72(3), Online first.
45. Wilson, C.A., Laing, R.M. & Carr, D.J. (2002). Air and air spaces-the invisible addition to thermal resistance. *Journal of the Human Environmental System*, 5(2), 69-77.
46. Weedall, P.J. & Goldie, L. (2001). The objective measurement of the 'cool feeling' in fabrics. *Journal of The Textile Institute*, 92(1, Part 4), 379-386.

47. Li, Y. (2001). The science of clothing comfort. *Textile Progress*, 31(1-2), 1-135.
48. Havenith, G., Holmér, I. & Parsons, K. (2002). Personal factors in thermal comfort assessment: Clothing properties and metabolic heat production. *Energy and Buildings*, 34(6), 581-591.
49. Knížek, R., Tunák, M., Tunáková, V. & Honzíková, P. (2024). Effect of membrane morphology on the thermo-physiological comfort of outdoor clothing. *Journal of Engineered Fibers and Fabrics*, 19.
50. Khakpour, A., Gibbons, M. & Chandra, S. (2021). Effect of moisture condensation on vapour transmission through porous membranes. *Journal of Industrial Textiles*, 51(2_suppl), 1931S-1951S.
51. He, H. & Yu, Z. (2018). Effect of air gap entrapped in firefighter protective clothing on thermal resistance and evaporative resistance. *Autex Research Journal*, 18(1), 28-34.



Çukurova Üniversitesi Mühendislik Fakültesi Dergisi Yazım Kuralları

Çukurova Üniversitesi Mühendislik Fakültesi Dergisi mühendislik bilimlerindeki güncel bilimsel ve teknik gelişmeleri araştırmacılara ve mühendislere ulaştırmayı hedeflemektedir. Tüm mühendislik alanları ile ilgili teorik/deneysel ve uygulamaya yönelik çalışmalar Dergi'nin ilgi alanındadır. Çukurova Üniversitesi Mühendislik Fakültesi Dergisi daha önce başka hiçbir yerde yayımlanmamış, özgün ve güncel bilimsel araştırma ve uygulamaları içeren makale ve kısa bildirileri ile Dergi'de daha önce yayımlanmış makalelerle ilgili tartışmaları kabul etmektedir. Çeviri ve derleme makaleler kabul edilmemektedir. Makaleler Türkçe veya İngilizce olabilir. Dergi'ye gönderilen yazılar yayımlansın veya yayımlanmasın iade edilmez. Makaleleri yayınlama yetkisi dergi yayın kuruluna aittir.

Makalenin tamamı bilgisayar ortamında hazırlanmalı baskıya hazır biçimde iki adet (birisinde yazar adları ve adresleri olmamak üzere) **Dergi web sayfasında verilen örnek şablona uygun hazırlanmış olarak elektronik ortamda veya e-posta ekinde gönderilmelidir.**

Makale, A4 (210x297 mm) boyutlu kağıda, Windows uyumlu Word ortamında, Times New Roman font kullanılarak, birer aralıkla yazılmalı, ana başlıklar 12 Punto (koyu), yazar adları 12 Punto (koyu), alt başlıklar ve metin 10 punto olmalıdır.

Çizimler ve çizelgelerle birlikte, makaleler 15 sayfadan, kısa bildiriler 4 sayfadan daha uzun olmamalıdır. **Makale Türkçe ise: Türkçe Başlık, İngilizce Başlık, Öz, Anahtar Kelimeler, Abstract** (İngilizce öz), **Keywords, Metin, Semboller** (gerekliyse), **Teşekkür** (gerekliyse) ve **Kaynaklar**,

Makale İngilizce ise: İngilizce Başlık, Türkçe Başlık, Abstract, Keywords, Öz (Türkçe öz), **Anahtar Kelimeler**, Text, Notations (gerekliyse), **Acknowledgments** (gerekliyse) ve **References** sırası içinde düzenlenmelidir.

Başlık kısa ve açık olmalı, içeriği yansıtabilmelidir. **Öz** (ve **Abstract**) çalışmanın amacını, kapsamını, yöntemini ve ulaşılan sonuçları kısaca tanımlamalı ve 150 kelimeyi aşmamalıdır. En az üç, en fazla beş tane anahtar kelime verilmelidir. Başlık, Öz ve İngilizce başlıklı Abstract birinci sayfaya sığdırılmalı ve ana metin ikinci sayfadan başlatılmalıdır. **Metin** yalın bir dil ve anlatımla yazılmalı, makale Türkçe ise Türkçe, makale İngilizce ise İngilizce yazım kurallarına uygun olmalıdır.

Bölüm ve alt bölüm başlıkları numaralanmalıdır. Eşitlikler (denklemler) numaralanmalı ve bu numaralar satır sonunda parantez içinde gösterilmelidir. Yazılarda yalnızca SI birimleri kullanılmalıdır. Semboller uluslararası kullanıma uygun seçilmeli; her bir sembol ilk kullanıldığı yerde tanımlanmalıdır. Teşekkür olabildiğince kısa olmalı, çalışmaya katkısı ve desteği bulunan kişi ve kuruluşlar belirtilmelidir. Çizelgeler, çizimler ve fotoğraflar metin içine yerleştirilmeli, her birine numara ve başlık verilmeli ve başlıklar çizim ve fotoğrafların altına, çizelgelerin üstüne yazılmalıdır.

Kaynaklar metinde köşeli parantez içinde numaralanmalı ve kaynaklar listesinde metin içinde veriliş sırasına uygun biçimde belirtilmelidir. Etik kuralları gereğince, alıntılar tırnak içinde verilmeli ve bir referans numarası ile kaynak belirtilmelidir. Kaynaklarda şu bilgiler verilmelidir:

Makale ise: Yazar(lar)ın soyadı, adının baş harfi., yıl. makalenin tam başlığı. derginin adı, cilt, sayı, başlama ve bitiş sayfaları.

Kitap ise: Yazar(lar)ın soyadı, adının baş harfi., yıl. kitabın tam adı. cilt numarası, varsa editörü, yayınlandığı yer, son sayfa no.

Bildiri ise: Yazar(lar)ın soyadı, adının baş harfi., yıl. bildirinin tam başlığı. konferansın adı, yapıldığı yer.

Elektronik ortamdan alınmış ise: Bildiri sadece CD/DVD'de yayınlanış ise: Yazar(lar)ın soyadı, adının baş harfi., diğer yazarlar, yıl, bildirinin adı. konferansın adı (CD/DVD), yapıldığı yer, yıl. "Konferansın adından hemen sonra parantez içerisinde CD/DVD olduğu belirtilecektir."

Tez ise: Yazarın soyadı, adının baş harfi., yıl. tezin adı. derecesi, sunulduğu kurum, son sayfa no.

Rapor ise: Yazarın soyadı, adının baş harfi., diğer yazarlar, yıl. raporun adı. türü, numarası, kuruluşun adı, yayınlandığı yer.

Web Sitesi ise: Yazar(lar) belli ise; yazarın soyadı, adının baş harfi., diğer yazarlar, varsa yayınlandığı yıl. yazının başlığı, web adresi (<http://www.xxxxx.xxx.xxx>), Erişim Tarihi: gün.ay.yıl.

Kurum/Kuruluşlara ait ise: Kurum adı, varsa yayınlandığı yıl. yazının başlığı, web adresi (<http://www.xxxxx.xxx.xxx>), Erişim Tarihi: gün.ay.yıl.

Elektronik kitap/rapor ise: Yazar(lar)ın soyadı, adının baş harfi veya kitap/raporu yayımlayan kurum veya kuruluşun adı, varsa yayınlandığı yıl, konunun tam adı, Erişim Tarihi: gün.ay.yıl.