



22nd Series by MATDER

ICMME²⁰²⁵

INTERNATIONAL CONFERENCE ON MATHEMATICS
AND MATHEMATICS EDUCATION

MATHEMATICS
MATHEMATICS EDUCATION
COMPUTER SCIENCE
MATHEMATICAL PHYSICS
ENGINEERING MATHEMATICS
MATHEMATICS AND ART EXHIBITION

PROCEEDINGS BOOK

Mathematics in Istanbul, Bridge Between Continents

SEPTEMBER 11-13, 2025



ISTANBUL MEDENIYET UNIVERSITY

Istanbul/Türkiye

www.theicmme.org | www.matder.org.tr | www.medeniyet.edu.tr



*International Conference on Mathematics and
Mathematics Education (ICMME-2025), İstanbul Medeniyet
University, İstanbul, Türkiye, September 11-13, 2025*
"Mathematics in İstanbul, Bridge Between Continents"



International Conference on Mathematics and Mathematics Education
(ICMME - 2025)

İstanbul Medeniyet University, İstanbul, Turkey, 11-13 September 2025

©All rights reserved. Throughout this book, or part, without the permission of the author, electronic, mechanical, photocopying or reproduced in any recording system, transmitted, stored.

All kinds of information in this book is the responsibility of the authors.

Editors

Nurettin DOĞAN

Selçuk University, Turkey

İsmail EKİNCİOĞLU

İstanbul Medeniyet University, Turkey

Assistant Editor

Bilal İNAN

Sakarya University of Applied Sciences, Turkey



PREFACE

The International Conference on Mathematics and Mathematics Education (ICMME-2025) was held in İstanbul, Türkiye, from September 11-13, 2025. The conference brought together researchers, educators, and practitioners to exchange ideas, present recent research findings, and discuss future challenges in mathematics and related disciplines. The scientific program included invited plenary lectures and short oral presentations, providing a comprehensive overview of current developments in the field.

MATDER – Mathematicians Association was established in 1995 by mathematicians in Türkiye. Since its foundation, MATDER has organized 14 national and 7 international mathematics symposia. Over the years, these meetings have become major scientific events at both national and international levels. Covering a wide range of topics in pure and applied mathematics, mathematics education, engineering mathematics, and computer science, the ICMME conferences have consistently attracted significant interest from mathematicians and engineers in academia.

The previous ICMME conferences were held in Nevşehir (ICMME-2024), Denizli (ICMME-2022), Ankara (ICMME-2021), Konya (ICMME-2019), Ordu (ICMME-2018), Şanlıurfa (ICMME-2017), and Elazığ (ICMME-2016). This year, ICMME-2025 was hosted by İstanbul Medeniyet University, İstanbul, Türkiye, September 11–13, 2025, as an international scientific conference.

The main objective of ICMME-2025 is to contribute to the advancement of mathematical sciences, mathematics education, computer sciences, and their applications, as well as to provide a platform for international collaboration among researchers, educators, mathematicians, statisticians, and interdisciplinary scholars from around the world.

Abstracts of all accepted presentations are included in these conference proceedings. In addition, selected and peer-reviewed full papers will be published in the following journals:

- [Advanced Studies: Euro-Tbilisi Mathematical Journal](#)
- [Turkish Journal of Mathematics and Computer Science \(TJMCS\)](#)
- [International Journal of Maps in Mathematics](#)
- [Selcuk University Journal of Engineering Sciences](#)
- [Gazi Eğitim Bilimleri Dergisi](#)
- [Advances in Differential Equations and Control Processes](#)

The ICMME-2025 conference is jointly organized by MATDER – The Mathematicians' Association and İstanbul Medeniyet University. We sincerely thank



*International Conference on Mathematics and
Mathematics Education (ICMME-2025), İstanbul Medeniyet
University, İstanbul, Türkiye, September 11-13, 2025*
"Mathematics in İstanbul, Bridge Between Continents"



all participants, authors, reviewers, and invited speakers for their valuable contributions to the success of this conference.

On Behalf of The Organizing Committee
Hasan Hüseyin SAYAN



*International Conference on Mathematics and
Mathematics Education (ICMME-2025), İstanbul Medeniyet
University, İstanbul, Türkiye, September 11-13, 2025*
"Mathematics in İstanbul, Bridge Between Continents"



COMMITTEES

HONORARY AND ADVISORY BOARD

Davut GÜL	Governor of İstanbul
Gülfettin ÇELİK	Rector of İstanbul Medeniyet University
Ömer AKIN	Honorary President of MATDER

SYMPOSIUM PRESIDENTS

H. Hüseyin SAYAN	Chairman of the Board, Türkiye
İsmail EKİNCİOĞLU	Head of Mathematics Department, İstanbul Medeniyet University, Turkey

ORGANIZING COMMITTEE

H. Hüseyin SAYAN (President)	Chairman-MATDER President, Gazi University, Türkiye
İsmail EKİNCİOĞLU (President)	Chairman, İstanbul Medeniyet University, Türkiye
Abdülhamit KÜÇÜKASLAN (Vice President)	Vice Chairman, MATDER, Ankara Yıldırım Beyazıt University, Türkiye
Nurettin DOĞAN (Vice President)	Vice Chairman, MATDER, Selçuk University, Türkiye
Mustafa ÖZKAN	MATDER, Gazi University, Türkiye
Ali ÖZTURAN	MATDER, Türkiye
Amiran GOGATISHVILI	Ins. of Math. of The Czech Acad. of Sciences, Czech Rep.
Cengiz ÇINAR	MATDER, Gazi University, Türkiye
Ercan ÇELİK	Kırgızistan-Türkiye Manas University, Kyrgyzstan
Fatma AYAZ	MATDER, Gazi University, Türkiye
Gülay KORU	Gazi University, Türkiye
Hacı Mehmet BAŞKONUŞ	Harran University, Türkiye
Hasan BULUT	Fırat University, Türkiye



*International Conference on Mathematics and
Mathematics Education (ICMME-2025), İstanbul Medeniyet
University, İstanbul, Türkiye, September 11-13, 2025*
"Mathematics in İstanbul, Bridge Between Continents"



Hüseyin DEMİR	Samsun Ondukuzmayıs University, Türkiye
Hüseyin YILDIRIM	Kahramanmaraş Sütçü İmam University, Türkiye
Jan ZEMLICKA	Univerzita Karlova, Matematicko-fyzikální fakulta: Praha, CZ
Jehad ALZABUT	Prince Sultan University, Ostim Technical University, Türkiye
Maria Alessandra RAGUSA	Università Di Catania, Italy
Matthias HEIL	School of Mathematics, University of Manchester, UK
Mohamed İbrahim ABBAS	Alexandria University, EGYPT
Mohammad Esmael SAMEI	Bu Ali Sina University: Hamedan, IR
Naim TUĞLU	Gazi University, Türkiye
Nilay Akgönüllü PİRİM	MATDER, Başkent University, Türkiye
Ömer ÖZEN	MATDER, Türkiye
Sorina BARZA	KARLSTADS Universitet, Sweeden
Tahir FEYZİOĞLU	MATDER, Türkiye
Truong Cong QUYNH	The University of Danang: Danang, Vietnam
Vagif S. Guliyev	Baku State University, Azerbaijan, Ahi Evran University, Türkiye
Yasin DEMİRCİ	MATDER, Türkiye
Yusif GASİMOV	Azerbaijan University, Azerbaijan
Zakia HAMMOUCH	Moulay Ismail University, Morocco

LOCAL ORGANIZING COMMITTEE

Ahmet KILINÇ	İstanbul Medeniyet University
Ali DİNLER	İstanbul Medeniyet University
Betül HİÇDURMAZ	İstanbul Medeniyet University
Çiğdem KILIÇ	İstanbul Medeniyet University
Cansu KESKİN	İstanbul Medeniyet University



*International Conference on Mathematics and
Mathematics Education (ICMME-2025), İstanbul Medeniyet
University, İstanbul, Türkiye, September 11-13, 2025
"Mathematics in İstanbul, Bridge Between Continents"*



Ertan SÖNMEZ	İstanbul Medeniyet University
Emine CAN	İstanbul Medeniyet University
Eyüp SEVİMLİ	İstanbul Medeniyet University
Fatih Mehmet COŞKUN	İstanbul Medeniyet University
Hayriye SUNDU PAMUK	İstanbul Medeniyet University
İlknur GÜNDÜZ AYKAÇ	İstanbul Medeniyet University
İsmail EKİNCİOĞLU	İstanbul Medeniyet University
M. Nureddin TÜRKAN	İstanbul Medeniyet University
Mustafa COŞKUN	İstanbul Medeniyet University
Mustafa TEK	İstanbul Medeniyet University
Nursena GÜNHAN AY	İstanbul Medeniyet University
Pınar OBAKAN YERLİKAYA	İstanbul Medeniyet University
Rumeysa KİNSİZ ÖZTÜRK	İstanbul Medeniyet University
Rüya ŞEN	İstanbul Medeniyet University
Sare ŞENGÜL	Marmara University, Türkiye
Sefa Anıl SEZER	İstanbul Medeniyet University
Semran İPEK	İstanbul Medeniyet University
Uğur YİĞİT	İstanbul Medeniyet University

SCIENTIFIC COMMITTEES

MATHEMATICS

Abdulhamit KÜÇÜKASLAN	Ankara Yıldırım Beyazıt University, Türkiye
Adil Güler	Marmara University
Amiran GOGATISHVILI	Institute of Mathematics of the Czech Academy of Sciences, Czech Republic
Ayhan ŞERBETÇİ	Ankara University, Türkiye
Bayram ŞAHİN	Ege University
Cenap Özel	King Abdulaziz University, Saudi Arabia
David Cruze-Urbe	University of Alabama, USA
Emine Can	İstanbul Medeniyet University
Emran ALİ	Northern University, Bangladesh



*International Conference on Mathematics and
Mathematics Education (ICMME-2025), İstanbul Medeniyet
University, İstanbul, Türkiye, September 11-13, 2025*
"Mathematics in İstanbul, Bridge Between Continents"



Erhan COŞKUN	Karadeniz Technical University, Türkiye
Fuad Kittaneh	Mathematics, University of Jordan, Amman
Hacı Mehmet BAŞKONUŞ	Harran University
Hasan BULUT	Fırat University
Hayriye SUNDU PAMUK	İstanbul Medeniyet University
İsmail EKİNCİOĞLU	İstanbul Medeniyet University, Türkiye
Jan ZEMLICKA	Charles University
Jean-Christophe	Bourin University of Franche-Comté, France
Jehad ALZABUT	Prince Sultan University
Kazım İLARSLAN	Kırıkkale University
Kemal AYDIN	Selçuk University
Lars-Erik Persson	UiT Artic University of Norway, Norway
M. Tamer KOŞAN	Gazi University
Maria Alessandra RAGUSA	Università di Catania, Italy
Martin BOHNER	Missouri University Of Science And Technology
Matthias HEIL	School of Mathematics, University of Manchester
Mustafa ÖZKAN	Gazi University
Naim TUĞLU	Gazi University
Natalia Bebiano	University of Coimbra, Spain
Natasha Samko	Arctic University of Norway, Norway
Nureddin TÜRKAN	İstanbul Medeniyet University
Oktay MUHTAROĞLU	Gaziosmanpaşa University
Rabil AYAZOĞLU	Bayburt University, Türkiye
Ryskul Oinarov	Eurasian National University, Kazakhstan
Shigeru Furuichi	Mathematical Science, Nihon University, Japan
Siegfried M. Rump	Hamburg University of Technology, Germany
Sorina BARZA	Karlstads Universitet, Sweden
Murat TOSUN	Sakarya University
Varga KALANTAROV	Koç University
Vagif S. Guliyev	Baku State University, Azerbaijan, Ahi Evran



*International Conference on Mathematics and
Mathematics Education (ICMME-2025), İstanbul Medeniyet
University, İstanbul, Türkiye, September 11-13, 2025*
"Mathematics in İstanbul, Bridge Between Continents"



Yoshihiro Sawano	University, Türkiye
Yusuf S. GASIMOV	Chuo University, Japan
Zakia HAMMOUCH	Azerbaijan University
	Moulay Ismail University

COMPUTER SCIENCES

Ayhan ERDEM	Gazi University
Ercan Nurcan YILMAZ	Gazi University
Necaattin BARIŞÇI	Gazi University
Yusuf SÖNMEZ	Gazi University
Aydın ÇETİN	Gazi University
Murat YÜCEL	Gazi University
Uğur GÜVENÇ	Gazi University
Hüseyin Polat	Gazi University
Cemal KOÇAK	Gazi University
Cemal YILMAZ	Gazi University
Hasan Hüseyin SAYAN	Gazi University
Sinan TOKLU	Gazi University
Adem TEKEREK	Gazi University
İbrahim Alper DOĞRU	Gazi University
Tuncay Yiğit Süleyman	Demirel University
Osman ÖZKARACA	Muğla Sıtkı Koçman University
Gürcan ÇETİN	Muğla Sıtkı Koçman University
Nurettin DOĞAN	Selçuk University
Murat KÖKLÜ	Selçuk University
Metin KAPIDERE	İnönü University
Fatih BAŞÇİFTÇİ	Selçuk University
İlkay ÇINAR	Selçuk University
Adem Alpaslan ALTUN	Selçuk University



*International Conference on Mathematics and
Mathematics Education (ICMME-2025), İstanbul Medeniyet
University, İstanbul, Türkiye, September 11-13, 2025*
"Mathematics in İstanbul, Bridge Between Continents"



Humar KAHRAMANLI ÖRNEK	Selçuk University
Şakir TAŞDEMİR	Sinop University
Hamdi Tolga KAHRAMAN	Karadeniz Technical University
Sabri KOÇER	Necmettin Erbakan University

MATHEMATICS EDUCATION

Ali ERASLAN	Samsun Ondokuz Mayıs University
Ayşe Tuğba ÖNER	İstanbul Medeniyet University
Cemil İNAN	Artuklu University
Cengiz ÇINAR	Gazi University
Çağlar Naci HİDİROĞLU	Pamukkale University
Damla Solmaz GEDİK	İstanbul Medeniyet University
Deniz KAYA	İstanbul Medeniyet University
Çiğdem KILIÇ	İstanbul Medeniyet University
Derya Özlem YAZLIK	İstanbul Medeniyet University
Esra YILDIZ	İstanbul Medeniyet University
Eyüp SEVİMLİ	İstanbul Medeniyet University
Hamza Yaşar OCAK	Marmara University
Hasan ES	Gazi University
Hülya GÜR	Balıkesir University
İsmet AYHAN	Pamukkale University
Melek MASAL	Sakarya University
Nazan MERSİN	İstanbul Medeniyet University
Murat ALTUN	Uludağ University
Murat PEKER	Afyon Kocatepe University
Mustafa DOĞAN	Selçuk University
Mustafa KANDEMİR	Amasya University
Necdet GÜNER	Pamukkale University
Nurullah YAZICI	Karamanoğlu Mehmetbey University



*International Conference on Mathematics and
Mathematics Education (ICMME-2025), İstanbul Medeniyet
University, İstanbul, Türkiye, September 11-13, 2025*
“Mathematics in İstanbul, Bridge Between Continents”



Osman GÜRSOY	Maltepe University
Ömer DERELİ	Necmettin Erbakan University
Ömer Faruk TEMİZER	Fırat University
Özge Yiğitcan NAYIR	Başkent University
Rabia SAVAŞ	İstanbul Medeniyet University
Safure BULUT	Middle East Technical University
Sare ŞENGÜL	Marmara University
Sibel Yeşildere İMRE	Dokuz Eylül University
Süha YILMAZ	Dokuz Eylül University
Şenol DOST	Hacettepe University
Şenol KARTAL	İstanbul Medeniyet University
Şeref MİRASYEDİOĞLU	Başkent University
Tunay BİLGİN	Van Yüzüncü Yıl University
Yasemin KATRANCI	Kocaeli University
Zaki F.A. EL-RAHEM	Alexanderia University
Zeynep Çiğdem ÖZCAN	İstanbul Medeniyet University



CONTENTS

Leveraging Generated Synthetic Data to Preserve Data Properties and Validate Clusters Using Cluster Number Optimization Technique	1
Automatic Lesion Segmentation in Dental Bitewing Radiographs: Comparative Analysis of U-Net, Attention U-Net, Mask R-CNN, and YOLOv11n-seg Models	19
Examining the Norms of the Problem Solving Process within the Framework of Practical Rationality	41
A Mathematical Comparison of SLAM Families: Observability, Information-Theoretic Bounds, and Optimization Landscapes.....	63
Modeling Problem Solving Process Interaction Analysis: A Discourse Study on the Parking Lot Design Problem	79
Schur Stability Analysis of Second-degree Extended Legendre Polynomial Families	92
Formal Graphs: Graphing via Formal Variables and Formal Evaluation	99
Fuzzy Multi-Criteria Analysis of Ship Accident Causes: A Comparative Evaluation	111
An Instructional Activity Based on Hypothetical Learning Trajectories: Modeling of Identities.....	135
A new alpha power Rayleigh Weibull distribution with its applications to simulated and real data.....	165



Leveraging Generated Synthetic Data to Preserve Data Properties and Validate Clusters Using Cluster Number Optimization Technique

Mahdi Zarei¹, Hijran G. Mirzayeva² and Zari Dzalilov³

¹ Department of Bioengineering and Therapeutic Sciences University of California, San Francisco CA, 94143-2811, USA, mehdi.nn@gmail.com

² Baku Engineering University, Department of Information Technology and Programming, Baku, Azerbaijan, himirzayeva@beu.edu.az

³ Department of Applied Science, Federation University of Australia, zdzalilo@gmail.com

ABSTRACT

In unsupervised machine learning, clustering algorithms uncover hidden patterns within datasets. Their effectiveness, however, often depends on having substantial, representative data. Small datasets, common in many fields, can hinder performance by increasing overfitting risks and limiting generalizability. This paper explores synthetic data generation as a strategy to augment small datasets, improving the robustness and accuracy of clustering algorithms. We focus on using synthetic data to identify optimal cluster numbers in unsupervised clustering. Through experiments, we show that carefully generated synthetic data, preserving the statistical properties of original small datasets, creates an enriched data environment. This environment enhances the training, validation, and testing of clustering models. We discussed methods for generating synthetic data that mimic real-world distributions, enabling better model tuning and hypothesis testing in unsupervised learning. Our findings highlight synthetic data's potential to address small dataset limitations, advancing machine learning by providing a versatile framework for clustering tasks. This study guides researchers and practitioners in using synthetic data to improve unsupervised clustering, leading to more reliable outcomes across applications.

Key Words: Unsupervised machine learning, clustering algorithms, synthetic data generation, optimal cluster numbers.



1. INTRODUCTION

In the constantly changing field of data analytics, unsupervised learning algorithms are pivotal for discovering hidden structures and patterns in datasets (Wheeldon & Serb, 2023; Johnston, Jones, & Kruger, 2019; Hrebik & Kukal, 2023). These algorithms are vital for analyzing data without predefined labels, heavily relying on the data's quantity and quality (Bishop, 2006; Hinton & Sejnowski, 1999; Witten, Frank, Hall, Pal, & Data, 2005). A significant challenge in this area is managing small datasets, which are susceptible to overfitting and often lack the robustness required for generalization, thereby affecting the effectiveness of clustering results (Hastie, Tibshirani, Friedman, & Friedman, 2009; Fahad et al., 2014). To address these challenges, this paper explores the strategic use of synthetic data generation techniques, which aim to augment the size and quality of datasets by creating artificial data points that statistically reflect the original data. The dual focus of this investigation is to examine how synthetic data can enhance dataset size to improve the performance of clustering algorithms and to refine the process for determining the optimal number of clusters in unsupervised clustering scenarios (Goodfellow, Bengio, Courville, & Bengio, 2016; Chawla, Bowyer, Hall, & Kegelmeyer, 2002). Through detailed experiments and rigorous analysis, we demonstrate that synthetic data, carefully generated to maintain the statistical properties of original datasets, not only enhances the data environment but also optimizes conditions for training, validating, and testing clustering models. This enhanced data environment is crucial for reducing biases associated with small datasets and for improving the predictive capabilities of clustering algorithms. Additionally, this paper explores various advanced techniques for synthesizing data, including iterative proportional fitting and Gibbs sampling, ensuring that the synthetic data closely replicates real-world distributions and adheres to the inherent relationships and constraints found in the original data (Chawla et al., 2002; McLachlan & Peel, 2000). The insights presented highlight the potential of synthetic data in addressing the limitations of small datasets. By providing a more adaptable analytical framework, synthetic data generation can contribute to advancements in the field of machine learning. This approach allows data scientists to perform more accurate and insightful analyses in clustering tasks, potentially leading to improvements in various



applications. In this paper, we propose incorporating synthetic data generation into the analytical toolkit of researchers and practitioners. The use of synthetic data has the potential to improve the performance of unsupervised clustering algorithms and extend their applicability. We provide a detailed overview of data synthesis techniques, underscoring their importance and the different methodologies involved. We also examine two popular clustering algorithms: Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and k-means. These algorithms are instrumental in uncovering hidden patterns and structures within datasets. To identify the optimal number of clusters, we employed the Elbow method and Within-Cluster Sum of Squares (WCSS). This comprehensive approach improved the efficacy of unsupervised learning applications.

1.1 Synthesizing data

Data synthesis integrates various data sources to create a unified dataset, providing a comprehensive and insightful perspective on specific topics or problems. Commonly used in business, research, and data analysis, this method enhances decision-making by revealing patterns and trends not apparent in isolated datasets. It increases the accuracy of findings by merging data from multiple credible sources, validating results, and identifying gaps or inconsistencies for future data collection efforts. Moreover, data synthesis simplifies data management by consolidating multiple datasets into a single source (Park et al., 2018; Evans, 2002). Analysis and interpretation of synthesized data uncover valuable insights, which are clearly communicated. In statistical modeling, data synthesis prepares data, estimates marginal and joint distributions using methods like copula models or multivariate regression, generates synthetic data respecting original data constraints, and validates this synthetic data to ensure it accurately reflects the original datasets while preserving privacy and confidentiality. This approach streamlines data analysis and enhances the reliability and applicability of the synthesized information.



1.2 UCI Machine Learning Repository: Iris dataset

The Iris dataset is a widely used benchmark in machine learning and statistics, consisting of four features: sepal length, sepal width, petal length, and petal width, for 150 iris flowers classified into three species: setosa, versicolor, and virginica. Each row represents a flower, and each column corresponds to a feature, with an additional column indicating the species. Introduced by Ronald Fisher in 1936, the Iris dataset is structured as a tabular dataset, typically stored as a data frame in R. It is extensively used for evaluating classification algorithms and data visualization techniques. Commonly used to train and evaluate classification models, the Iris dataset enables the prediction of an iris flower's species based on its measurements. Various machine learning algorithms, such as decision trees, support vector machines, and neural networks, can be applied to classify the flowers accurately. The dataset is available at the UCI Machine Learning Repository (Dua & Graff, 2017).

1.3 The Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm

The Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm (Ester, Kriegel, Sander, Xu, et al., 1996) is an unsupervised machine learning technique for clustering data based on density (Schubert, Sander, Ester, Kriegel, & Xu, 2017). It groups data points based on density and proximity, identifying clusters of arbitrary shape and effectively handling noise. Unlike other clustering algorithms, DBSCAN does not require specifying the number of clusters in advance. Key parameters include epsilon (ϵ) for neighborhood distance and MinPts for the minimum points in a cluster.

DBSCAN's objective is to identify dense regions, defined as:

$$N_{\epsilon}(x) = \{y \in R^n \mid |x - y| \leq \epsilon\} \quad (1)$$



where $N_\epsilon(x)$ is the ϵ -neighborhood of point x . A point x is a core point if $|N_\epsilon(x)| \geq \text{MinPts}$. Points not meeting this criterion are classified as noise or border points.

Applying DBSCAN with $\epsilon = 0.5$ and $\text{MinPts} = 5$ to the iris dataset clusters flowers based on their features (sepal length, sepal width, petal length, and petal width).

1.4 k-mean clustering algorithm

The K-means algorithm is a widely used unsupervised machine learning technique for clustering data into K distinct clusters based on similarity (Lloyd, 1982; Yadav & Sharma, 2013). It aims to partition a dataset by minimizing the within-cluster sum of squared distances (inertia or distortion), defined as:

$$\arg \min_S \sum_{i=1}^k \sum_{x \in S_i} |x - \mu_i|^2 \quad (2)$$

where S represents the set of clusters $S = S_1, S_2, \dots, S_k$ k is the number of clusters x is a data point in cluster S_i μ_i is the mean (centroid) of points in cluster S_i . The algorithm iteratively assigns each data point to the nearest centroid and updates the centroids based on the mean of the assigned points until convergence or a maximum number of iterations is reached.

1.5 Selecting the Optimal Number of Clusters Using the Elbow Method

To determine the optimal number of clusters in an unsupervised learning algorithm, a common approach is to calculate the within-cluster sum of squares (WCSS) for various values of (k) (the number of clusters) (Shi et al., 2021; Nanjundan, Sankaran, Arjun, & Anand, 2019). The WCSS measures the compactness or cohesion of the clusters, with a lower WCSS indicating that the data points within each cluster are closer to their respective centroid. The Elbow Method for clustering, as detailed in Algorithm 1, is a heuristic used to determine the optimal number of

clusters, K , in a dataset. The procedure begins by applying k-means clustering for a range of cluster values from 1 to a specified maximum, max_K . For each value of K , the k-means algorithm partitions the data into K clusters and computes the Within-Cluster Sum of Squares (WCSS). The WCSS is a metric that quantifies the variance within each cluster and is calculated as the sum of the squared Euclidean distances between each data point and the centroid of its respective cluster:

$$WCSS_k = \sum_{i \in C_k} \|x_i - \mu_k\|^2$$

where x_i represents a data point in cluster C_k , and μ_k denotes the centroid of cluster C_k . The total WCSS for all clusters is then accumulated in a list. After computing the WCSS values for all K values, the next step is to identify the “elbow point” in the WCSS values, which is the point where the rate of decrease in WCSS sharply slows down. This point signifies diminishing returns on adding additional clusters and is considered the optimal number of clusters for the dataset. The Elbow Method thereby aids in balancing the trade-off between the number of clusters and the variance within each cluster (Rokach & Maimon, 2005).

Algorithm 1 Apply Elbow Method for Clustering

Require: data, max_K
Ensure: optimal number of clusters
 $K_values \leftarrow range(1, max_K)$
 $WCSS_values \leftarrow []$
for each K **in** K_values **do**
 Apply k-means clustering with K clusters
 Calculate Within-Cluster Sum of Squares (WCSS) as:

$$WCSS_k = \sum_{i \in C_k} \|x_i - \mu_k\|^2$$

 Append WCSS to $WCSS_values$
end for
 $optimal_K \leftarrow FindElbowPoint(WCSS_values)$
return $optimal_K$



2. METHODOLOGY

2.1 Generating synthetic data from iris data

The synthetic data generation process begins by randomly sampling the first variable from the observed dataset without considering its relationships with other variables. For each subsequent variable, Classification and Regression Trees (CART) models are employed to capture the dependencies between variables.

For regression trees, the primary objective is to minimize the Mean Squared Error (MSE) at each split, defined as:

$$MSE = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (3)$$

where y_i are the observed values, $\bar{y}$ is the mean of subset, and n is the number of observations.

For classification trees, the Gini index measures the impurity of a node and is given by:

$$I_g(S) = \sum_{i=1}^k p_i(1 - p_i) \quad (4)$$

where p_i is the fraction of data points of class i in subset S , and k is the number of classes.

To determine the best split, Information Gain (IG) is calculated as:

$$IG(S, S_1, S_2) = I(S) - \left(\frac{|S_1|}{|S|} I(S_1) + \frac{|S_2|}{|S|} I(S_2) \right) \quad (5)$$

where S is the parent set and S_1 and S_2 are the subsets after splitting.

The synthetic data generation method involves creating a synthetic data object and sampling the first variable from the observed data. For the subsequent variables, CART models are used to generate synthetic values, which are then added to the synthetic dataset. If multiple synthetic datasets are required, the process is repeated accordingly. The synthetic data object is then populated with the synthesized dataset, methods used, synthesis order, and predictor matrix, before returning the final synthetic data object.

This approach ensures that the synthetic data accurately captures the relationships between variables, maintaining the structural integrity and relational dynamics of the original dataset. For a detailed procedural explanation, refer to Algorithm 2.

Algorithm 2 Synthetic Data Generation

Require: Data $\mathcal{D} = \{X_1, X_2, \dots, X_n\}$, Number of Datasets M (default = 1)

Ensure: Synthetic Data $\mathcal{S} = \{S_1, S_2, \dots, S_M\}$

Step 1: Initialize
Initialize an empty structure for synthetic data $\mathcal{S}$

Step 2: Sample First Variable
Randomly sample $X_1^{(syn)}$ from X_1 independently of other variables
Add $X_1^{(syn)}$ to $\mathcal{S}$

Step 3: Generate Subsequent Variables
for each subsequent variable X_i where $i = 2$ to n do
 Step 3.1: Fit Predictive Model
 Fit a model $\hat{f}_i$ to X_i using $\mathcal{D}$ and appropriate predictors $\{X_1, X_2, \dots, X_{i-1}\}$
 Step 3.2: Generate Synthetic Values
 Generate synthetic values $X_i^{(syn)}$ by sampling from $\hat{f}_i$
 Add $X_i^{(syn)}$ to $\mathcal{S}$
end for

Step 4: Repeat for Multiple Datasets (if applicable)
if $M > 1$ then
 for each dataset $m = 2$ to M do
 Repeat Steps 2 and 3 to generate additional synthetic datasets S_m
 end for
end if

Step 5: Populate Synthetic Data Structure
Populate $\mathcal{S}$ with metadata including methods, synthesis order, and predictor matrix

Step 6: Return Result
return $\mathcal{S}$

3. RESULTS AND DISCUSSION

This section presents our analysis using synthetic data generation to enhance unsupervised clustering algorithms. By generating synthetic datasets that emulate the statistical properties of small original datasets, we assessed the algorithms' effectiveness in more robust data environments. Our core analysis focused on determining the optimal number of clusters, a crucial factor for unsupervised learning models. The results quantify improvements in clustering performance, measured by metrics such as cluster cohesion and separation, when synthetic data is integrated. We also explore how synthetic data more accurately identifies the optimal number of clusters compared to smaller original datasets. Additionally, the results demonstrate the versatility and adaptability of synthetic data across various data characteristics and clustering algorithm specifications. The DBSCAN algorithm's clustering of the Iris dataset, shown in Figure 1, effectively identifies distinct groups corresponding to the three iris species. The color-coded clusters highlight DBSCAN's capability to manage varied densities and separate noise.

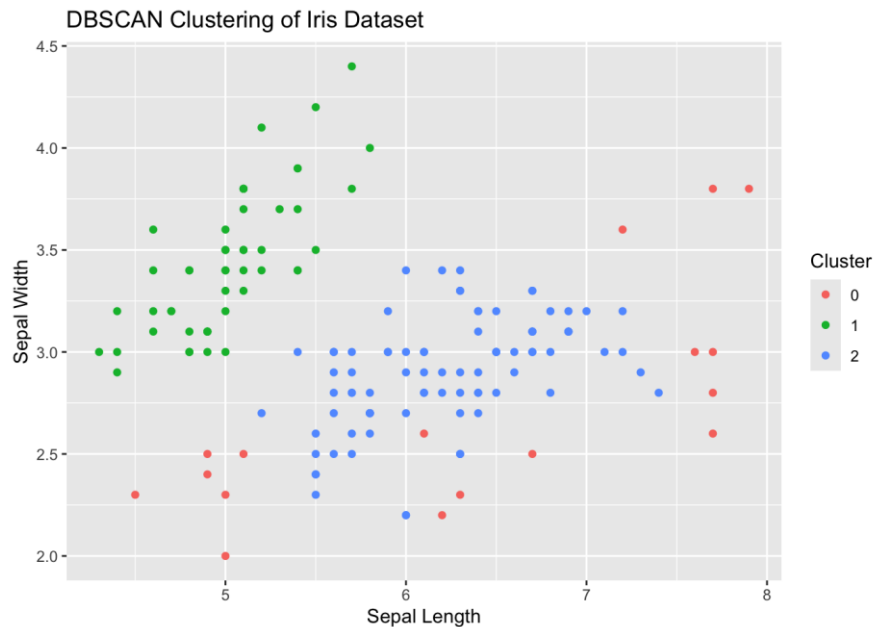


Figure 1. Visualization of DBSCAN clustering applied to the Iris dataset, showcasing the algorithm's capability to identify distinct clusters based on data density and handle noise effectively

Illustrated in Figure 2, the K-means algorithm partitions the same dataset into clusters based on features such as sepal length and width. The distinct colors representing each cluster showcase K-means' proficiency in optimizing within-cluster sum of squares and forming coherent groups. Figure 3 provides a density plot of the Iris dataset's features—sepal length, sepal width, petal length, and petal width—across the three species. The color differentiation for each species underscores the distribution patterns and overlaps, offering insights into feature contributions to clustering. Comprehensive density plots for synthetic datasets generated based on the statistical properties of the Iris dataset are displayed in Figure 4. These plots visually compare the original and synthetic data, showcasing the synthetic data's ability to mimic the original distributions of sepal length, sepal width, petal length, and petal width across the iris species. Each color differentiated curve represents a set of synthetic data, demonstrating consistency and variability relative to the original data. This visualization is crucial for assessing the quality and accuracy of the synthetic data generation process, highlighting its effectiveness in replicating key statistical traits and ensuring reliable use for further analysis and model training.

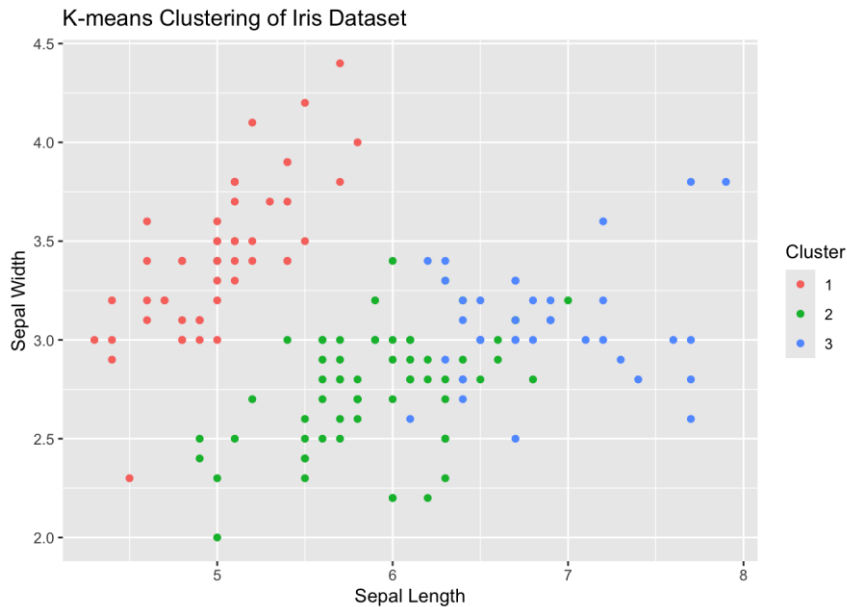


Figure 2. Representation of K-means clustering on the Iris dataset, illustrating the partitioning of data points into well-defined clusters corresponding to different iris species.

Table 1 validates the synthetic data generation process by comparing key statistical properties with the original data. The p-value, crucial in hypothesis testing, determines if observed differences in sepal length between original and synthetic datasets are significant. A p-value below 0.05 indicates a meaningful deviation of synthetic data from the original. The confidence interval further illustrates the range within which the true mean difference likely resides, providing insights into the variability and reliability of the synthetic data. This helps quantify the uncertainty around the mean difference estimate, ensuring the precision of synthetic data relative to real-world parameters.

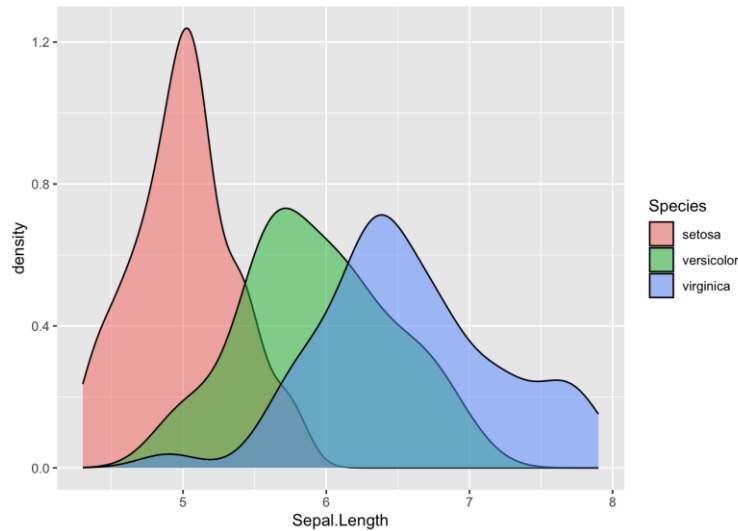


Figure 3. Density plot of the Iris dataset, highlighting the distribution of data points across different features and species.

Presenting these statistics prominently, the table evaluates the quality of synthetic data and highlights the robustness of the methodologies used. This ensures stakeholders can make informed decisions based on synthetic data, useful in scenarios where real data is limited due to privacy, logistical, or resource constraints. Ultimately, this statistical analysis fosters confidence in synthetic data as a reliable proxy for the original dataset, extending the applicability of data-driven strategies in research and industry. The figure titled “Elbow Method for Optimal k” illustrates the application of the Elbow Method to determine the optimal number of clusters for K-

means clustering. It plots the withincluster sum of squares (WCSS) against the number of clusters (k), with each point representing a specific k value. The Fig. 4: Density plot of synthetic datasets derived from the Iris dataset, illustrating the distribution of features across generated data points, reflecting the preservation of original data characteristics.

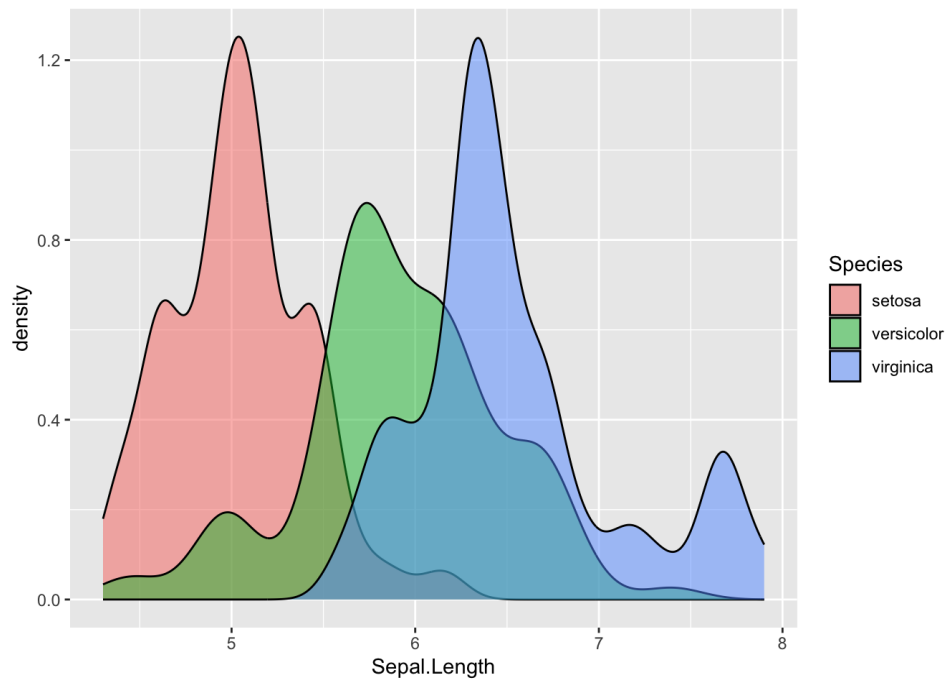


Figure 4. Density plot of synthetic datasets derived from the Iris dataset, illustrating the distribution of features across generated data points, reflecting the preservation of original data characteristics.

TABLE 1. Comparison of Sepal Lengths: Welch Two Sample t-test Results between Original Iris Data and Synthetic Data. Key statistics are highlighted, including the p-value and confidence interval bounds, to emphasize the statistical significance and the estimated range of difference in means.

Statistic	Value
t-value	0.61654
Degrees of Freedom (df)	286.76
p-value	0.538
95% Confidence Interval	
Lower Bound	-0.11035
Upper Bound	0.21102
Sample Estimates	
Mean of Original Data (x)	5.8433
Mean of Synthetic Data (y)	5.7930



Distinct ‘elbow’ point, where the decrease in WCSS significantly slows, indicates the optimal k . This point is highlighted to guide the selection of k , balancing between minimizing WCSS and avoiding overfitting. This visualization is crucial for justifying the choice of k , ensuring appropriate model complexity for the data structure.

Table 1 presents a quantitative analysis of WCSS values for different numbers of clusters determined by the elbow method. Each row corresponds to a different number of clusters and lists the WCSS value, measuring the total variance within each cluster. The highlighted row indicates the point where the reduction in WCSS becomes less pronounced, guiding the selection of the optimal number of clusters. This table is crucial for identifying the most efficient clustering cutoff, ensuring effective data segmentation by balancing between minimizing WCSS and avoiding unnecessary complexity.

Figure 6 graphically represents the outcomes of applying clustering algorithms to nine synthetic datasets derived from the Iris dataset. This visualization highlights how algorithms effectively group data points based on features such as sepal length, sepal width, petal length, and petal width.

Table 2. Within-cluster sum of squares (WCSS) values for different numbers of clusters determined by the elbow method. The row highlighted in gray indicates the optimal number of clusters, where a significant change in the rate of decrease in WCSS values suggests the most efficient clustering cutoff.

k (Number of Clusters)	WCSS
1	680.82440
2	152.36871
3	78.94084
4	57.31787
5	46.53558
6	38.93096
7	36.81647
8	29.95409
9	27.76542
10	26.12237

Each cluster is distinguished by a unique color, simplifying the identification of group boundaries and showcasing the algorithms' ability to discern distinct patterns. These plots are essential for evaluating the performance of clustering algorithms with synthetic data, reflecting their consistency in cluster formation despite variations in synthetic data characteristics. This illustration validates the use of synthetic datasets in machine learning, particularly when real data is limited or incomplete, affirming their utility in testing and improving clustering techniques.

The figure titled “Density of Synthetic Data Sets” presents density plots for each of the nine synthetic datasets derived from the Iris dataset. Each plot includes curves for key features—sepal length, sepal width, petal length, and petal width—color-coded for differentiation. These plots visually compare the synthetic distributions with those of the original dataset, highlighting the replication of the original data's statistical properties. This visualization is crucial for validating

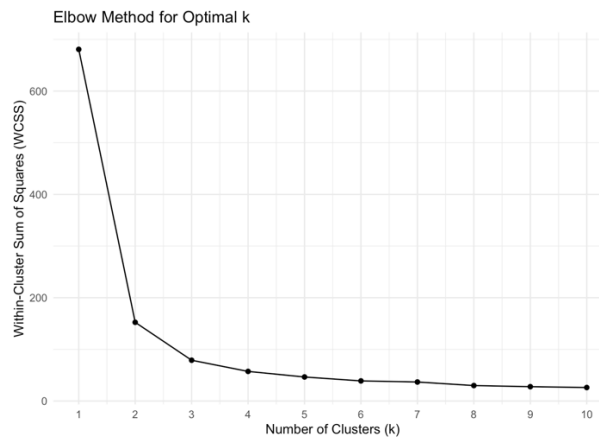


Figure 5: Elbow plot illustrating the determination of the optimal number of clusters (k) for K-means clustering using the within-cluster sum of squares (WCSS) across different k values.

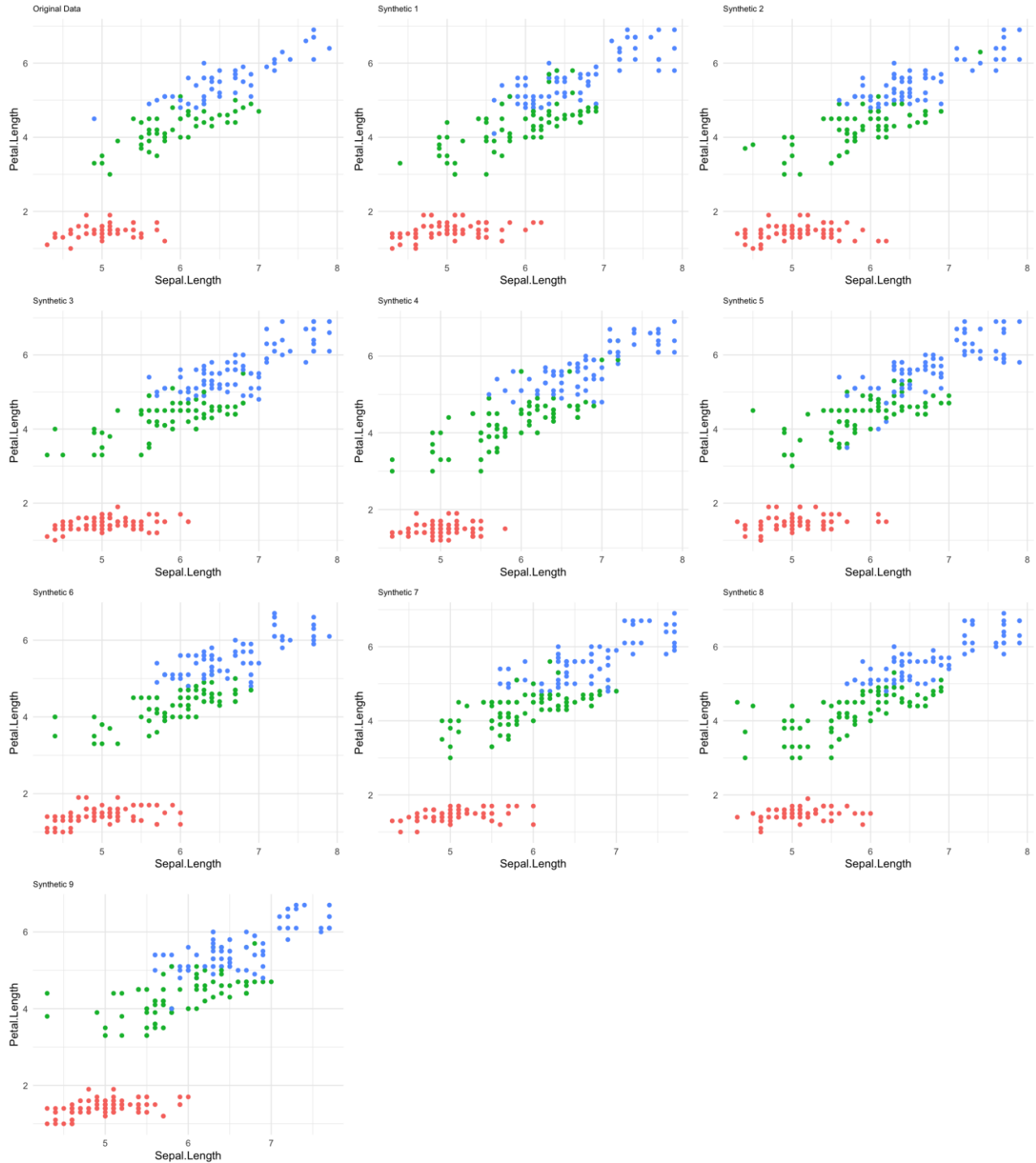


Figure 6. Cluster analysis across nine synthetic datasets generated from the Iris dataset, showcasing the adaptability and consistency of clustering algorithms in synthetic environments.

the quality of synthetic datasets, showcasing their consistency, variability, and utility in extending data analysis capabilities while preserving the intrinsic characteristics of the original data.

The figure titled “Elbow Method for Optimal k on Synthetic Data Sets” applies the Elbow Method across nine synthetic datasets, each with 300 data points, designed to mirror the statistical properties of the Iris dataset. The first plot represents the original Iris dataset, followed by synthetic datasets. Each plot shows the within-cluster sum of squares (WCSS) as a function of the number of clusters (k), highlighting the ‘elbow’ point where WCSS reduction plateaus, indicating the optimal number of clusters. This visualization allows for comparative analysis, demonstrating how synthetic data maintains clustering characteristics of the original dataset, enhancing robustness in model training and validation.

The generation of synthetic data from the Iris dataset has produced nine distinct synthetic datasets, each preserving the general structure and relationships of the original data while introducing some variability. Scatter plots show that the synthetic datasets maintain the clustering patterns of the three Iris species, with variations in cluster spread and overlap. Density plots of Sepal Length indicate that synthetic data largely preserves feature distributions across species, with minor deviations in peak heights or widths. Elbow plots for determining the optimal number of clusters consistently suggest a value around 3 for most synthetic datasets, aligning with the original Iris data. Slight variations in the elbow curves and inflection points reflect the randomness in the data generation process. These differences highlight the balance between maintaining the essential characteristics of the original data and introducing enough variation to create unique, representative synthetic samples.

4. CONCLUSION

This research systematically explored the use of synthetic data to preserve data properties and validate clusters through cluster number optimization techniques. Our findings demonstrate that synthetic data significantly enhances the robustness and accuracy of clustering models, particularly with limited original datasets like the Iris dataset. Using the Welch Two Sample t-test, we validated that synthetic data preserves key statistical properties of the original data, confirming its reliability for complex analysis tasks. The Elbow Method revealed that synthetic data effectively aids in determining the optimal number of clusters,



optimizing both computational resources and analytical accuracy. Visual tools such as density plots and clustering diagrams provided intuitive insights into the consistency between synthetic and original datasets, validating the synthetic data generation technique under various conditions. In summary, our research highlights the transformative potential of synthetic data in extending the capabilities of machine learning frameworks, particularly for unsupervised clustering algorithms. Integrating synthetic data generation into analytical workflows enhances the depth, reliability, and scalability of data analysis, empowering stakeholders to address complex problems with greater confidence. Future advancements in synthetic data generation will be crucial in refining cluster validation and optimization methods, bridging data gaps, and enabling sophisticated analyses across diverse disciplines, thereby advancing scientific discovery and business innovation.

REFERENCES

- Bishop, C. M. (2006). *Pattern recognition and machine learning by christopher m. bishop*. Springer Science+ Business Media, LLC.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16, 321–357.
- Dua, D., & Graff, C. (2017). *Uci machine learning repository*. <http://archive.ics.uci.edu/ml>. (Machine Learning Repository, University of California, Irvine, School of Information and Computer Sciences)
- Ester, M., Kriegel, H.-P., Sander, J., Xu, X., et al. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd* (Vol. 96, pp. 226– 231).
- Evans, D. (2002). Systematic reviews of interpretive research: interpretive data synthesis of processed data. *Australian Journal of Advanced Nursing*, 20(2).
- Fahad, A., Alshatri, N., Tari, Z., Alamri, A., Khalil, I., Zomaya, A. Y., ... Bouras, A. (2014). A survey of clustering algorithms for big data: Taxonomy and empirical analysis. *IEEE transactions on emerging topics in computing*, 2(3), 267–279.
- Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1) (No. 2). MIT press Cambridge.
- Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2). Springer.
- Hinton, G., & Sejnowski, T. J. (1999). *Unsupervised learning: foundations of neural computation*. MIT press.
- Hrebik, R., & Kukal, J. (2023). Concept of hidden classes in pattern classification. *Artificial Intelligence Review*, 56(9), 10327–10344.
- Johnston, B., Jones, A., & Kruger, C. (2019). *Applied unsupervised learning with python: Discover hidden patterns and relationships in unstructured data with python*. Packt Publishing Ltd.
- Lloyd, S. (1982). Least squares quantization in pcm. *IEEE transactions on information theory*, 28(2), 129–137.



**International Conference on Mathematics and
Mathematics Education (ICMME-2025), İstanbul Medeniyet
University, İstanbul, Türkiye, September 11-13, 2025**
“Mathematics in İstanbul, Bridge Between Continents”



- McLachlan, G. J., & Peel, D. (2000). *Finite mixture models*. John Wiley & Sons.
- Nanjundan, S., Sankaran, S., Arjun, C., & Anand, G. P. (2019). Identifying the number of clusters for k-means: A hypersphere density based approach. *arXiv preprint arXiv:1912.00643*.
- Park, N., Mohammadi, M., Gorde, K., Jajodia, S., Park, H., & Kim, Y. (2018). Data synthesis based on generative adversarial networks. *arXiv preprint arXiv:1806.03384*.
- Rokach, L., & Maimon, O. (2005). Clustering methods. *Data mining and knowledge discovery handbook*, 321–352.
- Schubert, E., Sander, J., Ester, M., Kriegel, H. P., & Xu, X. (2017). Dbscan revisited, revisited: why and how you should (still) use dbscan. *ACM Transactions on Database Systems (TODS)*, 42(3), 1–21.
- Shi, C., Wei, B., Wei, S., Wang, W., Liu, H., & Liu, J. (2021). A quantitative discriminant method of elbow point for the optimal number of clusters in clustering algorithm. *EURASIP journal on wireless communications and networking*, 2021, 1–16.
- Wheeldon, A., & Serb, A. (2023). A study on the clusterability of latent representations in image pipelines. *Frontiers in neuroinformatics*, 17, 1074653.
- Witten, I. H., Frank, E., Hall, M. A., Pal, C. J., & Data, M. (2005). Practical machine learning tools and techniques. In *Data mining* (Vol. 2, pp. 403–413).
- Yadav, J., & Sharma, M. (2013). A review of k-mean algorithm. *Int. J. Eng. Trends Technol*, 4(7), 2972–2976.



Automatic Lesion Segmentation in Dental Bitewing Radiographs: Comparative Analysis of U-Net, Attention U-Net, Mask R-CNN, and YOLOv11n-seg Models

Bilal İNAN¹, Nurettin DOĞAN²

¹ Information Technology Vocational School, Department of Computer Technology, Sakarya University of Applied Sciences, bilalinan@subu.edu.tr

² Faculty of Technology Department of Computer Engineering, Selçuk University, nurettin.dogan@selcuk.edu.tr

ABSTRACT

Bitewing radiographs provide high diagnostic value in the early diagnosis of dental diseases, particularly in the detection of caries and restorative materials (Lee et al., 2021). In this study, the performance of four different deep learning-based architectures was compared for automatic lesion segmentation in dental bitewing radiographs. The models used were U-Net, Attention U-Net, Mask R-CNN, and YOLOv11n-seg. The Dental Bitewing X-ray Dataset, which is openly available on Kaggle, was used, and color-coded mask annotations were converted to class ID format to make them suitable for model training. U-Net and Attention U-Net are pure segmentation models that perform pixel-level classification on the entire image, and the Dice coefficient and Intersection over Union (IoU) metrics were used to measure segmentation success. Similar approaches in the literature report high Dice scores in tooth segmentation (Dhar & Deb, 2022). Mask R-CNN and YOLOv11n-seg are object-based segmentation models that generate the location and pixel mask of each object and were therefore tested using COCO mAP@0.5:0.95 criteria.

The results show that the U-Net model achieved the highest pixel-based segmentation success with an average Dice = 0.8940 and IoU = 0.8104, while Attention U-Net ranked second with Dice = 0.8726 and IoU = 0.7766. Mask R-CNN achieved a mask mAP@0.5:0.95 of 0.6656 for mask segmentation, while the YOLOv11n-seg model achieved a box mAP@0.5:0.95 of 0.5313 and a mask mAP@0.5:0.95 of 0.4647, demonstrating high performance particularly in the filling-metal and dentin classes. However, it is noteworthy that both models showed low success in the cavity class. This situation may be due to factors such



as the loss of small lesion areas at low resolution, class imbalance, and the uncertainty of annotations. Furthermore, this low success rate in cavity detection from a clinical perspective may lead to missed opportunities for early diagnosis and delays in treatment planning. The findings show that U-Net-based architectures offer strong performance in pixel-level segmentation, while the YOLOv11n-seg model provides advantages in applications requiring real-time and object-based segmentation. In the literature, approaches similar to YOLOv11n-seg, such as YOLO-DentSeg, provide high mAP and FPS values in real-time dental diagnosis systems (Hua et al., 2025).

Key Words: Dental bitewing radiographs, Deep learning segmentation, Caries detection

1. INTRODUCTION

Dental caries represents one of the most common chronic health conditions globally, and when left undiagnosed at an early stage, it can lead to permanent tooth loss and more complicated treatment processes. Therefore, accurate diagnosis in the initial phase is essential not only for the success of treatment but also for maintaining patients' quality of life. Among the conventional diagnostic methods, dental bitewing radiographs stand out as an effective tool, as they allow for detailed evaluation of dental structures, existing restorations, and, in particular, early carious lesions located in the interproximal regions. Previous studies have highlighted that the diagnostic sensitivity of bitewing radiographs increases substantially when combined with artificial intelligence-based approaches (Grieco et al., 2022; Szabó et al., 2024; Lui et al., 2025).

Nevertheless, the interpretation of bitewing radiographs largely depends on the clinician's experience. Consequently, there is a significant risk of overlooking small or incipient lesions, which may lead to uncertainty and delays in the diagnostic process (Lee et al., 2021). This limitation underscores the necessity of developing AI-assisted systems that can enhance the reliability of diagnostic procedures and provide additional support to practitioners.

Deep learning, which has achieved groundbreaking success in numerous medical imaging tasks, also demonstrates considerable promise in the analysis of dental radiographs. Convolutional Neural Networks (CNNs), with their diverse architectures, are capable of



performing highly precise pixel-level analyses. Within this context, U-Net and its variants—such as Attention U-Net—have gained prominence owing to their high sensitivity in segmentation tasks (Dhar & Deb, 2022). In parallel, object-based approaches such as Mask R-CNN and YOLO not only identify the presence of lesions but also determine their locations and boundaries, thereby enabling both detailed and real-time analysis (Hua et al., 2025).

This study seeks to evaluate and compare the performance of four deep learning architectures—U-Net, Attention U-Net, Mask R-CNN, and YOLOv11n-seg—in the automatic segmentation of caries on dental bitewing radiographs. Comparable approaches have recently been applied with success in detecting apical lesions, caries, and other pathological conditions in dental images (Bayrakdar et al., 2022; Khattak et al., 2025). Through the assessment of both pixel-based and object-based segmentation approaches in terms of accuracy and clinical relevance, this study aims to make a substantive contribution to the advancement of AI-driven diagnostic systems in dentistry. In particular, U-Net and its derivatives have been reported to yield promising results in dental segmentation and caries detection (Bouali et al., 2024; Negi et al., 2024). The findings of this study are expected to provide valuable insights into how artificial intelligence can transform diagnostic workflows in dentistry. The evaluation will not only focus on diagnostic accuracy but also address clinical applicability, with the ultimate goal of supporting both individual treatment processes and broader public oral health. Large-scale investigations have already demonstrated that deep learning–based systems can significantly assist clinicians in the detection of proximal caries on bitewing radiographs (Pérez de Frutos et al., 2023; Rezaie et al., 2024). This study is therefore positioned to serve as a practical reference point for the integration of AI-assisted diagnostic tools into dental practice.

2. MATERIALS AND METHODS

2.1. Dataset

The dataset used in this study is the *Dental Bitewing X-ray Dataset*, which is publicly available on the Kaggle platform. It comprises a total of 1,099 bitewing radiographic images, each provided together with its corresponding mask file. The images are in JPEG format, while the masks are stored as PNG files. Within the masks, dental structures and pathological conditions are annotated with distinct color codes. The annotated classes are as follows: *background*,

enamel, dentin, pulp, filling-metal, filling-non-metal, cavity, crown, dental-implant, dental-implant-crown, root canal, sinus, and periapical radiolucency.

For the purposes of this study, particular attention was given to the clinically critical classes—*enamel, dentin, cavity, filling-metal, and filling-non-metal*. These classes are those most frequently examined in clinical settings and provide essential diagnostic information for early detection. However, the distribution of classes in the dataset is imbalanced. While the *dentin* and *enamel* classes are highly represented, the number of samples corresponding to the *cavity* class is relatively limited. This imbalance led to reduced accuracy in the cavity class, as small lesions that were few in number were more difficult to detect due to resolution loss and insufficient representation.

In this study, *enamel, dentin, cavity, filling-metal, and filling-non-metal* were evaluated as “clinically critical classes,” reflecting their importance for early diagnosis in practice. Nevertheless, the imbalance across these categories was a key factor underlying the reduced performance observed in cavity detection.

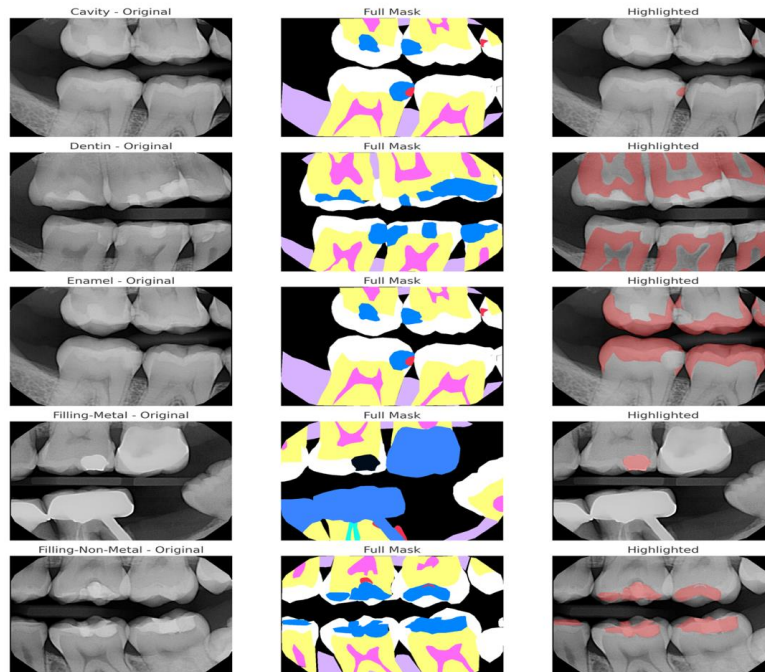


Figure 1. Examples of the five selected classes in the dataset: original image, complete mask, and the highlighted target class.



2.2. Preprocessing

Prior to model training, the color-coded masks were converted into class ID format, where each pixel was matched with its corresponding class label. This conversion enabled the masks to be directly processed by the deep learning architectures in numerical form.

The resizing of images varied according to the requirements of each model architecture. For the U-Net and Attn U-Net models, all images and masks were rescaled to 256×256 pixels. This choice was made to reduce training time and optimize memory usage. In the case of Mask R-CNN, the default parameters of the Detectron2 library were retained, with images resized to a range of 640–800 pixels and a maximum dimension of 1333 pixels. For YOLOv11n-seg, all images were reshaped to 640×640 pixels during training, thereby ensuring standardized input sizes tailored to each model.

Data augmentation strategies were employed to improve the models' ability to generalize across diverse imaging conditions. These included random horizontal and vertical flipping, 90-degree rotations, variations in brightness and contrast, as well as mild blurring. Such transformations were implemented to improve the robustness of the models against diverse imaging conditions.

2.3. Yolo-Seg

YOLOv11, one of the latest versions of the YOLO family, achieves substantially higher accuracy compared to its predecessors. This improvement is primarily attributed to optimized backbone and neck architectures, which enhance the model's feature extraction capacity. These architectural refinements enable more precise object detection, particularly in complex tasks. The general structure of the YOLO object detection architecture is presented in Figure 2.

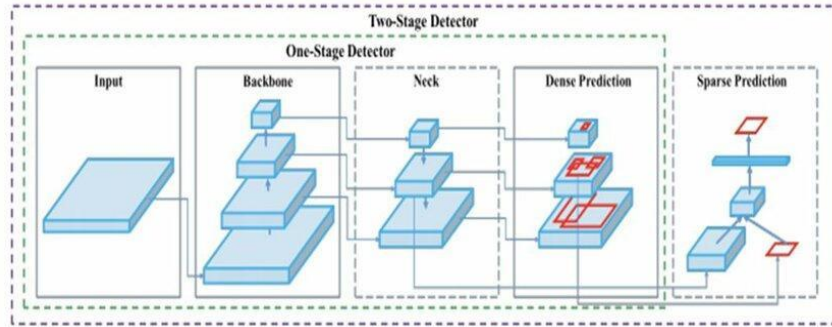


Figure 2. YOLO Object Detection Architecture (Ultralytic, 2025).

A comparative overview of different YOLO versions is provided in Figure 3.

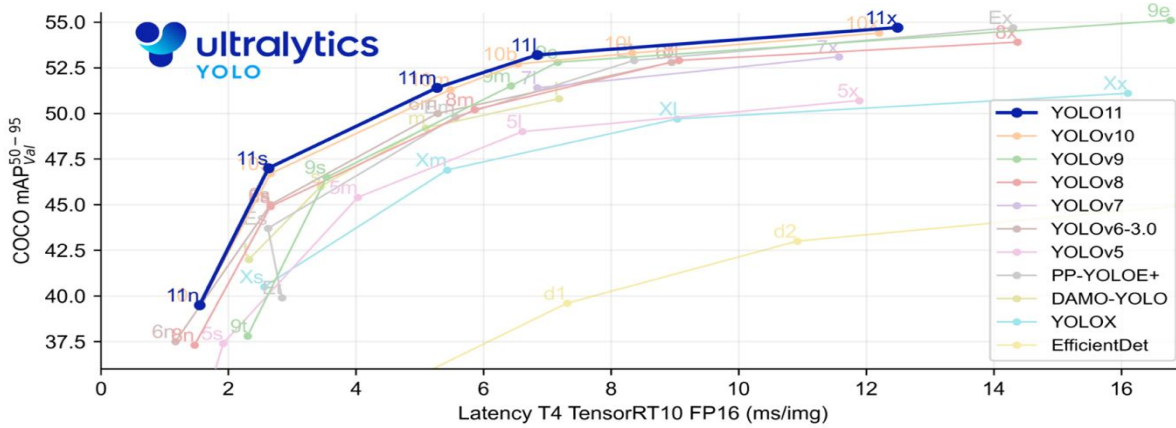


Figure 3. Accuracy comparison of YOLO versions (Ultralytic, 2025).

As shown in Figure 3, YOLOv11 surpasses earlier versions in terms of detection accuracy, demonstrating its capability for high-precision object recognition. YOLOv11-seg is available in five variants—YOLOv11n-seg, YOLOv11s-seg, YOLOv11m-seg, YOLOv11l-seg, and YOLOv11x-seg—supporting a wide range of computer vision tasks such as detection, segmentation, and classification. The architectural improvements allow the model to achieve a high mean Average Precision (mAP) score on the widely used COCO dataset. A notable advantage of YOLOv11 is its use of 22% fewer parameters compared to YOLOv8 while maintaining accuracy, thereby ensuring greater computational efficiency.

In this study, the YOLOv11-seg variant—specifically designed for segmentation—was employed. In addition to the detection head, YOLOv11-seg incorporates a dedicated segmentation head that generates pixel-level masks for each detected object. Through its

Feature Pyramid Network (FPN) design, the model effectively captures both small and large structures. Low-resolution mask predictions are subsequently refined through upsampling to approximate the original image size. Consequently, the model provides both bounding boxes and pixel-level boundaries in a single process.

The adaptability of YOLOv11-seg provides notable benefits in the analysis of dental radiographs, as it can detect caries, fillings, and dentin regions simultaneously through the use of both bounding boxes and segmentation masks. This combined functionality makes the model highly applicable to real-time clinical settings, offering efficient performance across various hardware platforms while preserving high accuracy (Hua et al., 2025).

2.4. U-Net and Attention U-Net

U-Net, a well-recognized architecture in segmentation tasks, has become prominent in biomedical image analysis owing to its effectiveness in identifying small anatomical structures. An overview of the U-Net design is illustrated in Figure 4.

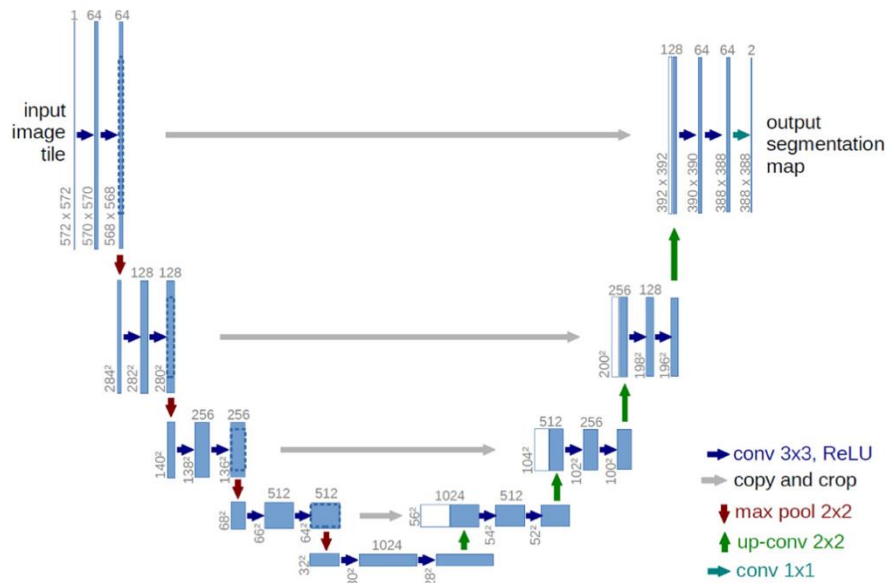


Figure 4. U-Net architecture. (Ronneberger et. al., 2015).

As illustrated in Figure 4, the U-Net architecture minimizes spatial resolution loss through its encoder–decoder design and skip-connection mechanism. This characteristic provides a critical advantage in areas requiring detailed analysis. For instance, in panoramic dental

radiographs, the application known as *Teeth U-Net* employed Squeeze-Excitation and dense skip connections, achieving a remarkably high Dice score of 94% (Hou et al., 2023).

The performance of U-Net has been further enhanced with the integration of attention mechanisms. The structure of the Attention U-Net architecture is presented in Figure 5.

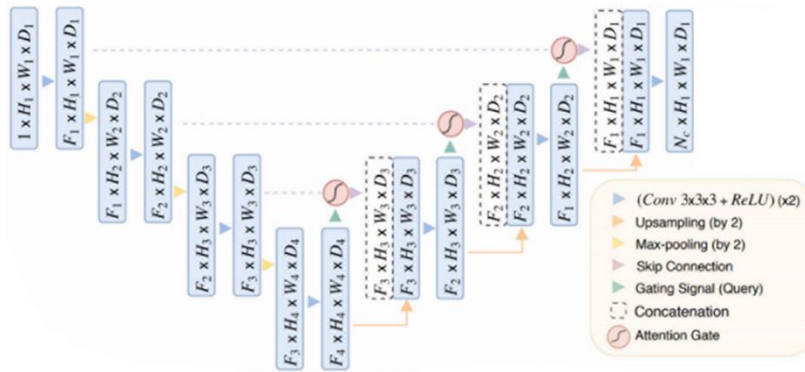


Figure 5. Attention U-Net architecture. (Kamnitsas et. al., 2017).

As shown in Figure 5, the Attention U-Net architecture allows the model to concentrate more effectively on clinically significant regions, such as lesion sites. An evaluation using the Tufts Dental X-Ray Dataset demonstrated that Attention U-Net delivered outstanding segmentation performance, achieving a Dice score of 95% and an IoU of 90% (Mahran et al., 2023).

In this study, both the architectural simplicity and robust performance of U-Net, as well as the enhanced ability of Attention U-Net to discriminate small and early-stage lesions, were utilized. Both models were configured for pixel-level caries segmentation in dental bitewing radiographs. During the preprocessing stage, all images were normalized to a resolution of 256×256 pixels, and data augmentation techniques such as rotation, flipping, and brightness-contrast variations were applied to improve the generalization capability of the models. The training process was carried out for 50 epochs using the Adam optimizer and the categorical cross-entropy loss function. This approach was designed to achieve robust pixel-based analysis of both larger structural segments and fine lesion regions.

2.5. Mask R-CNN

Mask R-CNN is a powerful and widely used architecture developed for object-based segmentation tasks. Built upon Faster R-CNN, this model not only detects and classifies objects while generating bounding boxes, but also incorporates an additional branch that produces precise pixel-level masks. By combining object-level and pixel-level information, Mask R-CNN offers a significant advantage in the analysis of complex images, particularly in medical and biomedical domains.

In dentistry, the effectiveness of Mask R-CNN has been demonstrated in numerous studies. For example, research on panoramic radiographs has achieved high accuracy in the automatic segmentation of teeth (Jader et al., 2018). In a similar manner, the application of this method to Cone-Beam Computed Tomography (CBCT) images has produced encouraging outcomes in identifying both dental structures and pathological entities. Such results demonstrate the model's capacity to differentiate complex anatomical areas as well as small lesions in dental radiographs (Cui et al., 2019). An overview of the Mask R-CNN architecture is provided in Figure 6.

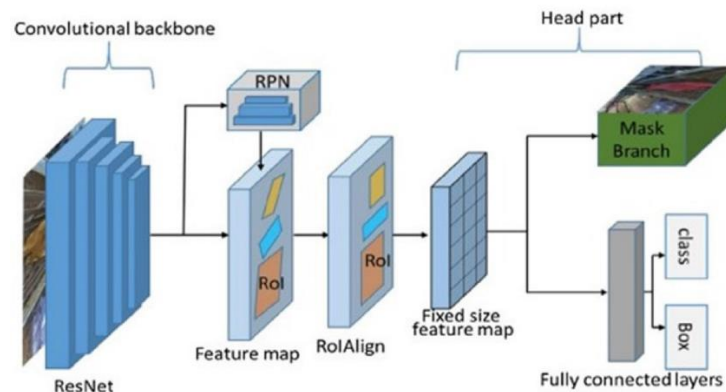


Figure 6. Mask R-CNN architecture. (Rubiu et. Al., 2023).

In this study, Mask R-CNN was applied to dental bitewing radiographs to detect both anatomical and pathological structures—such as enamel, dentin, metal fillings, non-metal fillings, and caries—at the object level while simultaneously generating pixel-based masks. During model training, the default parameters of the Detectron2 library were preserved, and the images were resized within a range of 640–800 pixels. To ensure robust performance under



varying conditions, data augmentation techniques including horizontal flipping and brightness–contrast adjustments were employed.

2.6. Performance Metrics

In this study, the different architectures were evaluated using performance metrics tailored to their respective approaches. For pixel-based segmentation models such as U-Net and Attention U-Net, which perform pixel-by-pixel classification across the entire image, performance was measured using the Dice Coefficient and Intersection over Union (IoU). These metrics quantify the extent to which the predicted segmentation masks overlap with the ground truth annotations.

In contrast, object-based segmentation models such as Mask R-CNN and YOLOv11n-seg not only detect objects but also generate bounding boxes and pixel-level masks, thus requiring a different evaluation strategy. The performance of these models was assessed using the COCO mean Average Precision (mAP@0.5:0.95) metric, which jointly evaluates the accuracy of object localization and segmentation masks. By accounting for both positional precision and mask quality, this metric provides a more comprehensive assessment of model performance (Ye et al., 2024).

The Dice Coefficient measures the similarity between the predicted mask (P) and the ground truth mask (G), and is defined as follows:

$$\text{Dice} = \frac{2|P \cap G|}{|P| + |G|} \quad (1)$$

Here, $|P|$ denotes the number of predicted pixels, $|G|$ represents the number of ground truth pixels, and $|P \cap G|$ indicates the number of pixels in their intersection. The IoU metric, on the other hand, measures the overlap between the predicted and ground truth masks, and is calculated as follows:

$$\text{IoU} = \frac{|P \cap G|}{|P \cup G|}$$

(2)

These metrics have been widely employed in the literature, particularly for the evaluation of segmentation performance in dental radiographs (Luo et al., 2021). For object-based segmentation models such as Mask R-CNN and YOLOv11n-seg, the COCO mAP@0.5:0.95 criterion was adopted. The mean Average Precision (mAP) reflects the average precision of the model across different IoU thresholds (e.g., 0.5, 0.75, 0.95), and is defined as follows:

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^N AP_i$$

(3)

Here, AP_i denotes the average precision calculated for each class, while N represents the total number of classes. In models such as Mask R-CNN, the COCO mAP metric is considered more meaningful since it simultaneously evaluates both the accuracy of object localization (bounding box precision) and the quality of pixel-level segmentation (He et al., 2017).

Through this approach, each architecture was assessed fairly and in alignment with its inherent design strengths. Pixel-based models were evaluated using Dice and IoU to provide a detailed overlap analysis, whereas object-based approaches were analyzed with the COCO mAP criterion, which reflects both localization accuracy and segmentation performance in an integrated manner.

3. FINDINGS

For this study, a dataset comprising a total of 1,099 bitewing images was used. In the YOLO-seg experiments, 879 images were allocated for training and 220 for validation. This model employed an object-based evaluation approach. In contrast, to assess the performance of pixel-based segmentation models such as U-Net and Attention U-Net, all images were resized to 256×256 pixels. These models were analyzed based on pixel-by-pixel classification.

Mask R-CNN and YOLO-seg, by comparison, were evaluated using object-based performance metrics, since they not only detect the presence of objects but also determine

their positions and boundaries. This strategy allowed for a fair comparison of the strengths of different architectures and provided insights into their potential for dental image analysis.

Among the pixel-based models, the performances of U-Net and Attention U-Net were compared. Analysis results indicated that U-Net achieved the highest performance in terms of mean Dice and IoU scores, while Attention U-Net ranked second. Both models yielded particularly high overlap values for the enamel and dentin classes. However, a noticeable decline in performance was observed for the cavity (caries) class in both models. This decrease was attributed to the small size of carious lesions and their tendency to disappear in lower-resolution images. Furthermore, class imbalance within the dataset was also considered a contributing factor to this decline. The class-based results of the pixel-based models are presented in Table 1.

Table 1. Class-based results of pixel-based models.

Class	U-Net Dice	U-Net IoU	Att U-Net Dice	Att U-Net IoU
0	0.925527	0.861377	0.919152	0.850399
1	0.913129	0.840145	0.902135	0.821718
2	0.812710	0.684509	0.774559	0.632065
3	0.920580	0.852847	0.889289	0.800648
4	0.897275	0.813688	0.883258	0.790924
5	0.892672	0.806150	0.882470	0.789662
6	0.835783	0.717893	0.857597	0.750695
Mean	0.885382	0.796658	0.872637	0.776587

* For U-Net and Attention U-Net, classes 0–6 represent the harmonized class IDs in the dataset.

An examination of Table 1 shows that the pixel-based segmentation models, U-Net and Attention U-Net, generally achieved high performance. However, the comparison indicates that U-Net performed better overall. U-Net obtained a mean Dice score of 0.885 and a mean IoU score of 0.797, surpassing the results of Attention U-Net. The average values for Attention U-Net were calculated as Dice = 0.873 and IoU = 0.777.

In the class-based analysis, U-Net reached Dice scores above 0.90 particularly in classes 0, 1, 3, and 4, which represent broader structures such as enamel and dentin. This outcome demonstrates the model’s reliability in distinguishing these structures. However, in the cavity (caries) class (Class 2), a noticeable decline in performance was observed, with Dice = 0.813

and $IoU = 0.685$. This reduction was associated with the small size of carious lesions, their tendency to disappear at lower resolutions, and the limited number of annotations.

A similar trend was identified for Attention U-Net. While the model also achieved high scores in enamel and dentin classes, it showed its lowest performance in the cavity class, with $Dice = 0.775$ and $IoU = 0.632$. Despite the attention mechanism's ability to highlight small regions, the overall performance of Attention U-Net remained behind that of U-Net. This limitation may be attributed to the imbalance among classes in the dataset, which could have restricted the effectiveness of the attention mechanism. To address this issue, strategies such as data augmentation focused on underrepresented classes (e.g., synthetic generation of carious regions) or the application of weighted loss functions are recommended. Figure 7 presents the training curves for U-Net and Attention U-Net, as well as a comparison between their ground truth and predicted masks.

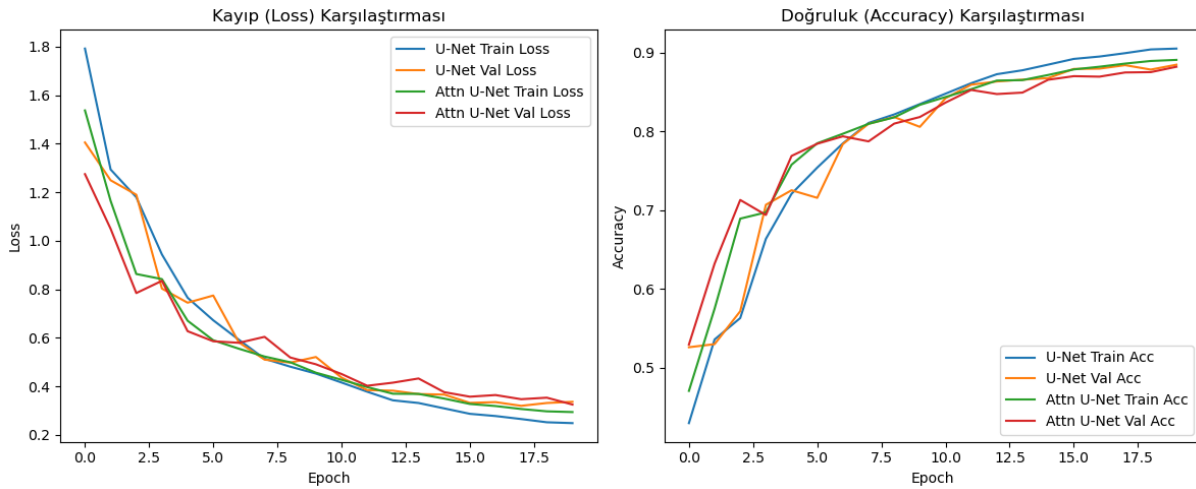


Figure 7. Loss and accuracy comparison of U-Net and Attention U-Net during training.

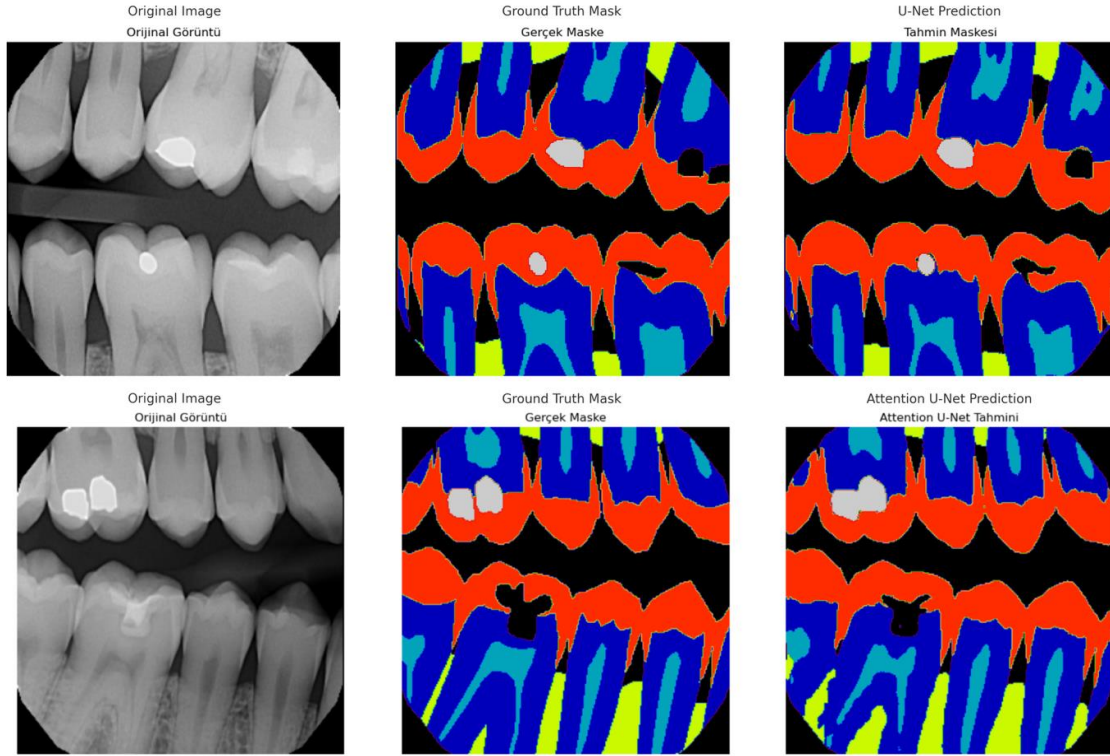


Figure 8. Comparison of ground truth and predicted masks for U-Net and Attention U-Net.

In conclusion, the results presented in Table 1, Figure 7, and Figure 8 demonstrate that the U-Net architecture provides a strong foundation for pixel-based segmentation in dental bitewing radiographs, whereas Attention U-Net, despite offering certain localized advantages, lags behind U-Net in terms of overall average performance. The class-based segmentation results of YOLOv11n-seg are presented in Table 2.

Table 2. Class-based segmentation results of YOLOv11n-seg.

Class	Box AP@0.5:0.95	Mask AP@0.5:0.95
Filling-metal	0.8318	0.7867
Filling-non-metal	0.5107	0.4854
Cavity	0.0903	0.0912
Enamel	0.4795	0.3931
Dentin	0.7443	0.5673
Mean	0.5313	0.4647

An examination of Table 2 presents the class-based box and mask mAP@0.5:0.95 values for the YOLOv11n-seg model. According to the results, the model achieved overall averages of

box mAP = 0.5313 and mask mAP = 0.4647. These values indicate that the model delivered balanced yet class-dependent performance in object-based segmentation.

At the class level, the highest performance was achieved for the *filling-metal* category. With box mAP = 0.8318 and mask mAP = 0.7867, these results suggest that metal filling regions were easily distinguished by the model due to their strong contrast differences. Similarly, the *dentin* class produced successful outcomes, with box mAP = 0.7443 and mask mAP = 0.5673.

In contrast, the *cavity* class exhibited the lowest performance, with box mAP = 0.0903 and mask mAP = 0.0912. This result indicates that small lesion areas were insufficiently detected due to their tendency to disappear in low-resolution images and the presence of class imbalance. For the *enamel* class, box mAP = 0.4795 and mask mAP = 0.3931 were obtained. The wide anatomical distribution of enamel regions, combined with their frequent proximity to neighboring structures, likely contributed to this moderate level of performance.

For the *filling-non-metal* class, the results were box mAP = 0.5107 and mask mAP = 0.4854, which can be attributed to the fact that non-metal fillings lack the visual distinctiveness of metallic ones.

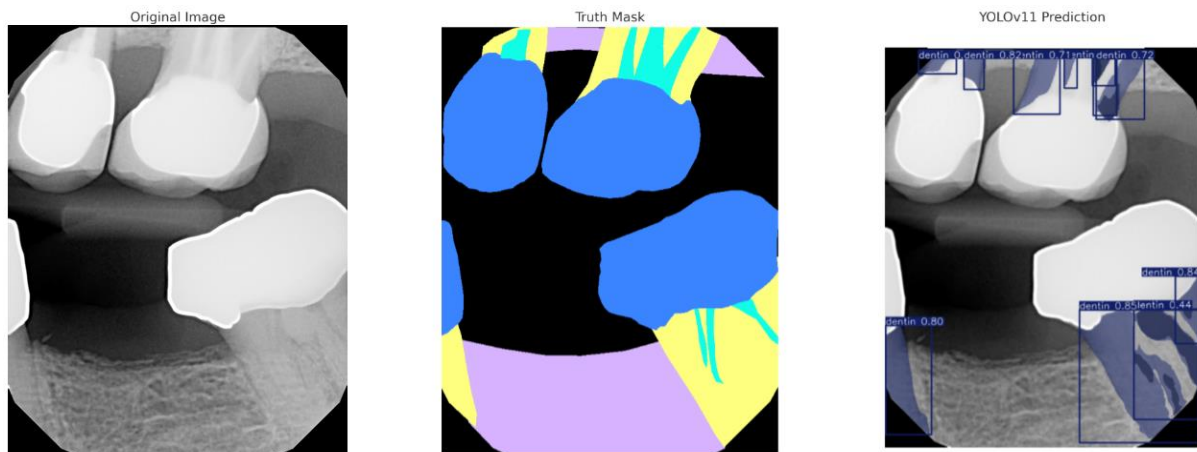


Figure 9. presents a comparison between ground truth and predicted masks.

Overall, YOLOv11n-seg demonstrated strong performance particularly in the *filling-metal* and *dentin* classes. However, its low performance in the *cavity* class highlights the model's limitations in detecting early-stage caries. This outcome represents a clinically critical finding,

indicating an area that could be improved through data augmentation or enhanced annotation quality.

For the Mask R-CNN model, a marked imbalance was observed in class-based results. While the *dentin* class achieved an acceptable success rate of approximately 33%, the *filling-metal*, *filling-non-metal*, *cavity*, and *enamel* classes yielded performances close to zero. This suggests that although the model is relatively more effective at capturing larger, high-contrast structures, it struggles to distinguish smaller or low-contrast regions. The near-zero performance in the *cavity* class is particularly concerning, as it indicates a failure to detect clinically critical lesions, thereby posing a significant limitation for early diagnosis.

Despite this, the overall mAP@0.5:0.95 value reported by Detectron2 appeared relatively high. However, this average score masked the class-level failures, presenting the overall performance as stronger than it truly was, even though certain classes achieved no success at all. This finding underscores that Mask R-CNN provides a statistically misleading global score in dental radiographs and is highly sensitive to class imbalance.

In conclusion, Mask R-CNN cannot be regarded as a reliable standalone method for dental radiographs. Its class-specific shortcomings—particularly the inability to detect small and poorly defined lesions—pose critical constraints for clinical use. Therefore, improving the performance of Mask R-CNN requires the integration of weighted loss functions, targeted data augmentation strategies for underrepresented classes, and hybrid approaches that combine complementary architectures.

Table 3. Summary of mAP results for object-based models

Model	mAP _{box} 0.5:0.95	mAP _{mask} 0.5:0.95
YOLOv11n-seg	0.5313	0.4647
Mask R-CNN	0.2246	0.6656*

In Table 3, the object-based average performance values of the YOLOv11n-seg and Mask R-CNN models are compared. The YOLOv11n-seg model achieved balanced results in object-based segmentation, with box mAP@0.5:0.95 = 0.5313 and mask mAP@0.5:0.95 = 0.4647.



The relatively higher box-based performance compared to the mask-based results suggests that the model was more successful in determining the locations of objects, while demonstrating comparatively lower accuracy in generating pixel-level masks.

For Mask R-CNN, the reported mask $mAP@0.5:0.95 = 0.6656$ appeared relatively high on average. However, class-based analysis revealed failures in the cavity, enamel, and filling classes. Due to these class-level inconsistencies, Mask R-CNN cannot be considered a reliable model. Thus, the values presented in Table 3 indicate that YOLOv11n-seg provided a more stable performance in practical applications, whereas Mask R-CNN, despite its seemingly high average score, fell short due to inconsistencies across classes.

Overall, U-Net-based models delivered strong pixel-level performance, while the YOLOv11n-seg model offered a balanced and practical solution for real-time applications. In contrast, Mask R-CNN was not evaluated as a clinically reliable method due to its class-level inconsistencies.

4. RESULTS

This study compared four different deep learning architectures for the automatic segmentation of caries and dental structures in bitewing radiographs. The models examined were U-Net, Attention U-Net, Mask R-CNN, and YOLOv11n-seg. Through this analysis, both the strengths and limitations of pixel-based and object-based approaches were evaluated, providing a comprehensive assessment in terms of accuracy and clinical applicability.

The results revealed that pixel-based approaches provided superior accuracy. Among them, the U-Net model demonstrated the highest performance, reaching a Dice coefficient of 0.894 and an IoU of 0.810. This outcome confirms the prominent role of U-Net in biomedical image segmentation (Ronneberger et al., 2015). In agreement with this finding, Hou et al. (2023) reported similarly high Dice scores in their *Teeth U-Net* study on panoramic radiographs, reinforcing the validity of the present results. While Attention U-Net possesses the capability to emphasize smaller, clinically important regions, its performance was hindered by class imbalance, resulting in lower accuracy compared to U-Net. Mahran et al. (2023) also observed that small lesions in the Tufts Dental X-Ray dataset were difficult to detect due to limited and uneven annotations.



Regarding object-based models, YOLOv11n-seg delivered noteworthy results, particularly in the filling-metal and dentin categories. The model achieved a mean box mAP@0.5:0.95 of 0.531 and a mean mask mAP of 0.465. Similarly, Hua et al. (2025) emphasized in their YOLO-DentSeg study that the real-time capability of YOLO architectures highlights their suitability for clinical practice, especially in terms of processing efficiency and hardware adaptability. Nevertheless, the limited performance in the cavity class pointed to ongoing challenges in the detection of early-stage caries. This shortcoming was primarily related to the small lesion size and the insufficient number of annotations (Ye et al., 2024).

Mask R-CNN, on the other hand, exhibited significant inconsistencies in class-based analysis. Although its overall mean mAP appeared relatively high, it achieved almost no success in cavity, enamel, and filling classes. Jader et al. (2018) reported a similar finding in panoramic radiographs, noting that the model performed well for large structures but was inadequate for fine details. Likewise, Miki et al. (2020) highlighted the limited accuracy of Mask R-CNN in complex anatomical regions of CBCT images. The results of this study are consistent with those findings.

In general, U-Net and Attention U-Net provide reliable solutions for pixel-based segmentation tasks, while YOLOv11n-seg stands out as a practical alternative thanks to its speed and real-time processing capability. Mask R-CNN, however, cannot be considered a reliable standalone method due to inconsistent performance across classes and low accuracy in small lesion detection. The main limitations of this study include the small number of annotations for the cavity class and the tendency of small lesions to disappear at lower resolutions. To address these limitations, more balanced datasets, higher-resolution images, and the integration of explainable AI methods (e.g., Grad-CAM++, SHAP) are recommended.

In conclusion, this study demonstrates under which conditions AI-based segmentation systems can be considered more reliable for dental radiographs. The findings contribute to the development of clinical diagnostic systems and provide an important foundation for hybrid approaches that can support early diagnosis and treatment planning.



5. DISCUSSION

In this study, four different deep learning architectures were evaluated on dental bitewing radiographs, each exhibiting varying levels of performance. The findings demonstrate that pixel-based models achieved higher accuracy compared to object-based approaches. The high Dice and IoU scores of U-Net are consistent with previous studies that identified this model as a reliable method in biomedical image segmentation (Ronneberger et al., 2015; Nishitani et al., 2021). Research conducted on panoramic radiographs has similarly reported that U-Net-based approaches deliver clinically acceptable accuracy (Hou et al., 2023).

Although Attention U-Net is an architecture designed to focus on small and clinically significant lesions, its performance in this study was limited due to class imbalance. Mahran et al. (2023) also emphasized that balanced datasets and comprehensive annotations are critical for the effective detection of small lesions. This finding highlights that the quality of training data, rather than model architecture alone, is a decisive factor in performance.

Among object-based models, YOLOv11n-seg yielded satisfactory results for high-contrast structures such as metal fillings. This trend aligns with the literature, which supports the adoption of YOLO-based methods in clinical applications where speed and hardware efficiency are required (Hua et al., 2025). However, the low performance observed in the cavity class underscores the ongoing challenge of detecting early-stage caries with object-based approaches. As reported by Ye et al. (2024), the disappearance of small lesions at low resolutions is one of the primary reasons for this limitation.

Mask R-CNN, in contrast, showed pronounced variations in performance across classes. While acceptable accuracy was achieved for broader structures such as dentin, performance was nearly negligible for small or low-contrast regions. Jader et al. (2018) reported similar limitations in panoramic radiographs, and Miki et al. (2020) identified comparable shortcomings in CBCT images. These findings make it difficult to consider Mask R-CNN a reliable standalone method for dental image analysis. Nevertheless, adopting hybrid strategies could partially overcome these limitations. For instance, integrating U-Net-based pixel segmentation with YOLO-based object detection may produce more balanced outcomes in terms of both accuracy and efficiency.



In recent years, the integration of attention-based Transformer architectures into medical image segmentation has emerged as a promising development for the detection of small lesions (Dosovitskiy et al., 2021; Chen et al., 2023). Adapting such approaches to dental radiograph analysis could enhance pixel-level precision while preserving the speed advantages of object-based models.

In light of the present findings, future research should focus on several key directions. First, the improvement of performance in the cavity class requires the development of balanced, high-quality annotated datasets. Additionally, strategies such as synthetic data generation and weighted loss functions could strengthen the contribution of small lesions to overall model performance (Lin et al., 2020). Training models in alignment with international classification systems (ICDAS, ICCMS) would further enhance clinical validity. Finally, the integration of explainable AI methods, such as Grad-CAM++ and SHAP, would not only improve accuracy but also support clinical trustworthiness by enabling practitioners to better understand model decisions.

REFERENCES

- Bayrakdar, I. S., Orhan, K., Çelik, Ö., Bilgir, E., Sağlam, H., Kaplan, F. A., ... & Różyło-Kalinowska, I. (2022). AU-net approach to apical lesion segmentation on panoramic radiographs. *BioMed Research International*, 2022(1), 7035367.
- Bouali, R., Mahboub, O., & Lazaar, M. (2024). Unleashing the potential of applied UNet architectures and transfer learning in teeth segmentation on panoramic radiographs. *Intelligenza Artificiale*, 18(2), 205-217.
- Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., ... & Zhou, Y. (2021). Transunet: Transformers make strong encoders for medical image segmentation. *arXiv preprint arXiv:2102.04306*.
- Cui, Z., Li, C., & Wang, W. (2019). ToothNet: automatic tooth instance segmentation and identification from cone beam CT images. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 6368-6377).
- Dhar, M. K., & Deb, M. (2024). S-R2F2U-Net: A single-stage model for teeth segmentation. *International Journal of Biomedical Engineering and Technology*, 46(2), 81-100.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.
- Grieco, P., Jivraj, A., Da Silva, J., Kuwajima, Y., Ishida, Y., Ogawa, K., ... & Ishikawa-Nagai, S. (2022). Importance of bitewing radiographs for the early detection of interproximal carious lesions and the impact on healthcare expenditure in Japan. *Annals of translational medicine*, 10(1), 2.



- He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 2961-2969).
- Hou, S., Zhou, T., Liu, Y., Dang, P., Lu, H., & Shi, H. (2023). Teeth U-Net: A segmentation model of dental panoramic X-ray images for context semantics and contrast enhancement. *Computers in Biology and Medicine*, 152, 106296.
- Hua, Y., Chen, R., & Qin, H. (2025). YOLO-DentSeg: A Lightweight Real-Time Model for Accurate Detection and Segmentation of Oral Diseases in Panoramic Radiographs. *Electronics*, 14(4), 805.
- Jader, G., Fontineli, J., Ruiz, M., Abdalla, K., Pithon, M., & Oliveira, L. (2018, October). Deep instance segmentation of teeth in panoramic X-ray images. In *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)* (pp. 400-407). IEEE.
- Jader, G., Fontineli, J., Ruiz, M., Abdalla, K., Pithon, M., & Oliveira, L. (2018, October). Deep instance segmentation of teeth in panoramic X-ray images. In *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)* (pp. 400-407). IEEE.
- Kamnitsas, K., Ledig, C., Newcombe, V. F., Simpson, J. P., Kane, A. D., Menon, D. K., ... & Glocker, B. (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical image analysis*, 36, 61-78.
- Khattak, O., Hashem, A. S., Alqarni, M. S., Almufarrij, R. A. S., Siddiqui, A. Y., Anis, R., ... & Agarwal, A. (2025, June). Deep Learning Applications in Dental Image-Based Diagnostics: A Systematic Review. In *Healthcare* (Vol. 13, No. 12, p. 1466). MDPI.
- Lee, S., Oh, S. I., Jo, J., Kang, S., Shin, Y., & Park, J. W. (2021). Deep learning for early dental caries detection in bitewing radiographs. *Scientific reports*, 11(1), 16807.
- Li, X., Lv, C., Wang, W., Li, G., Yang, L., & Yang, J. (2022). Generalized focal loss: Towards efficient representation learning for dense object detection. *IEEE transactions on pattern analysis and machine intelligence*, 45(3), 3139-3153.
- Lui, J. C. L., Lam, W. Y. H., Chu, C. H., & Yu, O. Y. (2025). Global research trends in the detection and diagnosis of dental caries: a bibliometric analysis. *International Dental Journal*, 75(2), 405-414.
- Mahran, A. M. H., Hussein, W., & Saber, S. E. D. M. (2023). Automatic teeth segmentation using attention U-Net.
- Miki, Y., Muramatsu, C., Hayashi, T., Zhou, X., Hara, T., Katsumata, A., & Fujita, H. (2017). Classification of teeth in cone-beam CT using deep convolutional neural network. *Computers in biology and medicine*, 80, 24-29.
- Negi, S., Mathur, A., Tripathy, S., Mehta, V., Snigdha, N. T., Adil, A. H., & Karobari, M. I. (2024). Artificial intelligence in dental caries diagnosis and detection: an umbrella review. *Clinical and Experimental Dental Research*, 10(4), e70004.
- Nishitani, Y., Nakayama, R., Hayashi, D., Hizukuri, A., & Murata, K. (2021). Segmentation of teeth in panoramic dental X-ray images using U-Net with a loss function weighted on the tooth edge. *Radiological Physics and Technology*, 14(1), 64-69.
- Pérez de Frutos, J., Holden Helland, R., Desai, S., Nymoen, L. C., Langø, T., Remman, T., & Sen, A. (2024). AI-Dentify: deep learning for proximal caries detection on bitewing x-ray-HUNT4 Oral Health Study. *BMC Oral Health*, 24(1), 344.



**International Conference on Mathematics and
Mathematics Education (ICMME-2025), İstanbul Medeniyet
University, İstanbul, Türkiye, September 11-13, 2025**
"Mathematics in İstanbul, Bridge Between Continents"



- Rezaie, S., Saberitabar, N., & Salehi, E. (2024). Improving dental diagnostics: enhanced convolution with spatial attention mechanism. *arXiv preprint arXiv:2407.08114*.
- Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Cham: Springer international publishing.
- Rubiu, G., Bologna, M., Cellina, M., Cè, M., Sala, D., Pagani, R., ... & Ali, M. (2023). Teeth segmentation in panoramic dental X-ray using mask regional convolutional neural network. *Applied Sciences*, 13(13), 7947.
- Szabó, V., Szabo, B. T., Orhan, K., Veres, D. S., Manulis, D., Ezhov, M., & Sanders, A. (2024). Validation of artificial intelligence application for dental caries diagnosis on intraoral bitewing and periapical radiographs. *Journal of Dentistry*, 147, 105105.
- Ye, Y., Chen, Y., Wang, R., Zhu, D., Huang, Y., Huang, Y., ... & Xiahou, J. (2024). Image segmentation using improved U-Net model and convolutional block attention module based on cardiac magnetic resonance imaging. *Journal of Radiation Research and Applied Sciences*, 17(1), 100816.



Examining the Norms of the Problem Solving Process within the Framework of Practical Rationality

Ahmet Salih Doğan¹ and Ahmet Şükrü Özdemir²

¹ Marmara University, Turkey, Doctoral Student,
asalihdogan@gmail.com, 0000-0002-0040-174X

² Istanbul Sabahattin Zaim University, Turkey,
Prof. Dr., asarzuhalci@gmail.com, 0000-0002-0597-3093

ABSTRACT

This research examines the norms that emerge during problem-solving processes within the framework of the Theory of Practical Rationality. The study aims to uncover the underlying norms behind the behaviors and decisions exhibited by participants during their problem-solving actions. The problem-solving process is based on Polya's (1945) phased approach, systematically analyzing how participants understand the problem, devise a plan, execute the plan, and evaluate the solution process.

The Theory of Practical Rationality, developed by Herbst and Chazan (2011), allows for understanding how rational decisions are made and how they guide solution processes in mathematical situations during mathematics education. According to this theory, individuals' actions are guided by norms specific to their social environment and professional practices. Therefore, problem-solving actions are not merely cognitive processes but are interwoven with these norms.

The study included mathematics teachers with varying professional experiences and academics working in the field of mathematics education. Two different problem contexts were used, and the participants' solution processes were collected using a qualitative method, namely focus group interviews (Krueger and Casey, 2015). Focus group interviews were chosen as an appropriate method to deeply understand participants' thought processes, strategy preferences, and the dilemmas they faced during the solution path. The data obtained from four different focus group interviews, each corresponding to a step in the problem-solving process, were analyzed using content analysis to determine the norms related to the problem-solving processes (Miles, Huberman and Saldaña, 2019). These interviews allowed



participants to collectively experience the problem-solving processes in two different complex and open-ended problem contexts. The analyses enabled the systematic presentation of the rational approaches exhibited by the participants in their solution processes.

As a result of the analyses, the main norms displayed by the participants during their problem-solving processes were determined. These norms were found to be closely related to Polya's problem-solving steps. For example, in the understanding the problem step, participants exhibited norms of checking prior knowledge and analyzing the given information in the problem. In the devising a plan phase, norms such as simplifying the problem, leveraging similar solved problems, and supporting the use of different strategies came to the forefront. In the executing the plan and evaluating the solution steps, norms like asking for explanations of solutions, self-evaluation, peer evaluation, and posing a problem that takes the problem-solving process a step further were observed. Furthermore, as a common decision-making norm for the participants, the practice of preparing before problem-solving was found to be significant.

The findings of the research demonstrated which norms were adopted by the participants in their problem-solving processes and how these norms shaped their solution strategies. In addition, the evaluations made within the framework of the Theory of Practical Rationality contributed to understanding the role of norms in instructional processes and teachers' pedagogical decisions. This study contributes to the mathematics education literature by determining the norms related to the problem-solving process within the scope of the Theory of Practical Rationality and the applicability of this approach in education. Furthermore, this work shows that the problem-solving process is not just a series of cognitive skills but also a product of professional practices and social interactions. These findings offer important implications for the design of training programs for the professional development of mathematics teachers and the teaching of problem-solving skills, suggesting that practical rationality norms should be considered in addition to cognitive processes.

Key Words: Practical Rationality, Norms, Problem Solving Process.

* This study was developed from a section of the doctoral dissertation title "Examining Preservice Mathematics Teachers' Practical Rationality of Problem Solving Process" prepared



by the second author under the supervision of the first author at the Marmara University Institute of Educational Sciences.

1. INTRODUCTION

The sustainability of the economic, social, and cultural development of societies largely depends on education systems equipping individuals with the knowledge, skills, and attitudes required by the era and the future. In this context, global and national educational reforms are structuring mathematics education to assume a central role in developing students' critical, creative, and analytical thinking skills. The skill of problem solving plays a key role in the reform efforts that have been implemented in mathematics education from the past to the present. This is because problem solving is regarded as a cognitive foundation that not only develops students' mathematical competencies but also their abilities to overcome challenges they may encounter in daily life.

This strategic importance of problem-solving is clearly emphasized by national and international official institutions. Within the scope of the Türkiye Century Education Model implemented by the Ministry of National Education (MEB) in 2024, problem solving is defined as one of the higher-order skills and one of the five fundamental mathematical area skills that need to be developed. Internationally, the PISA assessments conducted by the Organisation for Economic Co-operation and Development (OECD) relate problem solving ability to the competence of "solving problems encountered in daily life," stating that it is essential for individuals' high quality of life and productivity. Similarly, the National Council of Teachers of Mathematics (NCTM) considers problem-solving a fundamental ability of mathematical thinking and underlines the necessity of a broad problem-solving perspective to develop students' communication and discussion skills, prepare them for real-world problems, and support their flexible thinking. This institutional consensus indicates that problem-solving is one of the cornerstones of mathematical thinking skills.

The problem-solving process consists of four fundamental steps, as conceptualized by George Polya (1945): understanding the problem, devising a plan, carrying out the plan, and looking back (evaluating the solution process). Instantaneous decisions by the teacher, such as



how much guidance to give students throughout these steps, which strategy selection to permit, and when to intervene, are influenced by external and contextual factors like class size, curriculum intensity, and students' lack of motivation. This situation makes PS implementation a complex domain that requires a continuous balancing act between teachers' pedagogical ideals and practical necessities.

The decision-making actions of teachers, who assume a key role in implementing classroom practices, occur in highly social environments based on continuous and complex human interactions. Decision-making is generally defined as the process of consciously selecting an alternative to achieve a goal. However, unlike the 'rational actor' model in finance or economics, teacher decisions exhibit an instantaneous and volatile structure due to the influence of unstructured and dynamic classroom environments.

Mathematics education research has long focused on explaining teacher decisions by reducing them to individual and cognitive characteristics such as teachers' knowledge, beliefs, and values. These traditional approaches fall short in explaining teaching practices merely as an expression of individual competence, ignoring pressures from external conditions, administrative requirements, and social/organizational environments. Therefore, the focus shifts from whether a teacher's decision is 'pedagogically correct' to whether it is feasible, acceptable, and justifiable.

To overcome these traditional limitations, the Theory of Practical Rationality addresses instructional actions as a product of the socio-cultural context that lies outside individual characteristics. This theory, developed by Herbst and Chazan (2011), is a socio-cultural theoretical system that requires accounting for the resources offered and the contextual conditions encountered in practice to understand teachers' instructional actions and decisions. The theory conceptualizes mathematics teaching as the management of specific scenarios, called instructional situations, where traditional mathematical tasks (solving equations, proving, problem-solving, etc.) take place. The two foundational concepts of this theory are norms and professional obligations.

The concept of a norm is central to understanding the rationality of actions that occur in a social context, such as problem-solving. Norms are a set of implicit or explicit rules that guide



and justify the behaviour of individuals and groups. Evaluating the concept of a norm in the field of education is critical for understanding the complex social and cognitive dynamics of learning environments. In educational research, the concept of a norm refers to the standards of acceptable or expected behaviour, interaction, and reasoning in a learning environment. These standards are central to instructional processes and significantly influence how students construct and justify their mathematical knowledge. Norms serve as a fundamental building block in the formation of the micro-culture that supports individual and collective learning.

Guy Brousseau's Theory of Didactical Situations (1997) aims to enable a student to construct knowledge only through a problem situation called a "dimensional environment," without teacher intervention. This theory is based on the concept of the didactical contract. The didactical contract is a set of rules and norms implicitly existing between the teacher and the student, defining each one's roles, responsibilities, and expectations in the classroom environment. This contract allows the student to understand what the teacher wants and when they should do it. However, the rigidity or inappropriateness of the contract can impede the student's process of constructing personal knowledge. Therefore, in this theory, norms function as regulators of didactic relationships, generally accepted unconsciously, structuring the way knowledge is transmitted and constructed.

Paul Cobb and Erna Yackel (1996) made a significant contribution to this field by dividing norms in the context of mathematics education into two categories: general classroom norms and sociomathematical norms. General classroom norms are standards of interaction that are applicable in a mathematics classroom but are not specific to mathematical activity. For example, these are expectations regarding general classroom behaviours such as taking turns to speak, listening to each other, or requiring explanations to be understandable. Sociomathematical norms, on the other hand, are specific to mathematical discussion and activity, regulating students' mathematical thinking and justification. Examples include shared expectations regarding mathematical judgments such as "what constitutes a mathematically different solution," "when a mathematical explanation is considered sufficient," or "characteristics of an efficient solution".



In another study concerning norms, Stigler and Hiebert (1997) show how norms are influenced by a broader cultural framework and shape the implicit expectations shared by teachers and students. According to this view, teaching is a cultural activity, and it is suggested that every culture has its own, often unconscious, script. These cultural scripts function as high-level norms that determine the typical structure of a lesson, the roles of the teacher and students, and fundamental beliefs about learning.

The Theory of Practical Rationality, developed by Patricio Herbst and Daniel Chazan, is a mid-level theory focused on examining the rationale behind mathematics teaching practice and why teacher actions are considered "reasonable" in a certain way. This theory addresses the teacher's actions not simply as individual preferences or pedagogical competence but as being shaped by the normative expectations of an institutional and professional context. Within this framework, the concept of a norm not only explains the observed regularities of classroom interactions but also reveals the professional justification that forms the basis for sustaining these regularities. Herbst and Chazan use the term norm in a sociological sense as the tacitly expected normal or unmarked behaviour in an environment. The teacher's sense of what to do and what not to do during a lesson is largely determined by these norms. These norms make the teacher's choice of a specific action reasonable or justify its rejection. For instance, even if new pedagogical approaches that allow for student discovery are suggested while teaching a theorem, the established institutional norm might require the teacher to ensure that the knowledge related to that theorem is presented correctly and officially. This norm includes the expectation that the teacher formally sanctions that knowledge, and that students are subsequently held responsible for it. According to Herbst and Chazan, norms explain the distinction between what is possible and what is feasible in classroom interaction. Innovative teaching visions or actions aimed at enabling students to do more authentic mathematical work, if they fall outside existing norms, may be seen as "unfeasible" or "unreasonable" actions by experienced practitioners and therefore rejected. This theory particularly emphasizes the decisive role of norms in the management of instructional exchanges. The interaction between teachers and students adheres to consistent sets of norms specific to certain instructional situations.



When Polya's (1954) work on the problem-solving process in mathematics education and the theory of practical rationality are considered together, Polya's model is a fundamental reference point for structuring mathematical thinking. However, how students experience these steps in classroom practice and how teachers guide this process is related not only to cognitive frameworks but also to socio-cultural and normative frameworks. In this context, the theory of practical rationality, put forward by Herbst and Chazan (2003, 2012), allows for the explanation of how teachers make decisions according to specific pedagogical norms, classroom expectations, and contextual constraints.

Evaluating this research within this scope highlights it as a study that brings together two powerful dimensions. First, it positions Polya's problem-solving steps not only as a theoretical framework but also as a tool for the analysis of classroom practices. This makes it possible to examine the roles of students and teachers in the problem-solving process more systematically. Second, the Theory of Practical Rationality offers a powerful explanatory framework for understanding why teachers make certain choices in problem-solving processes. Teachers' decisions in the classroom are shaped not only by individual pedagogical preferences but also by professional norms, curriculum expectations, and classroom dynamics. Therefore, this research aims to reveal how normative structures and pedagogical rationalities in mathematics teaching are reflected in problem-solving processes.

The main objective of this research is to deeply examine and characterize the norms related to the four steps constituting Polya's problem-solving process within the framework of the theory of practical rationality. The study particularly aims to reveal by which norms teachers justify their actions in the problem-solving process and how these norms regulate the potential for students to engage in original mathematical work. In this context, the research seeks an answer to the question: "What are the norms related to the problem-solving process within the framework of the Theory of Practical Rationality?".

This study is expected to make important contributions to the mathematics education literature and practices. The first is that, by integrating two different theories of mathematics education, a comprehensive analytical framework is presented that explains both what should be taught and why the teacher teaches in a certain way regarding the problem-solving process.



This integration will fill the theoretical gap between a cognitive ideal model and educational feasibility. Another contribution is that the explicit definition of implicit norms and the understanding of justification mechanisms will offer practitioners operational ways for feasible, gradual, and sustainable improvements in problem-solving instruction that will enable students to do more authentic mathematical work. Understanding why it is acceptable to violate a norm can aid the adoption of new practices that promote desired pedagogical outcomes.

In conclusion, this research makes an important contribution to developing a more realistic, context-sensitive, and normatively focused understanding of problem-solving pedagogy by maintaining the theoretical power of Polya's model while showing how it is transformed or constrained in classroom practice. The strength of such a study is its emphasis on the centrality of normative structures in mathematics education by proving that teacher actions are not merely personal choices but also adaptations to collective and institutional expectations.

2. METHOD

This research was conducted using a qualitative method to deeply examine the problem-solving process based on George Polya's (1945) four steps. Due to the necessity of in-depth examination of a specific cognitive phenomenon (problem-solving according to Polya's steps) within a real-life context, a case study design, a qualitative research approach, was used in the research. Qualitative research was preferred because it offers an interpretative process focusing on understanding the 'why' of human and group behaviours.

The research was designed as a holistic multiple-case study, as defined by Stake (2005), to increase methodological rigor. This design allowed for the comparative examination of the problem-solving processes of experts with different professional roles. Furthermore, in accordance with Yin's (1984/1994) classification, each of Polya's problem-solving stages was treated as a separate focus group session, and this approach permitted the separation of units of analysis as an embedded multiple-case design.

The study group consisted of a total of 8 experts determined using a purposive sampling strategy. This group was composed of 5 mathematics teachers working in different high



schools and 3 academics working in the field of mathematics education. The main objective of this selection was to observe the potential interactions between pedagogical practice knowledge and theoretical mathematics education knowledge, and to provide a rich perspective on problem-solving processes. The size of the study group (8 people) provided a balance that is suggested by Krueger (1994) for focus group interviews, where complex and controversial issues need to be addressed in depth (not more than seven), and which preserves the richness of opinions.

Data were collected through focus group interviews to reveal the participants' interactive and collaborative reasoning processes. The focus group interviews were conducted as sessions managed by a non-directive moderator and recorded via audio or audio-visual means with the participants' permission. A total of 4 group interviews were conducted, based on George Polya's (1945) four classic steps, for the systematic observation of problem-solving processes. Each session was structured by focusing on the stages of understanding the problem, devising a plan, carrying out the plan, and looking back (evaluating the solution process), respectively.

The basis of each interview session was formed by norm violation scenarios, which are the methodological tool of Herbst and Chazan's (2006) theory of practical rationality. In the study, each step of Polya was identified as a moment of decision for the participants, and situations involving mismanaged or implicitly violated norms in classroom interaction were presented. These scenarios were verbally elaborated in detail by the moderator and opened for discussion, aiming to draw the participants' attention to the existence of the norm, which is normally implicit.

Two different non-routine problems on Exponential Equations and Logarithms were used to reveal the experts' high-level reasoning abilities. The use of non-routine problems required the experts to go beyond mere procedural skills and employ higher-order thinking skills such as organizing data, seeing relationships, and performing sequential activities. The reason for preferring such problems is their potential to make the participants' normative tendencies more visible in the problem-solving process. The problems enabled the analysis of



norms that emerged in the participants' processes of explanation, justification, and persuasion by triggering both cognitive and social interactions.

The obtained data were analyzed in accordance with qualitative data analysis. The analysis of the data was carried out within the framework of two main approaches. First, the data analysis process proposed by Miles, Huberman, and Saldaña (2019) was followed. This process includes the systematic coding of data, categorization, theme formation, and the interpretation of these themes by transforming them into visual tables or networks. Thus, the norms related to the problem-solving process were revealed systematically. Second, Braun and Clarke's (2006) thematic analysis approach was used. Thematic analysis is a flexible method focused on extracting meaningful patterns (themes) from data and allowed for the visibility of different dimensions of the norms in line with the research objectives. By using these two approaches together, both a systematic and flexible analysis process was followed; the norms emerging in the problem-solving processes were examined both in depth and holistically.

The data analysis was conducted in six stages within the framework of these approaches. In the first stage, all recorded focus group interview transcripts, interview notes, and experts' worksheets were examined to become familiar with the data. In the second stage, transcripts were analyzed line by line, and both deductive coding according to Polya's steps and inductive coding according to the experts' spontaneous strategies were performed. In the third stage, the codes created were grouped into sub-themes and main themes representing patterns of meaning. In the fourth stage, the consistency of the themes with both the raw data set and the theoretical frameworks was systematically checked. In the fifth stage, the final themes were defined to best answer the research question. In the final stage, the findings obtained were reported, supported by rich quotes and detailed descriptions.

Various strategies were implemented to enhance the validity and reliability of the research. Firstly, audio recordings were taken during the data collection process, and all interviews were transcribed. In the coding process, two researchers worked independently, comparisons were made on the emerging codes and themes, and consensus was reached. Furthermore, member checking was conducted to verify the findings obtained from the



participants. These strategies were used as measures to increase the reliability and credibility of the research findings.

3. FINDINGS

Considering the data obtained from the research, the findings include an in-depth analysis of the metacognitive and pedagogical approaches exhibited by participating teachers and academics during the classroom implementation of non-routine problems. The obtained data structure the problem-solving process not only as a sequential series of the four-stage model but also as implicit cognitive norms that guide and facilitate these steps. Below, the norms that emerged for each stage are presented along with participant statements, and descriptive comments on how these norms operate in the problem-solving process are included. The data were collected in the context of non-routine problems related to the topics of exponential equations and logarithms; the specific conceptual requirements of this context made some norms more visible.

3.1 Before the Problem-Solving Process Begins: Teachers' Pre-Preparation Activities

Although Polya's problem-solving process model addresses the "understanding the problem" stage as the first phase, the research showed that teachers carry out an active preparation phase before this official starting point. Throughout the focus group interviews, participants stated that pre-preparation is a professional practice carried out by the teacher. Participants stated that the teacher's level of preparation determines both the way problems are presented and the normative expectations within the classroom. For example, one teacher expressed: "Before presenting the problem to the students, I analyze the possible solution obstacles and strategic errors. My pre-preparation as a teacher allows me to anticipate where students will get stuck throughout the process and intervene at the right moment". An academic, on the other hand, stated: "Pre-preparation for me also means checking the appropriateness of the problem. The problem should enable students to better understand their



personal world and related concepts. I do not bring a problem that doesn't ensure this to the class". Furthermore, the situations of how the classroom application will be carried out (group work or individual work), anticipating situations that may be encountered in the later stages of the process, and how the evaluation process will be handled at the end of the problem are shaped by the teacher around the pre-preparation norm.

3.2 Norms Related to the "Understanding the Problem" Stage

Two main norms stand out in the "understanding the problem" stage. These are checking prior knowledge and analyzing the information given in the problem. Participants placed great importance on verifying prerequisite knowledge, especially in logarithm-exponential problems. One teacher stated: "It is not possible for students to attempt to solve a logarithm problem without remembering the definition of a logarithm. Prior knowledge must be checked so that they do not proceed on incorrect foundations". This quote clearly demonstrates the content of the norm: the understanding of the problem should be supported by the teacher's guidance, the recollection of relevant concepts, and, if necessary, short conceptual activities. In this stage, participants encourage students to actively recall past knowledge related to the current problem, remember critical concepts, and question their relevance to the problem. On this matter, a teacher said: "After projecting the problem onto the board, the first thing I ask is for students to ask themselves, 'Have we solved a problem like this before? Which formula would be appropriate?'. I prevent them from just diving into the solution; I first ask them to question the knowledge base". An academic participant expressed their view on this: "I ask guiding questions to help students activate their past experiences related to the problem. Metacognitive inquiries like, 'What was the physical principle underlying this question?' necessitate checking prior knowledge. This is a deep metacognitive control mechanism".

Analyzing the given information was also frequently emphasized; teachers preferred to use representational tools such as tables, graphs, or symbolic transformations to show which components of the problem are critical. A participant's statement reflects this practice: "When we put the given information into a table or show it on a figure, students grasp the problem more clearly". This norm ensures that the problem context is established on a clear and common ground. Thus, a consensus is formed in classroom communication about which



elements will be considered. This norm indicates that teachers guide their students to clearly distinguish and structure the relationship between the concrete givens and the unknowns to be reached in the problem. Regarding this, one teacher stated: "I ask students to note down the 'given' and 'required' parts of the problem and, if possible, mark them on a graph. This systematic separation reduces understanding errors and strategic blockage, as having clear data is essential for a good plan". An academic said: "For us, the success of the 'Understanding' stage means that students can define not only what they read but also the relationships between those data. Being unable to define the problem is one of the solution obstacles; we aim to overcome the obstacle by having them do this analysis".

3.2 Norms Related to the "Devising a Plan" Stage

Multiple norms emerged from the participants in the "devising a plan" stage. These were identified as simplifying the problem, utilizing similarly solved problems, adhering to the hierarchy among questions, and supporting different strategies. Regarding the simplifying the problem norm, participants emphasized that complex expressions should be divided into sub-problems. "Students' progress more comfortably when they break down the complex problem into smaller sub-problems". This norm aims to reduce cognitive load and also establishes the set of accepted methods in the classroom (e.g., simple examples first, then generalization). At this stage, teachers encourage students to break down complex problems into smaller parts or to proceed using simple cases to create a general solution model. The following statement from a teacher was presented as evidence of this situation.

Regarding the norm of utilizing similarly solved problems, an academic participant pointed to the nature of mathematics, saying, "A new problem in mathematics is often like a re-staged version of a previously solved problem. It is very important for students to notice these similarities". This norm supports teachers' practices of selecting past classroom examples and framing the problem presentation; accordingly, that is, the teacher shapes the classroom norms by determining which similarities will be emphasized. Within the framework of this norm, teachers state that the skill of successfully transferring previously learned strategies to new and current contexts should be developed in their students. Regarding this, a teacher stated: "When solving every new problem, I force students to connect with past experiences by asking,



'Where have you used this strategy before?'. Analogical thinking ensures the effective use of the strategic repertoire".

Regarding the norm of adhering to the hierarchy among questions, participants emphasized that in situations involving multiple questions, the solution should be carried out according to the given order. On this matter, one teacher mentioned telling her students to first solve the initial question in the presented problem and then follow the sequence. She underlined that this is important for the student's development and success in the problem-solving process if the questions related to the same problem situation are given in a specific order. Another teacher expressed this situation as follows: "I emphasize to the students that the solution must always follow a structured path. I want them to adopt an approach like, 'I realized I couldn't move on to part B unless I solved part A first, and I sequenced accordingly.' This hierarchical arrangement ensures the systematic management of complex problems".

The last norm in this stage, supporting the use of different strategies, is considered important for the effective and efficient execution of the problem-solving process. This norm expresses the educator's willingness to encourage students to keep alternative paths mentally ready or to try them, instead of getting stuck on a single strategy. This requires high-level meta-strategic awareness. At this point, an academic stated: "I require them to solve a problem with at least two different methods or to investigate whether a different strategy (e.g., working backward or making a systematic list) is valid. This breaks the obsession with focusing on a single correct answer and increases high-level thought flexibility". Teachers defined classroom practices aimed at breaking the habit of solving with a single method as a norm. "The fact that students are not limited to only one method... strengthens their mathematical self-confidence". This norm demonstrates how teachers reward alternative strategies in evaluation and feedback processes. When these norms in the planning stage are considered together, they reveal the normative preferences that determine the teacher's guiding role in this stage and what planning behaviours will be considered acceptable within the classroom.

3.3 Norms Related to the "Carrying Out the Plan" Stage

This stage involves testing the pre-determined strategic roadmap and performing the calculations. The teacher's main role in this stage is to ensure the traceability and discipline of



the process. When the data obtained from the focus group interviews in this stage are examined, the norm of requiring the explanation of the solution is seen to emerge. This critical norm expresses the teacher's obligation to request students to detail their thought process verbally or in writing while the solution is being carried out. Participants commonly stated that merely providing the result is insufficient; step-by-step justified explanations are the main condition for classroom acceptability. At this point, one teacher's emphasis is clear: "In mathematics, it is the justified process, not the result, that is important. Another teacher said: "While doing the solution process, I ask them to note down every step and justify themselves with questions like, 'Why did you use this formula?', 'What are you calculating in this step?'. Just finding the answer is not enough; justifying how you reached that answer is essential." If the student does not explain the solution path, that solution actually does not become acceptable to the community. The operation of this norm sets the standards of persuasion in the class: which justifications are considered sufficient, which intermediate steps are expected to be shown, and what types of evidence are accepted as convincing. In the context of exponential/logarithm problems, teachers specifically requested that intermediate transformations be clearly shown.

3.4 Norms Related to the Evaluating the Solution Process Stage

When the obtained data are analyzed, three different norms are identified in this stage: self-assessment, peer assessment, and problem posing. These norms represent the transition from a simple result check to the creative expansion of knowledge phase, which is the main objective of the problem-solving process. The self-assessment norm was recorded as an expected behaviour in which students critically check their own solution. A teacher's comment, "When the student doesn't check their own solution, they don't realize they made a mistake. Self-assessment makes them more careful," emphasizes the practical effect of self-assessment. This norm indicates that it is supported by the short self-assessment questions that teachers place at the end of the lesson or within the solution steps. An academic also stated the existence of this norm by saying: "After the solution is finished, I always ask students: 'Is this answer a reasonable value in the context of the problem? Is it consistent with



the initial conditions?'. They should make this internal questioning a habit; this is also the way to overcome math anxiety".

The peer assessment norm involves the teacher structuring the classroom environment for the validation, discussion, and collection of alternative views of students' solutions by others. An academic expressed the following regarding this: "In class discussions, I ask everyone to explain their solution to the person next to them and to find errors mutually. Peer feedback opens individual blind spots and engages collective metacognition. The teacher's role here is not just to encourage; it mandates this social learning". According to this norm, teachers aimed to make visible which justifications are shared and which ones remain controversial by opening different solutions for comparison in the problem-solving process.

The problem posing norm indicates that problem generation activities by students after the solution are evaluated by the participants as a sign of competence. A teacher said: "When the student starts posing a problem, we understand that they have truly grasped the subject". This norm reflects the teachers' tendency to reward high-level understanding in evaluation. This norm encompasses the educators' effort to encourage students to produce new problems like the solved problem, to generalize and extend the existing solution. An academic's emphasis on this matter: "Problem posing is the level of synthesizing knowledge. Students producing new math problems from a given problem challenges their skills in proving, reversing, and generalizing the solution. This is proof that strategies have been learned deeply and can be transferred" was evaluated as evidence of this situation.

When the obtained findings are considered generally, they show that participants addressed the problem-solving process not only within the framework of the four steps proposed by Polya (1945) but also by extending it to include the process before it. Participants emphasized that the teacher's pre-preparation before starting the process, checking prior knowledge in understanding the problem, strategic diversity and hierarchical progression in the planning stage, justification of the solution in the carrying out the plan stage, and self- and peer assessment as well as problem posing in the evaluation stage are normatively important. These findings reveal that not only students but also teachers' normative behaviours are



decisive in the problem-solving process, and they show that problem-solving involves a multilayered normative structure in terms of classroom mathematical practices.

4. CONCLUSION, DISCUSSION AND SUGGESTIONS

In this research, the normative orientations of mathematics teachers and academics regarding problem-solving processes were examined within the framework of Polya's (1945) four-step model and Herbst and Chazan's (2011) theory of practical rationality. The findings of the research showed that the problem-solving process cannot be explained only by Polya's cognitive steps, and that the teachers' "pre-preparation" activities carried out before starting the process also emerged as a normative category. Furthermore, it was found that certain norms (checking prior knowledge, supporting different strategies, requiring explanation of the solution, self/peer assessment, and problem posing) were emphasized by teachers and academics at every stage of the problem-solving process.

The obtained findings show strong alignment with the literature on sociomathematical norms. As stated in Yackel and Cobb's (1996) work, what counts as a mathematically acceptable solution in the classroom, and which explanations students find "sufficient" or "convincing," are determined by norms. In this research, the prominence of the norms of "requiring the explanation of the solution" and "peer assessment" confirms the regulatory power of sociomathematical norms on classroom mathematical discourse. Similarly, Kazemi et al.'s (2007) research reported that opening different solution methods for discussion and evaluating peers' justifications deepens mathematical understanding. The findings of "supporting different strategies" and "peer assessment" obtained in this study overlap with the discussion-focused norms emphasized by Kazemi. Therefore, it can be said that norms in the problem-solving process serve not only to produce solutions but also to discuss the solution within the community and to verify it collectively.

A significant contribution of the study is the revelation of the norm of "teacher pre-preparation," which has not been sufficiently emphasized in the literature. Existing studies have generally addressed sociomathematical norms by focusing on students' solution processes. However, in this research, teachers' self-initiated preparation activities before the problem-solving process—solving the problem in different ways, anticipating possible student errors,



planning strategic questions—were defined as a normative behaviour. In the context of Herbst and Chazan's (2011) practical rationality approach, this finding shows that teachers construct rational justifications in advance to make their classroom activities "appropriate" and "reasonable". Thus, teacher preparation reveals that teacher practices also operate on normative foundations, in addition to the student dimension of sociomathematical norms.

A close parallel is also seen between the norms of "problem posing" and "utilizing errors" reported in Gülburnu's (2019) study and the finding of "evaluating the solution process" that emerged in this research. The participants' encouragement of students to pose new problems at the end of the problem-solving process supports the approach called the "problem posing norm" in the literature. However, in this study, problem posing was evaluated not only as a student-centered activity but also as an activity that the teacher anticipates during the preparation process and normatively guides at the end of the lesson. In conclusion, this research is largely consistent with the findings in the sociomathematical norms literature, but it offers a unique contribution to the literature by revealing teachers' "pre-preparation" practices as a normative dimension.

This study also reveals noteworthy results when evaluated in the context of the theory of practical rationality (Herbst and Chazan, 2003, 201), in addition to sociomathematical norms. Practical rationality emphasizes that what makes teachers' decisions in the classroom "rational" is the normative justifications accepted by the community, rather than individual preferences. The activities that teachers carried out in the pre-preparation process in this research—solving the problem in different ways, anticipating possible student errors, planning guiding questions—precisely support the approach defined by Herbst (2003) as "instructional actions being carried out under normative frameworks". In other words, through pre-preparation, teachers pre-construct which solution paths will be supported, and which justifications will be found convincing in the classroom.

This finding also shows parallels with Herbst, Nachlieli, and Chazan's (2011) studies on "teachers' ways of presenting mathematical tasks". In their research, it was emphasized that the choices teachers make while presenting tasks directly shape classroom mathematical practices. Similarly, in this research, the strategies determined by teachers in the preparation



process influenced which solution paths students would normatively consider acceptable. In Herbst and Kosko's (2014) study, "what types of explanations teachers request" and "which representations they prioritize" in problem-solving activities were analyzed in the context of practical rationality. The norm of "requiring the explanation of the solution" that stood out in our findings is consistent with Kosko and Herbst's observations. By planning which steps of the solution, they would request students to explain during the pre-preparation process, teachers determined the boundaries of classroom norms.

All these comparisons show that our findings are consistent with the practical rationality literature. However, an important contribution is revealed here: while practical rationality is generally associated with teachers' in-class decisions in the literature, in this study, teachers' preparation activities before class were also defined as a normative category. This shows that teacher practices are shaped not only in the classroom but also before the class by normative justifications considered reasonable by the community. Thus, the theory of practical rationality contributes to the understanding of teacher preparation processes as well, extending beyond the explanation of in-class actions.

Several suggestions can be developed in line with the research findings. Firstly, problem-solving in teacher education programs should be addressed not only through Polya's cognitive steps but also through teachers' pre-preparation processes. Pre-service teachers should be enabled to acquire the skills of trying different solution methods in advance, anticipating possible student errors, and planning guiding questions as a normative professional standard. Furthermore, in classroom practices, students should be systematically encouraged to engage in self- and peer assessment, discuss justification processes, and expand mathematical thinking through problem-posing activities. Future research should more comprehensively examine how the student-centered operation of sociomathematical norms and the practical rationality foundations that become visible in teachers' preparation processes interact.

This research makes a significant contribution to the literature. Firstly, the problem-solving process has been mostly examined in the context of Polya's cognitive steps or sociomathematical norms until now; however, studies addressing this process holistically within



the framework of the theory of practical rationality have remained limited. Yet, practical rationality has the power to explain teachers' in-class actions and decisions by linking them to normative justifications accepted by the community. This study fills this gap by addressing the problem-solving process not only in terms of students' cognitive orientations and sociomathematical norms but also in terms of teachers' pre-class preparations and in-class decisions. Secondly, the research offers a new perspective by linking discussions about norms with the theory of practical rationality. Practices such as teachers' pre-preparation, diversifying solution methods, and planning which explanations to request have become visible not just as pedagogical preferences but as normatively justified actions within the framework of practical rationality. Thus, this research reveals how norms regarding the problem-solving process are constructed not only as student-centered but also as teacher-centered, contributing to the understanding of problem-solving as a holistic normative practice.

REFERENCES

- Herbst, Patricio and Chazan, Daniel (2011) "Research on Practical Rationality: Studying the justification of actions in mathematics teaching," *The Mathematics Enthusiast*: Vol. 8 : No. 3 , Article 2.
- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389–407. <https://doi.org/10.1177/0022487108324554>
- Bishop, A. J., Clarke, B., Corrigan, D., & Gunstone, D. (2003). Values in mathematics and science education: Similarities and differences. *The Mathematics Educator*, 13(2), 4–11.
- Blackley, S., Wilson, S., Sheffield, R., & Howell, J. (2021). Teachers' decision-making in complex classrooms. *Australian Journal of Teacher Education*, 46(1), 58–76. <https://doi.org/10.14221/ajte.2021v46n1.5>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brousseau, G. (1997). *Theory of didactical situations in mathematics*. Kluwer Academic Publishers.
- Brown, T. (2018). Curriculum reform and mathematics teachers' practices. *Journal of Curriculum Studies*, 50(3), 317–335. <https://doi.org/10.1080/00220272.2018.1435723>
- Chazan, D., Herbst, P., & Clark, L. M. (2016). Research on mathematics teachers' use of curriculum resources. In G. Kaiser (Ed.), *Proceedings of the 13th International Congress on Mathematical Education* (pp. 483–497). Springer.
- Cobb, P., & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist*, 31(3–4), 175–190. <https://doi.org/10.1080/00461520.1996.9653265>
- Creswell, J. W. (2018). *Qualitative inquiry & research design: Choosing among five approaches* (4th ed.). SAGE.



**International Conference on Mathematics and
Mathematics Education (ICMME-2025), İstanbul Medeniyet
University, İstanbul, Türkiye, September 11-13, 2025**
“Mathematics in İstanbul, Bridge Between Continents”



- Çokluk, Ö., Yılmaz, K., & Oğuz, E. (2011). Nitel bir görüşme yöntemi: Odak grup görüşmesi. *Kuramsal Eğitimbilim Dergisi*, 4(1), 95–107.
- Fenstermacher, G. D., & Soltis, J. F. (2004). *Approaches to teaching* (4th ed.). Teachers College Press.
- Gülburnu, M. (2019). Matematik öğretiminde problem kurma ve hata kullanımına ilişkin normların incelenmesi. *Eğitim ve Bilim*, 44(199), 341–357. <https://doi.org/10.15390/EB.2019.8100>
- Herbst, P. (2003). Using novel tasks to teach mathematics: A case of what is possible. *Educational Studies in Mathematics*, 54(1), 1–24. <https://doi.org/10.1023/A:1024414112520>
- Herbst, P. (2011). The role of theory in mathematics education scholarship. *Educational Studies in Mathematics*, 77(1), 91–105. <https://doi.org/10.1007/s10649-011-9294-x>
- Herbst, P. (2018) Theory and methods for research on the practical rationality of mathematics teaching. *Educ. mat.* [online]. vol.30, n.1, pp.9-46. Epub 07-Ene-2022. ISSN 2448-8089. <https://doi.org/10.24844/em3001.01>.
- Herbst, P., & Chazan, D. (2003). Exploring the practical rationality of mathematics teaching through conversations about videotaped lessons. *For the Learning of Mathematics*, 23(1), 2–14.
- Herbst, P., & Chazan, D. (2011). On the instructional triangle and sources of justification for actions in mathematics teaching. *ZDM – The International Journal on Mathematics Education*, 43(5), 601–612. <https://doi.org/10.1007/s11858-011-0349-0>
- Herbst, P., & Chazan, D. (2012). Research on practical rationality: Studying the justification of actions in mathematics teaching. *The Mathematics Enthusiast*, 9(3), 433–468.
- Herbst, P., & Chazan, D. (2023). Norms, obligations, and rationality in mathematics teaching. In P. Herbst & D. Chazan (Eds.), *Advances in mathematics education research* (pp. 15–32). Springer.
- Herbst, P., Kosko, K. W. (2014). Using representations to reason about mathematics teaching. *Mathematics Education Research Journal*, 26(3), 527–554. <https://doi.org/10.1007/s13394-013-0097-2>
- Herbst, P., Nachlieli, T., & Chazan, D. (2011). Studying the practical rationality of mathematics teaching: What goes into “installing” a theorem in geometry? *Cognition and Instruction*, 29(2), 218–255. <https://doi.org/10.1080/07370008.2011.555008>
- Johnson, A. (2019). Class size and its impact on student achievement. *Educational Review*, 71(4), 495–511. <https://doi.org/10.1080/00131911.2018.1427574>
- Johnson, A., & Brown, T. (2018). Motivation and learning in mathematics classrooms. *Mathematics Education Research Journal*, 30(2), 125–140. <https://doi.org/10.1007/s13394-018-0234-7>
- Kazemi, E., Franke, M. L., & Lampert, M. (2007). Developing pedagogies in teacher education to support novice teachers’ ability to enact ambitious instruction. In D. Tirosh & T. Wood (Eds.), *International handbook of mathematics teacher education: Vol. 2. Tools and processes in mathematics teacher education* (pp. 13–36). Sense Publishers.
- Krueger, R. A. (1994). *Focus groups: A practical guide for applied research*. SAGE.
- Leinhardt, G., & Greeno, J. G. (1986). The cognitive skill of teaching. *Journal of Educational Psychology*, 78(2), 75–95.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE.



**International Conference on Mathematics and
Mathematics Education (ICMME-2025), İstanbul Medeniyet
University, İstanbul, Türkiye, September 11-13, 2025**
“Mathematics in İstanbul, Bridge Between Continents”



- Lopez, A., & Allal, L. (2007). Sociomathematical norms in collaborative problem solving. *Educational Studies in Mathematics*, 66(2), 123–146. <https://doi.org/10.1007/s10649-006-9070-6>
- Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. Jossey-Bass.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2019). *Qualitative data analysis: A methods sourcebook* (4th ed.). SAGE.
- Millî Eğitim Bakanlığı [MEB]. (2024). Türkiye Yüzyılı Maarif Modeli matematik dersi öğretim programı. Millî Eğitim Bakanlığı Yayınları.
- National Council of Teachers of Mathematics [NCTM]. (2024). *Principles to actions: Ensuring mathematical success for all*. NCTM.
- Organisation for Economic Co-operation and Development [OECD]. (2024). *PISA 2024 assessment and analytical framework*. OECD Publishing. <https://doi.org/10.1787/pisa-2024-en>
- Özenç, M. (2022). Öğretmenlerin problem çözme süreçlerinde pedagojik uygulama bilgisi: Bir durum çalışması. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, 42(1), 355–382. <https://doi.org/10.17152/gefad.1000000>
- Polya, G. (1945). *How to solve it: A new aspect of mathematical method*. Princeton University Press.
- Polya, G. (1954). *Mathematics and plausible reasoning*. Princeton University Press.
- Schoenfeld, A. H. (2010). How we think: A theory of human decision-making, with a focus on teaching. *ZDM – The International Journal on Mathematics Education*, 42(3–4), 457–465. <https://doi.org/10.1007/s11858-010-0205-6>
- Schoenfeld, A. H. (2011). *How we think: A theory of goal-oriented decision making and its educational applications*. Routledge.
- Schoenfeld, A. H. (2016). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. *Journal of Education*, 196(2), 1–38. <https://doi.org/10.1177/002205741619600202>
- Simon, H. A. (1997). *Administrative behavior: A study of decision-making processes in administrative organizations* (4th ed.). Free Press.
- Stake, R. E. (2005). Qualitative case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (3rd ed., pp. 443–466). SAGE.
- Stigler, J. W., & Hiebert, J. (1997). Understanding and improving classroom mathematics instruction: An overview of the TIMSS video study. *Phi Delta Kappan*, 79(1), 14–21.
- Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education*, 27(4), 458–477. <https://doi.org/10.2307/749877>
- Yin, R. K. (1994). *Case study research: Design and methods* (2nd ed.). SAGE. (Original work published 1984)



A Mathematical Comparison of SLAM Families: Observability, Information-Theoretic Bounds, and Optimization Landscapes

Burak Ağgöl¹ and Hüseyin Gökal²

¹ Sakarya University of Applied Sciences

Vocational School of Information Technologies, Sakarya, Karasu, Türkiye, burakaggul@subu.edu.tr

² Sakarya University of Applied Sciences

Vocational School of Information Technologies, Sakarya, Karasu, Türkiye, huseyingokal@subu.edu.tr

ABSTRACT

Simultaneous Localization and Mapping (SLAM) is a fundamental capability in autonomous robotics, enabling agents to estimate their trajectories while building maps of unknown environments. Several algorithmic families have been developed, including filter-based methods, particle-based approaches, optimization-based formulations, and density-based techniques.

Most comparative studies rely on trajectory-specific metrics such as Absolute Pose Error (APE) and Relative Pose Error (RPE). While useful, these do not fully explain the theoretical factors shaping performance. This study introduces a unified mathematical framework for comparison. First, observability and consistency are examined, showing how gauge freedoms cause inconsistency and how invariant formulations preserve stability. Second, information-theoretic bounds are explored through Fisher Information Matrices and Posterior Cramér–Rao Lower Bounds, revealing dataset-independent accuracy limits and the impact of graph connectivity. Third, the optimization landscape is analyzed, demonstrating how sparse solvers achieve scalability, while semidefinite relaxations provide certifiable global optimality under realistic noise conditions.

The analysis highlights trade-offs across SLAM families: particle-based methods remain effective in sparse measurement regimes, invariant filtering ensures robustness on Lie-group representations, and optimization-based approaches excel when loop closures provide strong conditioning. Density-based methods are robust in repetitive environments but remain highly sensitive to parameterization choices.



By bridging empirical evaluation with rigorous mathematical analysis, this framework provides principled guidelines for algorithm selection and the design of active SLAM strategies.

Key Words: SLAM, Observability, Fisher Information, Cramér–Rao Bound, Pose Graph Optimization, Autonomous Robotics.

1. INTRODUCTION

Simultaneous Localization and Mapping (SLAM) has emerged as a central problem in robotics, enabling autonomous agents to construct environmental representations while concurrently estimating their own trajectories. This dual-estimation capability is essential in domains such as autonomous driving, service robotics, aerial surveying, and planetary exploration, and continues to motivate new advances from deep learning–based pipelines (Teed & Deng, 2021; Yan et al., 2024; Tosi et al., 2024) to distributed multi-robot frameworks (Tian et al., 2023).

Over the past two decades, several algorithmic families have been introduced. Filter-based approaches, such as the Extended Kalman Filter (EKF-SLAM), established the probabilistic formulation of the problem (Bailey & Durrant-Whyte, 2006a, 2006b). However, EKF-SLAM is known to suffer from inconsistency due to linearization errors, which has been rigorously analyzed using observability-based methods (Huang, Mourikis, & Roumeliotis, 2010). Extensions such as the First-Estimate Jacobian EKF and the Invariant EKF have been proposed to improve consistency and stability (Barrau & Bonnabel, 2017).

Particle-based approaches, introduced with FastSLAM, exploit Rao–Blackwellization to factorize the posterior distribution, thereby improving scalability. FastSLAM 1.0 demonstrated this factored solution (Montemerlo, Thrun, Koller, & Wegbreit, 2002), while FastSLAM 2.0 refined the proposal distribution to reduce variance and improve accuracy (Montemerlo et al., 2003).

Optimization-based formulations, commonly known as Graph-SLAM or Pose Graph Optimization (PGO), leverage factor-graph representations and sparse linear algebra for large-scale efficiency. Key advances include Square Root SAM (Dellaert & Kaess, 2006), which applied sparse algebra to smoothing, and iSAM2, which introduced incremental relinearization

for near real-time updates (Kaess, Johannsson, Roberts, Ila, Leonard, & Dellaert, 2012). More recently, semidefinite relaxations such as SE-Sync have provided certifiable global optimality guarantees (Rosen, Carlone, Bandeira, & Leonard, 2018).

Building on this, **certifiably correct range-aided SLAM (CORA)** (Papalia et al., 2024) and its distributed extension (Thoms, Rosen, & Carlone, 2025) provide stronger guarantees for scenarios with ranging sensors and multi-robot coordination. Recently, deep-learning based pipelines such as DROID-SLAM (Teed & Deng, 2021) and NeRF/3D Gaussian Splatting approaches (Tosi et al., 2024; Gaussian-SLAM, 2024; SplatMap, 2025) have reshaped dense SLAM, complementing these classical families.

Density-based techniques, such as the Normal Distributions Transform (NDT-SLAM), employ continuous occupancy representations for scan alignment (Biber & Straßer, 2003). Later work on 3D NDT-OM extended this method to large-scale dynamic mapping (Saarinen, Andreasson, Stoyanov, & Lilienthal, 2013). These density-based strategies now coexist with learning-based dense representations such as Gaussian Splatting, suggesting a convergence between geometric and neural mapping approaches (Tosi et al., 2024; Yan et al., 2024). Figure 1 provides an overview of the major SLAM algorithm families, illustrating their representative methods from classical filter- and particle-based approaches to recent optimization, density, and learning-based pipelines.

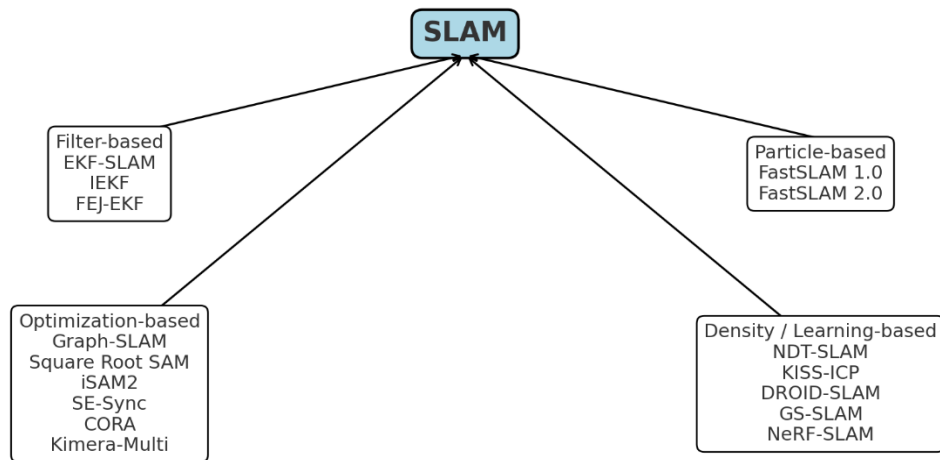


Figure 1. Overview of major SLAM algorithm families



Most comparative studies of SLAM rely on empirical evaluation, using metrics such as Absolute Pose Error (APE), Relative Pose Error (RPE), and map consistency. While indispensable, these measures remain highly dependent on dataset-specific characteristics and experimental conditions, thereby limiting their ability to explain fundamental theoretical trade-offs across algorithmic families. To address this gap, the present study develops a unified mathematical framework for comparing SLAM algorithms. The proposed framework is organized around three complementary perspectives:

(i) Observability and estimator consistency, analyzing gauge freedoms and their impact on filter stability. (ii) Information-theoretic performance bounds, based on Fisher Information Matrices and Cramér–Rao limits. (iii) Optimization landscape properties, focusing on scalability, conditioning, and certifiability in graph-based formulations.

2. RELATED WORK

The SLAM problem has been extensively investigated, giving rise to several distinct algorithmic families. Table 1 summarizes their representations, strengths, weaknesses, and complexities. Here, we briefly review the most influential contributions that provide the foundation for our theoretical comparison.

Filter-based methods established the probabilistic formulation of SLAM. The original EKF-SLAM formulation was introduced in early surveys (Bailey & Durrant-Whyte, 2006a, 2006b). However, EKF-SLAM suffers from inconsistency due to linearization errors, a limitation that was formally analyzed through observability-based rules (Huang, Mourikis, & Roumeliotis, 2010). Improved variants include the First-Estimate Jacobian EKF, which preserves the nullspace for consistency (Huang et al., 2010), and the Invariant EKF, which leverages Lie-group formulations to enhance stability (Barrau & Bonnabel, 2017). **Recent distributed frameworks such as Kimera-Multi further extend these formulations to multi-robot settings** (Tian et al., 2023).

Particle-based approaches introduced Rao–Blackwellized particle filters for scalable SLAM solutions. FastSLAM 1.0 applied this factorization to improve efficiency (Montemerlo, Thrun, Koller, & Wegbreit, 2002), while FastSLAM 2.0 refined the proposal distribution for



reduced variance and improved convergence (Montemerlo et al., 2003). Despite scalability, these approaches remain vulnerable to particle degeneracy under sparse measurement updates.

Optimization-based methods reformulated SLAM as a smoothing and mapping problem. Square Root SAM demonstrated the use of sparse linear algebra for large-scale optimization (Dellaert & Kaess, 2006), while iSAM2 introduced incremental relinearization and Bayes tree factorization for near real-time performance (Kaess, Johannsson, Roberts, Ila, Leonard, & Dellaert, 2012). More recently, semidefinite programming relaxations such as SE-Sync have provided certifiable global optimality under bounded noise conditions (Rosen, Carlone, Bandeira, & Leonard, 2018). Extensions to range-aided and distributed scenarios further strengthen certifiability (Papalia et al., 2024; Thoms et al., 2025).

Density-based techniques extend SLAM to continuous representations. The Normal Distributions Transform (NDT) was proposed as a robust scan-matching framework for repetitive geometries (Biber & Straßer, 2003). Later extensions such as 3D NDT-OM enabled large-scale online mapping in dynamic environments (Saarinen, Andreasson, Stoyanov, & Lilienthal, 2013). More recent pipelines such as KISS-ICP demonstrate robust LiDAR odometry without extensive parameter tuning (Vizzo et al., 2022/2023), while NeRF/3DGS-based methods enable dense photorealistic mapping (Tosi et al., 2024; Gaussian-SLAM, 2024; SplatMap, 2025).

Hybrid and sensor-fusion approaches combine multiple modalities to overcome individual sensor limitations. Visual-Inertial Odometry (VIO) improves scale observability in monocular systems, while LiDAR-IMU fusion enhances six-degree-of-freedom state estimation. Although these methods improve robustness, accurate extrinsic calibration remains a major challenge in practice (Khosoussi, Huang, & Dissanayake, 2016, 2019).

In summary, each SLAM family demonstrates unique strengths and limitations. Yet, most prior comparisons remain empirical. In contrast, the present study introduces a unified mathematical framework that clarifies theoretical trade-offs beyond dataset-specific evaluations.

Table 1. Taxonomy of SLAM Families and Their Theoretical Characteristics

Family	Representation	Strengths	Weaknesses	Complexity	References
EKF-SLAM	State-space filter	Simple recursive update; direct covariance estimate	Inconsistency due to linearization; fragile gauge handling	$O(N^2)$	Bailey & Durrant-Whyte (2006a, 2006b); Huang et al. (2010)
FEJ-EKF	EKF w/ fixed Jacobians	Preserves nullspace; improves consistency	More complex Jacobian management	$O(N^2)$	Huang et al. (2010)
Invariant EKF	Lie-group observer	Gauge-invariant error; improved stability	Requires Lie-group machinery/calibration	$O(N^2)$	Barrau & Bonnabel (2017)
FastSLAM 1.0	RBPF + per-landmark EKFs	Factorization enables scalability	Particle degeneracy under sparse updates	$O(M \log N)$	Montemerlo et al. (2002)
FastSLAM 2.0	Improved RBPF	Better proposal; lower variance	Still particle-intensive; tuning sensitive	$O(M \log N)$	Montemerlo et al. (2003)
Square Root SAM	Batch graph optimization	Sparse algebra; good scaling	Global relinearization may be needed	$O(N^{1.5})$ (typical)	Dellaert & Kaess (2006)
iSAM2	Incremental smoothing	Near real-time; reordering via Bayes tree	Memory growth if unchecked	$\approx O(N)$ amortized	Kaess et al. (2012)
SE-Sync	SDP relaxation (PGO)	Certifiable global optimum (below noise thresholds)	Tightness depends on noise/graph	Polynomial-time (relax.)	Rosen et al. (2018)
NDT (2D)	Gaussian grid	Robust matching repetitive geometry	scan in Parameter (voxel/resolution) sensitive	Moderate	Biber & Straßer (2003)
3D OM	NDT- 3D occupancy (dynamic)	Large-scale online mapping	Heavier computation; memory intensive	Moderate	Saارينen et al. (2013)

3. PRELIMINARIES AND NOTATION

We consider the SLAM problem in discrete-time state-space form. The system dynamics and observation models are defined as:

$$x_{k+1} = f(x_k, u_k) \oplus w_k, \quad w_k \sim \mathcal{N}(0, Q_k) \quad (1)$$

$$z_k = h(x_k, m) + v_k, \quad v_k \sim \mathcal{N}(0, R_k) \quad (2)$$

where x_k is the robot state at time k , u_k the control input, and m denotes the map landmarks. The operator $\oplus$ indicates Lie-group composition on $SE(2)$ or $SE(3)$ depending on 2D/3D formulation.

In the factor-graph formulation, the posterior distribution factorizes as a product of motion and measurement factors. Applying Gauss–Newton linearization yields the normal equations:

$$H \delta x = J^T R^{-1} r \quad (3)$$

where $H = J^T R^{-1} J$ is the system Hessian (or information matrix), J is the stacked Jacobian of residuals, R is the block-diagonal covariance, and r is the residual vector.

This formulation provides the mathematical basis for analyzing observability, Fisher information, and optimization properties in later sections.

4. THEORETICAL FRAMEWORK

In this section we introduce the theoretical principles used for the comparison of SLAM families. The framework is organized along three complementary perspectives: (i) observability and consistency, (ii) information-theoretic bounds, (iii) optimization landscape analysis.

4.1 Observability and Consistency

In Simultaneous Localization and Mapping (SLAM), certain components of the system state are intrinsically unobservable unless appropriate reference anchors are introduced. Specifically, in two-dimensional formulations, the global pose (x, y, θ) is unobservable, whereas in three-dimensional settings the complete configuration $(x, y, z, \text{roll}, \text{pitch}, \text{yaw})$ is affected by gauge freedoms.

Let $\mathcal{O}$ denote the observability matrix of the linearized system around a nominal trajectory. The condition for local observability is expressed as;

$$\text{rank}(\mathcal{O}) = n - d_g \quad (4)$$

where n is the dimension of the state vector and d_g denotes the dimension of the gauge subspace. Violation of this condition results in rank deficiency, the presence of unobservable subspaces, and consequently ill-posed covariance estimates in the estimator.

Extended Kalman Filter (EKF)-based SLAM frequently exhibits inconsistency due to repeated linearization, which produces overconfident covariances (Huang, Mourikis, & Roumeliotis, 2010). To address this, the First-Estimate Jacobian EKF (FEJ-EKF) preserves the correct nullspace by freezing Jacobians at their initial linearization (Huang et al., 2010). Alternatively, the Invariant Extended Kalman Filter (IEKF) redefines the estimation error on the Lie group SE(2) or SE(3), ensuring that gauge freedoms remain properly unobservable and improving estimator stability in both theory and practice (Barrau & Bonnabel, 2017).

In conclusion, observability analysis not only formalizes the role of gauge freedoms in SLAM but also provides principled guidelines for the development of consistent and reliable state estimators.

4.2 Information-Theoretic Bounds

The Fisher Information Matrix (FIM) quantifies the information content of measurements. Under Gaussian noise:

$$F = \sum_k J_k^\top R_k^{-1} J_k \quad (5)$$

where $J_k = \partial h / \partial \theta$ is the measurement Jacobian. The Cramér–Rao Lower Bound (CRLB) satisfies:

$$F^{-1} = F^\dagger \quad (6)$$

with $F^\dagger$ denoting the Moore–Penrose pseudoinverse (required when gauge freedoms make J rank-deficient).

For dynamic filtering, the Posterior CRLB (PCRLB) evolves via the Tichavský recursion:

$$F_{k+1} = Q_k^{-1} + F_k^\top F_k + H_{k+1}^\top R_{k+1}^{-1} H_{k+1} \quad (7)$$

where $F_k = \partial f / \partial x$ and $H_{k+1} = \partial h / \partial x$. Q_k is the process noise covariance. Gauge-aware projection must be applied before inversion to ensure consistency.

In pose-graph SLAM, blocks of the information matrix align structurally with a weighted graph Laplacian; determinants and eigenvalues relate to weighted spanning tree counts and

algebraic connectivity. Consequently, D-optimality criteria and reliability admit graph-theoretic interpretations. These insights align with graph-theoretic analyses linking Laplacian eigenvalues and spanning tree counts to SLAM reliability (Khosoussi, Huang, & Dissanayake, 2016, 2019).

This analysis implies that higher loop-closure density (stronger connectivity) improves information gain ($\Delta \log \det F$), reduces uncertainty ($\text{trace}(F^{-1})$) and yields better conditioning of the estimation problem.

4.3 Optimization Landscape and Certifiability

Graph-based SLAM (PGO) estimates poses $\{R_i, t_i\}$ from relative measurements by minimizing a nonconvex objective due to $SO(d)$ constraints:

$$\min_{\{R_i, t_i\}} \sum_{(i,j) \in \mathcal{E}} \| z_{ij} \ominus ((R_j, t_j))^{-1} \oplus (R_i, t_i) \|_{\Sigma_{ij}^{-1}}^2 \quad (8)$$

Sparse linear algebra, square Root SAM exploits sparsity of the information matrix for efficient batch optimization (Dellaert & Kaess, 2006). iSAM2 performs incremental relinearization and variable reordering using the Bayes tree, achieving near real-time updates (Kaess, Johannsson, Roberts, Ila, Leonard, & Dellaert, 2012).

Certifiable relaxations, semidefinite relaxations such as SE-Sync provide convex surrogates to PGO; below problem-dependent noise thresholds, the relaxation is *tight* and returns a solution with an a posteriori optimality certificate. This connects SLAM with certifiable nonconvex optimization and clarifies when global optimality can be guaranteed.

Hence, with sufficient loop closures and well-behaved noise, the optimization landscape becomes well-conditioned for Gauss–Newton or Levenberg–Marquardt methods, and semidefinite relaxations (e.g., SE-Sync) can certify global optimality. Recent work extends these relaxations to range-aided (Papalia et al., 2024) and distributed multi-robot settings (Thoms et al., 2025), enhancing practical certifiability.

5. COMPLEXITY AND THEORETICAL COMPARISON

This section compares SLAM families from two perspectives: (i) computational complexity, and (ii) scenario-based theoretical performance. The analysis is deliberately

dataset-independent and based only on model Jacobians, symbolic metrics, and asymptotic properties.

5.1 Asymptotic Complexity

The asymptotic computational cost of the main SLAM families is summarized in Table 2. The theoretical framework thus unifies three perspectives (Montemerlo, Thrun, Koller, & Wegbreit, 2002; Montemerlo et al., 2003):

Table 2. Computational Complexity of Major SLAM Families

Method	Update Cost	Notes
EKF-SLAM	$O(N^2)$	Dense covariance; poor scalability
FastSLAM (RBPf)	$O(M \log N)$	Depends on particle count (M) and proposal
Graph-SLAM / PGO	$O(N) - O(N^{1.5})$	Sparse solvers; scalable with iSAM2
NDT-SLAM	Moderate	Scan alignment cost depends on resolution/window

The discussion highlights several important points. EKF-SLAM is conceptually simple but its computational complexity grows quadratically with the number of landmarks due to dense cross-covariances. FastSLAM addresses scalability by factorizing the problem through Rao–Blackwellization, although its performance strongly depends on the particle count and the quality of the proposal distribution. Graph-based methods achieve near-linear runtime by exploiting sparsity, particularly when implemented with incremental solvers such as iSAM2. NDT-SLAM, on the other hand, provides moderate scalability, with performance largely influenced by voxel resolution and map window size (Biber & Straßer, 2003; Saarinen, Andreasson, Stoyanov, & Lilienthal, 2013).

5.2 Scenario-Based Theoretical Evaluation

Instead of empirical trajectories, we evaluate algorithm families under canonical SLAM scenarios using symbolic metrics:

- i. $\Delta \log \det(F)$ - information gain
- ii. $\text{trace}(F^{-1})$ - uncertainty lower bounds
- iii. $\kappa(H)$ - Hessian condition number (problem conditioning)

Several canonical scenarios can be distinguished when evaluating SLAM families. In the absence of loop closures, EKF-SLAM typically drifts over time, Graph-SLAM becomes ill-conditioned, and FastSLAM suffers from particle degeneracy when measurements are scarce (Montemerlo et al., 2002, 2003). Under sparse loop closures, FastSLAM demonstrates relative robustness because factorization reduces particle variance, whereas EKF continues to exhibit inconsistency over long horizons (Huang, Mourikis, & Roumeliotis, 2010). When loop closures are frequent, Graph-SLAM and PGO achieve strong conditioning, and semidefinite relaxations such as SE-Sync can provide certifiable optimality (Dellaert & Kaess, 2006; Kaess et al., 2012). Finally, in multi-sensor configurations, inertial aiding enhances observability; the Invariant-EKF maintains consistency by respecting Lie-group structure, while LiDAR-IMU fusion further strengthens six-degree-of-freedom (6DOF) observability (Khosoussi, Huang, & Dissanayake, 2016, 2019). Distributed frameworks such as Kimera-Multi demonstrate resilience to communication dropouts in multi-robot scenarios (Tian et al., 2023). Algorithm 1 summarizes the symbolic PCRLB evaluation procedure, which propagates Fisher information under gauge-aware projection.

Algorithm 1. Gauge-aware PCRLB Evaluation (Model-Only)

- 1: **Input:** Trajectory $\{\hat{x}_k\}$, models f, h , noise Q_k, R_k
 - 2: initialize $J_0 \leftarrow \epsilon I$ $\blacktriangleright$ small regularization or anchor to remove gauge
 - 3: **for** $k = 0$ to $K - 1$ **do**
 - 4: compute $F_k = \partial f / \partial x, H_{k+1} = \partial h / \partial x$
 - 5: $J_{k+1} \leftarrow Q_k^{-1} + F_k^\top J_k F_k + H_{k+1}^\top R_{k+1}^{-1} H_{k+1}$
 - 6: project onto observable subspace (remove gauge freedoms)
 - 7: **end for**
 - 8: **Output:** $J_k^{-1} = J_k^\dagger$ information gains, condition numbers
-

The complexity and scenario-based evaluation reveal several structural trade-offs among different SLAM families. EKF-SLAM is simple and suitable for online operation, but it suffers from quadratic computational scaling and inconsistency issues. FastSLAM performs favorably in sparse measurement regimes, although particle degeneracy remains a critical limitation. Graph-based methods and PGO are scalable and robust when loop closures are



present, and semidefinite relaxations further enable certifiability under suitable conditions. NDT-SLAM provides robustness in scan alignment, yet its performance is highly sensitive to parameterization choices and map resolution. Collectively, these comparisons offer theoretical guidance for selecting an appropriate SLAM family under specific operating regimes.

6. DISCUSSION AND LIMITATIONS

The proposed framework offers several advantages compared to purely empirical benchmarks. By analyzing observability, information-theoretic bounds, and optimization landscapes, we can explain *why* certain SLAM families perform better under specific conditions, without relying on dataset-specific outcomes. This provides a principled basis for algorithm selection, sensor placement, and the design of active SLAM strategies.

6.1 Advantages

- (i) Dataset-independence: The framework relies only on system models, Jacobians, and noise assumptions, avoiding dataset-specific biases.
- (ii) Unified perspective: Filtering, particle, graph-based, and density-based families are compared under common mathematical tools.
- (iii) Design guidance: The framework highlights how loop closure density, gauge freedoms, and sensor fusion impact the fundamental performance limits of SLAM.

6.2 Limitations

Despite its strengths, the framework has several limitations:

- (i) Abstracted data association: Correct correspondences between measurements and landmarks are assumed. In practice, outliers and wrong associations significantly affect SLAM.
- (ii) Static-world assumption: The current analysis excludes dynamic environments with moving objects or non-rigid scenes.



(iii) Relaxation tightness conditions: Certifiable guarantees from semidefinite relaxations depend on noise thresholds that are only partially characterized. An emerging limitation is that classical SLAM families do not fully exploit photorealistic dense representations. Recent NeRF and Gaussian Splatting-based approaches address this gap, though real-time scalability remains an open challenge (Tosi et al., 2024; Gaussian-SLAM, 2024; SplatMap, 2025).

6.3 Future Work

This study establishes a solid foundation for several future research directions:

- i. Developing new mathematical performance metrics that combine FIM determinants, condition numbers, and uncertainty bounds into unified criteria.
- ii. Extending the framework to dynamic SLAM with moving objects and online data association.
- iii. Validating the theoretical predictions with empirical experiments on real-world datasets and simulation platforms.
- iv. Integrating the framework with simulation environments such as Gazebo and Carla to empirically validate theoretical findings.

7. CONCLUSION

This paper presented a unified mathematical framework for comparing major SLAM families, including EKF-SLAM, particle-based FastSLAM, optimization-based Graph-SLAM/PGO, and density-based NDT-SLAM. Unlike empirical benchmarks that depend on datasets and experimental conditions, our approach focused on three theoretical perspectives: (i) observability and estimator consistency, (ii) information-theoretic limits via Fisher Information and Cramér–Rao bounds, and (iii) optimization landscape properties with certifiable relaxations.

The analysis revealed clear theoretical trade-offs. EKF-SLAM is simple but suffers from inconsistency due to linearization. FastSLAM achieves scalability through Rao–Blackwellization but is sensitive to particle degeneracy. Graph- and optimization-based SLAM provide scalable and certifiable solutions under sufficient loop closures, while NDT methods offer robustness in scan alignment but depend strongly on parameterization.



The main contribution of this work is to bridge the gap between empirical evaluations and rigorous theoretical analysis. By relying on symbolic metrics such as information gain, PCRLB, and Hessian conditioning, the framework clarifies when each SLAM family is preferable, independent of dataset-specific biases.

In future work, this framework will serve as a foundation for developing new mathematical performance models and unified metrics, enabling stronger theoretical guarantees for SLAM. Moreover, extending the analysis to dynamic environments and validating the predictions on real-world datasets will further strengthen the connection between theory and practice. Recent advances suggest several promising directions, including certifiably correct distributed SLAM (Papalia et al., 2024; Thoms et al., 2025), parameter-free LiDAR odometry pipelines such as KISS-ICP (Vizzo et al., 2022/2023), and dense neural scene representations leveraging NeRF and 3D Gaussian Splatting (Tosi et al., 2024; Gaussian-SLAM, 2024; SplatMap, 2025). Incorporating these paradigms within the proposed theoretical framework will help unify classical probabilistic methods with emerging data-driven approaches, advancing the next generation of SLAM systems.

REFERENCES

- Bailey, T., & Durrant-Whyte, H. (2006a). Simultaneous localization and mapping (SLAM): Part I. *IEEE Robotics & Automation Magazine*, 13(2), 99–110.
- Bailey, T., & Durrant-Whyte, H. (2006b). Simultaneous localization and mapping (SLAM): Part II. *IEEE Robotics & Automation Magazine*, 13(3), 108–117.
- Barrau, A., & Bonnabel, S. (2017). The invariant extended Kalman filter as a stable observer. *IEEE Transactions on Automatic Control*, 62(4), 1797–1812. <https://doi.org/10.1109/TAC.2016.2603153>
- Biber, P., & Strasser, W. (2003). The normal distributions transform: A new approach to laser scan matching. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 2743–2748). IEEE. <https://doi.org/10.1109/IROS.2003.1249285>
- Dellaert, F., & Kaess, M. (2006). Square root SAM: Simultaneous localization and mapping via square root information smoothing. *The International Journal of Robotics Research*, 25(12), 1181–1203. <https://doi.org/10.1177/0278364906072768>



- Huang, G., Mourikis, A. I., & Roumeliotis, S. I. (2010). Observability-based rules for designing consistent EKF SLAM estimators. *The International Journal of Robotics Research*, 29(5), 502–528. <https://doi.org/10.1177/0278364909353640>
- Kaess, M., Johannsson, H., Roberts, R., Ila, V., Leonard, J., & Dellaert, F. (2012). iSAM2: Incremental smoothing and mapping using the Bayes tree. *The International Journal of Robotics Research*, 31(2), 216–235. <https://doi.org/10.1177/0278364911430419>
- Khosoussi, K., Huang, S., & Dissanayake, G. (2016). Tree-connectivity: Evaluating the graphical structure of SLAM. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)* (pp. 1310–1317). IEEE. <https://doi.org/10.1109/ICRA.2016.7487259>
- Khosoussi, K., Huang, S., & Dissanayake, G. (2019). Reliable graphs for SLAM. *The International Journal of Robotics Research*, 38(2–3), 260–286. <https://doi.org/10.1177/0278364918823086>
- Montemerlo, M., Thrun, S., Koller, D., & Wegbreit, B. (2002). FastSLAM: A factored solution to the SLAM problem. In *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 593–598). AAAI Press.
- Montemerlo, M., Thrun, S., Koller, D., & Wegbreit, B. (2003). FastSLAM 2.0: An improved particle filtering algorithm for SLAM that provably converges. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)* (pp. 1151–1156). Morgan Kaufmann.
- Rosen, D. M., Carlone, L., Bandeira, A. S., & Leonard, J. J. (2018). SE-Sync: A certifiably correct algorithm for synchronization over the special Euclidean group. *The International Journal of Robotics Research*, 37(4–5), 439–465. <https://doi.org/10.48550/arXiv.1612.07386>
- Saارينen, J., Andreasson, H., Stoyanov, T., & Lilienthal, A. (2013). Normal distributions transform occupancy maps: Application to large-scale online 3D mapping. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)* (pp. 2233–2238). IEEE. <https://doi.org/10.1109/ICRA.2013.6630878>
- Tichavský, P., Muravchik, C. H., & Nehorai, A. (1998). Posterior Cramér–Rao bounds for discrete-time nonlinear filtering. *IEEE Transactions on Signal Processing*, 46(5), 1386–1396. <https://doi.org/10.1109/78.668800>
- Papalia, A., Fishberg, A., O'Neill, B. W., How, J. P., Rosen, D. M., & Leonard, J. J. (2024). Certifiably correct range-aided SLAM. *IEEE Transactions on Robotics*. <https://doi.org/10.1109/TRO.2024.3454430>
- Thoms, A., Rosen, D. M., & Carlone, L. (2025). Distributed certifiably correct range-aided SLAM. *arXiv preprint arXiv:2503.03192*. <https://arxiv.org/abs/2503.03192>
- Vizzo, I., Behley, J., & Stachniss, C. (2022). KISS-ICP: In defense of point-to-point ICP. *arXiv preprint arXiv:2209.15397*. <https://arxiv.org/abs/2209.15397>



- Vizzo, I., Guadagnino, T., Mersch, B., Wiesmann, L., Behley, J., & Stachniss, C. (2023). KISS-ICP: In defense of point-to-point ICP — Simple, accurate, and robust registration if done the right way. *IEEE Robotics and Automation Letters*, 8(2), 872–879. <https://doi.org/10.1109/LRA.2023.3236571>
- Teed, Z., & Deng, J. (2021). DROID-SLAM: Deep visual SLAM for monocular, stereo, and RGB-D cameras. In *Advances in Neural Information Processing Systems (NeurIPS)* (pp. xxxx–xxxx). Curran Associates. <https://arxiv.org/abs/2108.10869>
- Tosi, F., Zhang, Y., Gong, Z., Sandström, E., Mattoccia, S., Oswald, M. R., & Poggi, M. (2024). How NeRFs and 3D Gaussian Splatting are reshaping SLAM: A survey. *arXiv preprint arXiv:2402.13255*. <https://doi.org/10.48550/arXiv.2402.13255>
- Yan, C., Qu, D., Xu, D., Zhao, B., Wang, Z., Wang, D., & Li, X. (2024). GS-SLAM: Dense visual SLAM with 3D Gaussian splatting. *arXiv preprint arXiv:2311.11700*. <https://doi.org/10.48550/arXiv.2311.11700>
- Hu, Y., Liu, R., Chen, M., Beerel, P., & Feng, A. (2025). SplatMAP: Online dense monocular SLAM with 3D Gaussian splatting. *arXiv preprint arXiv:2501.07015*. <https://arxiv.org/abs/2501.07015>
- Tian, Y., Khosoussi, K., Rosen, D. M., & How, J. P. (2023). Kimera-Multi: A distributed metric-semantic SLAM system. *IEEE Transactions on Robotics*, 39(6), 4567–4583. <https://doi.org/10.1109/ICRA48506.2021.9561090>



Modeling Problem Solving Process Interaction Analysis: A Discourse Study on the Parking Lot Design Problem

Elif KARA EZAN¹, Nazan SEZEN YÜKSEL²

1 Ondokuz Mayıs University, Faculty of Education, Department of Mathematics and Science Education,
Samsun/Turkey, elif.kara@omu.edu.tr

2 Hacettepe University, Faculty of Education, Department of Mathematics and Science Education,
Ankara/Turkey, nsezen@hacettepe.edu.tr

ABSTRACT

This study aims to analyze the discourse moves exhibited by pre-service mathematics teachers while solving a mathematical modeling problem in a collaborative setting. The participants consisted of three third-year students from the elementary mathematics education program at a public university. They worked as a group to solve the “Parking Lot Design” modeling problem during a 90-minute session. All verbal interactions were audio-recorded and transcribed. The data were analyzed using discourse analysis across three dimensions: discourse channel, discourse level, and discourse type.

A total of 985 discourse moves were identified. The findings revealed that interpersonal discourse was dominant although personal discourse also played a notable role. Non-object level discourse was more prevalent than object-level discourse indicating a strong focus on process organization and group coordination beyond pure mathematics. In terms of discourse types, information sharing was the most frequent followed by questioning and confirmation which supports earlier findings that collaborative learning improves problem-solving through peer interaction and support. Participants adopted distinct roles: Participant 1 was primarily an information provider, Participant 2 acted as a supporter, and Participant 3 as a questioner and critic, reinforcing the importance of role diversity for productive mathematical discourse.

Overall, this study reinforces the critical role of discourse in mathematical problem-solving processes and underscores the value of structured collaboration. It further suggests that integrating discourse-rich modeling tasks into teacher education programs can enhance pre-service teachers’ mathematical reasoning, communication, and reflective thinking skills (). These insights contribute to the growing body of research advocating for discourse-based pedagogies within mathematics education.



Key Words: Mathematics education, discourse analysis, modeling, pre-service teachers, group interaction

1. INTRODUCTION AND PURPOSE OF THE STUDY

Understanding students' thinking and exploring their intellectual worlds are among the primary goals of mathematics education research. One of the most appropriate methods for achieving this goal is undoubtedly to consult students' discourse. Discourse is a process that encompasses all communicative activities an individual engages in with others or with oneself, and it involves students' participation in classroom discussions as well as their communication through written and visual representations (Sfard, 2001; NCTM, 2014). In the context of mathematical problem-solving, discourse analysis is crucial for understanding the dynamics of student interactions and the development of mathematical reasoning. For example, Robinson (2020) emphasizes the complexity of mathematical discourse and suggests that understanding these interactions is the key to fostering higher-order thinking in students. This complexity is important in educational settings because it shapes how students express and negotiate mathematical ideas.

The role of discourse in collaborative settings is also noteworthy. Studies show that collaborative learning models enhance mathematical problem-solving skills by encouraging peer interaction and discussion. Klang et al. (2021) state that peer acceptance and friendships significantly influence students' participation and success in solving mathematical problems. This finding aligns with those of Purba and Ramadhani (2020), who revealed that collaborative learning can significantly enhance students' problem-solving abilities and help them express their thought processes more effectively. Moreover, focusing on discourse not only aids in understanding mathematical concepts but also fosters critical and independent thinking (Syafri et al., 2020).

During the process of solving mathematical modelling problems, students go through various stages. Mardianti et al. (2018) emphasize that students with strong mathematical skills can effectively use the modelling process, highlighting the importance of conceptual frameworks in guiding students through complex problem-solving tasks. Furthermore, Jiang and Jia-Qing (2024) point out how the high school curriculum emphasizes the practical



application of mathematical concepts through modelling activities, serving to establish a connection between theoretical mathematics and real-world applications. Such connections are vital for students to develop a deeper understanding of the significance of mathematics in their everyday lives.

Mathematical discourse observed during modelling activities is crucial for student engagement and understanding. Problem-solving requires students to participate in a “simultaneous mathematization process,” in which they organize their understanding of mathematical concepts through structured discourse (Biza, 2017). This process is essential for constructing conceptual models that effectively solve specific problems. Additionally, Çelik and Baki (2023) highlight the fundamental moves in mathematical discourse, noting that the interactions between teachers and students during explanations shape the overall understanding of mathematical ideas. Such structured interactions form the basis for developing a solid mathematical discourse in classrooms and ultimately foster a culture of communication necessary for effective learning.

While the complex student interactions and discourse dynamics in the modelling process are critical for understanding the development of mathematical thinking, the existing literature offers limited studies that deeply examine group interactions among pre-service mathematics teachers. However, such interactions provide essential insights into how future teachers construct their problem-solving skills and communicate with their students. In this context, the main purpose of this study is to analyze the discourse moves that emerge in the group interactions of pre-service mathematics teachers while solving the “Parking Lot Design” problem and to examine how these moves influence the problem-solving process. In line with this aim, the study seeks to answer the following question:

“What discourse moves do pre-service mathematics teachers use during group interactions while solving a modelling problem?”

This research not only sheds light on how pre-service teachers think during the modelling process but also offers important implications for designing more effective collaborative learning environments in mathematics education.



2. METHODOLOGY

This section provides a detailed explanation of the research design, the characteristics of the participants, the data collection process, the data analysis method, and the measures taken to ensure the validity and reliability of the study.

2.1. Research Design

In this study, a case study design, one of the qualitative research methods, was employed to examine in depth the interactions of pre-service mathematics teachers during the process of solving a modelling problem. The case study method allows a specific phenomenon to be investigated in detail within its natural context (Creswell, 2018). The selection of this design provided an appropriate approach to interpret the complex dynamics of participants' interactions by seeking answers to the questions of “*how*” and “*why*.”

2.2. Participants

The study was conducted with three volunteer pre-service teachers (PST1, PST2, PST3) enrolled in the third year of the Elementary Mathematics Education undergraduate program at a public university. At this stage of their education, participants had already completed several subject-matter and pedagogy courses such as Mathematics Teaching Methods and Instructional Technologies, which provided them with basic competencies in problem-solving, communication, and group work. However, they had not yet taken any course specifically focused on mathematical modeling. This condition provided the opportunity to observe how participants structured the modeling process through natural interaction without relying on previously acquired modeling knowledge.

2.3. Data Collection

Data were collected during a 90-minute session in which the participants collaboratively solved a mathematical modelling problem titled “*Parking Lot Design*.” The session was conducted in a classroom setting to ensure that participants could interact comfortably in a natural environment. To enhance the depth of the data, a single data collection tool was used:



Audio Recordings:

Throughout the entire session, all verbal communications among participants were recorded using a high-quality microphone to capture speech clearly and accurately.

2.4. Data Analysis

Data analysis was carried out using an interaction analysis approach and qualitative content analysis techniques. The analysis process consisted of three main stages:

Transcription:

The audio recordings were transcribed verbatim, including all details of the conversations. The transcripts were carefully examined to identify key discourse moves that emerged during participant interactions. These moves were categorized according to discourse analysis frameworks proposed by Sfard (2001), and three main inductively derived coding dimensions:

- Discourse Channel: Specifies between whom the interaction occurs.
- Discourse Level: Defines the nature of the mathematical content in the discourse.
- Discourse Type: Identifies the communicative purpose of the discourse.

2.5. Coding and Analysis of the Data

Using the established coding scheme, each segment of the conversation was labeled with the appropriate codes. The coding process was conducted meticulously, focusing on who performed each discourse move and in what context. The coded data were transformed into findings supported by numerical distributions (frequency tables) and direct quotations selected from participants' dialogues.

The following table (Table 2.1) presents the definitions and examples of discourse moves identified in group interactions based on channel, level, and type.

Table 2.1. Code Description Table

Dimension	Code Name	Definition	Example Dialogue
Discourse Channel	Interpersonal	Moments when one group member directly interacts with other members.	PST1: “Should we use the data we have now?” PST2: “I think we should first draw a sketch.”
	Personal	Moments when a student thinks individually or works on the problem alone.	PST3 (muttering to herself): “A car is 2.5 meters...”
Discourse Level	Object Level	Instances where the discourse focuses directly on mathematical objects, concepts, operations, or numbers related to solving the parking lot problem.	PST1: “Given that the parking area is 1200 square meters, how many cars can fit into a 100-square-meter section?”
	Non-object Level	Instances where the discourse regulates the problem-solving process or social dynamics.	PST2: “Who will take notes? You can start from there.”
Discourse Type	Information Sharing	When a group member presents data, ideas, strategies, or results related to the problem.	PST1: “We can assume the average length of a car is 4.5 meters.”
	Affirmation	When a member expresses agreement, support, or approval of another’s idea or statement.	PST2: “Yes, that makes sense. Let’s continue that way.”
	Questioning	When a group member asks a question to seek information or clarify an idea.	PST3: “So, what percentage of this area will we use?”
	Disagreement / Rejection	When a group member expresses disagreement with another’s idea, strategy, or conclusion.	PST1: “But I think we can’t reach the right result using this method.”
	Strategy Suggestion	When a member proposes a new approach or method for solving the problem.	PST2: “I think we should first draw a sketch and then calculate the areas.”

2.6. Validity and Reliability

To ensure the validity and reliability of the qualitative research, the following measures were taken:

- Thick Description:

The data collection and analysis processes were described in detail so that readers could interpret the context and findings of the study independently.



- Inter-rater Reliability:

To calculate the reliability coefficient of the study, an additional expert other than the researcher was asked to independently code a selected portion of the data. The percentage of agreement between the two sets of codes was calculated. This percentage indicated that the coding scheme was consistent and reliable.

3. RESULTS

This section presents the analysis of the discourse moves exhibited by pre-service mathematics teachers during the process of solving the “*Parking Lot Design*” problem. The findings reveal the overall distribution of discourse moves, the distinct roles of participants within the group, and how these roles influenced the problem-solving process.

3.1. General Distribution and Dynamics of Discourse Moves

Table 3.1: Distribution of Discourse Channels

Discourse Channel	Frequency	Percentage (%)
Interpersonal	804	82
Personal	181	18

The analysis revealed a total of 985 discourse moves throughout the session. As shown in Table 3.1, the majority of these moves (82%) were interpersonal interactions. This indicates that participants tended to solve the problem through intensive collaboration rather than through individual efforts. On the other hand, the presence of personal discourse (18%) suggests that moments of individual thinking and calculation were also an essential part of the solution process.

Table 3.2: Distribution of Discourse Levels

Discourse Level	Frequency	Percentage (%)
Object Level	413	42
Non-object Level	572	58

The distribution of discourse levels is presented in Table 3.2. The findings indicate that non-object level discourse (58%) was more prevalent than object-level discourse (42%). This suggests that the group not only focused on mathematical calculations but also devoted considerable attention to organizing the problem-solving process, managing task distribution, and coordinating the overall flow of the activity.

Table 3.3: Distribution of Discourse Types

Discourse Type	Frequency	Percentage (%)
Information Sharing	402	41
Affirmation	186	19
Questioning	221	22
Disagreement / Rejection	83	8
Strategy Suggestion	93	9

As shown in Table 3.3, the most frequently observed discourse move was information sharing, indicating that students continuously generated and shared data, ideas, and strategies throughout the problem-solving process. The high occurrence of questioning suggests that participants actively sought clarification, validation, and guidance to direct the flow of the solution process.

3.2. Relationship Between Discourse Levels and Types

The analysis revealed which discourse types were more frequently associated with specific discourse levels. Object-level discourse was often related to information sharing and questioning, showing that students tried to solve the problem by sharing or examining mathematical ideas and calculations.

The following dialogue exemplifies this interaction:



PST1: “There’s an area measuring 19 by 13 here.”

PST2: “Then twelve cars would fit there.”

PST1: “Where’s the 13 then?”

PST3: “So how will we calculate this area?”

In contrast, non-object level discourse was commonly observed in strategy suggestion and affirmation moves. This indicates that the group used such discourse when making process-related decisions or reinforcing collaboration.

Example dialogue supporting this observation:

PST1: “I think we should first draw a sketch and then calculate the areas.”

PST2: “Yes, that makes sense. Let’s continue that way.”

3.3. Participant Roles and Interaction in the Problem-Solving Process

The analysis showed that each of the three participants assumed distinct yet complementary roles within the group. These roles were closely associated with the dominant discourse types used by each participant.

- **PST1 – Information Provider and Idea Generator:**
PST1 stood out as the participant who most frequently shared information. They continuously proposed ideas and strategies necessary for solving the problem, laying the foundation for object-level discourse essential for group progress.
- **PST2 – Supportive and Process Regulator:**
PST2 maintained group harmony by frequently using affirmation expressions such as “Yes, *that makes sense*” or “*Exactly, let’s continue.*” Most of PST2’s discourse occurred at the non-object level, indicating a focus not only on the problem but also on ensuring the smooth functioning of the process.
- **PST3 – Questioning and Guiding Participant:**
PST3 was the participant who most frequently asked questions and expressed disagreement. Through questions such as “*So what percentage of this area will we use?*”, PST3 encouraged peers to reflect and clarify their ideas. Disagreement



statements (e.g., “No, 4-meter cars won’t fit.”) promoted critical evaluation of ideas and strategies within the group.

The distribution of these roles can be explained not only by individual tendencies but also by the group’s internal dynamics and social relationships. Participants’ familiarity, prior friendships, and shared academic experiences likely fostered comfort and trust in communication, allowing these roles to emerge naturally and harmoniously. Therefore, if the same activity were applied to a different group, a similar role distribution might not necessarily occur, as group relations and social atmosphere significantly influence discourse dynamics.

4. DISCUSSION AND CONCLUSION

This study analyzed the discourse moves exhibited by pre-service mathematics teachers while solving the “Parking Lot Design” modelling problem, providing significant insights into the nature of group interactions and their influence on the problem-solving process.

The findings revealed that participants predominantly conducted the solution process through interpersonal discourse. This result aligns with Sfard’s (2001) discourse approach, which posits that mathematical meaning-making occurs through both social interaction and individual thinking. It also supports the findings of Klang et al. (2021) and Purba and Ramadhani (2020), who reported that collaborative learning enhances problem-solving skills. The participants’ strong tendency toward collaborative problem-solving is a positive indicator for their future professional practice.

In terms of discourse levels, the dominance of non-object level discourse is noteworthy. This finding demonstrates that students were not only engaged in mathematical calculations but also in organizing processes, developing strategies, and coordinating group dynamics. It confirms that modelling problems require not only mathematical knowledge but also process management skills.

Throughout the study, discourse moves were observed to evolve over time. Early in the session, organizational discourse (e.g., “Where should we start?”, “Should someone take notes?”) was more prominent. After understanding the problem and performing initial



calculations, object-level discourse became dominant. This dynamic indicates an increasing focus on conceptual reasoning and a rise in the “mathematical discourse intensity” of the group. Such temporal transitions in discourse align with Cobb and Yackel’s (1996) concept of intertwined *social construction* and *individual reasoning*. Early stages featured discourse aimed at social consensus, while later stages integrated more individual computation and strategic reasoning.

The analysis of discourse types revealed that while information sharing was dominant, questioning and affirmation played critical roles in facilitating understanding. This supports the views of Syafri et al. (2020) and Robinson (2020), who argue that mathematical discourse is a crucial tool for developing critical thinking and higher-order reasoning. Even less frequent discourse moves, such as strategy suggestion and disagreement, were significant for diversifying solution approaches and fostering critical evaluation.

The predominance of information sharing among participants indicates, in line with constructivist learning theory, that mathematical knowledge was constructed through social interaction within the group. However, another critical finding from the discourse analysis is the relative scarcity of disagreement moves. This may suggest limited opportunities for constructive debate and hesitation among students to challenge peers’ ideas. As emphasized by Syafri et al. (2020), to promote the development of critical inquiry, such disagreement discourse should be more encouraged in classroom contexts. Otherwise, excessive information sharing might evolve into a “*consensual routine*” rather than creative thinking.

Regarding participant roles, PST1 functioned as the information provider, PST2 as the supporter, and PST3 as the inquirer. Consistent with Çelik and Baki (2023), this diversity of roles within group discourse enriched mathematical thinking and the learning process. Particularly, PST3’s role in deepening the group’s thought processes through questioning and disagreement strongly underscores the value of critical inquiry.

5. RECOMMENDATIONS

This study comprehensively examined the discourse moves exhibited by pre-service mathematics teachers while solving a modelling problem and clearly revealed the role of group interactions in the problem-solving process. The findings showed that pre-service teachers



reached solutions primarily through collaborative discourse, while individual thinking and calculations also constituted an integral part of the process. Furthermore, the predominance of non-object level discourse highlighted that modelling problems require not only mathematical knowledge but also effective process management and collaboration skills.

In this context, based on the findings of this study, the following recommendations can be made:

- Teacher education programs should support modelling activities with group work and include more practices focused on discourse analysis.
- Activity designs should aim to balance individual and group discourse to enhance both individual reasoning and collaborative thinking skills among pre-service teachers.
- Critical inquiry and disagreement discourse should be explicitly encouraged during instructional processes. This will allow students to engage in deeper discussions, evaluate ideas critically, and develop alternative solutions.
- Future research can conduct similar analyses using different modelling problems or larger participant groups to increase the generalizability of the findings.

In conclusion, this study demonstrates that the discourse analysis approach in mathematics education is a powerful tool for understanding pre-service teachers' problem-solving and communication skills. The findings provide valuable implications for both teacher education programs and classroom teaching practices.

REFERENCES

- Bakhtin, M. M. (1981). *The dialogic imagination: Four essays* (M. Holquist, Ed.; C. Emerson & M. Holquist, Trans.). University of Texas Press.
- Biza, I. (2017). Youngsters solving mathematical problems with technology: The results and implications of the Problem@Web project. *Research in Mathematics Education*, 19(3), 331–335. <https://doi.org/10.1080/14794802.2017.1365008>
- Cobb, P., & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist*, 31(3–4), 175–190. <https://doi.org/10.1080/00461520.1996.9653265>



**International Conference on Mathematics and
Mathematics Education (ICMME-2025), İstanbul Medeniyet
University, İstanbul, Türkiye, September 11-13, 2025**
“Mathematics in İstanbul, Bridge Between Continents”



- Creswell, J. W. (2016). *Nitel araştırma yöntemleri: Beş yaklaşıma göre nitel araştırma ve araştırma deseni* (M. Bütün & S. B. Demir, Çev.). Siyasal Kitabevi.
- Çelik Demirci, S., & Baki, A. (2023). Characterizing mathematical discourse according to teacher and student interactions: The core of mathematical discourse. *Journal of Pedagogical Research*, 7(4), 144–164. <https://doi.org/10.33902/jpr.202321852>
- Jiang, Y., & Jia-qing, X. (2024). Exploration and reflection of mathematical modeling activity course in high school teaching. *International Journal of New Developments in Education*, 6(1). <https://doi.org/10.25236/ijnde.2024.060103>
- Klang, N., Karlsson, N., Kilborn, W., Eriksson, P., & Karlberg, M. (2021). Mathematical problem-solving through cooperative learning: The importance of peer acceptance and friendships. *Frontiers in Education*, 6, 710296. <https://doi.org/10.3389/feduc.2021.710296> □
- Mardianti, I., Masriyah, & Wijayanti, P. (2018). Students' creative thinking process based on the Wallas stage in solving mathematical modeling problems. In *Proceedings of the 2018 Mathematics, Informatics, Science, and Education International Conference (MISEIC 2018)*. Atlantis Press. <https://doi.org/10.2991/miseic-18.2018.38>
- National Council of Teachers of Mathematics. (2014). Principles to actions: Ensuring mathematical success for all. NCTM.
- Purba, A., & Ramadhani, S. (2020). The improvement of students' mathematical problem-solving ability by implementing a cooperative learning model in SMP Tunas Pelita Binjai. *Journal of Research on Mathematics Instruction (JRMI)*, 1(2), 31–37. <https://doi.org/10.33578/jrmi.v1i2.20>
- Robinson, E. (2020). Understanding meaningful exchanges: Mathematics discourse analysis and complexity thinking. *in education*, 26(1), 103–138. <https://doi.org/10.37119/ojs2020.v26i1.457>
- Sfard, A. (2001). There is more to discourse than meets the ears: Looking at thinking as communicating to learn more about mathematical learning. *Educational Studies in Mathematics*, 46(1–3), 13–57. <https://doi.org/10.1023/A:1014097416157>
- Syafri, M., Zulkarnain, Z., & Maimunah, M. (2020). The effect of SSCS learning model on the mathematical problem-solving ability of junior high school students, Kampar Regency. *Journal of Educational Sciences*, 4(2), 309–317. <https://doi.org/10.31258/jes.4.2.p.309-317>



Schur Stability Analysis of Second-degree Extended Legendre Polynomial Families

Güner Topcu¹

¹ Mathematics, Selcuk University, guner.ozturk@selcuk.edu.tr

ABSTRACT

In this paper, matrix families constructed from second-degree Legendre polynomials are examined. During the construction of these families, three different perturbation matrices E_{21} , E_{22} and $E_{21} + E_{22}$ corresponding to the second-degree Legendre polynomial were considered. Thus, stability intervals for three different matrix families were determined with the aid of continuity theorems. Subsequently, these families were extended using algorithms from the literature while preserving their Schur stability properties. As a result, three different second-degree Legendre-centered extended interval polynomial families were obtained. Finally, illustrative examples related to the topic were presented.

Key Words: Schur stability, Legendre polynomial, matrix family.

1. INTRODUCTION

Stability analysis is a fundamental tool for understanding the long-term behavior of dynamical systems. In discrete-time systems, Schur stability corresponds to the condition that all eigenvalues of the system matrix lie within the unit circle (Akın and Bulgak, 1998; Elaydi, 1999). This ensures that the system responses remain bounded over time and converge to a stable equilibrium. However, for non-symmetric matrices, the eigenvalue problem is an ill-conditioned problem (Wilkinson, 1965; Bulgak, 1999). Therefore, for Schur stability analysis, it becomes necessary to solve the Lyapunov matrix equation (Akın and Bulgak, 1998; Elaydi, 1999; Aydın et al., 2000). Thus, the concepts of stability quality and sensitivity gain importance (Bulgak and Godunov, 1988; Akın and Bulgak, 1998; Bulgak, 1999).

On the other hand, Legendre orthogonal polynomials provide a powerful analytical framework in system theory due to their well-structured properties (Rainville, 1960). In this study, the Schur stability analysis of a matrix family centered on the Legendre polynomials has been conducted. In particular, second-degree Legendre polynomials have been selected and examples have been provided. The fundamental literature related to this topic is presented below.



1.1 Second-Degree Legendre Polynomials

Orthogonal polynomials play a central role in approximation theory, numerical analysis, and mathematical physics. Among them, Legendre polynomials P_N form a particularly important class due to their orthogonality on the interval $[-1,1]$. Moreover, they are significant due to their wide range of applications in mathematical physics.

Legendre polynomials $\{P_N(x)\}_0^\infty$ can be defined by

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x),$$

with initial conditions

$$P_0(x) = 1, \quad P_1(x) = x.$$

The first few members of this sequence are

$$P_0(x) = 1,$$

$$P_1(x) = x,$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1),$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x),$$

$$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$$

(Rainville, 1960). In this study, second-degree Legendre polynomials P_2 have been considered.

1.2 From Polynomials to Companion Matrices

While Legendre polynomials are primarily studied for their orthogonality properties, it is often insightful to analyze them from an algebraic perspective.

In particular, the companion matrix provides a bridge between polynomial equations and linear algebra: the roots of a polynomial can be recovered as the eigenvalues of its associated companion matrix.

Given a monic polynomial of degree n ,

$$p_n(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0,$$

its companion matrix C_n is defined as

$$C_n = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ -a_0 & -a_1 & \cdots & -a_{n-2} & -a_{n-1} \end{pmatrix}$$

(Elaydi, 1999; Akın and Bulgak, 1998). Now we apply the construction to the second-degree Legendre polynomials.

Example 1. The second-degree Legendre polynomial is

$$P_2(x) = \frac{1}{2}(3x^2 - 1).$$

In monic form:

$$p_2(x) = x^2 - \frac{1}{3}.$$

The corresponding companion matrix:

$$C_2 = \begin{pmatrix} 0 & 1 \\ \frac{1}{3} & 0 \end{pmatrix}.$$

1.3 Matrix Families Centered on the Second-Degree Legendre Polynomial

Building upon this framework, matrix families centered on the second-degree Legendre polynomial are introduced. In this construction, the matrix A is obtained as the companion

matrix of the second-degree Legendre polynomial P_2 . The perturbation matrices B are defined as E_{21}, E_{22} and $E_{21} + E_{22}$. Based on these definitions, the families centered on the second-degree Legendre polynomial are expressed as

$$\mathcal{L}_2^1 = \mathcal{L}_2^1(A, B_1) = \{A(r_1) = A + r_1 B_1 \mid A = C_2, B_1 = E_{21}\},$$

$$\mathcal{L}_2^2 = \mathcal{L}_2^2(A, B_2) = \{A(r_2) = A + r_2 B_2 \mid A = C_2, B_2 = E_{22}\},$$

$$\mathcal{L}_2^3 = \mathcal{L}_2^3(A, B_3) = \{A(r_3) = A + r_3 B_3 \mid A = C_2, B_3 = E_{21} + E_{22}\}.$$

Here, E_{21}, E_{22} and $E_{21} + E_{22}$ are real matrices and defined as follows:

$$E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, E_{21} + E_{22} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}.$$

2. STABILITY OF THE MATRIX FAMILIES CENTERED ON THE SECOND-DEGREE LEGENDRE POLYNOMIAL

In this section, before giving the stability of the matrix families centered on the second-degree Legendre polynomial, give the continuity theorems for Schur stability. This theorem, which exists in the literature, determines the sensitivity of stability. Let's remember the family of Schur stable matrices as follows:

$$S_N = \{A \in M_N(\mathbb{C}) \mid \omega(A) < \infty\}.$$

Theorem 1. Let $A \in S_N$. If $\|B\| < \sqrt{\|A\|^2 + \frac{1}{\omega(A)}} - \|A\|$ then the matrix $A + B \in S_N$ and

$$\omega(A + B) \leq \frac{\omega(A)}{1 - (2\|A\| + \|B\|)\|B\|\omega(A)}$$

holds (Aydın et al., 2002; Duman and Aydın, 2011).

Now, considering Theorem 1, the following continuity theorem, obtained by Topcu and Aydın, is given below.

Theorem 2. If $A \in S_N$ and $B \in M_N(\mathbb{C})$ and $r \in \mathcal{J}_\mathcal{L} = [\underline{r}, \bar{r}]$ then $\mathcal{L}(A, B) \subset S_N$, where

$$-l = u = -\frac{\|A\|}{\|B\|} + \frac{1}{\|B\|} \sqrt{\|A\|^2 + \frac{1}{\omega(A)}}, \quad l \leq \underline{r} \leq \bar{r} \leq u$$

(Topcu and Aydın, 2023; Topcu and Aydın, 2024).

Example 2. Let's take the matrix families centered on the second-degree Legendre polynomial $\mathcal{L}_2^1, \mathcal{L}_2^2$ and $\mathcal{L}_2^3$, respectively. With the application of the Theorem 2, the following Schur stability intervals are obtained, respectively:

$$r_1 \in \mathcal{J}_1 = [-0.1998, 0.1998],$$

$$r_2 \in \mathcal{J}_2 = [-0.1998, 0.1998],$$

$$r_3 \in \mathcal{J}_3 = [-0.1413, 0.1413].$$

3. EXTENDING THE SCHUR STABILITY INTERVAL AND OBTAINING THE INTERVAL POLYNOMIALS

In the last example, the obtained intervals preserve the Schur stability of the matrix families $\mathcal{L}_2^1, \mathcal{L}_2^2$ and $\mathcal{L}_2^3$. However, upon further analysis, it has been observed that there also exist points outside these intervals that ensure the Schur stability. To include these points into the obtained intervals, algorithms have been given by Topcu and Aydın (Topcu and Aydın, 2023; Topcu and Aydın, 2024). By means of these algorithms, extended intervals $\mathcal{J}_1^e, \mathcal{J}_2^e$ and $\mathcal{J}_3^e$ have been obtained for the extended matrix families $\mathcal{L}_2^{1e}, \mathcal{L}_2^{2e}$ and $\mathcal{L}_2^{3e}$ while preserving the Schur stability. Subsequently, second-degree Legendre-centered interval polynomial families were obtained. Let us examine the following example.

Example 3. Let's take the second-degree Legendre polynomial $P_2(x) = \frac{1}{2}(3x^2 - 1)$ and the perturbation matrices B as follows,

$$A = C_2 = \begin{pmatrix} 0 & 1 \\ \frac{1}{3} & 0 \end{pmatrix},$$

$$B_1 = E_{21}, B_2 = E_{22}, B_3 = E_{21} + E_{22}.$$

An examination of Table 1 is given below. The matrices A, B and the parameter r^* is the input element selected by the users. l and u are the lower and upper bounds which are calculated using the continuity theorems. l^e and u^e are the extended lower and upper bounds obtained by the Algorithm for the Schur Stability. N^l and N^u indicate how many steps the algorithms stopped for lower and upper bounds, respectively.

Table 1. The values l, l^e, u, u^e for the data A, B, r^* .

A	B	r^*	l	l^e	N^l	u	u^e	N^u
C_2	B_1	0.01	-0.1998	-1.3167	9	0.1998	0.6514	6
		0.001		-1.3322	13		0.6647	9
	B_2	0.01	-0.1998	-0.6540	6	0.1998	0.6540	6
		0.001		-0.6650	9		0.6650	9
	B_3	0.01	-0.1413	-1.2102	21	0.1413	0.3265	3
		0.001		-1.3181	51		0.3329	5

For example, according to Table 1, let's take the matrices C_2 and B_3 :

- For $r^* = 0.001, l^e = -1.3181$ with 51 steps and $u^e = 0.3329$ with 5 steps.
- Extended interval is $J_3^e = [-1.3181, 0.3329]$.
- Interval companion matrix for the Schur stable matrix family $\mathcal{L}_2^{3e}$ is

$$C_2^{3e} = \begin{pmatrix} 0 & 1 \\ [-0.9847, 0.6662] & [-1.3181, 0.3329] \end{pmatrix}.$$

- Legendre-centered Schur stable interval polynomial is

$$\mathcal{P}_2^{3e}(x) = x^2 + [-0.3329, 1.3181]x + [-0.6662, 0.9847].$$

5. CONCLUSION

The main objective of this study is to construct second-degree Legendre-centered Schur stable interval polynomial families. First, the companion matrix corresponding to the second-degree Legendre polynomial is constructed. Using this companion matrix, matrix families $\mathcal{L}_2^1, \mathcal{L}_2^2$ and $\mathcal{L}_2^3$ are formed based on the linear combination of the companion matrix and the perturbation matrix $B_1 = E_{21}, B_2 = E_{22}$ and $B_3 = E_{21} + E_{22}$. With the aid of the continuity theorems, the intervals J_1, J_2 and J_3 that ensure the stability of these matrix families are



determined. Subsequently, these intervals are extended through the algorithms while preserving the Schur stability property. The extended interval companion matrices corresponding to the extended intervals are constructed. Finally, by means of the extended interval companion matrices, second-degree Legendre-centered extended Schur stable interval polynomial families have been obtained.

REFERENCES

- Akın, Ö. and Bulgak, H. (1998). *Linear difference equations and stability theory*. Selçuk University Applied Mathematics Research Center.
- Aydın, K., Bulgak, H. and Demidenko, G. V. (2000). *Numeric characteristics for asymptotic stability of solutions to linear difference with periodic coefficients*. Siberian Math. J., 41 (6), 1005-1014.
- Aydın, K., Bulgak, H. and Demidenko, G. V. (2002). *Asymptotic stability of solutions to perturbed linear difference equations with periodic coefficients*. Siberian Mathematical Journal, 43 (3), 389-401.
- Bulgakov, A. Ya. and Godunov, S. K. (1988). *Circle dichotomy of the matrix spectrum*. Siberia Math. J., 29, No: 5, pp. 59-70.
- Bulgak, H. (1999). *Pseudoeigenvalues, spectral portrait of a matrix and their connections with different criteria of stability*. Error Control and Adaptivity in Scientific Computing, pp. 95-124.
- Duman, A. and Aydın, K. (2011). *Sensitivity of Schur stability of systems of linear difference equations with constant coefficients*. Scientific Research and Essays, 6 (28), 5846-5854.
- Elaydi, N. (1999). *An introduction to difference equations*. Springer.
- Rainville, E. D. (1960). *Special Functions*. The Macmillan Company, New York.
- Topcu, G. (2023). *Stability of polynomial family and sensitivity of stability*. Ph.D. Thesis. Selçuk University.
- Topcu, G. and Aydın, K. (2024). *A new approach to construct and extend the Schur stable matrix families*. Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics, 73(2), 538-553.
- Wilkinson, J. H. (1965). *The Algebraic Eigenvalue Problem*. Clarendon Press.



Formal Graphs: Graphing via Formal Variables and Formal Evaluation

Anthony J. Ripa¹

1 Stony Brook University, 100 Nicolls Road,
Stony Brook, NY 11790, aripa@cs.stonybrook.edu

ABSTRACT

Classical Cartesian graphs are usually obtained by viewing y as a function of a real variable x , where the function is defined by *substituting* real numbers for x inside a syntactic expression. This approach is convenient but fragile: removable singularities (e.g. $y = (x^2 - 1)/(x - 1)$ at $x = 1$) and similar “indeterminate forms” arise from naive substitution into arbitrary expressions. These indeterminate forms, such as $0/0$, traditionally require limits or calculus for resolution. This paper introduces *formal graphs*, a lightweight algebraic method using *formal variables* and *formal evaluation* to handle such forms without invoking calculus. By treating expressions as elements of an algebraic structure (e.g. a polynomial ring $\mathbb{k}[x]$, a fraction field $\mathbb{k}(x)$, etc.) and graphing via formal evaluation, formal graphs provide a robust, exception-free visualization that aligns with intuitive algebraic simplifications. We demonstrate the method through examples, and show how standard calculus facts (e.g. the derivative as the simplified difference quotient at $h = 0$) fall out algebraically.

Key Words: Graphing, Expressions, Formal Variables.

1. INTRODUCTION

In classical mathematical graphing, functions are typically defined by syntactic substitution of real values into expressions. This approach, while intuitive, introduces artifacts known as *removable discontinuities* when unreduced syntactic forms yield indeterminate expressions like $0/0$. For instance, the expressions

$$\frac{x^2 - 1}{x - 1} \text{ at } x = 1, \quad \frac{(1 + h)^2 - 1}{h} \text{ at } h = 0$$

exhibit such discontinuities despite being algebraically equivalent to polynomials $x + 1$ and $2 + h$, respectively. These artifacts obstruct straightforward algebraic computation in calculus problems, particularly in difference quotients for derivatives.

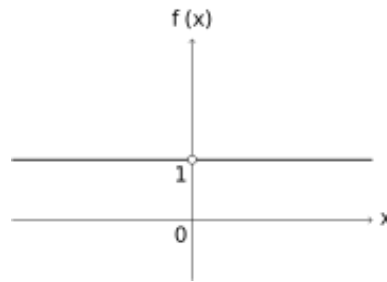


Figure 1. A traditional graph of x/x vs x , with undesirable hole at $x = 0$.

The core issue lies in the traditional evaluation homomorphism, which enforces pointwise substitution before algebraic simplification. This paper presents a paradigm shift: by treating symbols as formal variables (indeterminates) rather than traditional variables, and by defining evaluation algebraically instead of arithmetically, we obtain graphs free of removable singularities. Our method leverages the construction of the field of rational expressions $\mathbb{k}(x)$ as the fraction field of the polynomial ring $\mathbb{k}[x]$, ensuring that evaluation acts only as arithmetic substitution on canonical forms.

1.1 Contributions

- A rigorous framework for graphing based on formal variables and formal evaluation, eliminating removable discontinuities.
- Two equivalent constructions: (1) standard-form pairs with Euclid's algorithm for operations, and (2) equivalence classes of pairs.
- Pedagogical adaptations for high-school, undergraduate, and graduate levels.
- Applications to calculus, particularly in deriving derivatives algebraically via difference quotients.



2. RELATED WORK

The problem of *removable discontinuities* in graphs of rational functions is well known in both classical analysis and algebra. In analysis, a standard approach to “fixing” a hole in a function’s graph is to define a new function that agrees with the original formula except by filling in the undefined point with the limiting value. For example, if $f(x) = \frac{p(x)}{q(x)}$ and $x = a$ makes both p and q zero, one can define $\tilde{f}(a) := \lim_{x \rightarrow a} \frac{p(x)}{q(x)}$ (provided the limit exists) to obtain a continuous extension of f at a (Stewart, 2016; Apostol, 1967). In essence, analysis treats $\frac{p}{q}$ and its algebraically simplified form as the “same” function everywhere except the singular point, which is then patched by continuity. This idea appears throughout calculus textbooks and exercises, where students are asked to simplify expressions and evaluate limits to find the value that would make f continuous at a removable discontinuity (Stewart, 2016).

From an algebraic perspective, the treatment of rational expressions as formal objects inherently resolves removable singularities. In commutative algebra and algebraic geometry, a rational function is an element of a field of fractions $\mathbb{k}(x)$, meaning that two formulas like $\frac{x^2-1}{x-1}$ and $x+1$ represent the same element of $\mathbb{k}(x)$ (Dummit and Foote, 2004; Lang, 2002). Any factor that would cause a “hole” (such as $(x-1)$ in the denominator) is cancelled out in the field of fractions—indeed, $\frac{x^2-1}{x-1} = x+1$ in $\mathbb{k}(x)$ because one takes the reduced representative after cancelling the common factor. In this formal setting, x is an indeterminate rather than a varying real number, so questions of domain do not arise; there is no substitution into the raw expression until after it is reduced. This is precisely the insight employed by the Formal Graphs approach, and it has been implicitly used by mathematicians for a long time (e.g., when we say two rational expressions are “identical” if they differ only on a finite set of points in their naive evaluation).

Techniques for eliminating removable discontinuities also appear in computer algebra systems and advanced graphing tools. Many computer algebra systems (CAS) automatically or upon command simplify rational expressions by canceling common factors. For instance, SymPy’s `cancel` routine transforms a given rational expression into a canonical form p/q with



no common polynomial factors (SymPy Development Team, 2025). By working with such a simplified form, one can evaluate or graph the function at points that were formerly “holes” without encountering a $0/0$ error. In practice, however, most numerical graphing utilities do not perform this kind of symbolic preprocessing unless explicitly instructed; they typically sample points numerically and thus either omit the point or simply draw the curve without marking the removable discontinuity. The principle of Formal Graphs—symbolically simplify first, then evaluate—can be viewed as bringing a CAS-like step into the graphing process by default (Von Zur Gathen and Gerhard, 2013).

In mathematics education, the idea of “simplify before substitution” is a familiar theme that closely parallels the Formal Graphs philosophy. In limit computations and derivative calculations, students are taught to algebraically simplify expressions to avoid indeterminate forms. A classic example is the difference quotient in calculus: to find $f'(a)$ via $\lim_{h \rightarrow 0} \frac{f(a+h)-f(a)}{h}$, one simplifies $\frac{f(a+h)-f(a)}{h}$ to cancel h before setting $h = 0$ (Stewart, 2016; Apostol, 1967).

Likewise, in precalculus courses students learn that when graphing a rational function, they should factor and cancel common factors in the formula, then use the simplified form to identify any hole’s coordinates (e.g., graph $y = x + 1$ and mark an open point at $x = 1, y = 2$ for $y = \frac{x^2-1}{x-1}$ (Sullivan, 2016).

The concept of treating algebraic expressions in a domain where cancellation is done prior to evaluation is inherent in the algebra of rational functions (Dummit and Foote, 2004; Lang, 2002) and is the theoretical reason removable discontinuities can be bypassed. Analysis and calculus education provide the interpretation of filling in holes via limits or algebraic simplification (Stewart, 2016; Apostol, 1967). Computer algebra systems offer practical tools (e.g., cancel) that achieve the same outcome symbolically (SymPy Development Team, 2025; von zur Gathen and Gerhard, 2013). The contribution of the Formal Graphs approach is in unifying these insights into a coherent framework and explicitly applying them to graphing: it ensures that what might traditionally be seen as a “hole” in a graph is handled seamlessly by using the simplified representation of the function. In this sense, Formal Graphs build directly on well-established practices, but to our knowledge no prior work has articulated it as a general graphing methodology.



3. BACKGROUND

We recall standard facts; see, e.g. (Dummit and Foote, 2004; Lang, 2002).

3.1 Variables: Traditional vs. Formal.

3.1.1 Traditional Variables.

Traditional variables, as used in elementary algebra, are entities that vary and can be solved for in equations. For example, in $x^2 - 1 = 0$, the variable x varies until the equation balances, yielding solutions $x = 1$ and $x = -1$.

Unfortunately, traditional variables do not support unique factorization in polynomial expressions due to potential pointwise exceptions. While one factorization of $x^2 - 1$ is $(x - 1)(x + 1)$, this factorization is not unique. For example, consider the piecewise expression:

$$q(x) = \begin{cases} 1 & \text{if } x \neq 1 \\ 3 & \text{if } x = 1 \end{cases}$$

Another factorization of $x^2 - 1$ is $(x - 1)(x + 1)q(x)$.

This is precisely why with traditional variables we cannot simplify $\frac{x^2-1}{x-1}$ to $x + 1$. Another simplification is $(x + 1)q(x)$. Had we assumed $x + 1$ was the only solution, and evaluated at $x = 1$, we would get $1 + 1 = 2$ as the only solution. Knowing $(x + 1)q(x)$ is another solution, and evaluating at $x = 1$, we get $(1 + 1)3 = 6$. So, with traditional variables, we say $\frac{x^2-1}{x-1}$ is undefined at $x = 1$, because the answer is not unique.

3.1.2 Formal Variables.

Formal variables (indeterminates) are constants in abstract algebra, not subject to variation or solution. They enable unique factorization in polynomial rings (e.g. $\mathbb{k}[x]$ the ring of formal polynomials in x with coefficients in $\mathbb{k}$). For example, $x^2 - 1$ factors uniquely as $(x - 1)(x + 1)$ in $\mathbb{k}[x]$. Crucially, we do not need to worry if $x = 1$. This is because x is not a variable, but a constant, and that constant x is not equal to 1 (or any other real number). This is

not unlike the case for complex numbers, where $i = 1$ is always false. Similarly, $x^2 - 1 = 0$ is a false statement for formal variable x , as the left-hand side is not the zero polynomial.

3.2 Construction of $\mathbb{k}(x)$.

Let $\mathbb{k}$ be a field (for concreteness, $\mathbb{k} = \mathbb{R}$ suffices). The field of formal polynomial ratios $\mathbb{k}(x) = \text{Frac}(\mathbb{k}[x])$ can be constructed in two equivalent ways:

3.2.1 Standard-Forms Construction (V1).

Define $\mathbb{k}(x)$ to be the set of reduced pairs (p, q) where $p, q \in \mathbb{k}[x]$ are coprime polynomials (i.e. $\gcd(p, q) = 1$), and $q \neq 0$. Field operations use Euclid's algorithm to maintain reduced forms:

$$\begin{aligned}(a, b) + (c, d) &:= \text{Euclid}(ad + bc, bd), \\(a, b) - (c, d) &:= \text{Euclid}(ad - bc, bd), \\(a, b) \cdot (c, d) &:= \text{Euclid}(ac, bd), \\(a, b) \div (c, d) &:= \text{Euclid}(ad, bc).\end{aligned}$$

3.2.2 Equivalence-Class Construction (V2).

Define $\mathbb{k}(x)$ to be the set of equivalence classes under $\sim$. Define $\sim$ to be an equivalence relation on pairs (p, q) where $p, q \in \mathbb{k}[x]$ and $q \neq 0$:

$$(a, b) \sim (c, d) \Leftrightarrow ad = bc.$$

We denote the equivalence class that contains (a, b) as $[a, b]$. Field operations are defined concisely:

$$\begin{aligned}[a, b] + [c, d] &:= [ad + bc, bd], \\[a, b] - [c, d] &:= [ad - bc, bd], \\[a, b] \cdot [c, d] &:= [ac, bd], \\[a, b] \div [c, d] &:= [ad, bc].\end{aligned}$$

3.3 Traditional Evaluation.

Traditional evaluation ev is a special map called a *homomorphism*.

Definition 3.1 (Homomorphism). A map f is a homomorphism with respect to operation $\circ$ if $f(a \circ b) = f(a) \circ f(b)$ for all a, b .

Since traditional evaluation is a homomorphism, this leads to problems such as undefined values. For example, if $\text{ev}(x) = 0$, then

$$ev(x/x) = ev(x)/ev(x) = 0/0,$$

an indeterminate form, despite x/x being algebraically equivalent to 1 (for $x \neq 0$).

3.3.1 Evaluation Families.

Instead of the overhead of writing “if $ev(x) = a$ ” for some constant a , we can subscript ev , resulting in a family of evaluations, where $ev_a(x)$ equals a .

Restating the previous example:

$$ev_0(x/x) = ev_0(x)/ev_0(x) = 0/0.$$

3.4 Traditional Graphs.

Traditional graphs are defined by Traditional evaluations ev (subscripted or not).

3.4.1 Implicit Graphs.

Definition 3.2 (Implicit Graph). If y is an expression in x , and evaluation $ev: \mathbb{k} \rightarrow \mathbb{k}$, the *implicit graph* is

$$G(x, y) := \{(ev(x), ev(y)) \in \mathbb{k}^2 : ev(x) \in \mathbb{k}\}.$$

3.4.2 Explicit Graphs.

Definition 3.3 (Explicit Graph). If y is an expression in x , and evaluation $ev_a: \mathbb{k} \rightarrow \mathbb{k}$ parameterized by $a \in \mathbb{k}$, the *explicit graph* is

$$G(x, y) := \{(ev_a(x), ev_a(y)) \in \mathbb{k}^2 : a \in \mathbb{k}\}.$$

Since $ev_a(x) = a$, this simplifies to $\{(a, ev_a(y)) \in \mathbb{k}^2 : a \in \mathbb{k}\}$.

4. FORMAL GRAPH METHOD

4.1 Formal Evaluation.

Formal evaluation is a map from $\mathbb{k}(x)$ to $\mathbb{k}$. Formal evaluation is not a homomorphism; it violates $f(a \circ b) = f(a) \circ f(b)$ when $a \circ b$ is not in standard form. We define Formal evaluation for each construction of $\mathbb{k}(x)$.

Definition 4.1 (Formal Evaluation for V1).

In the standard-forms construction, elements are coprime pairs (p, q) . Formal evaluation ev_a acts directly on these pairs:

$$\text{ev}_a((p, q)) := \begin{cases} \frac{p(a)}{q(a)} & \text{if } q(a) \neq 0, \\ \text{undefined} & \text{if } q(a) = 0. \end{cases}$$

This avoids removable discontinuities as pairs are reduced.

Definition 4.2 (Formal Evaluation for V2).

In the equivalence-class construction, elements are classes $[p, q]$. Formal evaluation acts on the unique coprime representative:

$$\text{ev}_a([p, q]) := \text{ev}_a((p', q')) \text{ where } (p', q') \sim (p, q) \text{ and } \gcd(p', q') = 1.$$

Lemma 4.1 (No spurious 0/0).

If (p, q) is reduced and $q(a) = 0$, then $p(a) \neq 0$.

Proof. By the factor theorem, $q(a) = 0$ implies $(x - a) \mid q$ in $\mathbb{k}[x]$. If also $p(a) = 0$, then $(x - a) \mid p$, hence $(x - a)$ is a common nonconstant factor of p and q , contradicting $\gcd(p, q) = 1$.

Proposition 4.1 (Well-definedness on classes).

Let $r = [p, q] = [p', q'] \in \mathbb{k}(x)$. Then $\text{ev}_a(r)$ as in Definition 4.2 is independent of the choice of representatives. Moreover, ev_a is defined iff the denominator of the *reduced* representative is nonzero at a .

Proof. Replace both (p, q) and (p', q') by the common reduced representative $(\tilde{p}, \tilde{q})$. Then both definitions compute the same $\tilde{p}(a)/\tilde{q}(a)$ when $\tilde{q}(a) \neq 0$, and both are undefined when $\tilde{q}(a) = 0$ by Lemma 4.1.

4.2 Formal Graphs.

Definition 4.3 (Formal Graph).

For a rational expression $r \in \mathbb{k}(x)$ and formal evaluation $\text{ev}_a: \mathbb{k}(x) \rightarrow \mathbb{k}$, the *formal graph* is

$$FG(r) := \{(\text{ev}_a(x), \text{ev}_a(r)) \in \mathbb{k}^2 : a \in \mathbb{k}\}.$$

Since $\text{ev}_a(x) = a$, this simplifies to $\{(a, \text{ev}_a(r)) \in \mathbb{k}^2 : a \in \mathbb{k}\}$.

Theorem 4.1 (Formal graphs fill holes and preserve poles).

Let $r \in \mathbb{k}(x)$ have reduced representative (p, q) . Then

$$FG(r) = \{(a, p(a)/q(a)) : a \in \mathbb{k}, q(a) \neq 0\}.$$

In particular,

- if r admits a syntactic presentation with a removable discontinuity at a , then $(a, p(a)/q(a)) \in FG(r)$;
- if $q(a) = 0$, then $(a, \cdot)$ is absent from $FG(r)$ (a genuine pole).

Proof. The equality is just unwinding Definitions [4.1](#) and [4.3](#). The two bullets follow immediately.

Example 4.1 (Difference Quotient without limits).

Consider the difference quotient for $f(x) = x^2$ at $x = 1$:

$$E(h) = \frac{(1+h)^2 - 1}{h}.$$

Algebraic simplification yields:

$$E(h) = \frac{2h + h^2}{h} = 2 + h.$$

Using formal evaluation ev_a with $a \in \mathbb{k}$:

$$\text{ev}_a(E(h)) = \text{ev}_a(2 + h) = 2 + a.$$

The formal graph is $FG(E) = \{(a, 2 + a) : a \in \mathbb{k}\}$, the full line $y = x + 2$ with no hole at $h = 0$. The derivative at $x = 1$ is 2.

4.3 Pedagogical Considerations.

The method adapts to different educational levels:

- **Graduate/advanced undergraduate:** Build $\mathbb{k}(x)$ via equivalence-class construction of $\text{Frac}(\mathbb{k}[x])$ using Definition 3.2.2; define ev_a by reduced representatives using Definition 4.2; prove Lemma 4.1, Proposition 4.1, Theorem 4.1.
- **Undergraduate:** Build $\mathbb{k}(x)$ via standard-forms construction of $\text{Frac}(\mathbb{k}[x])$ using Definition 3.2.1; define ev_a directly using Definition 4.1; prove Lemma 4.1; state Theorem 4.1.
- **High-school / computational:** Teach simplification before substitution as a procedural rule; state Lemma 4.1. In plotting software, normalize rational expressions before sampling to avoid graphical holes.

For notational simplicity, consider light-weight alternatives to $\text{ev}(\cdot)$ for the evaluation operator. We propose notating with a vinculum operator (i.e. underlining) to represent formal evaluations. For example, to graph a parabola, we may graph $\underline{x^2}$ vs. $\underline{x}$.

5. DISCUSSION: EVALUATION AS A TYPED STAGE

Traditional homomorphic evaluation requires symmetric equality $\text{ev}(x \circ y) = \text{ev}(x) \circ \text{ev}(y)$. Formal evaluation relaxes this: $\text{ev}(x \circ y)$ is defined even if $\text{ev}(x) \circ \text{ev}(y)$ is not, by first simplifying $x \circ y$ to a standard form.

We might consider the operational view to be a naive elementary approach.

Remark 5.1 (Operational View).

Substitution has lower precedence than algebraic simplification.



However, this operational view scales to higher mathematics.

Remark 5.2 (Type-Theoretic View).

Evaluation is a two-stage process: (1) normalization $g : T_1 \rightarrow T_2$

(algebraic simplification), then (2) evaluation $f : T_2 \rightarrow T_3$. The evaluator f is not defined on raw syntax T_1 .

Remark 5.3 (Parsing View).

Asymmetric parsing allows backtracking: if $ev(x \circ y)$ parsed to $eval(x) \circ eval(y)$ is undefined, the system backtracks and simplifies $x \circ y$ before re-applying evaluation. Limit-laws in calculus follow this parsing pattern.

6. LIMITATIONS AND FUTURE WORK

What is fixed, what remains. Removable singularities are removed automatically. Equivalent expressions (in the fraction field) generate identical graphs. Difference quotients become algebraic, recovering standard derivatives for polynomials without limits. Genuine poles persist: if p/q is reduced and $q(a) = 0$, evaluation at a is not defined in the field, hence the point is absent from $FG(p/q)$. Thus $FG(r)$ faithfully encodes the “true” graph of a rational expression as an element of $\mathbb{k}(x)$.

Beyond rational expressions. The present development lives in $\mathbb{k}(x) = \text{Frac}(\mathbb{k}[x])$. Extensions to formal power series $\mathbb{k}[[x]]$ (with evaluation restricted to suitable a), or to rational functions in several indeterminates $\mathbb{k}(x_1, \dots, x_n)$, are straightforward in principle but require care with domains of definition when evaluating.

7. CONCLUSION

The traditional method of graphing by direct substitution is inherently limited by the arithmetic problem of indeterminate forms. This paper has presented Formal Graphs, an algebraic framework that circumvents this limitation. By re-framing expressions as elements of



a formal field of rational expressions and evaluating them via formal evaluations, we construct graphs that are continuous and intuitive where traditional graphs fail. This method provides a rigorous, lightweight alternative to calculus, grounding a key concept of precalculus and calculus firmly in modern algebra. Future work could explore extending this method to multivariate expressions and its implications for teaching the concepts of continuity and derivatives.

REFERENCES

- Stewart, J. (2016). *Calculus: Early Transcendentals* (8th ed.). Cengage Learning.
- Apostol, T. M. (1967). *Calculus, Vol. 1: One-Variable Calculus, with an Introduction to Linear Algebra* (2nd ed.). Wiley.
- Dummit, D. S., & Foote, R. M. (2004). *Abstract Algebra* (3rd ed.). Wiley.
- Lang, S. (2002). *Algebra* (Rev. 3rd ed.). Springer.
- SymPy Development Team. (2025). *SymPy Documentation*. Retrieved from <https://docs.sympy.org/>
- von zur Gathen, J., & Gerhard, J. (2013). *Modern Computer Algebra* (3rd ed.). Cambridge University Press.
- Sullivan, M. (2016). *Precalculus* (10th ed.). Pearson.



Fuzzy Multi-Criteria Analysis of Ship Accident Causes: A Comparative Evaluation

Serpil Gerdan¹, Murat Ramazan İLTAR², Emine CAN³

¹Kocaeli University, İzmit Vocational School, Kartepe, Kocaeli, Türkiye. sgerdan@kocaeli.edu.tr

²Istanbul Medeniyet University, Occupational Health and Safety Department, Istanbul, Türkiye. muratiltar@hotmail.com

³Istanbul Medeniyet University, Faculty of Engineering and Natural Sciences, Physics Engineering, Istanbul, Türkiye.
emine.can@medeniyet.edu.tr

ABSTRACT

This study examines the complex causes of accidents occurring in the maritime industry, which is a crucial component of global trade, and explores the interrelationships among these causes. Since maritime accidents result from the interaction of multiple factors, analyzing cause-and-effect relationships is vital for developing effective preventive strategies. For this purpose, the study employs multi-criteria decision-making (MCDM) methods—specifically Fuzzy DEMATEL, Fuzzy AHP, and Fuzzy TOPSIS—to identify the influences among risk factors and to determine their priority rankings. According to the analysis results, Structural Risks and Navigation-Related Factors emerged as the main causal criteria influencing other risk groups. At the sub-criteria level, Technical Failures and Inadequate Maintenance were identified as the most influential causal factors. In the prioritization analyses, the Education and Experience criterion was found to be the most critical factor in preventing accidents. The most significant sub-criterion was determined to be Technical Failures. These findings indicate that maritime safety can be ensured not only through technical improvements but also through continuous training of personnel and regular inspections.

Keywords: AHP, Fuzzy Logic, DEMATEL, Maritime Accidents, Human Factor, Risk Management, TOPSIS

1. INTRODUCTION

Maritime transportation is one of the main pillars of global trade; however, it also poses significant risks in terms of human life, environmental impact, and economic loss. Since maritime accidents often occur as a result of the interaction of multiple factors, understanding



the complex cause-and-effect relationships among these factors is of great importance. This study aims to systematically examine the factors that lead to ship accidents.

The existing literature highlights the effectiveness of Multi-Criteria Decision-Making (MCDM) methods in identifying the root causes of accidents. In this thesis, Fuzzy DEMATEL, Fuzzy AHP, and Fuzzy TOPSIS methods were specifically applied. These methods comprehensively address elements such as structural risks related to the ship, external factors, lack of training, and navigation errors, thereby revealing in detail the interactions and priority rankings among the factors contributing to maritime accidents.

2. LITERATURE REVIEW

There are numerous comprehensive studies examining the causes of maritime accidents and risk management (İnan & Baba, 2020; Olgaç, 2021; Uğurlu, 2022). These studies generally focus on the roles of human factors, technical failures, and environmental conditions in maritime accidents (Buber & Köseoğlu, 2020; Büyük & Bayer, 2022; Tantan & Yorulmaz, 2023). Previous research in this field has attempted to uncover the fundamental dynamics behind ship accidents using various analytical methods. Use of Multi-Criteria Decision-Making (MCDM) Methods:

In the literature, MCDM methods are widely used to solve complex decision-making problems (Tamer, Barlas & Günbeyaz, 2021). In particular, the Analytic Hierarchy Process (AHP) is frequently employed to determine the relative importance of multiple criteria (Kahraman, Cebeci & Ulukan, 2003). This method provides a systematic weighting based on pairwise comparisons of the criteria (Seçme, Bayrakdaroğlu & Kahraman, 2009).

DEMATEL Method:

The Decision-Making Trial and Evaluation Laboratory (DEMATEL) method was developed to analyze direct and indirect relationships—i.e., cause-and-effect interactions—among the criteria within a system (Chang, Chang & Wu, 2011). This approach allows for a clearer understanding of the system's structure and the interdependencies among criteria (Demirci, Canımoğlu & Elçiçek, 2023).



TOPSIS Method:

The Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) ranks alternatives by selecting the one closest to the ideal solution and farthest from the anti-ideal solution (Hwang & Yoon, 1981). This method serves as an effective tool for evaluating alternatives based on the weighted importance of defined criteria (Ashtiani, Haghighirad, Makui & Montazer, 2009).

This study adopts an integrated approach by applying these established methods to determine the cause-and-effect relationships and significance levels among ship accident risk factors. In particular, by integrating Fuzzy Logic, it aims to reduce the uncertainty inherent in expert judgments, thereby contributing to the existing literature (Özdemir, 2016; Shukla, Dixit & Agarwal, 2014; Yildiz et al., 2022).

3. METHODOLOGY

In this study, the risk factors were classified into five main criteria and seventeen sub-criteria. The analysis was conducted using data collected from three experts with ten years of experience in the maritime industry.

Main Criteria Analysis

Fuzzy DEMATEL Analysis: The results indicate that Structural Risks and Navigation-Related Factors fall into the “cause” group, meaning they significantly influence other factors. In contrast, Education and Experience, External Factors, and Psychosocial Risks belong to the “effect” group, representing outcomes influenced by the causal criteria.

Fuzzy AHP and TOPSIS Analysis: According to the ranking results, Education and Experience emerged as the most important main criterion for preventing maritime accidents. This finding highlights the critical role of the human factor in ensuring ship safety.

Sub-Criteria Analysis

Fuzzy DEMATEL Analysis: Among the sub-criteria, Technical Failures and Inadequate Maintenance were identified as the primary causal factors influencing all other sub-factors.



Fuzzy AHP and TOPSIS Analysis: The most significant sub-criterion was determined to be Technical Failures. This finding is consistent with the DEMATEL results, confirming that technical reliability represents the most critical area of risk.

4. RESULTS AND DISCUSSION

In this study, the Fuzzy AHP (F-AHP) method was used to determine the relative importance of factors influencing maritime accidents. The analysis focused on five main types of accidents: Collision/Allision, Fire/Explosion, Grounding, Fatal Occupational Accident, and Injury-Related Occupational Accident. Pairwise comparison matrices were constructed based on the opinions of three experts, and these matrices were then converted into fuzzy pairwise comparison matrices using fuzzy numbers. From the calculated fuzzy synthetic values, fuzzy weights were derived and subsequently defuzzified to obtain precise weight values.

The results showed that the most significant accident types were Fatal Occupational Accidents and Fire/Explosion, together accounting for approximately 80% of the total weight. These two types represent the primary areas that must be prioritized to prevent maritime accidents. The dominance of Fatal Occupational Accidents highlights that the protection of human life remains the foremost priority in maritime transport. The ranking of Fire/Explosion in second place indicates that these accidents have serious consequences for human life, the environment, and the economy. While the remaining accident types—Injury-Related Occupational Accident, Grounding, and Collision/Allision—had lower weights, they still play a vital role in maritime safety and should not be overlooked.

According to the F-TOPSIS analysis results, the most important main criterion for preventing maritime accidents was identified as Education/Experience. This finding emphasizes the critical role of the human factor in maritime transport and indicates that a significant portion of accidents stem from human errors. Enhancing education and experience levels is therefore a key element in reducing such errors. Consequently, the study suggests that the maritime sector should give greater importance to professional training, regular drills, and practical experience programs for personnel.



The second-ranked criterion, External Factors, refers to risks that are difficult to control, such as adverse weather conditions, heavy maritime traffic, and long waiting periods at anchorage. Although these factors cannot be completely eliminated, their effects can be mitigated through early warning systems, advanced weather forecasting technologies, and effective traffic management measures. The remaining three criteria—Psychosocial, Structural, and Navigation—though less significant in weight, should not be ignored. Improving the psychological and social conditions of crew members, strengthening structural maintenance and safety standards of ships, and ensuring accurate and up-to-date voyage planning all contribute significantly to maritime safety. In conclusion, while Education/Experience holds the highest priority, all criteria must be considered collectively.

After examining ship accident reports and conducting interviews with seafarers, 17 sub-factors and their associated risks were identified as follows:

A1: Adaptation Problems to the Ship – Difficulties faced by crew in adapting to the ship and working environment.

A2: Social Isolation – Separation from social life due to long voyages.

A3: Insufficient Provisions – Lack of adequate and quality food on board.

A4: Year of Construction – The age and construction year of the ship.

A5: Inadequate Maintenance – Lack of regular and sufficient maintenance of the vessel.

A6: Technical Failures – Malfunctions and faults in the ship's technical systems.

A7: Insufficient Inspection – Inadequate supervision of the vessel and crew.

A8: Adverse Meteorological Conditions – Storms, fog, and other unfavorable weather conditions.

A9: Heavy Maritime Traffic – High density of ship traffic in navigation zones.

A10: Waiting/Loading-Unloading Times – Duration of anchorage or port operations.

A11: Lack of Training/Drills – Insufficient training or drills for the crew.

A12: Lack of Experience – Insufficient operational or navigational experience of personnel.

A13: Unqualified Employees – Personnel lacking the necessary competence for assigned duties.

A14: Inadequate Voyage Planning – Poor or incomplete voyage planning.

A15: Outdated Nautical Charts – Navigation charts not being regularly updated.

A16: Workload of Watchkeeping – Long and intensive duty periods.

A17: Lack of Communication Between Watches – Inadequate information transfer during watch changes.

According to the findings, the sub-criteria influencing maritime accidents play distinct roles in cause and effect groups. The criteria in the cause group exert greater influence over the system and play an active role. Notably, Technical Failures (A6) and Inadequate Maintenance (A5) emerged as the most significant causal factors with the highest net effect values. Technical failures directly trigger other risk factors due to system malfunctions, while inadequate maintenance increases the likelihood of such failures, negatively affecting the ship's performance and safety. Therefore, regular maintenance and technical inspections are crucial for accident prevention.

Additionally, Insufficient Inspection (A7) ranked third, as lack of oversight hinders early risk detection, negatively impacting safety. Other causal criteria include Lack of Experience, Year of Construction, Communication Deficiencies, Adaptation Problems, and Unqualified Employees, all of which exert secondary influences and must be addressed in accident prevention efforts.

Conversely, criteria in the effect group are more influenced by the system and play a passive role. Heavy Maritime Traffic (A9) and Adverse Meteorological Conditions (A8) had the lowest net effect values, meaning they are the most affected by other factors. Enhancing voyage planning and technical equipment can help mitigate their adverse impacts. Other effect-group criteria—such as Inadequate Voyage Planning, Heavy Workload, Waiting Times, Insufficient Provisions, Lack of Training, Social Isolation, and Outdated Charts—are shaped by the influence of other factors.

Based on D+R values, Technical Failures (A6), Unqualified Employees (A13), and Inadequate Maintenance (A5) are positioned at the center of the system, demonstrating high interactivity. In contrast, Waiting Times (A10), Heavy Traffic (A9), and Social Isolation (A2) show lower interaction, occupying more isolated positions within the system. Consequently, in



order to prevent maritime accidents, priority should be given to managing the strong causal factors within the cause group.

According to the F-TOPSIS results, the top five sub-criteria in preventing maritime accidents are ranked as follows: Technical Failures, Inadequate Maintenance, Unqualified Employees, Lack of Training/Drills, and Lack of Experience. This ranking identifies the key intervention areas for accident prevention.

The most significant sub-criterion, Technical Failures, reveals that malfunctions and errors in ship systems are among the primary causes of accidents. This underscores the importance of regular inspection, timely maintenance, and early fault detection as critical components in accident prevention. Given the high potential of technical deficiencies to cause accidents, systematic maintenance and control procedures are essential.

The Inadequate Maintenance criterion, ranked second, shows that insufficient or irregular maintenance schedules are a major factor contributing to accidents. Therefore, meticulous adherence to maintenance programs and systematic record-keeping are crucial. Regular maintenance not only prevents technical failures but also directly enhances ship safety.

Human-related factors such as Unqualified Employees, Lack of Training/Drills, and Lack of Experience ranked third, fourth, and fifth, respectively. These results highlight the significant role of the human factor in the occurrence of maritime accidents. Thus, the maritime industry must prioritize merit-based personnel selection, ensure regular training and drills, and provide opportunities for experience development.

In conclusion, the effective management of technical systems, consistent maintenance practices, and enhancing crew competence are fundamental priorities for preventing maritime accidents. The analysis reveals that ensuring ship safety and reducing accidents requires a simultaneous focus on both technical and human factors.

According to the DEMATEL analysis results, the criteria Structural Risk and Navigation play influential roles over other factors and belong to the cause group, whereas Psychosocial Risk, External Factors, and Education and Experience are positioned in the effect group. This distinction provides an important perspective for prioritizing areas in risk management



strategies. The findings show that across all methods, Education and Experience and External Factors stand out as key criteria, emphasizing that improving human competence and taking effective measures against environmental risks are critical for preventing maritime accidents.

5. CONCLUSION

In this study, the factors influencing maritime accidents were analyzed using a comparative application of the DEMATEL and AHP-TOPSIS methods. Both methods systematically address complex decision-making problems by evaluating the relationships and importance levels among factors, providing different perspectives for the prevention of maritime accidents.

According to the DEMATEL analysis, the main criteria “Structural Risk” and “Navigation” were classified as the cause group, while “Risk Arising from External Factors,” “Education and Experience,” and “Psychosocial Risk” were categorized as the effect group. Among the sub-criteria, “Technical Failures,” “Inadequate Maintenance,” and “Insufficient Supervision” had the highest net effect values and were identified as causal factors. In the AHP-TOPSIS analysis, the most frequent accident types were identified and weighted as fatal occupational accidents, fire/explosion, injury-related occupational accidents, grounding, and collision/allision.

The critical findings indicate that technical factors (Technical Failures, Inadequate Maintenance) and human factors (Unqualified Personnel, Lack of Training/Drills, Insufficient Experience) play a decisive role in the occurrence of accidents. While DEMATEL reveals the causal relationships among the criteria, AHP-TOPSIS focuses directly on their relative importance levels. Lack of education and experience emerged as high-priority risks in both methods, emphasizing that continuous training, simulator-based exercises, mentoring, and cross-cultural education can effectively mitigate these risks.

Risks arising from external factors ranked first in DEMATEL and second in AHP-TOPSIS; therefore, the negative impacts of meteorological conditions, marine traffic, and insufficient inspections should be minimized through effective use of technological systems and regular audits. Navigation-related and psychosocial risks should also be managed through systematic

preventive measures. Although structural risks were found to have relatively lower priority, they should still be controlled through regular maintenance and modernization programs.

At the sub-criteria level, “Technical Failures” ranked first, “Inadequate Maintenance” second, while “Unqualified Personnel” and “Lack of Training/Drills” were identified as high-priority risks. Managing these factors is crucial for preventing accidents. Overall, this study presents a comprehensive risk management approach that holistically addresses technical, human, and environmental factors in reducing maritime accidents.

REFERENCES

- Abdullayev, B. (2013). *Türkiye-Rusya arasındaki mevcut ulaştırma ağının analizi ve çoklu ulaştırma sistemleri alternatiflerinin geliştirilmesi*. (Yayınlanmamış doktora tezi). Dokuz Eylül Üniversitesi, Türkiye.
- Ahmad, N., & Qahmash, A. (2020). Implementing Fuzzy AHP and FUCOM to evaluate critical success factors for sustained academic quality assurance and ABET accreditation. *PLOS ONE*, 15(9), e0239140. <https://doi.org/10.1371/journal.pone.0239140>
- Ashtiani, B., Haghighirad, F., Makui, A., & Montazer, G. A. (2009). Extension of fuzzy TOPSIS method based on interval-valued fuzzy sets. *Applied Soft Computing*, 9(2), 457-461. <https://doi.org/10.1016/j.asoc.2008.05.005>
- Buber, M., & Köseoğlu, B. (2020). Gemi Kaynaklı Petrol Kirliliği Kazalarında Net Çevresel Fayda Analizine Yönelik Bir Literatür Çalışması. *Türk Denizcilik Dergisi*, (Bilinmiyor).
- Büyük, N. & Bayer, D. (2022). Tanker Gemilerinde Kargo İşlemleri Esnasındaki Yangın Risklerinin Kök Sebeplerinin Tespiti ve Bow-Tie ile Analizi. *Denizcilik Araştırmaları Dergisi: Amfora*, 1(1), 79-87. <https://doi.org/10.29228/jomaramphora.62290>
- Çalışkan, N. (2019). *Uluslararası denizyolu taşımacılığının Türkiye dış ticaretine etkisinin analizi ve diğer taşıma modlarıyla rekabeti*. (Yayınlanmamış yüksek lisans tezi). Sosyal Bilimler Enstitüsü, Türkiye.
- Çiçek, H., & Çağdaş, A. (2020). Ergonomik Faktörlerin Çalışan Performansına Olan Etkileri. *OHS ACADEMY*, 3(2), 135-143.
- Çulha, H. S. (2021). Beş Faktör Kişilik Envanterinin Türkçe ‘Ye Uyarlanması, Denizcilik Sektöründeki Liderlerin Kişilik Özelliklerinin Belirlenmesi ve Kazalara Olan Etkisi. (Yayınlanmamış yüksek lisans tezi).
- Chang, B., Chang, C. W., & Wu, C. H. (2011). Fuzzy DEMATEL method for developing supplier selection criteria. *Expert Systems with Applications*, 38(3), 1850–1858. <https://doi.org/10.1016/j.eswa.2010.07.114>
- Chamoli, S. (2015). Hybrid FAHP (fuzzy analytical hierarchy process)-FTOPSIS (fuzzy technique for order preference by similarity of an ideal solution) approach for performance evaluation of t²he V down perforated baffle roughened rectangular channel. *Energy*, 84, 432-442. <https://doi.org/10.1016/j.energy.2015.03.007>
- Demir, A., & Öz, A. (2018). Teolojik Açıdan İş Kazalarının İncelenmesi. *Avrupa Bilim ve Teknoloji Dergisi*, (14), 189-197. <https://doi.org/10.31590/ejosat.459848>
- Demirci, S. E., Canımoğlu, R., & Elçiçek, H. (2023). Analysis of causal relations of marine accidents during ship navigation under pilotage: A DEMATEL approach. *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*, 237(2), 308-321.
- Hwang, C., & Yoon, K. (1981). *Multiple attributes decision making methods and applications*, berlin, heidelberg. Springer Berlin, Heidelberg.
- İnan, T., & Baba, A. F. (2020). Gemi çarpışmalarının önlenmesi için melez algoritma tabanlı bir karar destek sisteminin oluşturulması. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 35(3), 1213-1230.
- Kahraman, C., Cebeci, U., & Ulukan, Z. (2003). Multi-criteria supplier selection using fuzzy AHP. *Logistics Information Management*.
- Kayıran, E. B. & Çaylan, D. Ö. (2019). İş Sağlığı ve Güvenliği Konulu Bilimsel Çalışmalarda Denizcilik Sektörünün Yeri. IV. Ulusal Liman Kongresi “Küresel Eğilimler-Yerel Stratejiler”.

- Kodak, G., & Acarer, T. (2021). İstanbul Boğazı'nda deniz trafik düzenlemelerinin kaza oranına etkisinin değerlendirmesi. *Aquatic Research*, 4(2), 181–207. <https://doi.org/10.3153/AR21015>
- Kutlu, A. C., & Ekmekçioğlu, M. (2012). Fuzzy failure modes and effects analysis by using fuzzy TOPSIS-based fuzzy AHP. *Expert Systems with Applications*, 39(1), 61-67. <https://doi.org/10.1016/j.eswa.2011.06.044>
- Lan, D., Liang, B., Bao, C., Ma, M., Xu, Y., & Yu, C. (2015). Marine Oil Spill Risk Mapping for Accidental Pollution and Its Application in A Coastal City. *Marine Pollution Bulletin*, 96(1-2), 220–225.
- Nazim, M., Mohammad, C. W., & Sadiq, M. (2022). A comparison between fuzzy AHP and fuzzy TOPSIS methods to software requirements selection. *Alexandria Engineering Journal*, 61(12), 10851-10870.
- Olgaç, T. (2021). Deniz Kazaları ve Deniz Olaylarını İnceleme Çalışmalarında Kullanılan Analiz Yöntemleri Üzerine Bir Değerlendirme. *Journal of Maritime Transport and Logistics*, 2(2), 101-112. <https://doi.org/10.52602/mtl.885205>
- Özdemir, Ü. (2016). Bulanık DEMATEL ve Bulanık TOPSIS Yöntemleri Kullanılarak Limanlarda Yaşanan İş Kazalarının İncelenmesi. *Journal of ETA Maritime Science*, 4(3), 235-247.
- Özdemir, A., & Yılmaz, B. (2019). Melez algoritmaların denizcilik sektöründeki güvenlik yönetim sistemlerine etkisi. *Journal of Maritime Safety*, 15(3), 45-60.
- Sankar, J., & Augustin, F. (2016). An improved fuzzy dematel technique and its applications. (Bilinmiyor).
- Seçme, N. Y., Bayrakdaroğlu, A., & Kahraman, C. (2009). Fuzzy performance evaluation in Turkish banking sector using analytic hierarchy process and TOPSIS. *Expert Systems with Applications*, 36(9), 11699-11709.
- Shukla, R. K., Dixit, G., & Agarwal, A. (2014). An integrated approach of Fuzzy AHP and Fuzzy TOPSIS in modeling supply chain coordination. *Production & Manufacturing Research*, 2(1), 415–437. <https://doi.org/10.1080/21693277.2014.919886>
- Tamer, S., Barlas, B., & Günbeyaz, S. A. (2021). Çok Kriterli Karar Verme Yöntemlerinin Gemi İnşaatı ve Gemi Makineleri Mühendisliğinde Uygulamaları. *Gemi ve Deniz Teknolojisi Dergisi*, 220, 54-65.
- Tantan, E. & Yorulmaz, M. (2023). Gemilerde Meydana Gelen İş Kazalarının Çok Kriterli Karar Verme Yöntemleriyle İncelenmesi. *Uluslararası Batı Karadeniz Sosyal ve Beşerî Bilimler Dergisi*, 7(2), 132-158.
- Topaloğlu, H. (2007). *Dış Ticaret Yüklerimizin Taşınmasındaki Terminal Durumları ve Liman Yeterliliklerinin Değerlendirilmesi*. (Yayınlanmamış yüksek lisans tezi). İstanbul Üniversitesi Deniz Bilimleri ve İşletmeciliği Enstitüsü, Türkiye.
- Uğurlu, H. (2022). Gemi Çatışma Kazalarının Nedenlerinin ve Sonuçlarının Sistematik İncelemesi. *4th International Eurasian Conference on Science, Engineering and Technology*.⁴
- Yildiz, A., Guneri, A. F., Ozkan, C., Ayyildiz, E., & Taskin, A. (2022). An integrated interval-valued intuitionistic fuzzy AHP-TOPSIS methodology to determine the safest route for cash in transistt operations: a real case in İstanbul. *Neural Computing and Applications*, 36(2), 1-16.
- Yorulmaz, M. (2009). *Deniz Taşımacılığı ve Deniz Sigortaları*. Akademi Denizcilik.
- Yorulmaz, M., & Öztürk, M. A. (2018). Tersanelerdeki İş Kazası Nedenlerinin Önem Düzeylerine Göre Belirlenmesi. *Avrupa Bilim ve Teknoloji Dergisi*, (41), 132-143.



A Study on the Performance of Deep Neural Networks for Camouflaged Military Object Detection

Zehra Korkmaz, *Department of Computer Engineering,
Graduate School of Natural and Applied Sciences,
Gazi University, Ankara, Turkey, 0009-0000-8172-0955*

Murat Dörterler, *Department of Computer Engineering, Faculty of Technology,
Gazi University , Ankara, Turkey, 0000-0003-1127-515X*

ABSTRACT

Today, computer vision techniques are successfully utilised in numerous critical tasks across various domains, most notably in the defence industry, where object detection applications play a prominent role. In military contexts, the accurate, fast, and reliable detection of strategic assets such as military personnel, tanks, aircraft, and warships through image data is of vital importance, as it enhances both the effectiveness of operational decision-support systems and the overall efficiency of military operations. In this regard, computer vision-based object detection systems have become an indispensable component in modern warfare, contributing significantly to situational awareness and threat analysis processes. However, the detection of heavily camouflaged objects remains a major challenge for existing methods due to factors such as visual similarity, low contrast, and complex backgrounds. Addressing this issue, the present study focuses on the detection of highly camouflaged military objects by employing the comprehensive MHCD2022 dataset, which comprises 3,000 raw images across five distinct classes. In the experimental analyses, the performance of the advanced object detection algorithm YOLOv11 was evaluated on this dataset. The results showed that the model achieved a mean Average Precision (mAP) of 67% and a recall rate of 60.13%, indicating that YOLOv11 outperforms contemporary approaches in the current literature in terms of both consistency and accuracy. Moreover, this study examines the strengths and limitations of the model in detail and aims to contribute to the development of next-generation deep learning models for scenarios involving heavily camouflaged object detection in the military domain.

Keywords: Computer vision, object detection, military object detection, deep neural networks, artificial intelligence



1. INTRODUCTION

The subject of object detection has been a fundamental area of research for many years, and has attracted significant interest from researchers in the fields of artificial intelligence and computer vision. This domain encompasses methods aimed at identifying the presence of objects in static images or video content, classifying their categories, and determining their spatial locations (Arulprakash & Aruldoss, 2022).

The employment of camouflage patterns has become a pervasive phenomenon, particularly in recent decades, across a wide range of military equipment, vehicles, and combat gear. The objective of this technique is to enhance military superiority by reducing the visual detectability of objects and making them difficult to distinguish from the background. The high level of visual harmony between the object in the foreground and its surroundings makes it difficult to detect, creating serious challenges for observation, surveillance, and detection activities on the battlefield. Visual indicators employed in the identification of objects (e.g. edge features, textures, contrast differences, scale variability, and colour tones) remain susceptible to manipulation through camouflage strategies (e.g. disruptive colouring, blending with the background) (Zheng et al., 2019).

The present studies have the objective of developing methods of increasing the detectability of camouflage for military personnel and unmanned aerial vehicle (UAV)-based reconnaissance systems. However, a review of the extant literature reveals a paucity of systematic or comprehensive studies on the subject of advanced military camouflage patterns. A primary factor contributing to this issue is the absence of a standardised, accessible, comparative object detection dataset. Such a dataset would facilitate the analysis of these patterns. Consequently, research in the field of camouflage detection is imperative not only for theoretical research, but also for military applications.

The development of deep learning-based object detection approaches has been instrumental in overcoming challenges, including the detection of camouflaged objects. These approaches facilitate the identification of specific visual features despite variations in the shape and size of military objects and the complexity of the background. One of the approaches, You Only Look Once (YOLO) (Redmon et al., 2016), is designed to work quickly and accurately and



offers higher accuracy rates compared to most similar object detection methods (Song et al., 2025).

The objective of this study is to apply the YOLO object detection algorithm to the MHCD2022 dataset (Liu & Di, 2023). The rationale behind the selection of the extended MHCD2022 dataset is to establish a distinctive and diversified dataset. During the process of data augmentation, a range of transformations were applied to the images and their corresponding labels. In the initial phase, the labels defined in YOLO format were read in conjunction with each training image. In the subsequent phase, a range of image processing techniques were employed to enhance data diversity. In this particular context, a series of operations were performed, including horizontal flipping, 90° rotation, random shifting, and mosaicking. Additionally, the label coordinates were recalculated to ensure compatibility with the applied transformations.

Consequently, this process resulted in the acquisition of a more diverse and balanced data set for model training. While augmented data was employed during the training phase, the test data was evaluated directly without undergoing any transformation. Consequently, a comparison was made between the model outputs and those from analogous studies cited in the literature. This comparison indicated that the YOLO algorithm exhibited the most optimal performance on the expanded MHCD2022 data set.

The remaining sections of this study are structured as follows: In Section II, the general structure of the research articles examined is presented as a literature review, and the objectives of these articles are explained. Section III delineates the methodological framework employed in this study. Section IV is dedicated to the presentation of experimental findings, encompassing transfer learning methodologies and comparative analyses. The subsequent section, Section V, discusses the results obtained.

2. RELATED WORK

The development of algorithms to improve the accuracy of camouflaged object detection is a significant area of research in the literature. The camouflaged object detection algorithms employed in contemporary studies are predominantly categorised into two primary approaches: traditional methods and deep learning-based methods. Conventional



methodologies depend on the employment of manually crafted characteristics to differentiate camouflaged objects.

In a study conducted by Bhajantri and Nagabhushan, a grey-scale co-occurrence matrix at the patch level was utilised to represent tissue characteristics. The combination of this matrix with watershed segmentation and clustering strategies resulted in the successful detection of camouflaged objects (Bhajantri & Nagabhushan, 2006). In a study conducted by Song and Geng, an artificial texture descriptor was developed which covered texture direction, brightness, and entropy. The reliability of the foreground texture was evaluated based on the weight structure of the features (Song & Geng, 2010). The powerful feature representation capabilities of deep learning-based methods, in conjunction with rapid advances in the field, have resulted in a significant improvement in the accuracy rates of camouflaged object detection algorithms. The FAP-net, BGM, MFAM, and CFPM modules proposed by Zhou et al. develop boundary features to create richer and more detailed combined feature representations, effectively discover inter-layer relationships, and thus significantly improve the detection performance of camouflaged objects (Zhou et al., 2022). Sharma and Mir developed a framework that combines the detection of camouflaged objects with semantic object detection (SOD). This framework utilises an adaptive learning strategy and advanced feature representation, thereby effectively addressing uncertainty in model predictions (Sharma & Mir, 2022). Rani et al. established a positioning and focusing network with the objective of detecting and eliminating dispersion in camouflage scenes. This network was supported by a dispersion mining scheme (Rani, Ghai, & Kumar, 2022). Fan et al. developed a model for feature enhancement incorporating neighbour connection coders and group-level attention subcomponents (Fan, Ji, Cheng, & Shao, 2022). Yang and colleagues set out a robust framework for dealing with ambiguity in camouflage scenes. This framework uses Transformers and a probabilistic representation model (Yang et al., 2021). The Single Shot MultiBox Detector (SSD) method, which performs object detection through a single deep neural network without requiring additional processing steps such as proposal generation and ensures high computational efficiency and speed, was developed by Liu and colleagues (Liu et al., 2015). A fully convolutional Region Proposal Network (RPN), capable of simultaneously predicting object bounding boxes and objectness scores, sharing convolutional feature maps with Fast R-CNN for joint training, and eliminating the need for explicit region proposal



computations to accelerate the detection process, was developed by Ren and colleagues (Ren, He, Girshick, & Sun, 2015). Carion and colleagues proposed the Detection Transformer (DETR), an innovative approach to object detection that frames the problem as a set prediction task, thereby removing the need for anchor generation and handcrafted components, and simplifying the overall detection process (Carion et al., 2020). AdaMixer, a query-based detector architecture that enables adaptive feature sampling according to the spatial and scale offsets of queries, dynamically decodes these features using a query-specific MLP-Mixer, and achieves high-accuracy object detection without relying on complex attention mechanisms, was developed by Gao and colleagues (Gao, Wang, Han, & Guo, 2022). CrossDet, a novel object detection method that represents objects using expanding cross lines along horizontal and vertical axes, reduces background noise, effectively captures continuous object appearance information, and achieves precise anchor-free object detection, was proposed by Qiu and colleagues (Qiu et al., 2023). Liu and Di developed the Military High-Level Camouflage Detection (MHCD) scheme with the objective of achieving more accurate positioning and correct boundary determination of objects containing high-level military camouflage. The objective of this scheme is to enhance the detection performance of camouflaged objects (Liu & Di, 2023).

3. METHODOLOGY

In light of the paucity of research in the domain of military camouflage object detection, compounded by the absence of official studies in this area, Liu and Di undertook the creation of the MHCD2022 dataset (Liu & Di, 2023). This initiative was undertaken with the explicit objective of facilitating the detection of highly camouflaged military objects. The dataset under consideration consists of images collected from a variety of websites and covers a range of realistic scenes, including forests, deserts, towns, oceans, and snow. The dataset is divided into five categories: human, military vehicle, tank, warship, and aircraft. The dataset consists of a total of 3,000 highly camouflaged object images, each captured at 27 fps. The categories are represented by images taken from a variety of backgrounds and angles, ensuring that the models are adaptable to real-world conditions. There is an imbalance in the data distribution, with the human category being heavily represented, while the military vehicle and warship categories contain fewer examples.



The dataset was divided into 80% training and 20% testing subsets, with data augmentation applied to 2,400 images in the training subset. During the augmentation process, various image processing techniques such as horizontal flipping, rotation, random shifting, and mosaicking were employed. Through the generation of multiple derivatives from each training image, the diversity of the dataset was enhanced, increasing the total number of samples to 12,000.

3.1 METHOD

In this study, the MHCD2022 dataset was utilised as the primary benchmark for evaluating the detection of camouflaged military objects. To overcome the inherent challenges of camouflage, such as low visual contrast and background similarity, a deep learning model based on the YOLOv11 architecture was employed.

The YOLOv11 model comprises 181 layers and 2,590,815 parameters, and has been developed for tasks such as real-time object detection, image classification, segmentation, and pose estimation. The functionality of feature learning has been enhanced with the C2PSA module (Khanam & Hussain, 2024), and the efficiency of parameters has been increased with the C3k2 modules (Khanam & Hussain, 2024). The Darknet53-based backbone (Khanam & Hussain, 2024) of the model facilitates the effective learning of low- and high-level features, thereby ensuring high accuracy for small and distant objects. This is achieved through an anchor-free structure and optimised multi-task learning strategies (Sorour et al., 2025).

The model was trained using the Ultralytics library (Jegham et al., 2025), with various hyperparameters being meticulously configured during the training process. In the experimental setup, the pre-trained YOLOv11n network was utilised as a foundation and retrained on the dataset obtained within the scope of the study. The dataset was defined for the model via the data.yaml file, and the input image size was set to 768 pixels. The training process was executed for 300 epochs, and the batch size was optimised to 16.

A range of data augmentation techniques were employed to enhance the model's generalisation capacity. In this context, operations such as mosaic (100%), mixup (50%),

colour-based transformations (changes in hue, saturation, and brightness in the HSV colour space), rotation ($\pm 10^\circ$), scaling (90%), translation (10%), and shear (2°) were applied. The learning rate was modified by employing the cosine annealing strategy, with an initial value of 0.001 and a final value of 0.01. Furthermore, the early stopping strategy was employed to avert overfitting, and the training process was terminated when no further improvement was observed. The final phase of the study entailed training with the YOLOv11 model, following which the outcomes were compared with those of analogous studies documented in the extant literature.

3.2 EXPERIMENTAL SETUP

All experiments and visualisations were performed in a Google Colab Pro+ environment with 83.5 GB RAM, 40 GB GPU memory, and 235.7 GB disk space. The following libraries were utilised in this process: NumPy, Pandas, OpenCV, Matplotlib, Seaborn, scikit-learn, and Ultralytics YOLO. The execution of deep learning tasks, with a particular focus on image processing and model training, was facilitated by the Ultralytics YOLO framework.

4. EXPERIMENTAL RESULTS

As demonstrated in Figure 1, the training process metrics for the YOLOv11 model are as follows:



Figure 1. YOLOv11 Model Training Process Metrics Graph

As illustrated in Figure 1, the YOLOv11 model demonstrates a notable shift in precision, recall, mAP@0.5, and mAP@0.5:0.95 values over the course of 300 epochs. In the initial 50 epochs, the metrics of precision and recall exhibited a rapid increase, reaching a plateau at

approximately 0.95, while the metric of mAP@0.5 attained 0.95 around the 100th epoch. The mAP@0.5:0.95 metric, which evaluates a stricter IoU range, increased gradually and reached approximately 0.85 by the 300th epoch. The findings suggest that the model attained stable learning with a high degree of accuracy, thereby indicating a minimal risk of overfitting. The confusion matrix values of the YOLOv11 model are shown in Figure 2.

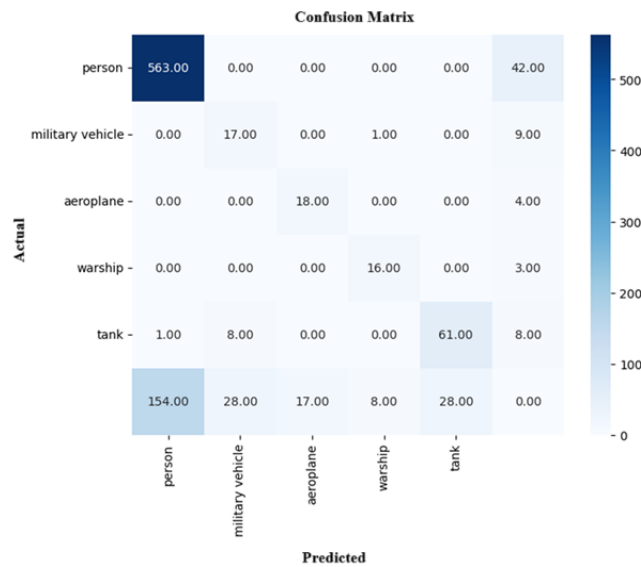


Figure 2. Confusion Matrix Values of the YOLOV11 Model

As demonstrated in Figure 2, the complexity matrix provides insight into the classification performance of the YOLOv11 model. A high level of accuracy was achieved, with 563 correct classifications in the Person class, but 154 and 42 examples were incorrectly classified as unknown/other. In the military vehicle class, 28 instances were erroneously categorised as 'unknown/other', despite 17 accurate classifications. In the aeroplane class, 17 instances were inaccurately classified as 'unknown/other', despite 18 precise classifications. In the warship class, 9 cases were incorrectly classified as 'unknown/other', despite 16 accurate classifications. The tank class demonstrated a high level of accuracy, with 61 instances of correct classification, though it is noteworthy that 28 cases exhibited misclassification. The findings indicate that the model demonstrates a high degree of accuracy, particularly in the person class. However, the military vehicle class is frequently misclassified as other classes. The sensitivity and recall distribution graph of the YOLOv11 model is shown in Figure 3.

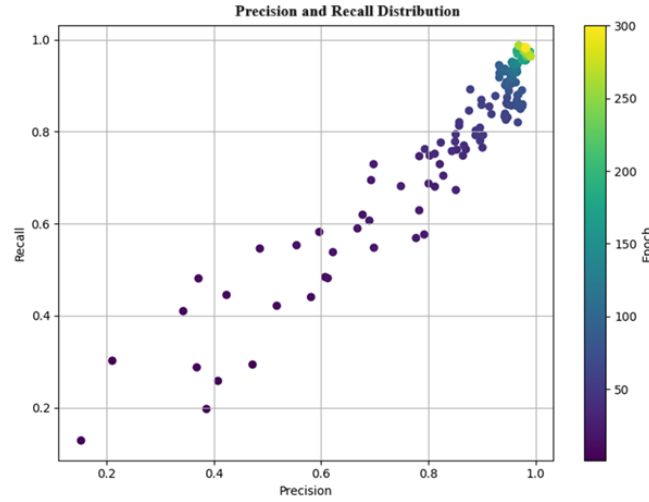


Figure 3. The following figure illustrates the sensitivity and recall distribution graph of the YOLOv11 model.

As demonstrated in Figure 3, the precision and recall values obtained per epoch during the training process of the YOLOv11 model are presented. The colour scale indicates that each point corresponds to the relevant epoch, with dark purple representing low epochs and yellow representing high epochs. A thorough examination of the graph reveals a consistent increase in both precision and recall values throughout the model's training process. It is evident that precision and recall values exceed 0.95 in the final epochs, attaining elevated levels of performance. This finding suggests that as the training progresses and the learning process is completed, the model's classification accuracy and detection capability will increase. The mAP scatter graph of the YOLOv11 model is shown in Figure 4.

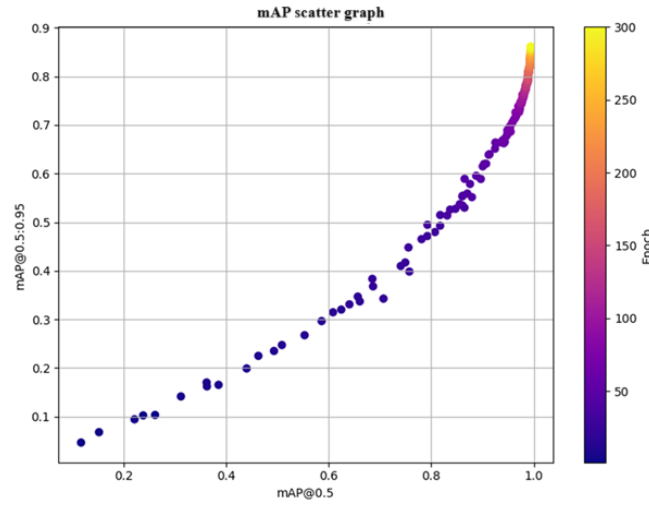


Figure 4. mAP scatter graph YOLOv11 Model

Figure 4 illustrates that both metrics remain at low levels during the initial epochs, indicating that the model exhibits limited discriminative capability at the outset. As training progresses, the points increasingly concentrate in the upper-right region; notably, after the 250th epoch, $mAP@0.5$ surpasses 0.95 and $mAP@0.5:0.95$ exceeds 0.85, demonstrating high accuracy under both loose and strict IoU thresholds. The observed curve indicates a strong correlation between the two metrics and confirms that the YOLOv11 model achieves robust and generalisable object detection performance. The Area Under the Curve (AUC) and the Receiver Operating Characteristic (ROC) curve graphs of the YOLOv11 model are presented in Figure 5.

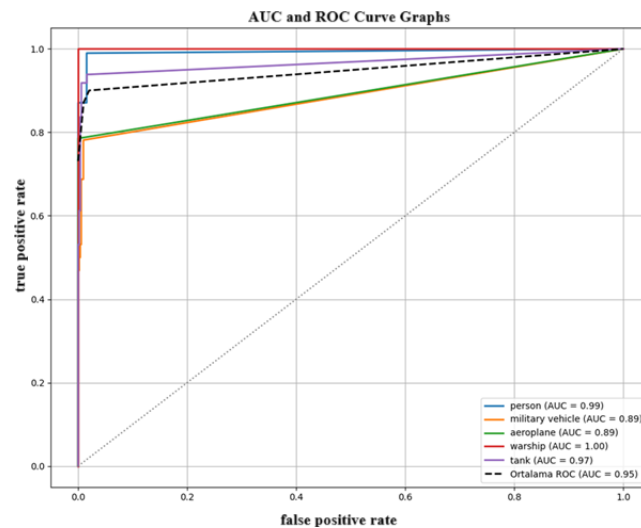


Figure 5. AUC and ROC Curve graphs YOLOv11 Model

The ROC curve (see Figure 5) demonstrates the classification performance of the YOLOv11 model, with an average AUC value of 0.95, thus confirming the model's strong discrimination capabilities in multi-class classification. The following values are reported for the AUC by class: warship 1.00, human 0.99, tank 0.97, military vehicle 0.89, and aircraft 0.89. The findings demonstrate that the model attains high accuracy, particularly in the 'warship,' 'human,' and 'tank' classes, while maintaining robust performance in other classes. The ROC analysis demonstrates that the model is a reliable and effective object detection system overall. Table 1 presents the class-based performance measures of the YOLOv11 model on the MHCD2022 test dataset.

TABLO 1. Class Based Performance Measures (%)

Class	Evaluation Metrics			
	<i>Precision</i>	<i>Recall</i>	<i>mAP @50</i>	<i>mAP@50-95</i>
all	79.2	60.1	67.0	42.9
person	91.2	77.9	87.1	56.6
military vehicle	71.3	37.5	44.8	26.3
aeroplane	74.7	51.4	54.9	33.3
warship	82.3	64.0	72.4	45.7
tank	76.2	69.6	75.7	52.4

Upon examination of Table 1, it was observed that the YOLOv11 model demonstrates a generally reasonable detection capability. Among the classes, the highest accuracy was observed in the “person” class (R 0.78, mAP@0.5 0.871), while notable performance was also achieved for the “tank” (R 0.697, mAP@0.5 0.757) and “warship” (R 0.64, mAP@0.5 0.724) classes. In contrast, the “military vehicle” class exhibited lower accuracy and detection rates (R 0.376, mAP@0.5 0.448), whereas the “aeroplane” class demonstrated moderate performance. The observed class-based differences are considered to arise from the number of samples in the dataset and the visual characteristics of the classes. Table 2 presents a quantitative comparison of state-of-the-art object detection methods on the MHCD2022 dataset.

TABLO 2. Comparison of Related Works (%)

Methods	Compared Features	
	<i>mAP @50</i>	<i>mAP@50-95</i>
MHNet [5]	56.7	36.6
SSD [13]	48.0	25.3
Faster R-CNN [14]	53.5	32.0
DETR [15]	56.5	34.7
AdaMixer [16]	56.8	36.5
CrossDet++ [17]	56.4	36.6
Ours	67.0	42.9

Upon examination of Table 2, it can be observed that, considering the $mAP@0.5$ and $mAP@0.5:0.95$ values, the proposed model achieves superior performance compared to all existing methods, with scores of 67.0 and 42.9, respectively. Notably, a significant improvement is observed relative to classical methods such as SSD ($mAP@0.5=48.0$, $mAP@0.5:0.95=25.3$) and Faster R-CNN ($mAP@0.5=53.5$, $mAP@0.5:0.95=32.0$). Furthermore, the model demonstrates a considerable increase in mAP values compared to modern approaches such as DETR and AdaMixer. These findings indicate that the proposed model provides superior detection accuracy and overall performance on the MHCD2022 dataset.

4 CONCLUSIONS

This study investigated the detection of highly camouflaged military objects. The performance of the YOLOv11 model was evaluated using the comprehensive MHCD2022 dataset, which comprises 3,000 images across five categories. Experimental results indicate that the model achieved a mAP of 67% and a recall of 60.13%. These findings demonstrate that YOLOv11 delivers consistent and superior performance compared to the most recent studies in the literature. The results highlight the strengths and limitations of the model and provide significant contributions to the development of military object detection systems. YOLOv11 exhibited robust and superior performance in detecting highly camouflaged objects compared to existing methods, offering valuable insights for the design of innovative models for military object detection. Future research should aim to further improve model performance



by addressing class imbalance and enhancing generalization capabilities. In this context, additional data should be collected for underperforming classes, and efforts should be made to balance the dataset.

REFERENCES

- Arulprakash, E., & Aruldoss, M. (2022). *A study on generic object detection with emphasis on future research directions. Journal of King Saud University - Computer and Information Sciences*, 34(9), 7347–7365. <https://doi.org/10.1016/j.jksuci.2021.08.001>
- Bhajantri, N. U., & Nagabhushan, P. (2006, December). *Camouflage defect identification: A novel approach*. In *9th International Conference on Information Technology (ICIT'06)* (pp. 145–148). IEEE. <https://doi.org/10.1109/ICIT.2006.34>
- Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020). *End-to-end object detection with transformers*. arXiv. <https://doi.org/10.48550/arXiv.2005.12872>
- Fan, D.-P., Ji, G.-P., Cheng, M.-M., & Shao, L. (2022). *Concealed object detection*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10), 6024–6042. <https://doi.org/10.1109/TPAMI.2021.3085766>
- Gao, Z., Wang, L., Han, B., & Guo, S. (2022). *AdaMixer: A fast-converging query-based object detector*. arXiv. <https://doi.org/10.48550/arXiv.2203.16507>
- Jegham, N., Koh, C. Y., Abdelatti, M., & Hendawi, A. (2025, March 17). *YOLO evolution: A comprehensive benchmark and architectural review of YOLOv12, YOLOv11, and their previous versions*. arXiv. <https://doi.org/10.48550/arXiv.2411.00201>
- Khanam, R., & Hussain, M. (2024, October 23). *YOLOv11: An overview of the key architectural enhancements*. arXiv. <https://doi.org/10.48550/arXiv.2410.17725>
- Liu, M., & Di, X. (2023). *Extraordinary MHNet: Military high-level camouflage object detection network and dataset*. *Neurocomputing*, 549, 126466. <https://doi.org/10.1016/j.neucom.2023.126466>
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2015). *SSD: Single shot multibox detector*. arXiv. <https://doi.org/10.48550/arXiv.1512.02325>
- Qiu, H., Li, J., Zhang, X., Zhu, H., Chen, L., Jiang, J., ... & Liu, S. (2023). *CrossDet++: Growing crossline representation for object detection*. *IEEE Transactions on Circuits and Systems for Video Technology*, 33(3), 1093–1108. <https://doi.org/10.1109/TCSVT.2022.3211734>
- Rani, S., Ghai, D., & Kumar, S. (2022). *Object detection and recognition using contour-based edge detection and fast R-CNN*. *Multimedia Tools and Applications*, 81(29), 42183–42207. <https://doi.org/10.1007/s11042-021-11446-2>
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016, May 9). *You only look once: Unified, real-time object detection*. arXiv. <https://doi.org/10.48550/arXiv.1506.02640>
- Ren, S., He, K., Girshick, R., & Sun, J. (2015). *Faster R-CNN: Towards real-time object detection with region proposal networks*. arXiv. <https://doi.org/10.48550/arXiv.1506.01497>
- Sharma, V. K., & Mir, R. N. (2022). *Saliency guided faster-RCNN (SGFr-RCNN) model for object detection and recognition*. *Journal of King Saud University - Computer and Information Sciences*, 34(5), 1687–1699. <https://doi.org/10.1016/j.jksuci.2019.09.012>



**International Conference on Mathematics and
Mathematics Education (ICMME-2025), İstanbul Medeniyet
University, İstanbul, Türkiye, September 11-13, 2025**
“Mathematics in İstanbul, Bridge Between Continents”



- Song, B., Chen, J., Liu, W., Fang, J., Xue, Y., & Liu, X. (2025). *YOLO-ELWNet: A lightweight object detection network*. *Neurocomputing*, 636, 129904. <https://doi.org/10.1016/j.neucom.2025.129904>
- Song, L., & Geng, W. (2010, October). *A new camouflage texture evaluation method based on WSSIM and nature image features*. In *2010 International Conference on Multimedia Technology* (pp. 1–4). IEEE. <https://doi.org/10.1109/ICMULT.2010.5631434>
- Sorour, S. E., Aljaafari, M., Alarfaj, A. A., AlMusallam, W. H. A., & Aljoqiman, K. S. (2025). *Fine-tuned vision transformers and YOLOv11 for precise detection of pediatric adenoid hypertrophy*. *Alexandria Engineering Journal*, 128, 366–393. <https://doi.org/10.1016/j.aej.2025.05.038>
- Yang, F., Fan, D.-P., Ji, G.-P., Sun, T., Xu, D., & Cheng, M.-M. (2021, October). *Uncertainty-guided transformer reasoning for camouflaged object detection*. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)* (pp. 4126–4135). IEEE. <https://doi.org/10.1109/ICCV48922.2021.00411>
- Zheng, Y., Zhang, X., Wang, F., Cao, T., Sun, M., & Wang, X. (2019). *Detection of people with camouflage pattern via dense deconvolution network*. *IEEE Signal Processing Letters*, 26(1), 29–33. <https://doi.org/10.1109/LSP.2018.2825959>
- Zhou, T., Zhou, Y., Gong, C., Yang, J., & Zhang, Y. (2022). *Feature aggregation and propagation network for camouflaged object detection*. *arXiv*. <https://doi.org/10.48550/arXiv.2212.00990>



An Instructional Activity Based on Hypothetical Learning Trajectories: Modeling of Identities

Nehir Keyik

Hacettepe University, Faculty of Education, Ankara, Türkiye

Email: nehirkeyik@gmail.com

ABSTRACT

This study investigates the learning processes of 8th grade students in the context of teaching identities through instructional activities designed with a hypothetical learning trajectory framework. The central research problem was defined as: *"How do 8th grade students' learning processes develop when modeling identities in an instructional activity based on hypothetical learning trajectories?"* The study was conducted with one 8th grade student attending a public middle school in Samsun, Türkiye during the 2023–2024 academic year. The participant was selected through criterion sampling, while the school was chosen using convenience sampling.

Prior to implementation, prerequisite learning outcomes necessary for modeling identities were identified, and a readiness test was administered. Based on the results, a hypothetical learning trajectory was developed and applied step by step. The instructional design included activity worksheets, individual worksheets, observations, interviews, and an achievement test administered at the end of the process. The data were analyzed using qualitative content analysis, and expert opinions were sought to ensure validity and reliability.

Findings revealed that the student was able to reach the target learning outcome of modeling algebraic identities. Although no major obstacles were encountered, the student initially struggled to recognize that algebraic expressions could also be treated as common factors. Additional tasks were designed and integrated into the trajectory to overcome this difficulty. Consequently, the trajectory was revised to include this potential misconception, resulting in a final version of the learning pathway.

This study contributes to the field of mathematics education by demonstrating that instructional activities structured around hypothetical learning trajectories provide effective guidance for revealing students' developmental levels and addressing their misconceptions.



Furthermore, the integration of visual models such as algebra tiles supported students' conceptual understanding and helped bridge the gap between concrete and abstract reasoning. The results highlight the importance of designing mathematics instruction with careful attention to students' prior knowledge, anticipated learning steps, and possible learning obstacles. It is suggested that hypothetical learning trajectories can serve as a practical tool for mathematics teachers to improve lesson planning, anticipate student difficulties, and foster meaningful algebra learning.

Key Words: Hypothetical learning trajectory, algebra, identities, modeling, mathematics education

1. INTRODUCTION

The teaching and learning of mathematics play a crucial role in today's rapidly changing world. While mathematics learning varies depending on students' interests, abilities, and cognitive levels, the essential aim remains to ensure that mathematical knowledge is taught meaningfully. Meaningful learning enables students to apply knowledge in different contexts and to establish relationships both between concepts and across learning domains (Akkan, Baki, & Çakıroğlu, 2011). One of the distinctive features of research in mathematics education is its focus on uncovering students' learning processes. By examining these processes, researchers and educators can identify the most effective learning pathways that allow students to construct meaningful knowledge.

Hypothetical learning trajectories (HLTs) offer a framework that links students' ways of thinking and learning with instructional goals. Although students' understandings and reasoning processes cannot be directly observed, HLTs define observable goals, structures, skills, and behaviors (Bardsley, 2006). Learning trajectories not only inform teachers about students' developmental levels (Barrett et al., 2012; Clements & Sarama, 2004; Simon, 2006a) but also guide the evaluation of learning goals and processes (Battista, 2004; Confrey, Maloney, & Corley, 2014). In addition, they serve as a useful tool for teacher education (Mojica, 2010; Simon, 2006b). Research suggests that learning trajectories help trace how students' mathematical thinking develops over time (Confrey et al., 2008). Simon and Tzur (2004) emphasized that there are still gaps in the teaching and learning of many mathematical



concepts. Daro, Mosher, and Corcoran (2011) identified algebra as one of the domains with significant gaps in understanding learning trajectories.

Algebra is a powerful mode of thinking that enables students to generalize and model mathematical situations and to analyze them systematically. It allows individuals to investigate relationships that describe and make sense of the world. Developing these skills takes time and should begin early, continuing from kindergarten through secondary education. Algebraic awareness enriches students' mathematical thinking, supports future learning, and provides opportunities across a wide range of careers (NCTM, 2008). However, students often find algebra difficult to learn and understand, which results in anxiety and negative attitudes (Dede & Argün, 2003; Herscovics & Linchevski, 1994; Macgregor & Stacey, 1997; Wagner, 1981a; Wagner, 1981b). Numerous studies have confirmed that students face significant challenges in learning algebra (Stacey & Macgregor, 1997; Ersoy & Erbaş, 2003; Dede, 2004; Baki, 2008). Such difficulties can negatively affect their overall mathematical achievement.

One of the key topics in algebra instruction is **identities**, defined as equalities that hold true for all values of the variables involved. Effective teaching of identities can help students understand the notion of mathematical generalization, support inductive reasoning, and foster mathematical literacy. Moreover, this topic enables students to see algebraic expressions not only as abstract symbols but also as representations with geometric meaning. For this reason, the meaningful teaching of identities is critical in preparing students for further algebraic learning.

Learning trajectories provide a promising framework for teaching identities, as they help teachers anticipate students' possible misconceptions, design activities appropriate to developmental levels, and plan effective instruction. By focusing on how students' thinking develops step by step, teachers can design lessons that both reveal students' reasoning processes and support their conceptual growth. Therefore, the aim of this study is to investigate the learning processes of 8th grade students in modeling identities through instructional activities designed with a hypothetical learning trajectory.

2. EXPERIMENTAL DETAILS



This study was designed as a **qualitative case study** and was conducted in the 2023–2024 academic year with an 8th grade student in a public middle school located in the central district of Samsun, Türkiye. Case study methodology was chosen because it allows an in-depth and holistic examination of phenomena that cannot be controlled by the researcher and focuses on answering the “how” and “why” questions (Yıldırım & Şimşek, 2006).

2.1 Participants

The study group consisted of a single 8th grade student attending a public school in Samsun during the 2023–2024 academic year. The school was selected using **convenience sampling**, while the participant was chosen through **criterion sampling**. The criteria included high classroom interaction, good communication skills, and active participation in mathematics lessons, with the researcher also being the student’s mathematics teacher. The use of both sampling strategies was preferred due to the flexible nature of qualitative research. To ensure confidentiality, the student was coded as **ÖY**. The participant was a male student with strong performance in mathematics, good self-expression, and active engagement in the lessons.

2.2 Data Collection Tools

Following the principle that multiple data sources strengthen case studies (Hartley, 1995, as cited in Yıldırım & Şimşek, 2006), the following instruments were employed:

- *Readiness Test*: Designed to assess prerequisite knowledge required for modeling identities. It consisted of 15 open-ended questions aligned with curriculum outcomes. Table 1 presents the outcomes covered in the test.
- *Activity Worksheets*: Three sets of worksheets were prepared to support modeling of $(a+b)^2$, $(a-b)^2$, and a^2-b^2 through guided discovery.
- *Observation Notes*: The researcher systematically recorded the student’s behaviors during the activities, which were later categorized and analyzed.
- *Individual Worksheets*: Additional tasks were provided to reinforce the discovered identities and ensure transfer of learning.
- *Interviews*: Semi-structured interviews were conducted after each activity, enabling deeper exploration of the student’s reasoning.



- **Achievement Test:** An open-ended test with 8 items was administered at the end of the instructional sequence to evaluate whether the targeted learning outcomes were achieved.

Expert reviews were obtained prior to implementation to strengthen the content validity of the instruments, and necessary revisions were made accordingly.

Table 1. Learning Outcomes of the Readiness Test

Learning Outcomes	Number of Questions
Understands that simplifying and expanding do not change the value of a fraction and creates equivalent fractions (M.5.1.3.4).	1
Performs four operations with natural numbers considering the order of operations (M.6.1.1.2).	1
Applies factoring and distributive property to natural numbers (M.6.1.1.3).	1
Determines the prime factors of natural numbers (M.6.1.2.4).	1
Multiplies and divides integers (M.7.1.1.3).	1
Adds and subtracts integers and solves related problems (M.7.1.1.1).	2
Performs multi-step operations with rational numbers (M.7.1.3.3).	1
Multiplies and divides rational numbers (M.7.1.3.2).	1
Multiplies a natural number by an algebraic expression (M.7.2.1.2).	1
Understands the principle of equality preservation (M.7.2.2.1).	1
Calculates integer powers of integers (M.8.1.2.1).	1
Understands the basic rules of exponents and creates equivalent expressions (M.8.1.2.2).	1
Understands and writes simple algebraic expressions in different forms (M.8.2.1.1).	1
Multiplies algebraic expressions (M.8.2.1.2).	1

2.3 Data Collection

Prior to the implementation, a hypothetical learning trajectory (HLT) was developed to map the milestones the student needed to reach and the possible misconceptions that might arise. After the application, the trajectory was revised and finalized. The initial version of the HLT is presented in **Figure 1**.

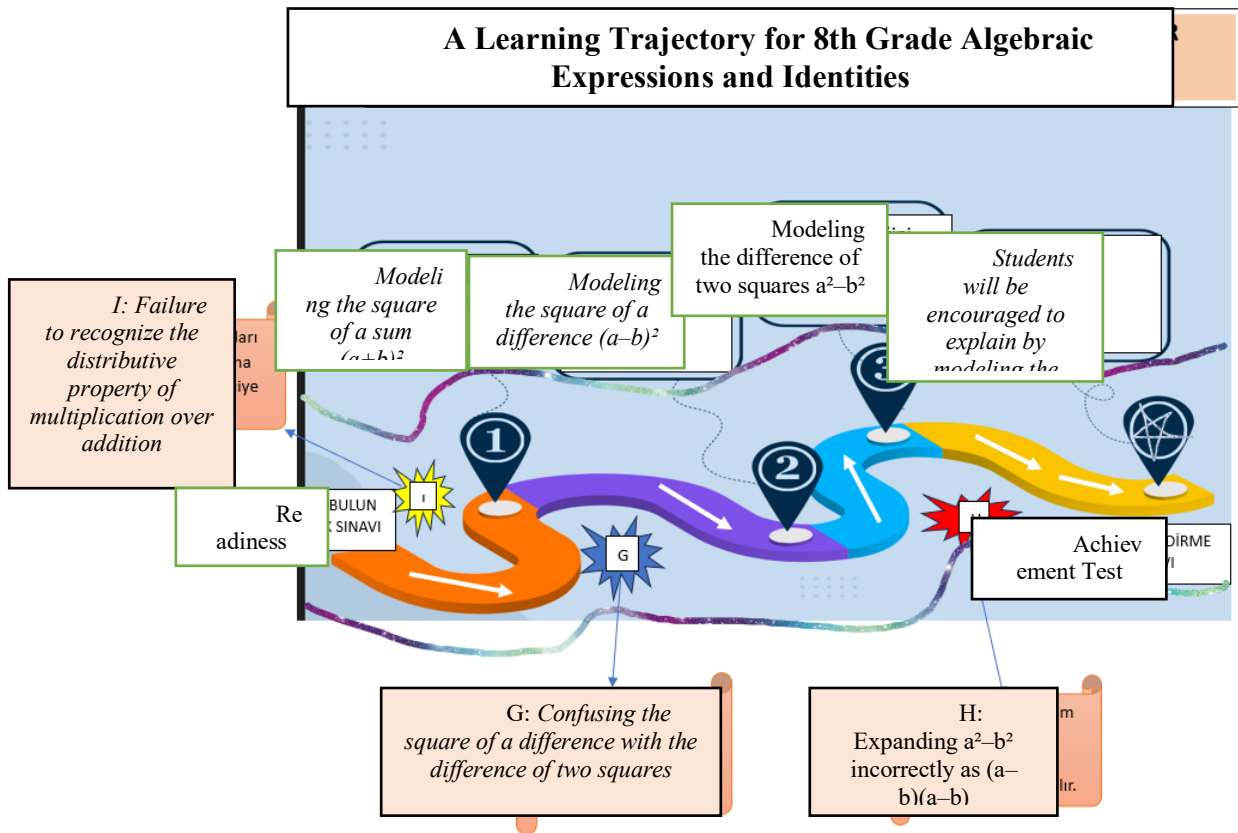


Figure 1. Hypothetical learning trajectory for modeling algebraic identities

According to the hypothetical learning trajectory in Figure 1, a readiness test was planned at the beginning of the instructional activity in order to predict the prior knowledge that the student was expected to have. The prepared trajectory included three milestones, each consisting of learning outcomes that the student needed to achieve in order to reach the target concept. For each milestone, activity worksheets were designed to guide the student in learning the intended outcomes.

In the first milestone, activities were included to enable the student to discover and model the identity of the square of a sum $(a+b)^2$. In the second milestone, activities were designed to help the student model the identity of the square of a difference $(a-b)^2$. In the third milestone, tasks were prepared to support the discovery and modeling of the difference of two squares identity a^2-b^2 .

Along the trajectory, potential misconceptions were anticipated and coded:

- On the way to the **first milestone**, the possible misconception was coded as I;



- On the way to the **second milestone**, the possible misconception was coded as **G**;
- On the way to the **third milestone**, the possible misconception was coded as **H**.

If the student encountered one of these misconceptions, supplementary worksheets were provided. After completing all milestones and the final worksheet including the target learning outcome, an achievement test was administered to evaluate whether the student had acquired the intended knowledge.

Misconceptions reported in the literature and their codes are presented in Table 2.

- Confusing the square of a difference with the difference of two squares (G)
- Expanding $a^2 - b^2$ incorrectly as $(a-b)(a-b)$ due to the minus sign (H)
- Failure to recognize the distributive property of multiplication over addition (I)

Table 2. Coding Scheme of Misconceptions Reported in the Literature

Misconceptions	Code
Confusing the square of a difference with the difference of two squares	G
Expanding $a^2 - b^2$ incorrectly as $(a-b)(a-b)$ due to the minus sign	H
Failure to recognize the distributive property of multiplication over addition	I

2.4 Validity and Reliability

In qualitative research, different terms and approaches exist regarding the processes of ensuring validity and reliability (Creswell, 2014). According to Creswell (2014), validity in qualitative research is described as the evaluation of the accuracy of the findings, while reliability refers to the consistency of the researcher's approach across different projects and among different researchers.

In this study, **detailed description** and **purposeful sampling** were employed to enhance validity. Creswell (2014) emphasizes that when qualitative researchers provide rich descriptions of the research setting, the results become more realistic and meaningful. Accordingly, in the present study, detailed descriptions were provided in the presentation of findings, direct quotations from interviews and documents were included, and the participant was selected



purposefully as an 8th grade student in line with the aim of the study. Furthermore, expert opinion was sought in the development of data collection instruments to ensure validity.

To establish reliability, efforts were made to ensure the **objectivity of the data**. The data obtained from the student's individual interviews were first analyzed by the researcher, and then independently recoded by another researcher who was knowledgeable about the topic but not involved in the study context. The coding results were compared and showed a high level of agreement. For codes that did not match, the coders came together to reach consensus. In addition, during the think-aloud protocols, the researcher periodically checked whether the participant's statements and behaviors were correctly understood. In cases of contradiction between verbal and behavioral data, the participant was given the opportunity to re-express his thoughts.

2.5 Teaching Process

First, the readiness test was prepared and administered to the participant (ÖY). Then, algebra tiles, paper, and scissors were provided to facilitate model construction. In the first lesson, the student completed Activity Worksheet 1, which contained four questions. The researcher observed the student during the activity and recorded notes. After completing Worksheet 1, an interview was conducted to examine the student's responses.

In the following lesson, the student was given Activity Worksheet 2. With the support of algebra tiles, paper, and scissors, the student constructed the required models and answered the four questions. Another individual interview was then conducted to analyze the reasoning process.

In the third stage, the student was given Activity Worksheet 3, along with the same materials, to complete the modeling tasks. After finishing this activity, the student participated in another interview.

After all milestone-related activities were completed, an achievement test was administered to determine whether the student had reached the targeted learning outcome. Finally, an interview was held in which the student was asked to explain and justify his answers on the test.



2.6 Data Analysis

The data obtained from the readiness test, activity worksheets, individual worksheets, observations, interviews, and the achievement test were analyzed using content analysis, a qualitative data analysis method. Content analysis requires a deep examination of data and allows the emergence of themes and codes that may not be initially apparent. The aim of such analysis is to present the findings in an organized and interpreted manner (Yıldırım & Şimşek, 2006).

Before the analysis, the researcher's observation notes were carefully reviewed, transcribed into digital format, and prepared as text suitable for analysis. Similarly, qualitative data from the interviews were transcribed according to the interview questions, and the achievement test responses were also digitized.

In order to facilitate analysis, data were coded using short, simple, and clear symbols (Karasar, 1998). Observation notes, interviews, and test data were repeatedly read and examined line by line, leading to the creation of a **code list** (Appendix 6). While generating the codes, both the related literature and the data collected from the field were considered. After coding, related codes were grouped, and themes that represented the main structure of the findings were formed. In the final stage of the analysis process, the codes and the conceptual framework consistent with these codes were defined and prepared for interpretation in the findings section.

3. RESULTS AND DISCUSSION

The findings related to the learning outcome of modeling algebraic identities were classified under **three main themes**:

- findings on modeling the identity of the square of a sum,
- findings on modeling the identity of the square of a difference, and
- findings on modeling the identity of the difference of two squares.

In this section, these findings are presented in detail.

3.1 Findings on Modeling the Identity of the Square of a Sum

To explore the student's discovery of the square of a sum identity, Activity Worksheet 1 was administered. In the first task, the student was asked to model a square with side length

$(a+b)$ using algebra tiles. The student attempted to construct the required shapes with algebra tiles and produced drawings similar to those shown in Figure 2.

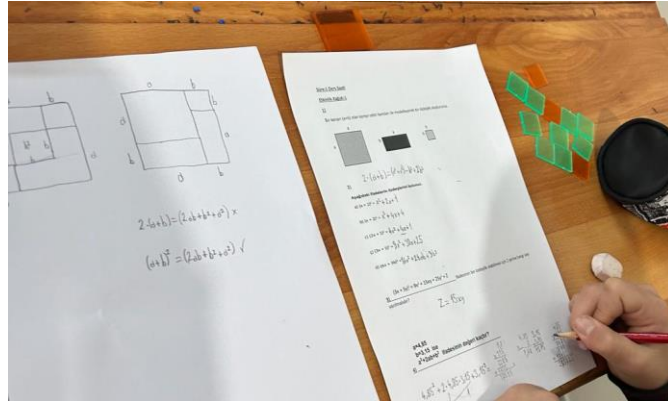


Figure 2. The student's drawing of a square with side length $(a+b)$ using algebra tiles

The participant (ÖY) initially modeled the required square as shown in Figure 3. However, with this model, he was unable to construct an identity related to its area. Upon this, the researcher asked: "I want you to construct a square with side length $(a+b)$, but how else could you create this square in a way that makes its area easier to calculate?" In response, ÖY created the model shown in Figure 4.

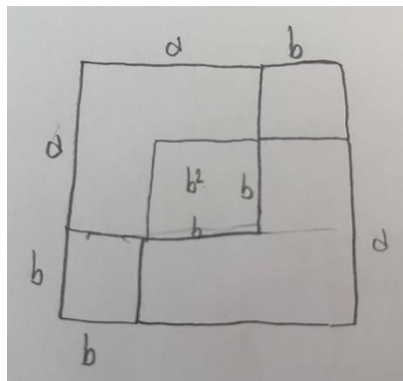


Figure 3. The student's first attempt to model the square with side length $(a+b)$

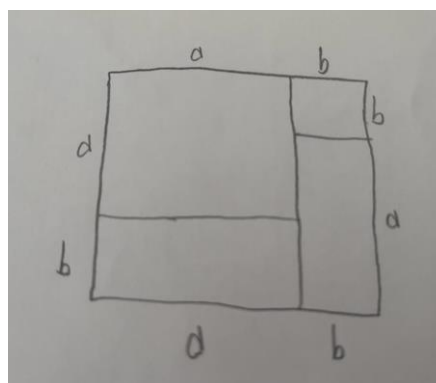


Figure 4. The student's revised model of the square with side length $(a+b)$ based on the researcher's guidance

When the researcher asked ÖY to construct the expression representing the area of the large square in Figure 4, the student wrote the areas one by one and expressed it as shown in Figure 5. When the researcher further asked, "How do we calculate the area of a square with side length $(a+b)$?" the student replied, "Sorry, teacher," and then reached the result shown in Figure 6.

$$2 \cdot (a+b) = (2ab + b^2 + a^2) \times$$

Figure 5. The student's expression of the area by summing the parts of the square

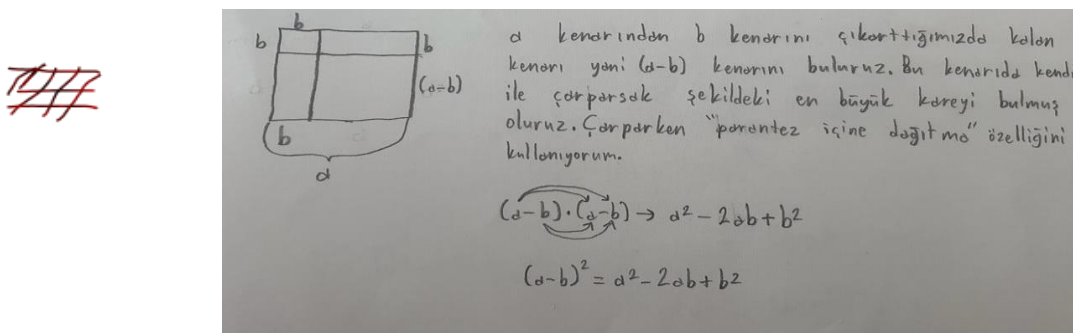
$$(a+b)^2 = (2ab + b^2 + a^2) \checkmark$$

Figure 6. The student's final expression for the area of a square with side length $(a+b)$

After constructing this identity, ÖY also completed the questions on the individual worksheet provided. In the achievement test, it was observed that ÖY answered all the questions requiring the modeling of the identity of the square of a sum correctly.

Findings on Modeling the Identity of the Square of a Difference

To enable ÖY to model the identity of the square of a difference, Activity Worksheet 2 was administered. In this worksheet, the student was asked to find the area of the shaded region in Figure 7 and construct an identity accordingly. ÖY solved the task as shown in Figure 7.



a kenarından b kenarını çıkarttığımızda kalan kenarı yani $(a-b)$ kenarını buluruz. Bu kenarıda kendi ile çarparsak şekildedeki en büyük kareyi bulmuş oluruz. Çarparken "parantez içine dağıtma" özelliğini kullanıyorum.

$$(a-b) \cdot (a-b) \rightarrow a^2 - 2ab + b^2$$

$$(a-b)^2 = a^2 - 2ab + b^2$$

Figure 7. The student's solution for modeling the identity of the square of a difference using the shaded region

Here, ÖY stated that the remaining square would have a side length of $(a-b)$, and by multiplying these two algebraic expressions, he constructed the identity of the square of a difference. Following Activity Worksheet 2, an individual worksheet was given to reinforce this

identity. ÖY answered all the questions on this worksheet correctly. It was also observed that, in the achievement test, he correctly solved all the questions requiring the modeling of the identity of the square of a difference.

Findings on Modeling the Identity of the Difference of Two Squares

To enable ÖY to construct the identity of the difference of two squares, scissors, paper, and a pencil were provided. He was asked to form an identity by calculating the area of the remaining region when a square with side length b is removed from a square with side length a . After modeling the required shape with paper, ÖY also drew the construction steps on paper, as illustrated in Figure 8.

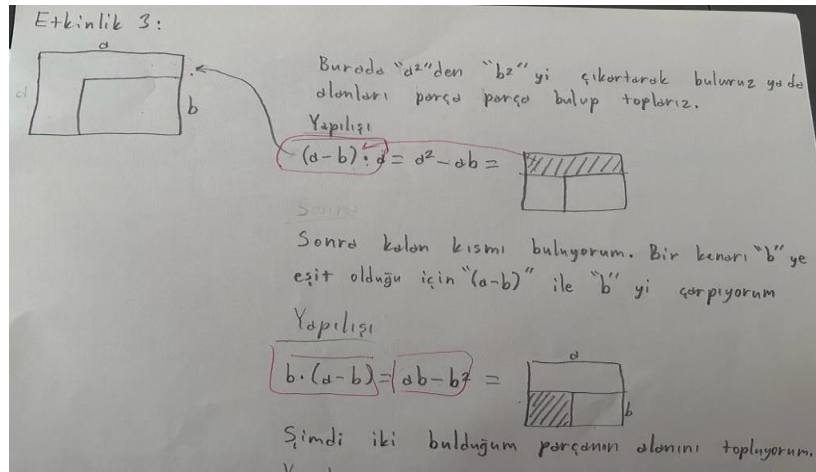


Figure 8. The student's modeling process for the identity of the difference of two squares ($a^2 - b^2$)

ÖY demonstrated that he could find the total area by adding the expressions $(a-b) \cdot a$ and $b \cdot (a-b)$, as shown in Figure 9.

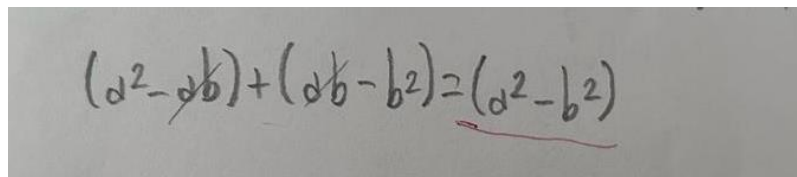
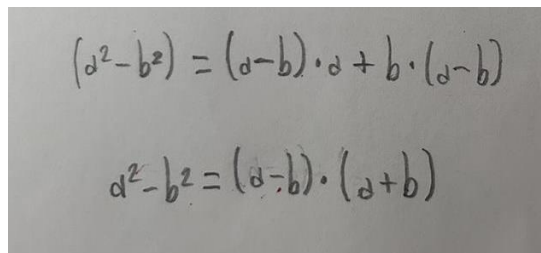


Figure 9. The student's representation of the total area by combining $(a-b) \cdot a$ and $b \cdot (a-b)$

Here, ÖY applied the distributive property of multiplication over subtraction by distributing the factors inside the parentheses and then summing the resulting expressions. The researcher then asked: "Well, without distributing the factors into the parentheses, how could we perform this operation in a shorter way?" in order to encourage the student to think about whether he could factor out $(a-b)$. The student replied, "Nothing else comes to mind, teacher."

At this point, the researcher realized that the student lacked sufficient knowledge about factoring and therefore introduced a supplementary activity. In this task, the researcher asked: "In the expression $8 \cdot 3 + 8 \cdot 7 = 8 \cdot (3+x)$, how do we find the value of x ?" The student correctly responded that $x = 7$. Then the researcher asked: "In the expression $x \cdot y^2 + z \cdot y^2 = y^2 \cdot (\dots\dots\dots)$, what should be written inside the parentheses?" The student answered $x+z$. Finally, when asked: "How would you factorize $(x+1) \cdot 3 + (x+1) \cdot 8$?" the student correctly gave the response $(x+1) \cdot (3+8) = (x+1) \cdot 9$.

Seeing that the student had reached the correct understanding, the researcher returned to the question in Activity Worksheet 3 and asked ÖY to express the problem by factoring. As a result, the student constructed the identity of the difference of two squares, as illustrated in Figure 10.



$$(a^2 - b^2) = (a-b) \cdot a + b \cdot (a-b)$$

$$a^2 - b^2 = (a-b) \cdot (a+b)$$

Figure 10. The student's construction of the identity of the difference of two squares by factoring

In one of the questions on the individual worksheet, it was observed that ÖY stated the identity $(4y^2 - 9)$ was equal to $(4y + 3)(4y - 3)$. Upon this, the researcher provided supplementary activities. By asking ÖY to multiply $(4y + 3)$ and $(4y - 3)$, the researcher enabled him to realize that the result was not equal to $4y^2 - 9$. Afterwards, the student was given similar questions to reinforce the concept.

At the end of the activities, it was observed that ÖY correctly answered all the questions in the achievement test that required modeling the identity of the difference of two squares.

5 RESULTS

The findings related to the achievement of modeling algebraic expressions were categorized under three main themes:

- i) Results on modeling the identity of the square of a sum
- ii) Results on modeling the identity of the square of a difference,



iii) Results on modeling the identity of the difference of two squares.

5.1 Results on Modeling the Identity of the Square of a Sum

ÖY initially modeled the identity of the square of a sum as shown in *Figure 3*. From the expression “the square of the sum of two terms,” the student understood that a square with side length equal to the sum of the two terms should be constructed. In *Figure 3*, the side of the square was indeed represented as $(a+b)$.

However, the student neglected the fact that the interior shapes of the square should be geometrical figures whose areas could be calculated. With a guiding question from the researcher, ÖY drew a new model as shown in *Figure 4*. In this model, the side of the square remained $(a+b)$, but the decomposition of the area included rectangles and smaller squares, making the calculation more explicit.

While finding the area of the larger square, ÖY added the areas of two smaller squares and two rectangles. Nevertheless, when expressing the square of the sum, he mistakenly wrote the exponent incorrectly as 2 due to carelessness rather than a conceptual misunderstanding. In both the individual worksheet and the achievement test, ÖY correctly answered all questions requiring the modeling of the square of a sum. Thus, the first milestone of the hypothetical learning trajectory was successfully completed, and it was confirmed that the student did not encounter the misconception coded as “I.”

5.2 Results on Modeling the Identity of the Square of a Difference

ÖY modeled the identity of the square of a difference accurately. He recognized that the side length of the relevant square should be the difference of two terms. Instead of subtracting the areas of smaller shapes from the larger square, ÖY calculated the area directly by multiplying $(a-b)$ by $(a-b)$. During the interview, when asked how else the area could be determined, he correctly explained that the shaded area could be obtained by subtracting the areas of two rectangles and a smaller square from the larger square. In both the individual worksheet and the achievement test, all related questions were answered correctly. Consequently, ÖY completed the second milestone successfully without encountering the misconception coded as “G.”

5.3 Results on Modeling the Identity of the Difference of Two Squares



According to ÖY's model in *Figure 8*, he understood that the identity of the difference of two squares represents the remaining area when the area of a smaller square is subtracted from that of a larger square. The student combined the remaining areas to calculate the total. When asked how to represent the process by factoring rather than adding the areas, ÖY stated that he did not understand. The researcher, suspecting a knowledge gap in factoring out common terms, administered supplementary questions. The activities revealed that ÖY struggled specifically with expressions where the common factor was algebraic. Through guidance, he recognized that algebraic expressions can also be treated as common factors and applied the factoring operation successfully. He then correctly identified the difference of two squares identity, as shown in *Figure 9*.

Following the activities, the student completed the individual worksheet and achievement test questions accurately, thereby completing the third milestone successfully without showing the misconception coded as "H." Moreover, an additional potential barrier—"failure to interpret algebraic expressions as common factors"—was identified and coded as "M," and the hypothetical learning trajectory was finalized accordingly.

6 DISCUSSION AND SUGGESTIONS

The hypothetical learning trajectory (HLT) prepared for this study helped to determine the prerequisite knowledge students must possess in order to model algebraic identities, which is considered a fundamental achievement in algebra. This awareness provides significant benefits for teachers in ensuring that a concept described as "the heart of mathematics" is built on solid foundations. By following the milestones step by step, the HLT revealed at which stages students might experience knowledge gaps or misconceptions (Myers, 2014). This was critical for supporting smooth progression through subsequent stages.

In this study, modeling activities were designed with the aid of algebra tiles to facilitate the modeling of identities. The teaching process revealed common errors and misconceptions, while simultaneously supporting conceptual understanding through the transition between visual and algebraic representations. The achievement test demonstrated that the student's performance improved positively throughout the process. Similar findings in the literature also highlight that the use of algebra tiles enhances students' understanding of algebraic concepts



and increases engagement in lessons (Aktaş, 2017; Çaylan, 2018; Karataş & Bahadır, 2018; Sharp, 1995).

Learning trajectories are considered valuable tools for teachers striving to raise every child to a high level of proficiency (Daro et al., 2011). Therefore, mathematics educators and institutions should recognize research on learning trajectories as an important area of study in mathematics education.

In the present study, the hypothetical learning trajectory was applied to a single 8th-grade student whose initial readiness level was high, as determined by the preparedness test. However, such a trajectory can also be applied to students with different readiness levels. Future research could design and implement learning trajectories for various topics across different grade levels, which would broaden the applicability and impact of this approach.

REFERENCES

- Akkan, Y., Baki, A., & Çakıroğlu, Ü. (2011). Aritmetik ile cebir arasındaki farklılıklar: Cebir öncesinin önemi. *İlköğretim Online*, 10(3), 812–823.
- Aktaş, G. S. (2017). Matematik öğretiminde somut materyaller ve tasarımları. *Pegem Atıf İndeksi*, 001–126.
- Baki, A., Akkan, Y., Atasoy, E., Çakıroğlu, Ü., & Güven, B. (2008, Ağustos 27–29). Aritmetikten cebire geçiş: Problem çözme stratejileri. 8. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi. Abant İzzet Baysal Üniversitesi, Bolu.
- Bardsley, M. E. (2006). Pre-kindergarten teachers' use and understanding of hypothetical learning trajectories in mathematics education (Doctoral dissertation). State University of New York, Buffalo, NY.
- Barrett, J. E., Sarama, J., Clements, D. H., Cullen, C., McCool, J., Witkowski-Rumsey, C., & Klanderma, D. (2012). Evaluating and improving a learning trajectory for linear measurement in elementary grades 2 and 3: A longitudinal study. *Mathematical Thinking and Learning*, 14(1), 28–54.
- Battista, M. T. (2004). Applying cognition-based assessment to elementary school students' development of understanding of area and volume measurement. *Mathematical Thinking and Learning*, 6(2), 185–204.
- Clements, D. H., & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning*, 6(2), 81–89.
- Confrey, J., Maloney, A., Nguyen, K., Wilson, P. H., & Mojica, G. (2008, Nisan). Synthesizing research on rational number reasoning. Working Session at the Research Presession of the National Council of Teachers of Mathematics, Salt Lake City, UT.
- Confrey, J., Maloney, A. P., & Corley, A. K. (2014). Learning trajectories: A framework for connecting standards with curriculum. *ZDM*, 46(5), 719–733.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative and mixed methods approaches* (4. baskı) (S. B. Demir, Çev. Ed.). Eğiten Kitap.
- Çaylan, B. (2018). Cebir karosu kullanımının altıncı sınıf öğrencilerinin cebir başarısı, cebirsel düşünceleri ve cebir karosu kullanımına ilişkin görüşleri üzerindeki etkileri (Yüksek lisans tezi). Orta Doğu Teknik Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.



- Daro, P., Mosher, F. A., & Corcoran, T. (2011). Learning trajectories in mathematics: A foundation for standards, curriculum, assessment, and instruction. Consortium for Policy Research in Education. http://repository.upenn.edu/cpre_researchreports/60
- Dede, Y., & Argün, Z. (2003). Cebir öğrencilere niçin zor gelmektedir? Hacettepe Eğitim Fakültesi Dergisi, 24, 180–185.
- Dede, Y. (2004). Değişken kavramı ve öğrenimindeki zorlukların belirlenmesi. Kuram ve Uygulamada Eğitim Bilimleri Dergisi, 4(1), 24–56.
- Ersoy, Y., & Erbaş, K. (2003). Kassel projesi cebir testinde bir grup Türk öğrencisinin başarısı ve öğrenme güçlükleri. İlköğretim Online Dergisi, 4(1), 18–39.
- Karasar, N. (2013). Bilimsel araştırma yöntemi. Nobel Yayıncılık.
- Karataş, C. G., & Bahadır, E. (2018). 8. sınıf cebirsel ifadeler ve özdeşlikler konusunun cebir gösterim kerosu materyali ile öğretilmesi ve materyalin kullanılabilirliğinin incelenmesi. Uluslararası Sosyal ve Eğitim Bilimleri Dergisi, 5(10), 209–224.
- Macgregor, M., & Stacey, K. (1997). Ideas about symbolism that students bring to algebra. The Mathematics Teacher, 90(2), 110–113.
- Mojica, G. F. (2010). Preparing pre-service elementary teachers to teach mathematics with learning trajectories (Doctoral dissertation). North Carolina State University, Raleigh, NC.
- Myers, M. (2014). The use of learning trajectory-based instruction (LTBI) in supporting equitable teaching practices in elementary classrooms: A multi-case study (Doctoral dissertation). North Carolina State University, Raleigh, NC.
- National Council of Teachers of Mathematics. (2008). Principles and standards for school mathematics. Reston, VA: Author.
- Sharp, J. M. (1995, Ekim). Results of using algebra tiles as meaningful representations of algebra concepts. Annual Meeting of the Mid-Western Education Research Association, Chicago, IL, ABD.
- Simon, M. A., & Tzur, R. (2004). Explicating the role of mathematical tasks in conceptual learning: An elaboration of the hypothetical learning trajectory. Mathematical Thinking and Learning, 6(2), 91–104.
- Simon, M. A. (2006a). Key developmental understandings in mathematics: A direction for investigating and establishing learning goals. Mathematical Thinking and Learning, 8(4), 359–371.
- Simon, M. A. (2006b). Pedagogical concepts as goals for teacher education: Towards an agenda for research in teacher development. In S. Alatorre, J. L. Cortina, M. Sáiz, & A. Méndez (Eds.), Proceedings of the Twenty-Eighth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education (Vol. 2, pp. 730–735). Universidad Pedagógica Nacional.
- Stacey, K. (1989). Finding and using patterns in linear generalising problems. Educational Studies in Mathematics, 20, 147–164.
- Tekin, H. H. (2006). Nitel araştırma yönteminin bir veri toplama tekniği olarak derinlemesine görüşme. İstanbul Üniversitesi Sosyoloji Dergisi, 3(13), 101–116.
- Wilson, P. H. (2009). Teachers' uses of a learning trajectory for equipartitioning (Doktora tezi). North Carolina State University, Raleigh, NC. <https://repository.lib.ncsu.edu/bitstream/handle/1840.16/2994>
- Wilson, P. H., Sztajn, P., Edgington, C., & Myers, M. (2015). Teachers' uses of a learning trajectory in student-centered instructional practices. Journal of Teacher Education, 66(3), 227–244.
- Yıldırım, A., & Şimşek, H. (2018). Sosyal bilimlerde nitel araştırma yöntemleri. Seçkin Yayıncılık.
- Wagner, S. (1981a). An analytical framework for mathematical variables. In C. Comti & G. Vernaud (Eds.), Proceedings of the Fifth Conference of the Psychology of Mathematics Education (pp. 165–170). Grenoble, France.
- Wagner, S. (1981b). Conservation of equation and function under transformations of variable. Journal for Research in Mathematics Education, 12(2), 107–118.



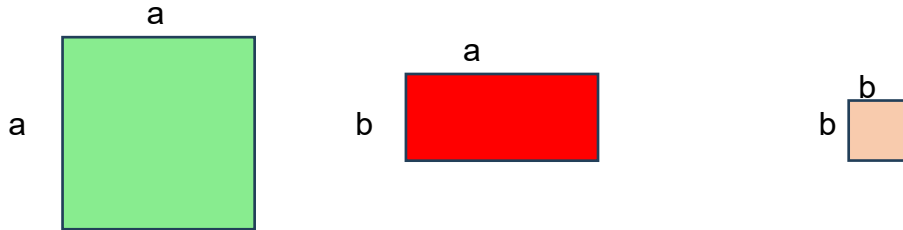
APPENDICES

APPENDIX-1:	Activity	Sheet	1
Duration:	1	Lesson	Hour

Activity Sheet-1

1)

Bir kenarı $(a+b)$ olan kareyi cebir karoları ile modelleyerek bir özdeşlik oluşturunuz.



2)

Aşağıdaki ifadelerin özdeşlerini bulunuz.

a) $(x + 1)^2 =$

b) $(x + 2)^2 =$

c) $(2x + 1)^2 =$

ç) $(3x + 5)^2 =$

d) $(4a + 3b)^2 =$

3)

$$(3x + 5y)^2 = 9x^2 + 15xy + 25y^2 + Z$$

ifadesinin bir özdeşlik olabilmesi için Z yerine hangi sayı yazılmalıdır?

$$a=4,85$$

$$b=3,15 \text{ ise}$$

$$a^2+2ab+b^2 \text{ ifadesinin değeri kaçtır?}$$

4)

Ek-2: Activity Sheet -2

Time: 1 Class Period



Bir kenar uzunluğu a br olan yandaki kareden bir kenar uzunluğu b br olan bir kare çıkarıldığında kalan bölgenin (yeşil kare) alanını göstererek bir özdeşlik oluşturunuz.

2)

Aşağıdaki cebirsel ifadelerin özdeşlerini yazalım.

I. $(a - 5)^2 =$

II. $(2x - 1)^2 =$

III. $(3x - 2y)^2 =$

a) $(x-5)^2 =$

d) $(2a-1)^2 =$

b) $(m-3)^2 =$

e) $(2e-3)^2 =$

c) $(5-2m)^2 =$

f) $(p-2a)^2 =$

3)

ÖRNEK

$$(Ax - B)^2 = 100x^2 - Cx + 9$$

Yukarıdaki özdeşliğe göre $C \div A - B$ ifadesinin değeri kaçtır?

4)

$$(2x - 3)^2 = 4x^2 + ax + 9$$

ifadesi bir özdeşlik olduğuna göre a kaçtır?

Appendix-3:

Activity

Sheet-3

Duration:

1

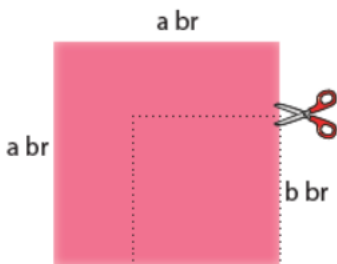
Class

Hour

Activity Sheet-3:

1)

Yandaki karenin alanı bir kenar uzunluğu a birimdir. Bu kareden bir kenar uzunluğu b birim olan kare çıkarılırsa kalan alanı gösteren cebirsel ifadeyi bulunuz.





2)

Aşağıdaki ifadelerin özdeşlerini iki kare farkı özdeşliğinden faydalanarak bulunuz.

a) $x^2 - 1 =$

b) $y^2 - 4 =$

c) $4z^2 - 9 =$

ç) $36a^2 - 144b^2 =$

d) $16n^2 - 9^2 =$

e) $64 - x^2 =$

f) $(9a)^2 - 25 =$

g) $(2k)^2 - (3p)^2 =$

3)

f. $(4 - a)^2 =$

g. $x^2 - 25 =$

h. $4a^2 - 16 =$

ı. $9x^2 - 49y^2 =$

j. $64y^2 - 1 =$

k. $100x^2 - 9 =$

l. $(4x)^2 - 1 =$

m. $(x + \frac{1}{x})^2 =$

n. $x^2 - \frac{1}{4} =$

o. $(-2x - 3y)^2 =$

Appendix-4:

Individual

Worksheet

Individual Worksheet

1)

Aşağıdaki ifadelerin özdeşlerini bulunuz.

a) $(x - 3).(x + 3) =$

b) $(2x + y).(2x - y) =$

c) $(3a - 5b).(3a + 5b) =$

ç) $(7 - x).(7 + x) =$

d) $(x + 6).(x + 6) =$

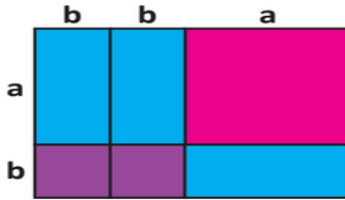
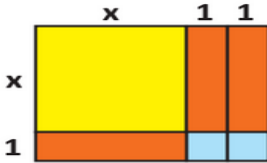
e) $(2n + 4).(2n + 4) =$

f) $(a - 2).(a - 2) =$

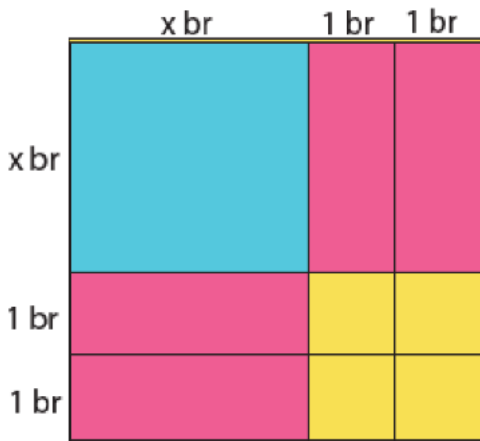
g) $(x + 9).(9 - x) =$

2)

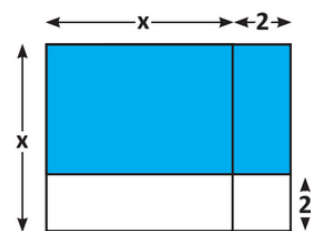
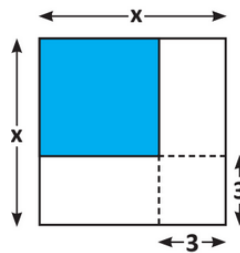
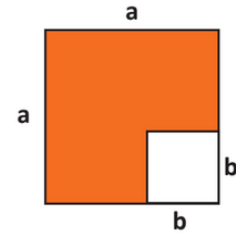
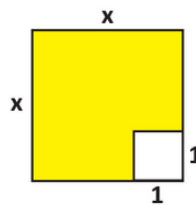
› Aşağıda modellenmiş özdeşlikleri yazınız.



3)



› Aşağıdaki taralı alanları temsil eden cebirsel ifadeleri ve bu cebirsel ifadelerin özdeşlerini yazınız.



4)

Yandaki şekli cebir karolarıyla oluşturarak alanını hesaplayınız.



- a) $x=1$ olursa şeklin alanı ne olur?
b) $x=3$ iken şeklin alanı ne olur?

5)

Kenar uzunlukları farkı 5 cm ve alanları farkı
45 cm² olan iki karenin çevreleri toplamı
kaçtır?

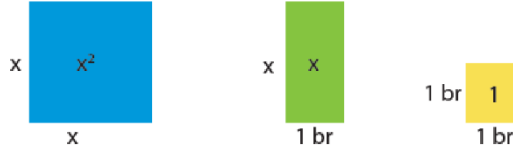
Appendix-5: Assessment Exam

Değerlendirme Sınavı

1. Aşağıda verilen çarpma işlemlerinin sonuçlarını bulunuz.

- a) $(b+1) \cdot (b+1)$ b) $m \cdot (3m-1)$ c) $y \cdot (-y+1)$ ç) $n \cdot 5n$

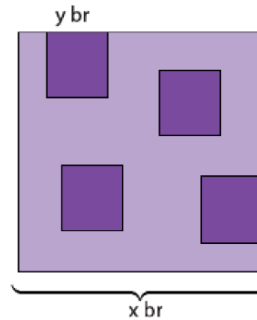
2. Yanda verilmiş olan cebir karolarını kullanarak kenar uzunlukları sırasıyla $(2x+1)$ ve $(x+2)$ olan iki karesel bölge modelleyiniz. Modellediğiniz karesel bölgelerin alanlarını özdeşlikler yardımıyla ifade ediniz.



3. Aşağıdaki tam kare özdeşliklerini modelleyiniz.

- a) $(5-b)^2 = 25 - 10b + b^2$ b) $(2a-3)^2 = 4a^2 - 12a + 9$

4. Yanda, bir kenar uzunluğu x br olan kareden kenar uzunluğu y br olan 4 kare kesilerek çıkarılıyor. Kalan bölgenin alanını özdeşlik yardımıyla ifade ediniz.



5. Aşağıdaki iki kare farkı özdeşliklerini modelleyiniz.

- a) $49x^2 - y^2 = (7x+y) \cdot (7x-y)$ b) $25 - b^2 = (5-b) \cdot (5+b)$

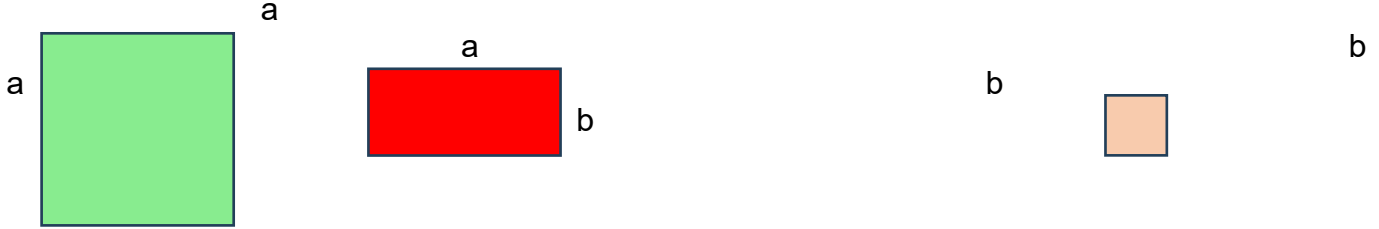
Başarılar





Lesson-1: The student is given algebra tiles and the following activity sheet, and is asked to complete the activity with the help of the algebra tiles.

Activity Sheet-1: Bir kenarı $(a+b)$ olan kareyi cebir karoları ile modelleyerek bir özdeşlik oluşturunuz.



NOT: $(a + b)^2 = a^2 + 2ab + b^2$ özdeşliğine **iki terimin toplamının karesi** özdeşliği veya **tam kare** özdeşliği denir.

Bilgileri öğrenci ile paylaşılır.

Öğrenciden aşağıdaki soruları cevaplandırması istenir.

Aşağıdaki ifadelerin özdeşlerini bulunuz.

a) $(x + 1)^2 =$

b) $(x + 2)^2 =$

c) $(2x + 1)^2 =$

ç) $(3x + 5)^2 =$

d) $(4a + 3b)^2 =$

$(3x + 5y)^2 = 9x^2 + 15xy + 25y^2 + Z$ ifadesinin bir özdeşlik olabilmesi için Z yerine hangi sayı yazılmalıdır?

Öğrenciye makas, kağıt ve aşağıdaki etkinlik kağıdı verilerek istenen modellemeyi yapması beklenir.

Activity Sheet

1)



Bir kenar uzunluğu a br olan yandaki kareden bir kenar uzunluğu b br olan bir kare çıkarıldığında kalan bölgenin (yeşil kare) alanını göstererek bir özdeşlik oluşturunuz.

NOT: $(a - b)^2 = a^2 - 2ab + b^2$ özdeşliğine **iki terimin farkının karesi** özdeşliği veya **tam kare** özdeşliği denir. Bilgisi öğrencilerle paylaşılır.



Aşağıdaki cebirsel ifadelerin özdeşlerini yazalım.

I. $(a - 5)^2 =$

II. $(2x - 1)^2 =$

III. $(3x - 2y)^2 =$

a) $(x-5)^2 =$

d) $(2a-1)^2 =$

b) $(m-3)^2 =$

e) $(2c-3)^2 =$

c) $(5-2m)^2 =$

f) $(p-2a)^2 =$

3)

ÖRNEK

$$(Ax - B)^2 = 100x^2 - Cx + 9$$

Yukarıdaki özdeşliğe göre $C \div A - B$ ifadesinin değeri kaçtır?

4)

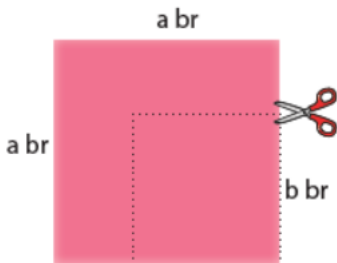
$$(2x - 3)^2 = 4x^2 + ax + 9$$

ifadesi bir özdeşlik olduğuna göre a kaçtır?

Activity Sheet -3

Öğrenciye kağıt, makas ve aşağıdaki etkinlik kağıdı verilerek etkinliği yapması istenir.

1)



Yandaki karenin alanı bir kenar uzunluğu a birimdir. Bu kareden bir kenar uzunluğu b birim olan kare çıkarılırsa kalan alanı gösteren cebirsel ifadeyi bulunuz.

Çözüm-1:

1. Cebirsel ifade:

Kenar uzunluğu a br olan karenin alanı: a^2

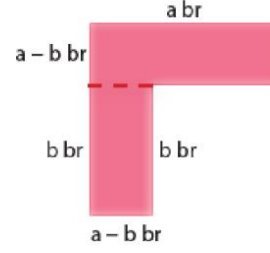
Kenar uzunluğu b br olan karenin alanı: b^2

Kalan alan : $a^2 - b^2$ olur.

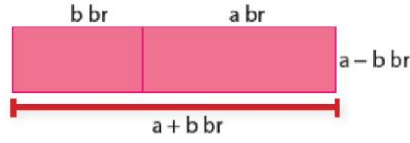
Çözüm-2:

2. Cebirsel ifade:

Kenar uzunluğu a br olan kareden kenar uzunluğu b br olan kare çıkarıldıktan sonra kalan alan yandaki gibidir. Bu şekli 2 dikdörtgene ayıralım.



Ayırdığımız dikdörtgenleri şekildeki gibi birleştirelim.



Elde ettiğimiz dikdörtgenin alanı $(a+b) \cdot (a-b)$ olur.

Birinci ve ikinci cebirsel ifadeler aynı alana ait olduğundan $a^2 - b^2 = (a+b) \cdot (a-b)$ olur.

NOT: $x^2 - y^2 = (x+y) \cdot (x-y)$ özdeşliğine iki kare farkı özdeşliği denir. (Öğrenci ile paylaşılır.)

2) Öğrencinin aşağıdaki soruları cevaplaması istenir.



2)

Aşağıdaki ifadelerin özdeşlerini iki kare farkı özdeşliğinden faydalananarak bulunuz.

a) $x^2 - 1 =$

b) $y^2 - 4 =$

c) $4z^2 - 9 =$

ç) $36a^2 - 144b^2 =$

d) $16n^2 - 9^2 =$

e) $64 - x^2 =$

f) $(9a)^2 - 25 =$

g) $(2k)^2 - (3p)^2 =$

3)

f. $(4 - a)^2 =$

g. $x^2 - 25 =$

h. $4a^2 - 16 =$

i. $9x^2 - 49y^2 =$

j. $64y^2 - 1 =$

k. $100x^2 - 9 =$

l. $(4x)^2 - 1 =$

m. $(x + \frac{1}{x})^2 =$

n. $x^2 - \frac{1}{4} =$

o. $(-2x - 3y)^2 =$

Etkinlik-3: Öğrencinin aşağıdaki soruları bireysel olarak cevaplandırması istenir.

Aşağıdaki ifadelerin özdeşlerini bulunuz.

a) $(x - 3).(x + 3) =$

b) $(2x + y).(2x - y) =$

c) $(3a - 5b).(3a + 5b) =$

ç) $(7 - x).(7 + x) =$

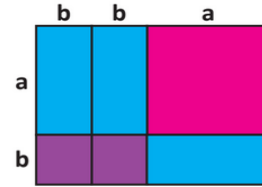
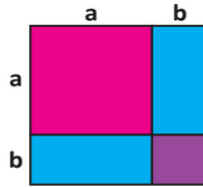
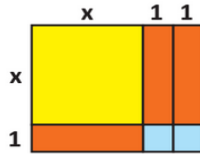
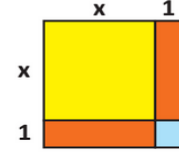
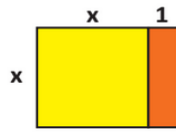
d) $(x + 6).(x + 6) =$

e) $(2n + 4).(2n + 4) =$

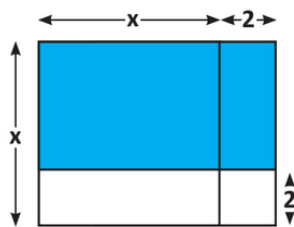
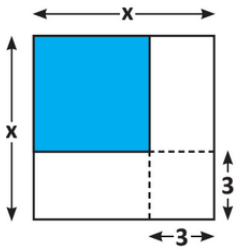
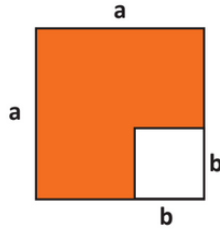
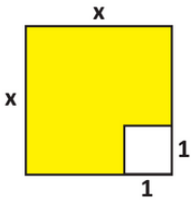
f) $(a - 2).(a - 2) =$

g) $(x + 9).(9 - x) =$

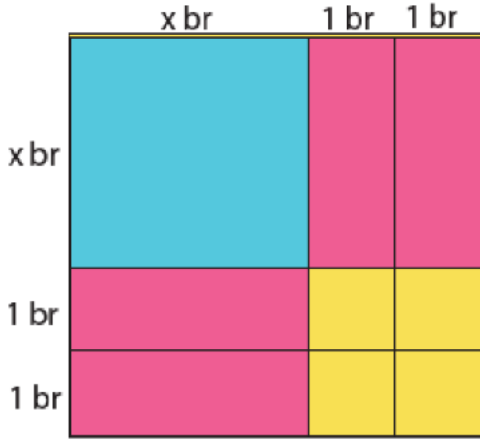
› Aşağıda modellenmiş özdeşlikleri yazınız.



› Aşağıdaki taralı alanları temsil eden cebirsel ifadeleri ve bu cebirsel ifadelerin özdeşlerini yazınız.



4)



Yandaki şekli cebir karolarıyla oluşturarak alanını hesaplayınız.

a) $x=1$ olursa şeklin alanı ne olur?

b) $x=3$ iken şeklin alanı ne olur?

Olası Cevaplar:

- 1) Bu bir kare olduğundan alanını $(x+2) \cdot (x+2)$ ile bulurum.
- 2) Ben cebir karolarının alanlarını tek tek bulup toplayarak $x^2 + 4x + 4$ bulurum.

5) Öğrencinin aşağıdaki soruyu cevaplandırması istenir.

- Kenar uzunlukları farkı 5 cm ve alanları farkı 45 cm^2 olan iki karenin çevreleri toplamı kaçtır?

Etkinlik-5:

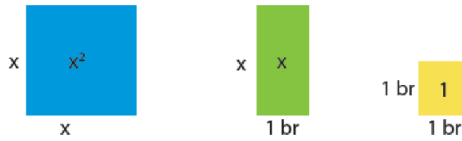
Öğrencilere aşağıdaki değerlendirme sınavı kağıtları dağıtılarak çözmeleri beklenir.

Değerlendirme Sınavı:

1. Aşağıda verilen çarpma işlemlerinin sonuçlarını bulunuz.

- a) $(b+1) \cdot (b+1)$ b) $m \cdot (3m-1)$ c) $y \cdot (-y+1)$ ç) $n \cdot 5n$

2. Yanda verilmiş olan cebir karolarını kullanarak kenar uzunlukları sırasıyla $(2x+1)$ ve $(x+2)$ olan iki karesel bölge modelleyiniz. Modellediğiniz karesel bölgelerin alanlarını özdeşlikler yardımıyla ifade ediniz.

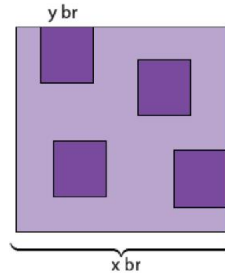


3. Aşağıdaki tam kare özdeşliklerini modelleyiniz.

- a) $(5-b)^2 = 25 - 10b + b^2$ b) $(2a-3)^2 = 4a^2 - 12a + 9$



4. Yanda, bir kenar uzunluğu x br olan kareden kenar uzunluğu y br olan 4 kare kesilerek çıkarılıyor. Kalan bölgenin alanını özdeşlik yardımıyla ifade ediniz.



5. Aşağıdaki iki kare farkı özdeşliklerini modelleyiniz.

a) $49x^2 - y^2 = (7x + y) \cdot (7x - y)$

b) $25 - b^2 = (5 - b) \cdot (5 + b)$



A new alpha power Rayleigh Weibull distribution with its applications to simulated and real data

Zainab Ali Ahmed ¹, Ali Talib Mohammed ² and Aytekin Enver³

^{1,2} *Department of Mathematics, College of Education for Pure Science (Ibn Al-Haitham),
University of Baghdad, Baghdad, Iraq.*

³Department of Mathematics, Gazi University, Türkiye

¹ Corresponding author: ali.t.m@ihcoedu.uobaghdad.edu.iq

² zainab.ali2203@ihcoedu.uobaghdad.edu.iq

³ aytekin.enver@gazi.edu.tr

Abstract

In this article, a statistically-enhanced lifetime density, with the name of the Alpha Power Rayleigh-Weibull Distribution (APRWD) is proposed. Emerging processes have been applied to the formulation of this distribution from a systematic generalization of the Rayleigh and Weibull families by employing the Alpha Power transformation technique. The parameter (α) as a tunable shape is introduced by this extension that enhances the distribution's responsiveness to diverse data structures and failure mechanisms. The inclusion of the alpha parameter grants the model flexibility to characterize a wide spectrum of hazard behaviors. Such as, monotonic increases or decreases, and configurations that are more complex, like the bathtub shape. The probability density function, the cumulative distribution function, and the survival function are the three main statistical functions that are included in the analytical development of the APRWD. We use both the Maximum Likelihood Estimation (MLE) and Ordinary Least Squares (OLS) methods to provide accurate parameter estimations. Additionally, meticulously planned simulation studies are used to test the estimators' accuracy and resilience. We apply the APRWD to actual datasets in order to validate the model's practicality. APRWD frequently produces a better fit and shows more modeling precision when handling lifespan and reliability data, according to comparisons with traditional models.

Keywords: moments about the origin, survival function, Rayleigh Weibull distribution, alpha-power family, and Monte Carlo simulation.



1. Introduction

In several scientific fields, including engineering and healthcare analytics, determining how long it takes particular processes to occur has grown in importance. As the systems under study continue to grow in complexity and interdependency, the statistical characteristics of event-time data have become more nuanced, prompting the need for more flexible and robust statistical methodologies. To meet these challenges, researchers have developed modified versions of classical probability distributions, tailoring them to better reflect real-world behavior and variability in system performance. Such modifications often result in models that can accommodate complex hazard rate structures and non-standard failure patterns. A key advancement in this area is the alpha power transformation, as introduced by Mahdavi and Kundu (2017), which modifies the baseline cumulative distribution function by raising it to a positive exponent, denoted by α . This transformation effectively generates a family of enhanced distributions capable of modeling diverse shapes and hazard rates. Because of this flexibility, the alpha power transformation has proven particularly useful in applications involving lifetime and reliability data, where standard distributions often fall short in addressing the asymmetry and variability inherent in empirical observations. The corresponding expressions for the cumulative and probability density functions under this transformation define what is known as:

$$G(x) = \begin{cases} \frac{\alpha^{F(x)} - 1}{\alpha - 1}, & x, \alpha, \vartheta, \rho > 0, \text{ and } \alpha \neq 1 \\ F(x) & \alpha = 1 \end{cases} \quad (1)$$

$$g(x) = \begin{cases} \frac{\log(\alpha)}{\alpha - 1} \alpha^{F(x)} f(x) & x, \alpha > 0, \text{ and } \alpha \neq 1 \\ f(x) & \alpha = 1 \end{cases} \quad (2)$$

This new approach of (APT) has been used by many researchers to produce new distributions (Mandouh et al., 2022) such as M. Nassar employed the APT with Weibull distribution (Nassar et al., 2017). The Alpha Power inverse Weibull (Basheer, 2019), alpha power lindly distribution (Hassan et al., 2019), alpha power pareto distribution (Ihtisham et al., 2019), alpha power teissier distribution (Eghwerido, 2021), alpha power Aradhana (Mohiuddin & Kannan, 2021), alpha power exponential Weibull distribution (Nassar et al., 2018). The Alpha Power method has received the attention of researchers and has undergone many developments and improvements (Kargbo et al., 2023). The Generalized Alpha Power Exponentiated, gull alpha power Weibull distribution (Ijaz et al., 2020), the new generalization of



Gull Alpha Power (Kilal et al., 2022), Alpha Power for Odd Generalized of Exponential Family of Distributions (Elbatal et al., 2022), The exponentiated generalized (APT) (ElSherpieny & Almetwally, 2022), The Extended of (APT) (Ahmad et al., 2019). While other researchers studied transmuted, Alpha power for exponentiated inverse Rayleigh distribution and weighted Rayleigh distribution (Ahmad et al., 2014; Ali et al., 2021; Zain & Hussein, 2020). We have investigated the Rayleigh and Weibull distributions in different cases with alpha power, which in turn is the building block in the presented research (Ganji et al., 2016; Ahmad et al., 2017; Nasiru, 2016; Zain & Hussein, 2021). Discussed some aspects of the properties and applications, characterization and estimation, and related function of weighted for Rayleigh Weibull distribution (Mohammed & Hussein, 2018). The two parameters Rayleigh Weibull distribution introduced by (Mohammed & Hussein, 2018) with cdf and pdf are given as follows:

$$F(x) = 1 - e^{-\left(\frac{\vartheta}{2}x^2 + x^\rho\right)}, x, \vartheta, \rho > 0 \quad (3)$$

$$f(x) = e^{-\left(\frac{\vartheta}{2}x^2 + x^\rho\right)} (\vartheta x + \rho x^{\rho-1}) \quad (4)$$

Researchers that used the Alpha Power method with the exponential Weibull distribution (Saleh & Mohammed, 2024; Mohammed, 2019; Mohammed & Hussein, 2019) introduced new lifetime Shanker-Weibull distribution and Comparison with many other distributions (Ashor & Mohammed, 2024; Mahdi, 2020). The next moves are summarized in the mathematical construction of (APRWD), the expansion of pdf, and the investigation of the mathematical and statistical properties of (APRWD).

Structure of New APRWD

If the random variable X is distributed APRWD with two scale parameters $(\vartheta \text{ \& } \rho)$, and one shape parameter α . Where the cdf and pdf respectively given as:

$$G(x) = \begin{cases} \frac{\alpha^{1-\varepsilon} - 1}{\alpha - 1} & x > 0; \alpha \neq 1; \alpha, \rho, \vartheta > 0 \\ 1 - e^{-\left(\frac{\vartheta}{2}x^2 + x^\rho\right)} & \alpha = 1 \end{cases} \quad (5)$$

$$g(x) = \begin{cases} \frac{\alpha \log(\alpha)}{\alpha - 1} \alpha^{-\varepsilon} e^{-\left(\frac{\vartheta}{2}x^2 + x^\rho\right)} (\vartheta x + \rho x^{\rho-1}) & \alpha \neq 1 \\ (\vartheta x + \rho x^{\rho-1}) e^{-\left(\frac{\vartheta}{2}x^2 + x^\rho\right)} & \alpha = 1 \end{cases} \quad (6)$$

The survival (reliability) and hazard function of Alpha Power Rayleigh Weibull Distribution can be obtained as

$$S(x) = \begin{cases} 1 - \left(\frac{\alpha^{1-\varepsilon} - \left(\frac{\vartheta}{2} x^2 + x^\rho \right) - 1}{\alpha - 1} \right) & \alpha \neq 1 \\ e^{-\left(\frac{\vartheta}{2} x^2 + x^\rho \right)} & \alpha = 1 \end{cases} \quad (7)$$

$$h(x) = \begin{cases} \frac{\alpha^{1-\varepsilon} - \left(\frac{\vartheta}{2} x^2 + x^\rho \right) \log(\alpha) e^{-\left(\frac{\vartheta}{2} x^2 + x^\rho \right)} (\vartheta x + \rho x^{\rho-1})}{\alpha - \alpha^{1-\varepsilon} - \left(\frac{\vartheta}{2} x^2 + x^\rho \right)} & \alpha \neq 1 \\ (\vartheta x + \rho x^{\rho-1}) & \alpha = 1 \end{cases} \quad (8)$$

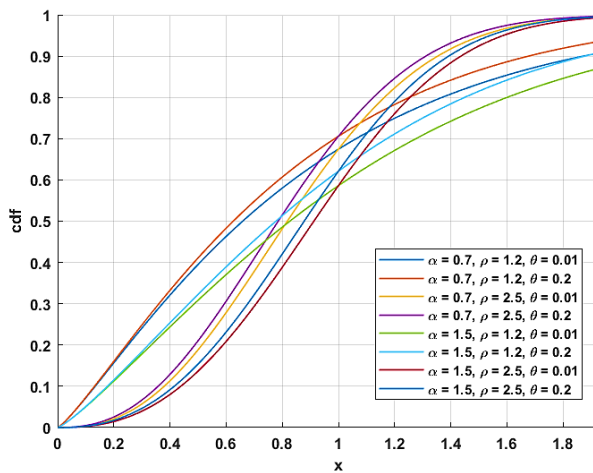


Figure 1. Plots of cdf

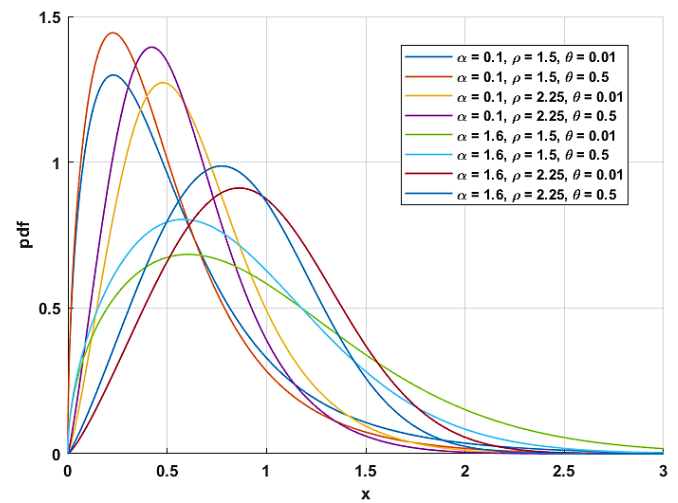


Figure 2. Plots of pdf

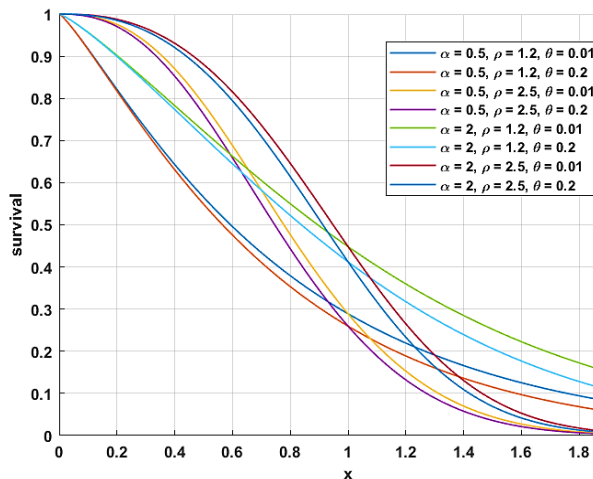


Figure 3. Plots of S(x)

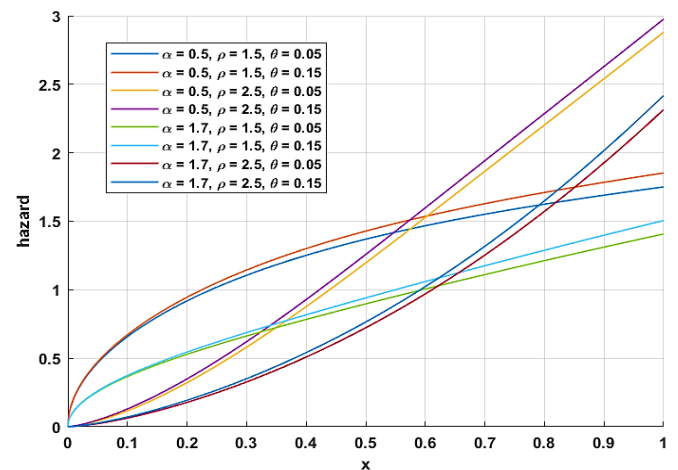


Figure 4. Plots of h(x)

Expanding the Probability Density Function

$$\alpha^{1-\varepsilon} - \left(\frac{\vartheta}{2} x^2 + x^\rho \right) = \alpha \cdot \alpha^{-\varepsilon} - \left(\frac{\vartheta}{2} x^2 + x^\rho \right) = \alpha \cdot (e^{\log(\alpha)})^{-\varepsilon} - \left(\frac{\vartheta}{2} x^2 + x^\rho \right) = \alpha \cdot e^{-\log(\alpha) \cdot \varepsilon} - \left(\frac{\vartheta}{2} x^2 + x^\rho \right)$$

$$\begin{aligned}
 e^{-\log(\alpha) \cdot \frac{\vartheta}{2} x^2 + x^\rho} &= \sum_{i=0}^{\infty} \frac{(-1)^i (\log(\alpha))^i}{i!} e^{-i \left(\frac{\vartheta}{2} x^2 + x^\rho \right)} \\
 g(x) &= \frac{\alpha \log(\alpha)}{\alpha - 1} (\vartheta x + \rho x^{\rho-1}) e^{-\log(\alpha)} e^{-\left(\frac{\vartheta}{2} x^2 + x^\rho \right)} e^{-\left(\frac{\vartheta}{2} x^2 + x^\rho \right)} \\
 &= \frac{\alpha}{\alpha - 1} \sum_{i=0}^{\infty} \frac{(-1)^i (\log(\alpha))^{i+1}}{i!} (\vartheta x + \rho x^{\rho-1}) e^{-(i+1) \frac{\vartheta}{2} x^2} e^{-(i+1) x^\rho} \\
 e^{-(i+1) x^\rho} &= \sum_{j=0}^{\infty} \frac{(-1)^j (i+1)^j}{j!} x^{j\rho} \\
 g(x) &= \frac{\alpha}{\alpha - 1} \sum_{i,j=0}^{\infty} \frac{(-1)^{i+j} (i+1)^j (\log(\alpha))^{i+1}}{i! j!} (\vartheta x^{j\rho+1} + \rho x^{\rho(j+1)-1}) e^{-(i+1) \frac{\vartheta}{2} x^2} \\
 \text{let } \xi_{i,j} &= \frac{\alpha}{\alpha - 1} \sum_{i,j=0}^{\infty} \frac{(-1)^{i+j} (i+1)^j (\log(\alpha))^{i+1}}{i! j!} \\
 g(x) &= \xi_{i,j} \cdot (\vartheta x^{j\rho+1} + \rho x^{\rho(j+1)-1}) e^{-(i+1) \frac{\vartheta}{2} x^2} \quad (9)
 \end{aligned}$$

Statistical and Mathematical Properties of (APRWD)

This section presents a detailed discussion of the fundamental mathematical properties of the distribution, including the moments, moment generating function, quantile function, as well as the coefficients of skewness and kurtosis.

The Quantile Function

The CDF of (APRWD) when $x > 0; \alpha \neq 1$

$$\begin{aligned}
 G(x) &= u \Rightarrow x = G^{-1}(u) \\
 u &= \frac{\alpha^{1-e^{-\left(\frac{\vartheta}{2} x^2 + x^\rho \right)}} - 1}{\alpha - 1} \\
 e^{-\left(\frac{\vartheta}{2} x^2 + x^\rho \right)} &= 1 - \left(\frac{\alpha - 1}{\log(u(\alpha - 1) + 1)} \right) \log(\alpha)
 \end{aligned}$$

$$x^\rho + \frac{\vartheta}{2} x^2 + \log \left(1 - \left(\frac{\log(u(\alpha - 1) + 1)}{\log(\alpha)} \right) \right) = 0 \quad (10)$$

Finding the roots of equation (10) represented by values of x requires using methods for solving nonlinear equations.

The median point at which the cumulative distribution function equal to 0.5 ($u = 0.5$).

$$x^\rho + \frac{\vartheta}{2} x^2 + \log \left(1 - \left(\frac{\log(0.5\alpha + 0.5)}{\log(\alpha)} \right) \right) = 0$$

Moments

Moments are fundamental descriptors of a probability distribution, providing valuable information about the central tendency and the dispersion of data relative to the origin. The r^{th} moment of a continuous random variable X about the origin is mathematically defined as follows:

$$M'_r = E(x^r) = \int_0^{\infty} x^r \cdot g(x) dx = \xi_{i,j} \int_0^{\infty} (\vartheta x^{j\rho+r+1} + \rho x^{\rho(j+1)+r-1}) e^{-(i+1)\frac{\vartheta}{2}x^2} dx$$

$$\text{let } y = (i+1)\frac{\vartheta}{2}x^2 \Rightarrow x = \sqrt{\frac{2y}{\vartheta(i+1)}} \Rightarrow dx = \frac{dy}{\sqrt{y}\sqrt{2\vartheta(i+1)}}$$

$$M'_r = \vartheta \xi_{i,j} \frac{2^{\frac{1}{2}(j\rho+r+1)-\frac{1}{2}}}{(\vartheta(i+1))^{\frac{1}{2}(j\rho+r+1)+\frac{1}{2}}} \cdot \Gamma\left(\frac{1}{2}(j\rho+r+1)+\frac{1}{2}\right) + \rho \xi_{i,j} \frac{2^{\frac{1}{2}(\rho(j+1)+r-1)-\frac{1}{2}}}{(\vartheta(i+1))^{\frac{1}{2}(\rho(j+1)+r-1)+\frac{1}{2}}} \cdot \Gamma\left(\frac{1}{2}(\rho(j+1)+r-1)+\frac{1}{2}\right) \quad (11)$$

The first moment is the mean of the distribution.

$$\mu_x = M'_1 = \vartheta \xi_{i,j} \frac{2^{\frac{1}{2}(j\rho+2)-\frac{1}{2}}}{(\vartheta(i+1))^{\frac{1}{2}(j\rho+2)+\frac{1}{2}}} \cdot \Gamma\left(\frac{1}{2}(j\rho+2)+\frac{1}{2}\right) + \rho \xi_{i,j} \frac{2^{\frac{1}{2}(\rho(j+1)-\frac{1}{2})-\frac{1}{2}}}{(\vartheta(i+1))^{\frac{1}{2}(\rho(j+1)-\frac{1}{2})+\frac{1}{2}}} \cdot \Gamma\left(\frac{1}{2}(\rho(j+1)-\frac{1}{2})+\frac{1}{2}\right)$$

The variance of the distribution is defined by

$$M'_2 = \vartheta \xi_{i,j} \frac{2^{\frac{1}{2}(j\rho+3)-\frac{1}{2}}}{(\vartheta(i+1))^{\frac{1}{2}(j\rho+3)+\frac{1}{2}}} \cdot \Gamma\left(\frac{1}{2}(j\rho+3)+\frac{1}{2}\right) + \rho \xi_{i,j} \frac{2^{\frac{1}{2}(\rho(j+1)+1)-\frac{1}{2}}}{(\vartheta(i+1))^{\frac{1}{2}(\rho(j+1)+1)+\frac{1}{2}}} \cdot \Gamma\left(\frac{1}{2}(\rho(j+1)+1)+\frac{1}{2}\right)$$

$$\text{Var}(x) = M'_2 - (M'_1)^2$$

Moments Generating Function

$$M_x(t) = E(e^{tx}) = \int_0^{\infty} e^{tx} g(x) dx = \int_0^{\infty} e^{tx} \cdot \xi_{i,j} \cdot (\vartheta x^{j\rho+1} + \rho x^{\rho(j+1)-1}) e^{-(i+1)\frac{\vartheta}{2}x^2} dx$$

$$e^{tx} = \sum_{k=0}^{\infty} \frac{t^k}{k!} x^k$$

$$M_x(t) = \xi_{i,j} \sum_{k=0}^{\infty} \frac{t^k}{k!} \int_0^{\infty} (\vartheta x^{j\rho+k+1} + \rho x^{\rho(j+1)+k-1}) e^{-(i+1)\frac{\vartheta}{2}x^2} dx$$

$$M_x(t) = \vartheta \xi_{i,j} \sum_{k=0}^{\infty} \frac{t^k}{k!} \frac{2^{\frac{1}{2}(j\rho+k+1)-\frac{1}{2}}}{(\vartheta(i+1))^{\frac{1}{2}(j\rho+k+1)+\frac{1}{2}}} \cdot \Gamma\left(\frac{1}{2}(j\rho+k+1)+\frac{1}{2}\right)$$

$$+ \rho \xi_{i,j} \sum_{k=0}^{\infty} \frac{t^k}{k!} \frac{2^{\frac{1}{2}(\rho(j+1)+k-1)-\frac{1}{2}}}{(\vartheta(i+1))^{\frac{1}{2}(\rho(j+1)+k-1)+\frac{1}{2}}} \cdot \Gamma\left(\frac{1}{2}(\rho(j+1)+k-1)+\frac{1}{2}\right) \quad (12)$$

Coefficients of Skewness (C.S) And Kurtosis (C.K)

$$M'_3 = \vartheta \xi_{i,j} \frac{2^{\frac{1}{2}(j\rho+4)-\frac{1}{2}}}{(\vartheta(i+1))^{\frac{1}{2}(j\rho+4)+\frac{1}{2}}} \cdot \Gamma\left(\frac{1}{2}(j\rho+4)+\frac{1}{2}\right) + \rho \xi_{i,j} \frac{2^{\frac{1}{2}(\rho(j+1)+2)-\frac{1}{2}}}{(\vartheta(i+1))^{\frac{1}{2}(\rho(j+1)+2)+\frac{1}{2}}} \cdot \Gamma\left(\frac{1}{2}(\rho(j+1)+2)+\frac{1}{2}\right)$$

$$C.S = \frac{M'_3}{(M'_2)^{\frac{3}{2}}} \quad (13)$$

$$M'_4 = \vartheta \xi_{i,j} \frac{2^{\frac{1}{2}(\rho+5)-\frac{1}{2}}}{(\vartheta(i+1))^{\frac{1}{2}(\rho+5)+\frac{1}{2}}} \cdot \Gamma\left(\frac{1}{2}(\rho+5)+\frac{1}{2}\right) + \rho \xi_{i,j} \frac{2^{\frac{1}{2}(\rho(j+1)+3)-\frac{1}{2}}}{(\vartheta(i+1))^{\frac{1}{2}(\rho(j+1)+3)+\frac{1}{2}}} \cdot \Gamma\left(\frac{1}{2}(\rho(j+1)+3)+\frac{1}{2}\right)$$

$$C.K = \frac{M'_4}{(M'_2)^2} - 3 \quad (14)$$

Table 1. The first - fourth moments, variance, skewness, and kurtosis for the distribution

α	ρ	ϑ	μ'_1	μ'_2	μ'_3	μ'_4	K	S	var
1.5	0.5	0.7	0.7315	1.0761	2.0581	4.6415	1.0080	1.8436	0.5411
	0.8	1.2	0.6584	0.7097	0.9762	1.5820	0.1410	1.6329	0.2762
2	0.8	0.5	0.8785	1.3159	2.5673	5.9767	0.4517	1.7008	0.5441
	1.5	1.2	0.7519	0.7478	0.8902	1.2123	-0.8319	1.3767	0.1824
3.5	2.5	1.5	0.8046	0.7545	0.7870	0.8916	-1.4340	1.2007	0.1071
	2	2.5	0.6993	0.5901	0.5666	0.6015	-1.2726	1.2500	0.1011

Order Statistics

The problems concerned with estimating parameters based on the use of order statistics are considered. Though simple, these examples show the relevance of order statistics in real world applications.

$$g_{j:n}(x) = \frac{n!}{(j-1)(n-j)} g(x)(G(x))^{j-1}(1-G(x))^{n-j}$$

Substituting the pdf and cdf of APRWD

$$(1-G(x))^{n-j} = \sum_{m=0}^{n-j} \binom{n-j}{m} (-1)^m (G(x))^m$$

$$g_{j:n}(x) = \sum_{m=0}^{n-j} \binom{n-j}{m} \frac{n!(-1)^m}{(j-1)(n-j)} g(x)(G(x))^{j+m-1}$$

$$g_{j:n}(x) = \sum_{m=0}^{n-j} \binom{n-j}{m} \frac{n!(-1)^m}{(j-1)(n-j)} \left(\frac{\alpha \log(\alpha)}{\alpha-1} \alpha^{-\log(\alpha)} e^{-\left(\frac{\vartheta}{2}x^2+x^\rho\right)} e^{-\left(\frac{\vartheta}{2}x^2+x^\rho\right)} (\vartheta x + \rho x^{\rho-1}) \right) \left(\frac{\alpha^{1-\varepsilon} e^{-\left(\frac{\vartheta}{2}x^2+x^\rho\right)} - 1}{\alpha-1} \right)^{j+m-1}$$

$$g_{j:n}(x) = \alpha \log(\alpha) \sum_{m=0}^{n-j} \left(\frac{1}{\alpha-1} \right)^{j+m} \binom{n-j}{m} \frac{n!(-1)^m}{(j-1)(n-j)} \left(\alpha^{-\log(\alpha)} e^{-\left(\frac{\theta}{2}x^2+x^\rho\right)} e^{-\left(\frac{\theta}{2}x^2+x^\rho\right)} (\vartheta x + \rho x^{\rho-1}) \right) \left(\alpha^{1-\varepsilon} e^{-\left(\frac{\theta}{2}x^2+x^\rho\right)} - 1 \right)^{j+m-1} \quad (15)$$

$$g_{1:n}(x) = \alpha \log(\alpha) \sum_{m=0}^{n-1} \left(\frac{1}{\alpha-1} \right)^{1+m} \binom{n-1}{m} n(-1)^m \left(\alpha^{-\log(\alpha)} e^{-\left(\frac{\theta}{2}x^2+x^\rho\right)} e^{-\left(\frac{\theta}{2}x^2+x^\rho\right)} (\vartheta x + \rho x^{\rho-1}) \right) \left(\alpha^{1-\varepsilon} e^{-\left(\frac{\theta}{2}x^2+x^\rho\right)} - 1 \right)^{+m}$$

$$g_{n:n}(x) = \alpha \log(\alpha) n \left(\frac{1}{\alpha-1} \right)^n \left(\alpha^{-\log(\alpha)} e^{-\left(\frac{\theta}{2}x^2+x^\rho\right)} e^{-\left(\frac{\theta}{2}x^2+x^\rho\right)} (\vartheta x + \rho x^{\rho-1}) \right) \left(\alpha^{1-\varepsilon} e^{-\left(\frac{\theta}{2}x^2+x^\rho\right)} - 1 \right)^{+m}$$

Renyi Entropy

Entropy is a measurable quantity that holds central importance in multiple scientific disciplines, such as thermodynamics, information theory, and statistical mechanics. Conceptually, it reflects the degree of randomness or uncertainty associated with a system, especially under perturbative conditions. Over time, several mathematical expressions of entropy have been proposed. Among them is Renyi entropy, which generalizes the classical entropy framework and is mathematically defined as follows:

$$I_R(\eta) = \frac{1}{1-\eta} \log \left(\int_0^\infty (g(x))^\eta dx \right)$$

$$I_R(\eta) = \frac{1}{1-\eta} \log \left(\left(\frac{\alpha \log(\alpha)}{\alpha-1} \right)^\eta \int_0^\infty (\vartheta x + \rho x^{\rho-1})^\eta e^{-\eta \log(\alpha)} e^{-\left(\frac{\theta}{2}x^2+x^\rho\right)} e^{-\eta \left(\frac{\theta}{2}x^2+x^\rho\right)} dx \right)$$

$$e^{-\eta \log(\alpha)} e^{-\left(\frac{\theta}{2}x^2+x^\rho\right)} = \sum_{s=0}^{\infty} \frac{(-1)^s (\eta)^s (\log(\alpha))^s}{s!} e^{-s \left(\frac{\theta}{2}x^2+x^\rho\right)}$$

$$I_R(\eta) = \frac{1}{1-\eta} \log \left(\left(\frac{\alpha}{\alpha-1} \right)^\eta \sum_{s=0}^{\infty} \frac{(-1)^s (\eta)^s (\log(\alpha))^{\eta+s}}{s!} \int_0^\infty (\vartheta x + \rho x^{\rho-1})^\eta e^{-(s+\eta) \left(\frac{\theta}{2}x^2\right)} e^{-(s+\eta)(+x^\rho)} dx \right)$$

$$e^{-(s+\eta)(x^\rho)} = \sum_{b=0}^{\infty} \frac{(-1)^b (s+\eta)^b}{b!} x^{b\rho}$$

$$I_R(\eta) = \frac{1}{1-\eta} \log \left(\left(\frac{\alpha}{\alpha-1} \right)^\eta \sum_{s=0}^{\infty} \sum_{b=0}^{\infty} \sum_{z=0}^{\infty} \binom{\eta}{z} (\vartheta)^{\eta-z} \rho^z \frac{(-1)^{s+b} (\eta)^s (s+\eta)^b (\log(\alpha))^{\eta+s}}{s! b!} \int_0^\infty x^{\eta+\rho z-2z+b\rho} e^{-(s+\eta) \left(\frac{\theta}{2}x^2\right)} dx \right)$$

$$\text{let } y = (s + \eta) \frac{\vartheta}{2} x^2 \Rightarrow x = \sqrt{\frac{2y}{\vartheta(s + \eta)}} \Rightarrow dx = \frac{dy}{\sqrt{y} \sqrt{2\vartheta(s + \eta)}}$$

$$I_R(\eta) = \frac{1}{1 - \eta} \log \left(\left(\frac{\alpha}{\alpha - 1} \right)^\eta \sum_{s=0}^{\infty} \sum_{b=0}^{\infty} \sum_{z=0}^{\infty} \binom{\eta}{z} (\vartheta)^{\eta-z} \rho^z \frac{(-1)^{s+b} (\eta)^s (s + \eta)^b (\log(\alpha))^{\eta+s} (2)^{\frac{1}{2} + \rho z - 2z + b\rho + \eta - \frac{1}{2}}}{s! b! (\vartheta(s + \eta))^{\frac{1}{2} + \eta + \rho z - 2z + b\rho - \frac{1}{2} + 1}} \int_0^\infty (y)^{\frac{1}{2} + \eta + \rho z - 2z + b\rho - \frac{1}{2} + 1 - 1} e^{-(y)} dx \right)$$

$$I_R(\eta) = \frac{1}{1 - \eta} \log \left(\pi_{s,b,z} \cdot \Gamma\left(\frac{1}{2} + \eta + \rho z - 2z + b\rho - \frac{1}{2} + 1\right) \right) \quad (16)$$

$$\pi_{s,b,z} = \left(\frac{\alpha}{\alpha - 1} \right)^\eta \sum_{s,b,z=0}^{\infty} \binom{\eta}{z} (\vartheta)^{\eta-z} \rho^z \frac{(-1)^{s+b} (\eta)^s (s + \eta)^b (\log(\alpha))^{\eta+s} (2)^{\frac{1}{2} + \rho z - 2z + b\rho + \eta - \frac{1}{2} + 1}}{s! b! (\vartheta(s + \eta))^{\frac{1}{2} + \eta + \rho z - 2z + b\rho - \frac{1}{2} + 1}}$$

Parameters Estimation

Maximum likelihood estimation

Let $x_1, x_2, \dots, x_n$ be a random samples from $\text{APRWD}(\alpha, \rho, \vartheta)$, then for estimating $(\hat{\alpha}, \hat{\rho}, \hat{\vartheta})$ of APRWD using MLE by maximizing the following likelihood function:

$$L(x_i; \vartheta, \rho, \alpha) = \prod_{i=1}^n g(x_i; \vartheta, \rho, \alpha) = \prod_{i=1}^n \left(\frac{\alpha \log(\alpha)}{\alpha - 1} \alpha^{-e^{-\left(\frac{\vartheta}{2} x_i^2 + x_i^\rho\right)}} e^{-\left(\frac{\vartheta}{2} x_i^2 + x_i^\rho\right)} (\vartheta x + \rho x_i^{\rho-1}) \right)$$

$$L = \left(\frac{\alpha \ln(\alpha)}{\alpha - 1} \right)^n \prod_{i=1}^n (\vartheta x + \rho x_i^{\rho-1}) \alpha^{-e^{-\sum_{i=1}^n \left(\frac{\vartheta}{2} x_i^2 + x_i^\rho\right)}} e^{-\sum_{i=1}^n \left(\frac{\vartheta}{2} x_i^2 + x_i^\rho\right)}$$

$$\hat{L} = n \ln(\alpha) + n \ln(\ln(\alpha)) - n \ln(\alpha - 1)$$

$$+ \sum_{i=1}^n \ln(\vartheta x + \rho x_i^{\rho-1}) + \ln(\alpha) \sum_{i=1}^n e^{-\sum_{i=1}^n \left(\frac{\vartheta}{2} x_i^2 + x_i^\rho\right)} - \sum_{i=1}^n \left(\frac{\vartheta}{2} x_i^2 + x_i^\rho \right)$$

$$\frac{\partial \ln L}{\partial \alpha} = \frac{n}{\alpha} + \frac{n}{\alpha \ln(\alpha)} - \frac{n}{(\alpha - 1)} - \frac{1}{\alpha} \sum_{i=1}^n e^{-\left(\frac{\vartheta}{2} x_i^2 + x_i^\rho\right)} \quad (17)$$

$$\frac{\partial \ln L}{\partial \rho} = \sum_{i=1}^n \frac{(\rho \ln x_i - 1) x_i^{\rho-1}}{(\vartheta x + \rho x_i^{\rho-1})} - \ln(\alpha) \sum_{i=1}^n e^{-\left(\frac{\vartheta}{2} x_i^2 + x_i^\rho\right)} x_i^\rho \ln(x_i) - \sum_{i=1}^n x_i^\rho \ln(x_i) \quad (18)$$

$$\frac{\partial \ln L}{\partial \vartheta} = \sum_{i=1}^n \frac{x}{(\vartheta x + \rho x_i^{\rho-1})} + \ln(\alpha) \sum_{i=1}^n e^{-\left(\frac{\vartheta}{2} x_i^2 + x_i^\rho\right)} \frac{x_i^2}{2} - \sum_{i=1}^n \frac{x_i^2}{2} \quad (19)$$

To solve the above equations, they must be set to zero to find the root, which represents the estimated value for each parameter. Numerical methods, such as the Newton-Raphson method, are used to solve nonlinear equations, and this can be implemented using MATLAB.

Ordinary Least Squares Estimation Method

Let $x_1, x_2, \dots, x_n$ be a random samples from $\text{APRWD}(\alpha, \rho, \vartheta)$, then for estimating $(\hat{\alpha}, \hat{\rho}, \hat{\vartheta})$ of APRWD using OLS by minimizing the following function:

$$W(x; \vartheta, \rho, \alpha) = \sum_{j=1}^n \varepsilon_j^2 = \sum_{j=1}^n \left(G(x) - \frac{j}{n+1} \right)^2 = \sum_{j=1}^n \left(\frac{\alpha^{1-\varepsilon}^{-\left(\frac{\vartheta}{2} x_j^2 + x_j^\rho\right)} - 1}{\alpha - 1} - \frac{j}{n+1} \right)^2$$

$$\frac{\partial W}{\partial \alpha} = \frac{2}{(\alpha - 1)^2} \sum_{j=1}^n \left(\frac{\alpha^{1-\varepsilon}^{-\left(\frac{\vartheta}{2} x_j^2 + x_j^\rho\right)} - 1}{\alpha - 1} - \frac{j}{n+1} \right) \left((\alpha - 1) \left(1 - e^{-\left(\frac{\vartheta}{2} x_j^2 + x_j^\rho\right)} \right) \alpha^{-\varepsilon}^{-\left(\frac{\vartheta}{2} x_j^2 + x_j^\rho\right)} - \left(\alpha^{1-\varepsilon}^{-\left(\frac{\vartheta}{2} x_j^2 + x_j^\rho\right)} \right) \right) \quad (20)$$

$$\frac{\partial W}{\partial \rho} = \frac{2}{(\alpha - 1)} \sum_{j=1}^n \left(\frac{\alpha^{1-\varepsilon}^{-\left(\frac{\vartheta}{2} x_j^2 + x_j^\rho\right)} - 1}{\alpha - 1} - \frac{j}{n+1} \right) \left(\alpha^{1-\varepsilon}^{-\left(\frac{\vartheta}{2} x_j^2 + x_j^\rho\right)} e^{-\left(\frac{\vartheta}{2} x_j^2 + x_j^\rho\right)} \ln(\alpha) x_j^\rho \ln(x_j) \right) \quad (21)$$

$$\frac{\partial W}{\partial \vartheta} = \frac{2}{(\alpha - 1)} \sum_{j=1}^n \left(\frac{\alpha^{1-\varepsilon}^{-\left(\frac{\vartheta}{2} x_j^2 + x_j^\rho\right)} - 1}{\alpha - 1} - \frac{j}{n+1} \right) \left(\alpha^{1-\varepsilon}^{-\left(\frac{\vartheta}{2} x_j^2 + x_j^\rho\right)} e^{-\left(\frac{\vartheta}{2} x_j^2 + x_j^\rho\right)} \ln(\alpha) \frac{x_j^2}{2} \right) \quad (22)$$

To obtain parameter estimates, the system of equations is transformed by equating them to zero, allowing numerical techniques to identify the roots. In cases involving nonlinear behavior, methods like Newton-Raphson are commonly applied. Such algorithms are often executed using platforms such as MATLAB for computational efficiency

Applications

This part of the study is structured into two sections. The first explores the simulation-based evaluation, while the second applies the proposed method to real data, with emphasis on computing relevant information criteria for model comparison.

Simulation

A Monte Carlo simulation was conducted to examine the statistical performance of parameter estimation methods for the Alpha Power Rayleigh Weibull Distribution (APRWD). The evaluation centered on two primary accuracy indicators: bias and mean squared error (MSE), following the framework outlined in reference (Shalan, 2024). To achieve this, random samples were generated based on Equation (10), using different sample sizes ($n = 10, 30, 50, 100$, and 150), and repeated across $L = 2000$ and $L = 1000$ replications. Parameter estimates were obtained via both Maximum Likelihood Estimation (MLE) and Least Ordinate Squares (OLS), with starting values specified as $\alpha = 1.5, \rho = 1.25, \vartheta = 0.75$. The simulations were implemented using MATLAB (version 2018a), and a summary of the outcomes is provided in Tables 2 and 3. Results indicated that the MLE consistently outperforms OLS across all sample sizes, yielding lower bias and smaller MSE values. These findings support the robustness and superior efficiency of the MLE method in estimating APRWD parameters under the simulated conditions.

Table 2. MSE and Bias values of MLE and OLS methods for $\alpha = 1.5, \rho = 1.25, \vartheta = 0.75$ with $L=2000$

n	Est. Par.	MLE	Bias	OLS	Bias	Best
10	$\hat{\alpha}$	0.6578	0.7326	0.8516	-0.4066	MLE



30	$\hat{\rho}$	0.0743	-0.0178	0.2151	0.2233	MLE
	$\hat{\theta}$	0.0453	-0.0681	0.2221	-0.0571	MLE
	$\hat{\alpha}$	0.5020	0.6473	0.5897	-0.1740	MLE
50	$\hat{\rho}$	0.0435	-0.0114	0.1241	0.1218	MLE
	$\hat{\theta}$	0.0208	-0.0228	0.1923	-0.0782	MLE
	$\hat{\alpha}$	0.4393	0.6121	0.5699	-0.1272	MLE
100	$\hat{\rho}$	0.0314	-0.0202	0.0954	0.0931	MLE
	$\hat{\theta}$	0.0140	-0.0067	0.1894	-0.1036	MLE
	$\hat{\alpha}$	0.3501	0.5627	0.5146	-0.0753	MLE
150	$\hat{\rho}$	0.0196	-0.0427	0.0639	0.0614	MLE
	$\hat{\theta}$	0.0070	0.0159	0.1744	-0.0930	MLE
	$\hat{\alpha}$	0.3245	0.5477	0.4780	-0.0554	MLE
150	$\hat{\rho}$	0.0158	-0.0476	0.0510	0.0529	MLE
	$\hat{\theta}$	0.0047	0.0252	0.1580	-0.0959	MLE

Table 3. MSE and Bias values of MLE and OLS methods for $\alpha = 1.5, \rho = 1.25, \theta = 0.75$ with L=1000

n	Est. Par.	MLE	Bias	OLS	Bias	Best
10	$\hat{\alpha}$	0.6827	0.7455	0.8471	0.3854	MLE
	$\hat{\rho}$	0.0710	0.0035	0.2063	0.2194	MLE
	$\hat{\theta}$	0.0437	-0.0624	0.2261	0.0640	MLE
30	$\hat{\alpha}$	0.5021	0.6468	0.6321	0.1819	MLE
	$\hat{\rho}$	0.0419	-0.0154	0.1236	0.1093	MLE
	$\hat{\theta}$	0.0250	-0.0324	0.2109	0.1044	MLE
50	$\hat{\alpha}$	0.4376	0.6114	0.5672	0.1550	MLE
	$\hat{\rho}$	0.0327	-0.0252	0.0889	0.0859	MLE
	$\hat{\theta}$	0.0155	-0.0071	0.1973	0.1193	MLE
100	$\hat{\alpha}$	0.3725	0.5742	0.5045	0.0514	MLE
	$\hat{\rho}$	0.0214	-0.0437	0.0605	0.0536	MLE
	$\hat{\theta}$	0.0064	0.0169	0.1636	0.0852	MLE
150	$\hat{\alpha}$	0.3137	0.5415	0.4628	0.0465	MLE
	$\hat{\rho}$	0.0139	-0.0495	0.0499	0.0520	MLE
	$\hat{\theta}$	0.0043	0.0259	0.1615	0.0955	MLE

A further Monte Carlo simulation was conducted to investigate the precision of parameter estimation using both MLE and OLS methods under the configuration $\alpha = 2, \rho = 0.5$, and $\theta = 0.2$, employing replication sizes of L = 2000 and 1000. The outcomes indicated that although MLE generally delivered consistent results, the OLS approach achieved better estimation accuracy in certain conditions. Specifically, for some parameters, OLS yielded lower bias and marginally smaller mean squared error (MSE), particularly when the true values of the parameters were relatively small. As observed in Tables 4 and 5, although MLE is widely regarded as the standard approach due to its favorable asymptotic behavior, the OLS method



can exhibit comparable, or in some cases even improved, performance under certain parameter settings or when dealing with limited sample sizes

Table 4. MSE and Bias values of MLE and OLS methods for $\alpha = 2, \rho = 0.5$, and $\vartheta = 0.2$ with $L=2000$

n	Est. Par.	MLE	Bias	OLS	Bias	Best
10	$\hat{\alpha}$	0.2786	0.3629	1.1131	-0.7714	MLE
	$\hat{\rho}$	0.0275	0.0763	0.0334	0.1222	MLE
	$\hat{\vartheta}$	0.0553	0.1320	0.0146	-0.0219	OLS
30	$\hat{\alpha}$	0.2344	0.3175	0.8646	-0.1842	MLE
	$\hat{\rho}$	0.0062	0.0384	0.0215	0.0814	MLE
	$\hat{\vartheta}$	0.0144	0.0598	0.0096	-0.0090	OLS
50	$\hat{\alpha}$	0.2183	0.3039	0.6657	-0.0838	MLE
	$\hat{\rho}$	0.0032	0.0260	0.0154	0.0660	MLE
	$\hat{\vartheta}$	0.0054	0.0385	0.0078	-0.0048	MLE
100	$\hat{\alpha}$	0.1672	0.2521	0.4957	0.0085	MLE
	$\hat{\rho}$	0.0012	0.0164	0.0097	0.0510	MLE
	$\hat{\vartheta}$	0.0025	0.0257	0.0055	-0.0001	MLE
150	$\hat{\alpha}$	0.1431	0.2280	0.3848	0.0154	MLE
	$\hat{\rho}$	0.0007	0.0121	0.0069	0.0452	MLE
	$\hat{\vartheta}$	0.0015	0.0203	0.0040	0.0007	MLE

Table 5. MSE and Bias values of MLE and OLS methods for $\alpha = 2, \rho = 0.5$, and $\vartheta = 0.2$ with $L=1000$

n	Est. Par.	MLE	Bias	OLS	Bias	Best
10	$\hat{\alpha}$	0.2763	0.3599	1.0794	-0.7443	MLE
	$\hat{\rho}$	0.0235	0.0751	0.0345	0.1243	MLE
	$\hat{\vartheta}$	0.0527	0.1253	0.0146	-0.0202	OLS
30	$\hat{\alpha}$	0.2417	0.3237	0.8413	-0.1607	MLE
	$\hat{\rho}$	0.0062	0.0367	0.0202	0.0801	MLE
	$\hat{\vartheta}$	0.0162	0.0631	0.0098	-0.0124	OLS
50	$\hat{\alpha}$	0.2126	0.2991	0.6611	-0.0718	MLE
	$\hat{\rho}$	0.0030	0.0264	0.0148	0.0621	MLE
	$\hat{\vartheta}$	0.0071	0.0422	0.0075	-0.0035	MLE
100	$\hat{\alpha}$	0.1683	0.2507	0.4796	0.0078	MLE
	$\hat{\rho}$	0.0013	0.0159	0.0097	0.0537	MLE
	$\hat{\vartheta}$	0.0024	0.0264	0.0051	-0.0017	MLE
150	$\hat{\alpha}$	0.1487	0.2364	0.3982	0.0304	MLE
	$\hat{\rho}$	0.0012	0.0132	0.0065	0.0426	MLE
	$\hat{\vartheta}$	0.0016	0.0207	0.0039	0.0035	MLE

Information Criteria

Information criteria are widely used statistical measures for comparing alternative models based on both their fit to the data and structural complexity. These metrics assist in assessing model adequacy, simplicity, and robustness, thus promoting more rigorous analysis and facilitating



evidence-based model selection. Commonly used measures such as AIC, BIC, and HQIC are applied to compare the performance of the APRWD distribution against alternative statistical distributions.

$$AIC = 2K - \hat{L}, AIC_c = AIC + \frac{2k(k+1)}{n-K-1}, BIC = k \ln(n) - 2\hat{L}, \text{ and } HQIC = -2\hat{L} + 2K \ln(\ln(n))$$

The first real data represents the failure time rates of turbochargers in 40 types of heavy machinery engines, with an average failure time of hours. These data were utilized by (Xu et al., 2003) and are presented as follows: (7.3, 2.0, 6.7, 3.0, 3.5, 6.3, 4.5, 8.0, 4.8, 5.0, 9.0, 5.3, 8.4, 5.8, 6.0, 8.7 6.0, 8.3, 3.9 , 6.5, 6.5, 2.6 , 7.0, 7.1, 5.4, 5.6 1.6, 7.3, 7.7, 5.1, 7.8, 7.9, 4.6, 8.1, 6.1 , 7.3 , 8.4, 8.5, 8.8, 7.7).

Table 6. Information Criteria for 1st Dataset

Dist.	Esti. Para.			$-\hat{L}$	AIC	AIC _c	BIC	HQIC
	$\hat{\alpha}$	$\hat{\rho}$	$\hat{\theta}$					
APRW	0.899	0.0129	0.0315	220.28	446.57	447.23	451.63	448.3
APEW	1.69	0.261	0.603	222.42	450.83	451.5	455.9	452.66
APW	1.333	0.526	0.152	240.22	486.44	487.1	491.5	488.27
APE	1.239		0.108	230.57	465.13	465.46	468.51	466.35
W		0.57	2.5	297.91	599.82	600.15	603.21	601.05

The second real data represents 38 wind-related disasters, where the observed or required values are classified into a value of two million dollars. These data were utilized by Mead (Mead et al., 2017) and are presented as follows: (2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 4, 4, 4, 5, 5, 5, 6, 6, 6, 6, 8, 8, 9, 15, 17, 22, 23, 24, 24, 25, 27, 32, 43).

Table 6. Information Criteria for 2nd Dataset

Dist.	Esti. Para.			$-\hat{L}$	AIC	AIC _c	BIC	HQIC
	$\hat{\alpha}$	$\hat{\rho}$	$\hat{\theta}$					
W APR	0.883	0.231	0.004	245.24	496.48	497.19	501.39	498.23
W APE	1.625	0.237	0.486	250.27	506.54	507.25	511.46	508.29
APW	1.432	0.672	0.023	251.93	509.87	510.58	514.78	511.62
APE	1.269		0.022	254.49	512.98	513.32	516.26	514.15
W		0.793	1.138	255.29	514.59	514.93	517.86	515.75

Based on the tables provided, the **APRWD NEW** model exhibits the lowest values across all the information criteria ($-\hat{L}$, AIC, AIC_c, BIC, and HQIC). This suggests that the APRWD NEW distribution provides the best fit for the data among the compared distributions. Lower values of these criteria indicate a better balance between model complexity and goodness of fit, reinforcing the superiority of the APRWD NEW model in representing the data effectively.



7 Conclusion

The proposed Alpha Power Rayleigh-Weibull Distribution (APRWD) offers a flexible generalization of the classical Rayleigh and Weibull distributions by introducing an additional shape parameter through the Alpha Power transformation. This enhancement allows the model to accommodate various shapes of hazard functions, improving its suitability for modeling lifetime and reliability data. The statistical properties of the distribution were systematically derived, and parameter estimation was performed using MLE and OLS methods. To assess the model's goodness-of-fit, several information-based criteria were employed, including AIC, AIC_c ,

BIC, and HQIC. The results indicate that the APRWD outperforms traditional models, demonstrating a superior fit to real-world data and highlighting its potential as a robust tool in applied reliability and survival analysis.

REFERENCES

- Ahmad, A., Ahmad, S. P., & Ahmed, A. (2014). *Transmuted inverse Rayleigh distribution: A generalization of the inverse Rayleigh distribution*. *Mathematical Theory and Modeling*, 4(7), 90–98.
- Ahmad, A., Ahmad, S. P., & Ahmed, A. (2017). *Characterization and estimation of Weibull-Rayleigh distribution with applications to lifetime data*. *Applied Mathematics & Information Sciences Letters*, 5(2), 71–79.
- Ahmad, Z., Ilyas, M., & Hamedani, G. G. (2019). *The extended alpha power transformed family of distributions: Properties and applications*. *Journal of Data Science*.
- Ali, M., Khalil, A., Ijaz, M., Saeed, N., & Manzoor, S. (2021). *Alpha-power exponentiated inverse Rayleigh distribution and its applications to real and simulated data*. *PLOS ONE*, 16(1), 1–17. <https://doi.org/10.1371/journal.pone.0245253>
- Ashor, N. A., & Mohammed, M. J. (2024). *Comparison between new lifetime Shanker-Weibull distribution and many other distributions*. In *AIP Conference Proceedings*. AIP Publishing.
- Basheer, A. M. (2019). *Alpha power inverse Weibull distribution with reliability application*. *Journal of Taibah University for Science*, 13(1), 423–432.
- Eghwerido, J. T. (2021). *The alpha power Teissier distribution and its applications*. *African Journal of Statistics*, 16(2), 2733–2747.
- Elbatal, I., Çakmakyapan, S., & Özel, G. (2022). *Alpha power odd generalized exponential family of distributions: Model, properties and applications*. *Gazi University Journal of Science*, 35(3), 1171–1188.
- ElSherpieny, E. A., & Almetwally, E. M. (2022). *The exponentiated generalized alpha power family of distributions: Properties and applications*. *Pakistan Journal of Statistics and Operation Research*, 18(2), 349–367.
- Ganji, M., Bevrani, H., Golzar, N. H., & Zabihi, S. (2016). *The Weibull-Rayleigh distribution, some properties, and applications*. *Journal of Mathematical Sciences*, 218(3), 269–277.
- Hassan, A. S., Elgarhy, M., Mohamd, R. E., & Alrajhi, S. (2019). *On the alpha power transformed power Lindley distribution*. *Journal of Probability and Statistics*, 2019(1), 8024769.
- Ihtisham, S., Khalil, A., Manzoor, S., Khan, S. A., & Ali, A. (2019). *Alpha-power Pareto distribution: Its properties and applications*. *PLOS ONE*, 14(6), e0218027.



- Ijaz, M., Asim, S. M., Alamgir, Farooq, M., Khan, S. A., & Manzoor, S. (2020). A gull alpha power Weibull distribution with applications to real and simulated data. *PLOS ONE*, 15(6), 1–19. <https://doi.org/10.1371/journal.pone.0233080>
- Kargbo, M., Waititu, A., & Wanjoya, A. (2023). *The generalized alpha power exponentiated inverse exponential distribution and its application to real data*. Authorea Preprints.
- Kilai, M., Waititu, G. A., Kibira, W. A., Alshanbari, H. M., & El-Morshedy, M. (2022). A new generalization of Gull Alpha Power family of distributions with application to modeling COVID-19 mortality rates. *Results in Physics*, 36, 105339. <https://doi.org/10.1016/j.rinp.2022.105339>
- Mahdavi, A., & Kundu, D. (2017). A new method for generating distributions with an application to exponential distribution. *Communications in Statistics - Theory and Methods*, 46(13), 6543–6557.
- Mahdi, G. J. (2020). A modified support vector machine classifier using stochastic gradient descent with application to leukemia cancer type dataset. *Baghdad Science Journal*, 17(4), 1255.
- Mandouh, R. M., Sewilam, E. M., Abdel-Zaher, M. M., & Mahmoud, M. R. (2022). A new family of distributions in the class of the alpha power transformation with applications to income. *Asian Journal of Probability and Statistics*, 19(1), 41–55.
- Mead, M. E., Afify, A. Z., Hamedani, G. G., & Ghosh, I. (2017). The beta exponential Fréchet distribution with applications. *Austrian Journal of Statistics*, 46(1), 41–63.
- Mohiuddin, M., & Kannan, R. (2021). Alpha power transformed Aradhana distributions, its properties and applications. *Indian Journal of Science and Technology*, 14(31), 2483–2493.
- Mohammed, A. T. (2019). Density estimation for right censored data using hybrid Breslow and semi-symmetric wavelet. *Journal of Engineering and Applied Sciences*, 14(1), 242–246.
- Mohammed, A. T., & Hussein, I. H. (2019). Nonparametric estimation for hazard rate function by wavelet procedures with simulation. In *IOP Conference Series: Materials Science and Engineering* (Vol. 518, p. 012013). IOP Publishing.
- Mohammed, M. J., & Hussein, I. H. (2018). Study of new mixture distribution. In *International Conference on Engineering and Applied Science (ICEMASP)*, Turkey.
- Nassar, M., Alzaatreh, A., Mead, M., & Abo-Kasem, O. (2017). Alpha power Weibull distribution: Properties and applications. *Communications in Statistics - Theory and Methods*, 46(20), 10236–10252.
- Nassar, M., Alzaatreh, A., Abo-Kasem, O., Mead, M., & Mansoor, M. (2018). A new family of generalized distributions based on alpha power transformation with application to cancer data. *Annals of Data Science*, 5, 421–436.
- Nasiru, S. (2016). Serial Weibull Rayleigh distribution: Theory and application. *International Journal of Computer Science and Mathematics*, 7(3), 239–244.
- Saleh, H. M., & Mohammed, A. T. (2024). Alpha power transformation family distributions: Properties and application to the exponential Weibull model. In *AIP Conference Proceedings*. AIP Publishing.
- Shalan, R. N. (2024). Compare some shrinkage Bayesian estimation method for Gumbel-Max distribution with simulation. *Iraqi Journal of Science*, 65(4), 2150–2159.
- Xu, K., Xie, M., Tang, L. C., & Ho, S. L. (2003). Application of neural networks in forecasting engine systems reliability. *Applied Soft Computing*, 2(4), 255–268.
- Zain, S. A., & Hussein, I. H. (2020). Some aspects of weighted Rayleigh distribution. *Ibn Al-Haitham Journal for Pure and Applied Science*, 33(2), 107–114.



**International Conference on Mathematics and
Mathematics Education (ICMME-2025), İstanbul Meder
University, İstanbul, Türkiye, September 11-13, 2025
"Mathematics in İstanbul, Bridge Between Continents"**



Zain, S. A., & Hussein, I. H. (2021). Estimate for survival and related functions of weighted Rayleigh distribution. *Ibn Al-Haitham Journal for Pure and Applied Science*, 34(1).



*International Conference on Mathematics and
Mathematics Education (ICMME-2025), İstanbul Meder
University, İstanbul, Türkiye, September 11-13, 2025*
"Mathematics in İstanbul, Bridge Between Continent"



SPONSORS

