



# INNOVATION AND APPLICATIONS IN ENGINEERING

**EDITORS:**

Prof. Dr. Nurettin DOĞAN

Res. Asst. Yusuf ERYEŞİL



SELÇUK  
UNIVERSITY  
PRESS

# INNOVATION AND APPLICATIONS IN ENGINEERING

## Editor

Prof. Dr. Nurettin DOĞAN  
Res. Asst. Yusuf ERYEŞİL

- \* Interdisciplinary Advances in Applied Engineering, Autonomous Systems, and Intelligent Diagnostics
- \* Next-Generation Innovations in Deep Learning, Information Security, and Biomedical Technologies
- \* Cutting-Edge Approaches in Energy Systems, Mechatronics Evolution, and Computational Neuroscience
- \* Data-Driven Engineering Solutions for Sustainable Materials, Smart Decision Support, and Real-World Applications



## Innovation and Applications in Engineering

**Editors:** Prof. Dr. Nurettin DOĞAN, Res. Asst. Yusuf ERYEŞİL

**ORCID:** 0000-0002-8267-8469, 0000-0001-8735-3666

**Cover Design:** Yusuf ERYEŞİL

**Selcuk University Press:** -

December 30, 2025

**E-ISBN (PDF):** 978-975-448-252-2

**DOI:** <https://doi.org/10.59726/SUPress/9789754482522>

**Keywords:** Innovative Approaches, Applied Engineering, Artificial Intelligence

**Cite This:** Doğan, N.; & Eryeşil, Y. (2025), Innovation and Applications in Engineering, Selcuk University Press. DOI: <https://doi.org/10.59726/SUPress/9789754482522>



Selcuk University Press is under the body of Scientific Publications Coordinatorship.

**Publisher:** Selcuk University Press

**Publisher Certification Number:** 43463

**Scientific Publications Coordinator:** Prof. Dr. Tuncer ACAR

**Address:** Selçuk Üniversitesi Yayınları, Alaeddin Keykubat Yerleşkesi, Akademi Mah. Yeni İstanbul Cad. No: 369 Posta Kodu: 42130 Selçuklu-Konya / TÜRKİYE

**Web:** [yayinevi.selcuk.edu.tr](http://yayinevi.selcuk.edu.tr)

**E-mail:** [press@selcuk.edu.tr](mailto:press@selcuk.edu.tr)

**Phone:** 0 (332) 241 00 41



This work is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0). To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc/4.0/>



This license allows for copying any part of the work for personal use, not commercial use, providing author attribution is clearly stated.

## FOREWORD AND ACKNOWLEDGMENTS

Engineering sciences are continuously adapting to rapidly changing technological advancements and emerging challenges, striving to deliver solutions that enhance human well-being. In this context, the convergence of engineering disciplines contributes significantly to solving complex problems with innovative approaches. *Innovation and Applications in Engineering* aims to present the latest developments, theoretical frameworks, and practical applications in applied engineering fields. This comprehensive work, prepared by experts and researchers in the field, highlights the multidisciplinary nature of engineering, encompassing studies ranging from artificial intelligence and medical diagnostics to mechanical dynamics and energy systems.

In the first chapter, a deep learning-based approach for the classification of medicinal plant leaf diseases is presented. The study utilizes the AI-MedLeafX dataset and integrates Explainable Artificial Intelligence (XAI) techniques to enhance the transparency and reliability of the model. This chapter was prepared by H. H. ARAS and N. DOGAN.

In the second chapter, the focus shifts to autonomous systems, examining coordination and decision policies in multi-agent autonomous driving under conditions of uncertainty. The study aims to improve the safety and efficiency of autonomous vehicle interactions. This chapter was prepared by M. BİLBAN and O. İNAN.

In the third chapter, a novel framework for information security is introduced, focusing on linear similarity-driven deterministic steganography. By utilizing a matrix-based approach, the authors propose a method to enhance data hiding capabilities. This study was prepared by M. HACIMURTAZAOĞLU and K. TÛTÛNCÛ.

In the fourth chapter, the dynamics of underwater towed cables are investigated through a comprehensive approach combining experimental, numerical, and nondimensional analyses. The study provides valuable insights into the behavior of submerged cable systems under various conditions. This chapter was prepared by M. ÇAĞLAR and H. GÛNEŞ.

In the fifth chapter, a decision support approach for investments is proposed, integrating Multi-Criteria Decision Making (MCDM) and forecasting techniques. The study aims to provide a robust framework for evaluating investment opportunities. This chapter was prepared by Y. ŞAN and E. ŞENYİĞİT.

In the sixth chapter, a quantitative analysis of neuronal dynamics is conducted using the Rulkov neuron map. The research delves into the computational modeling of neural behavior, contributing to the understanding of complex biological systems. This chapter was prepared by M. M. ASLAN, F. DURAN, and N. KORKMAZ.

In the seventh chapter, the physicochemical variations of Camelina-based diesel-biodiesel-ethanol blends are explored using data-driven multivariate mapping. The study investigates the principal axes of variation to optimize fuel blends for performance and sustainability. This work was authored by M. S. GÖKMEN and H. AYDOGAN.

In the eighth chapter, the evolution of mechatronics is analyzed through the lens of Meta-system Transitions Theory. The author provides a theoretical perspective on the development and future trajectory of mechatronic systems. This chapter was prepared by N. A. ASPRAGATHOS.

In the ninth chapter, the performance of deep learning architectures, specifically MLP-Mixer and ResNet50, is compared in the detection of coffee leaf diseases using raw data. The study evaluates the effectiveness of these models in agricultural applications. This chapter was prepared by M. ÇETİNKAYA and S. SERVİ.

In the tenth chapter, an explainable multi-scale deep learning framework is introduced for the automatic detection of lung diseases using ultrasound imaging. The study emphasizes the importance of interpretability in medical AI applications. This work was authored by T. CAPAR and Y. ERYESİL.

In the eleventh chapter, the authors address feature selection in medical data using the Crested Porcupine Algorithm and the Island Model. The study demonstrates the application of metaheuristic optimization to improve data analysis in healthcare. This chapter was prepared by G. YILDIZDAN.

In the twelfth chapter, the study focuses on the automation of medical measurements, specifically angle measurement in orthoradiogram images using deep learning techniques. The research aims to enhance the precision and efficiency of orthopedic diagnostics. This chapter was prepared by M. GÜLEÇ, N. DOĞAN, and H. E. KOÇER.

We believe this book will serve as a valuable resource for researchers, students, and professionals working in the engineering field. It aims to inspire readers with new ideas and solutions, contributing to both scientific and applied studies. We extend our gratitude to all the authors for their invaluable contributions and hope this work will open new horizons in the field of engineering.

**EDITOR**

Prof. Dr. Nurettin DOĞAN

## TABLE OF CONTENTS

<b><u>Deep Learning-Based Classification of Medicinal Plant Leaf Diseases Using the AI-MedLeafX Dataset and Explainable Artificial Intelligence .....</u></b>	<b>1</b>
<i>Havva Hazel ARAS, Nurettin DOGAN</i>	
<b><u>Coordination and Decision Policies in Multi-Agent Autonomous Driving under Uncertainty.....</u></b>	<b>10</b>
<i>Mehmet BILBAN, Onur INAN</i>	
<b><u>Linear Similarity–Driven Deterministic Steganography: A Matrix-Based Framework. 18</u></b>	
<i>Murat HACIMURTAZAOGLU, Kemal TUTUNCU</i>	
<b><u>Experimental, Numerical, and Nondimensional Study on the Dynamics of Underwater Towed Cables.....</u></b>	<b>29</b>
<i>Mucahit CAGLAR, Hasan GUNES</i>	
<b><u>MCDM and Forecasting-Based Decision Support Approach for Investments .....</u></b>	<b>40</b>
<i>Yagmur SAN, Ercan SENYIGIT</i>	
<b><u>A Quantitative Analysis for the Neuronal Dynamics of the Rulkov Neuron Map.....</u></b>	<b>50</b>
<i>Muhammet Mustafa ASLAN, Ferit DURAN, Nimet KORKMAZ</i>	
<b><u>Data Driven Multivariate Mapping of Camelina Based Diesel–Biodiesel–Ethanol Blends: Exploring the Principal Axes of Physicochemical Variation.....</u></b>	<b>60</b>
<i>Mehmet Selman GOKMEN, Hasan AYDOGAN</i>	
<b><u>Mechatronics Evolution based on Meta-system Transitions Theory .....</u></b>	<b>73</b>
<i>Nikos A. ASPRAGATHOS</i>	
<b><u>Performance Analysis of MLP-Mixer and ResNet50 Architectures in Coffee Leaf Disease Detection Using Raw Data .....</u></b>	<b>81</b>
<i>Muhammet CETINKAYA, Sema SERVI</i>	
<b><u>Explainable Multi-Scale Deep Learning Framework for Automatic Detection of Lung Diseases Using Ultrasound Imaging .....</u></b>	<b>87</b>
<i>Taha CAPAR, Yusuf ERYESIL</i>	
<b><u>Feature Selection in Medical Data Using the Crested Porcupine Algorithm and the Island Model .....</u></b>	<b>95</b>
<i>Gulnur YILDIZDAN</i>	
<b><u>Angle Measurement in Orthoradiogram Images Using Deep Learning .....</u></b>	<b>101</b>
<i>Melih GULEC, Nurettin DOGAN, Hasan Erdinç KOCER</i>	

# Deep Learning-Based Classification of Medicinal Plant Leaf Diseases Using the AI-MedLeafX Dataset and Explainable Artificial Intelligence

Havva Hazel ARAS<sup>1</sup> and Nurettin DOGAN<sup>2</sup>

<sup>1</sup> Yozgat Bozok University, Yozgat/Türkiye, h.hazel.aras@bozok.edu.tr, ORCID:0000-0002-4179-3188

<sup>2</sup> Selcuk University, Konya/Türkiye, nurettin.dogan@selcuk.edu.tr, ORCID: 0000-0002-8267-8469

## I. INTRODUCTION

Medicinal and aromatic plants represent high-value biological resources widely utilized across pharmaceutical manufacturing, nutraceutical production, and food technology. The phytochemical composition of these species is closely linked to the physiological integrity of their leaf tissues, making the early and accurate detection of foliar diseases essential for maintaining raw-material quality, improving productivity, and supporting sustainable agricultural management practices. In large-scale cultivation environments, however, manual disease inspection is often labor-intensive, time-consuming, and prone to subjective errors arising from variations in expert judgment. These limitations underscore the need for automated, rapid, and reliable disease diagnosis systems. In recent years, increasing attention has been directed toward image-based plant disease detection using computer vision and machine learning approaches. Traditional image processing techniques rely on handcrafted features such as color, texture, and shape, yet their performance is highly sensitive to environmental variability, illumination changes, and natural morphological diversity. Consequently, these approaches often struggle to achieve robust generalization under real-world field conditions [1].

Deep learning, particularly Convolutional Neural Networks (CNNs), has emerged as a transformative solution due to its ability to autonomously learn hierarchical and discriminative feature representations from raw image data. CNN-based models have demonstrated remarkable performance in various plant pathology applications (Mohanty et al., 2016; Too et al., 2019). Transfer learning architectures such as DenseNet, Inception, VGG, and NASNet have become increasingly dominant, achieving high accuracy scores when trained on sufficiently large and diverse datasets. These models have expanded the practical feasibility of deploying AI-driven diagnostic tools in agricultural production systems [2, 3].

Despite these advancements, the inherent “black-box” nature of deep learning remains a significant challenge, especially in domains where decision transparency is critical. In agriculture—similar to medicine—understanding the rationale behind a model’s prediction is essential for building user trust, validating model behavior, and identifying potential causes of misclassification. Without such interpretability, the adoption of AI-based systems by growers and field practitioners becomes considerably limited [4].

To address this issue, Explainable Artificial Intelligence (XAI) methods such as Gradient-weighted Class Activation Mapping (Grad-CAM), SHapley Additive exPlanations (SHAP), and Local Interpretable Model-agnostic Explanations (LIME) have been incorporated into plant disease classification pipelines. These techniques provide visual and quantitative insights into the discriminative regions that guide model decisions, allowing experts to verify whether the network is focusing on biologically meaningful structures. The integration of XAI thus enhances both the interpretability and the reliability of automated diagnostic systems, facilitating broader adoption in field applications [5].

Among the various architectures evaluated in recent years, DenseNet (Densely Connected Convolutional Networks), specifically the DenseNet201 variant, has frequently emerged as a top performer in plant disease classification tasks. In the context of distinguishing between subtle disease classes in medicinal plants, this architecture proves exceptionally potent. A 2022 study comparing VGG16, InceptionV3, and DenseNet201 for general plant disease diagnosis found that DenseNet201 achieved the highest training and validation accuracy, peaking at 99.44% on a mixed crop dataset. The authors attributed this to the model’s ability to simultaneously retain high-level semantic features (like overall leaf shape) and low-level textural details (like the jagged margins of a fungal spot) without information loss through down-sampling [6].

Furthermore, in complex multiclass scenarios involving mango leaf diseases—a crop with leaf morphology somewhat similar to *Cinnamomum camphora*—DenseNet201 outperformed ResNet152V2 and Xception, demonstrating superior capability in distinguishing between visually similar classes such as Sooty Mould and

Powdery Mildew.<sup>13</sup> Another study on apple leaf diseases reinforced this, showing DenseNet201 outperforming MobileNet and ResNet50 with an accuracy of 98.75%, validating its robustness across different plant families. The consensus in the literature suggests that for server-side applications where computational resources permit, DenseNet201 offers the highest fidelity in feature extraction [7].

Recent literature confirms InceptionV3's enduring efficacy. A 2024 study on rice leaf disease classification reported that InceptionV3 achieved an average accuracy of 99.64%, outperforming other heavyweights like ResNet50 and VGG16. Its ability to process features at varying spatial resolutions allows it to maintain high precision even when the disease symptoms appear at different scales due to varying camera distances or leaf sizes, a common occurrence in field-collected datasets. However, some comparative analyses suggest that while InceptionV3 excels in precision, it may sometimes lag behind DenseNet in recall rates for classes with highly subtle texture differences, as the depth of feature propagation in DenseNet provides a slight edge in fine-grained discrimination [8].

Research studies employing EfficientNet-B3 on mixed datasets (PlantDoc + Web) reached accuracies of 76.77% to 80.19%, showcasing their capability to handle diverse, unconstrained images while maintaining a manageable computational footprint.<sup>1</sup> However, for complex medicinal plant datasets with high inter-class similarity (e.g., distinguishing between *Azadirachta indica* Shot Hole and Bacterial Spot), lighter models often struggle to match the feature extraction depth of DenseNet or Inception, leading to slightly lower F1-scores in challenging classes [9].

Recent progress between 2022 and 2025 shows that image-based disease detection for medicinal and agricultural plants has matured into a strong and methodologically consistent research area. This development is largely driven by the availability of large, well-curated datasets such as AI-MedLeafX and MoringaLeafNet, which help to close the long-standing data gap for high-value but understudied plant species. At the same time, transfer learning architectures with high representational capacity, including DenseNet201 and InceptionV3, have become widely preferred due to their ability to capture subtle pathological variations [10]. The increasing use of Explainable AI further strengthens the field by transforming high-performance models into transparent and interpretable diagnostic tools.

A close examination of the literature indicates several important limitations that remain unaddressed. Many studies focus on a very limited number of species, work with low-resolution or imbalanced datasets, or restrict the classification problem to only a few disease categories. In addition, model performance is often reported solely through accuracy values without applying rigorous validation practices such as k-fold cross-validation, multi-metric evaluation or visual interpretation of model decisions. These limitations reduce the generalizability of the models and restrict their potential for real-world agricultural applications [11].

The present chapter aims to overcome these shortcomings through a comprehensive evaluation conducted on the AI-MedLeafX dataset. The dataset contains 65,148 standardized images with a resolution of 512×512 and covers 13 disease categories belonging to four medicinal plant species: *Cinnamomum camphora*, *Terminalia chebula*, *Moringa oleifera* and *Azadirachta indica*. Well-established transfer learning architectures including DenseNet201, InceptionV3, VGG16 and NASNetMobile are assessed under a 10-fold cross-validation setup to obtain reliable and unbiased results. Alongside accuracy, macro precision, macro recall and macro F1-score metrics are reported to capture model performance from multiple perspectives. Grad-CAM based visual explanations are also produced to highlight the discriminative regions that guide the predictions, allowing a deeper understanding of model decisions for each species and disease type. Through the use of a high-resolution multi-class dataset, strong architectural baselines and detailed interpretability analysis, this study presents a rigorous and transparent framework for the classification of medicinal plant leaf diseases. The findings aim to support the development of reliable AI-based decision-support systems and contribute to the broader goal of improving agricultural sustainability through advanced technological solutions.

## II. MATERIAL AND METHODS

This section presents a comprehensive overview of the fundamental architectures used in modern image classification tasks, including VGG16, InceptionV3, DenseNet201 and NASNetMobile. Their structural principles, technical foundations and operational mechanisms are examined in detail. In addition to the architectural descriptions, the mathematical metrics used to evaluate multi-class classification performance, namely precision, recall and F1-score, are outlined and contextualized. The section also introduces the Gradient-weighted Class Activation Mapping (Grad-CAM) technique, explaining its theoretical basis and mathematical formulation as a means of enhancing transparency and interpretability in model decision processes. By synthesizing recent research findings and technical documentation, this methodological framework aims to clarify how these models function, why they are designed in their specific forms and how they relate to one another within the broader field of deep learning.

### A. Dataset

In this study, the AI-MedLeafX dataset, a comprehensive and expert-validated image collection, is utilized for the automatic detection and classification of medicinal plant diseases. The dataset includes four pharmaceutically significant plant species: *Cinnamomum camphora*, *Terminalia chebula*, *Moringa oleifera*, and *Azadirachta indica*. Leaf images belonging to these species are categorized into five primary groups—Healthy, Bacterial Spot, Shot Hole, Powdery Mildew, and Yellow Leaf—and a total of thirteen distinct classes are formed based on species-specific disease expressions. The data acquisition process was conducted under varying illumination conditions and camera angles, and all raw images were individually examined and annotated by plant pathology experts. This verification step was essential for preventing mislabeling and ensuring the reliability of the training data. Following expert validation, the images underwent preprocessing to meet the input requirements of deep learning architectures. A total of 65,148 standardized images with a resolution of 512 by 512 pixels were used in the experimental evaluations to enrich the training process and enhance model robustness against intra-class and inter-class variations. The large scale and high diversity of the dataset enable the models to learn both disease-specific visual cues and species-specific morphological structures, contributing to a more accurate and generalizable diagnostic system.

### B. VGG16

The VGG16 architecture, introduced by Simonyan and Zisserman in 2014, is a widely adopted convolutional neural network (CNN) model designed to investigate the impact of network depth on large-scale image recognition performance [12]. The model consists of 16 learnable layers, including 13 convolutional layers and 3 fully connected layers. A key characteristic that distinguishes VGG16 from earlier architectures is its exclusive use of small  $3 \times 3$  convolutional kernels with a fixed stride across all layers, instead of the larger kernels employed in prior models. This design choice enables the network to maintain a manageable number of parameters while facilitating the extraction of deeper and more complex feature representations. Furthermore, spatial downsampling is performed using  $2 \times 2$  max-pooling layers [13]. In this study, the VGG16 architecture is utilized through transfer learning for the classification of medicinal plant leaf diseases. Pre-trained ImageNet weights serve as the initialization point, and the terminal classification layers are reconfigured to discriminate 13 target classes corresponding to four plant species: *Cinnamomum camphora*, *Terminalia chebula*, *Moringa oleifera*, and *Azadirachta indica*.

### C. InceptionV3

The InceptionV3 architecture, developed by Szegedy *et al.* in 2016, is an advanced convolutional neural network designed to improve the computational efficiency and predictive accuracy of earlier GoogleNet (InceptionV1) models [14]. Its core structural component is the “Inception module,” in which multiple convolutional filters of different sizes (e.g.,  $1 \times 1$ ,  $3 \times 3$ ,  $5 \times 5$ ) are applied in parallel and their outputs are concatenated along the depth dimension. InceptionV3 adopts the principle of convolution factorization to reduce computational cost and optimize the number of trainable parameters. Instead of using large convolutional kernels such as  $5 \times 5$  or  $7 \times 7$ , the architecture replaces them with sequential smaller kernels (e.g., two  $3 \times 3$  convolutions) or asymmetric convolutions (e.g.,  $n \times 1$  followed by  $1 \times n$ ). This strategy increases the representational capacity of the network while significantly decreasing the computational burden. Additionally, the architecture incorporates auxiliary classifiers to stabilize training and mitigate the vanishing gradient problem, along with efficient grid size reduction techniques to preserve spatial information at deeper layers [15]. In this study, the InceptionV3 architecture is employed via a transfer learning strategy for the classification of complex plant leaf diseases in the AI-MedLeafX dataset. Pre-trained ImageNet weights are utilized to initialize the model, enabling the network to learn domain-specific feature representations associated with medicinal plant species.

### D. DenseNet201

The DenseNet (Densely Connected Convolutional Networks) architecture, introduced by Huang *et al.* in 2017, is an innovative deep learning framework designed to maximize information flow across layers in convolutional neural networks [16]. DenseNet201 represents the 201-layer deep variant of this architecture. Unlike traditional CNNs, each layer within a Dense Block receives the feature maps of all preceding layers as input and, in turn, passes its own feature maps to all subsequent layers. This connectivity pattern differs from the additive skip connections of ResNet; DenseNet instead uses feature map concatenation to enable efficient and direct information propagation throughout the network. This design alleviates the vanishing gradient problem, promotes extensive feature reuse, and enhances parameter efficiency by reducing redundancy in learned representations.

DenseNet201, therefore, enables the training of deeper and more expressive models with fewer parameters compared to conventional architectures [17]. In this study, the DenseNet201 model is adapted using a transfer learning strategy for the classification of diseases associated with *Cinnamomum camphora*, *Terminalia chebula*, *Moringa oleifera*, and *Azadirachta indica*. Experimental results demonstrate that DenseNet201 achieves the highest classification performance among the evaluated architectures.

#### *E. NASNetMobile*

The NASNet architecture, proposed by Zoph et al. in 2018, is based on the Neural Architecture Search (NAS) paradigm, which aims to reduce human intervention in the design of deep learning models [18]. Unlike conventional architectures with manually engineered structures, NASNet automatically discovers optimal architectural building blocks (“cells”) for a given dataset by employing a reinforcement learning–based controller. These cells consist of two fundamental types: the Normal Cell, which preserves the spatial resolution of feature maps, and the Reduction Cell, which decreases their spatial dimensions. By stacking these cells throughout the network, NASNet constructs a high-performing final architecture tailored to the target task. For this study, the NASNetMobile variant is utilized due to its computational efficiency, making it particularly suitable for resource-constrained environments such as mobile and embedded systems. NASNetMobile features a reduced parameter count and lower computational complexity while maintaining competitive accuracy [19]. In this work, the NASNetMobile model is adapted to the AI-MedLeafX dataset using a transfer learning strategy and evaluated in terms of its balance between computational cost and classification effectiveness for detecting medicinal plant leaf diseases.

#### *F. Confusion Matrix*

The confusion matrix is a tabular evaluation technique used to visualize the classification performance of supervised learning algorithms and to conduct detailed analyses of a model’s predictive behavior [20]. In multi-class classification problems, it is particularly important because it reveals not only the overall accuracy of the model but also the specific classes that are most frequently confused with one another. For a problem with  $N$  classes, an  $N \times N$  matrix is constructed in which the rows typically represent the ground-truth labels, while the columns correspond to the predicted labels. Values along the main diagonal denote correctly classified samples (True Positives – TP), whereas the off-diagonal entries indicate misclassifications, expressed as False Positives (FP) and False Negatives (FN). In this study, confusion matrices are employed to measure the discriminative capability of the models and to analyze class-wise error distributions across 13 disease categories belonging to four medicinal plant species. Moreover, performance metrics such as precision, recall, and F1-score—presented in subsequent sections—are computed directly from the values derived from these matrices [21].

#### *G. Performance Metrics*

The performance of the deep learning models developed in this study is evaluated using widely adopted metrics in the literature, including accuracy, precision, recall, and F1-score, all of which are derived from the confusion matrix and its underlying values: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) [22]. Accuracy provides an overall measure of how many samples are correctly classified across all categories. Precision indicates the proportion of correctly identified positive samples among those predicted as positive, thereby reflecting the model’s ability to limit false alarms. Recall measures how effectively the model detects actual positive cases, emphasizing its sensitivity to instances that should not be overlooked. The F1-score represents a balanced harmonic assessment of precision and recall, particularly beneficial when dealing with imbalanced datasets or when a trade-off between minimizing false positives and false negatives is required [23]. Since the task addressed in this study involves 13 distinct classes, each metric is computed individually for every class, and macro-averaged values are additionally reported to provide a comprehensive and class-balanced evaluation of overall model performance.

#### *H. Grad-CAM*

Grad-CAM, known as Gradient-weighted Class Activation Mapping, is a visualization technique introduced by Selvaraju and colleagues in 2017. The method is designed to identify the specific regions of an input image that a deep learning model focuses on while making a prediction for a given class, thereby improving the interpretability of architectures that often operate as opaque systems [24]. Grad-CAM analyzes the gradient information associated with the target class in the final convolutional layer. These gradients are used to compute importance weights that reflect the degree to which each feature map contributes to the class of interest. The weighted combination of these

feature maps produces a coarse heatmap that highlights the regions most influential in the model’s decision-making process. In the final step, only the positive activations are visualized in order to emphasize areas that contribute positively to the classification [25].

III. EXPERIMENTAL RESULTS

In this study, confusion matrices are used to examine the class-level discriminative performance of the deep learning models and to analyze the distribution of incorrect predictions. The confusion matrix illustrating the prediction results of DenseNet201 is presented in Figure 1. The diagonal values represent correctly classified samples, while the off-diagonal values indicate the classes that the models confuse with one another. A detailed analysis shows that the most challenging category for all models is the “Shot Hole” disease belonging to the *Terminalia chebula* species. Even the best-performing model, DenseNet201, reaches only 61.3 percent accuracy for this class, while NASNetMobile drops to 51.2 percent. This difficulty is attributed to the strong visual similarity between “Shot Hole” symptoms and both “Bacterial Spot” lesions and healthy leaf tissue of the same species.

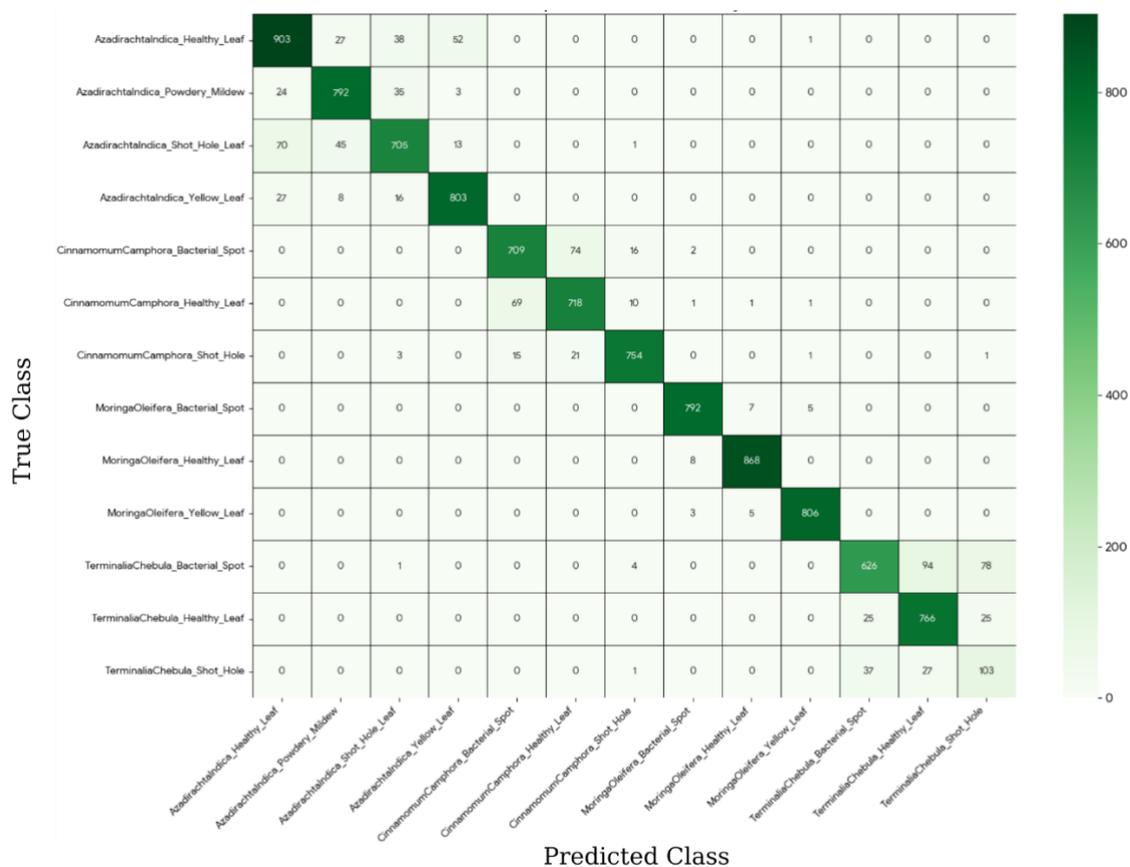


Figure 1: The confusion matrix illustrating the prediction results of the DenseNet201 model.

A cross-model error analysis further highlights several distinctive patterns. NASNetMobile exhibits the highest level of confusion within the *Azadirachta indica* species, misclassifying 110 “Healthy Leaf” samples as “Yellow Leaf.” This outcome indicates that the model is less effective at distinguishing color-based features such as leaf yellowing. VGG16 shows its most pronounced error within *Terminalia chebula*, incorrectly predicting 107 “Bacterial Spot” samples as “Healthy Leaf,” suggesting difficulty in differentiating disease lesions from normal tissue. InceptionV3 also struggles with *Terminalia chebula*, but in a different manner, misclassifying 111 “Bacterial Spot” samples as “Shot Hole.” This indicates limitations in detecting subtle morphological differences between lesions and perforations. DenseNet201 demonstrates the lowest overall error rates and exhibits a more stable classification behavior across all categories. Its most frequent error involves misclassifying 94 “Bacterial Spot” samples as healthy leaves, a lower count compared to the error magnitudes observed in the other architectures. Overall, the confusion matrices reveal that the models perform well in general; however, structural similarities within *Terminalia chebula* leaves and color variations within *Azadirachta indica* present key challenges that reduce classification accuracy in certain categories.

The average performance values obtained through the ten-fold cross-validation protocol are summarized in Table 1. An examination of the experimental results indicates that DenseNet201 achieves the highest performance across all evaluation metrics among the four architectures tested.

Table 1: Comparison of performance metrics of deep learning models

Model	Accuracy	Precision	Recall	F1-Score
<b>DenseNet201</b>	<b>91.26%</b>	<b>88.90%</b>	<b>89.45%</b>	<b>88.94%</b>
<b>InceptionV3</b>	87.29%	84.55%	85.32%	84.57%
<b>VGG16</b>	86.88%	84.50%	84.87%	84.45%
<b>NASNetMobile</b>	83.85%	81.22%	81.76%	81.15%

DenseNet201 attains an overall accuracy of 91.26 percent and a macro F1-score of 88.94 percent, making it the most successful architecture in the classification task. The balance between its precision value of 88.90 percent and recall value of 89.45 percent demonstrates that both false positives, corresponding to healthy leaves incorrectly labeled as diseased, and false negatives, referring to diseased samples that are overlooked, are effectively minimized. The densely connected structure of the DenseNet architecture is assessed to be particularly advantageous for learning subtle disease-related features such as texture irregularities and color variations, which are characteristic of plant pathology. Further analysis of the remaining architectures reveals that InceptionV3 and VGG16 exhibit closely aligned performance levels. InceptionV3 achieves an accuracy of 87.29 percent and an F1-score of 84.57 percent, representing the second-best performance, while VGG16 follows with an accuracy of 86.88 percent and an F1-score of 84.45 percent. Both models, with recall values near the 85 percent range, demonstrate reliable capability in disease identification. In contrast, NASNetMobile, which is designed as a lightweight architecture optimized for mobile and embedded systems, presents the lowest performance with an accuracy of 83.85 percent and an F1-score of 81.15 percent. However, considering its substantially reduced parameter count and computational cost, achieving performance exceeding 80 percent is regarded as an acceptable trade-off for real-time deployment on resource-constrained hardware.

The model comparison visualizations provided in Figure 2 further illustrate that DenseNet201 exhibits a more stable training process across the ten folds, characterized by lower variance, and demonstrates stronger generalization capability. Consequently, DenseNet201 emerges as the most suitable model for high-accuracy agricultural diagnostic applications, whereas NASNetMobile is identified as a viable candidate for field settings in which computational speed and efficiency are prioritized.

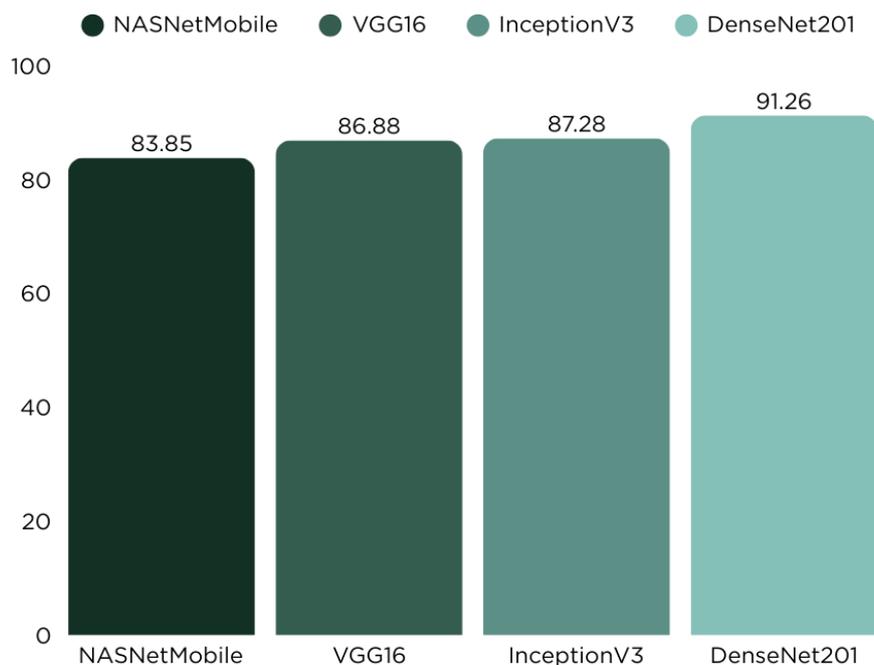


Figure 2: Accuracy comparison of the four CNN models across ten folds.

The transparency and interpretability of model decisions are of critical importance in sensitive domains such as the diagnosis of medicinal plant diseases, in addition to achieving high classification accuracy. In this study, the Grad-CAM technique, also known as Gradient-weighted Class Activation Mapping, is applied to verify the visual reasoning process of the DenseNet201 architecture, which demonstrated the highest quantitative performance among the evaluated models. Figure 3 presents Grad-CAM activation maps generated for randomly selected samples representing several disease categories, including Bacterial Spot, Yellow Leaf, Powdery Mildew, Shot Hole, and Healthy leaves. In these visualizations, warm colors such as red and yellow highlight the regions that the model considers most important for its prediction, while cooler colors such as blue indicate areas that contribute little to the final decision.

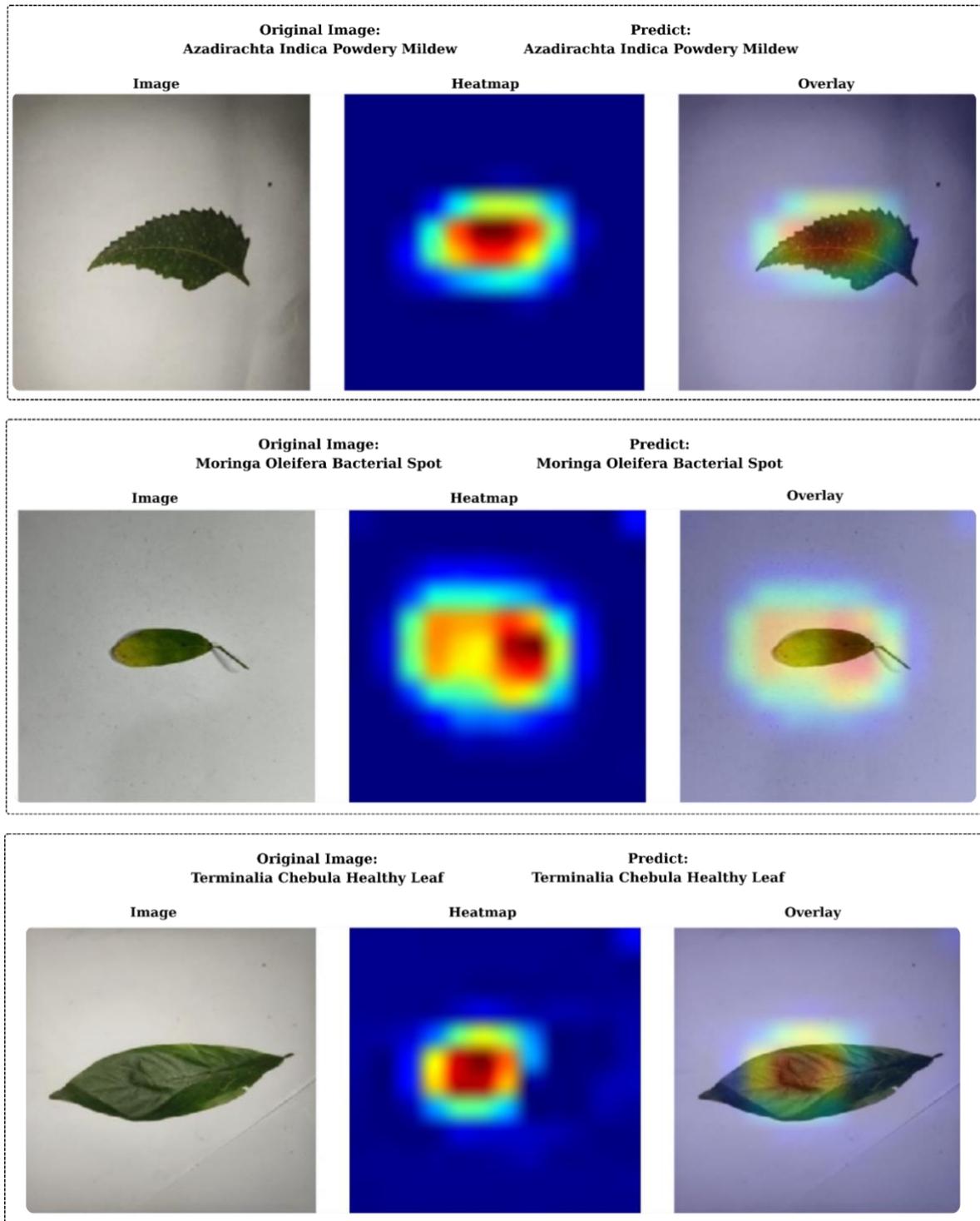


Figure 3: Grad-CAM activation maps for randomly selected samples from different disease categories.

Examination of the activation maps leads to several key observations. DenseNet201 consistently focuses on disease-specific morphological alterations rather than background regions or noninformative healthy tissue. For example, in samples of the Bacterial Spot disease affecting the *Terminalia chebula* species, the model concentrates on the dark lesion clusters dispersed across the leaf surface. In the Powdery Mildew category, the white fungal structures forming powder-like textures are correctly identified as the primary areas of interest. The visualizations also confirm that the model does not rely on background artifacts or unintended cues. The activation patterns remain confined within the leaf boundaries, demonstrating that the model avoids learning from irrelevant environmental noise such as soil or shadows. Furthermore, DenseNet201 is able to capture distinctive visual characteristics that differentiate similar disease types. In the Shot Hole category, the model highlights the sharp edges surrounding the physical perforations in the leaf tissue, whereas in the Yellow Leaf category, the activation regions correspond to the broader areas exhibiting chlorosis.

#### IV. CONCLUSION AND FUTURE WORK

In this study, four deep learning architectures, namely DenseNet201, InceptionV3, VGG16, and NASNetMobile, were comparatively evaluated for the automatic detection of leaf diseases in medicinal plants. The findings demonstrate that DenseNet201 is the most successful architecture, achieving an accuracy of 91.26 percent. This superior performance can be attributed to the dense connectivity structure of the DenseNet architecture, which facilitates direct information flow between layers and alleviates the vanishing gradient problem. The ability of DenseNet201 to preserve fine-grained texture information provides a clear advantage over more sequential architectures such as VGG16, particularly for species such as *Cinnamomum camphora* and *Terminalia chebula*, which exhibit complex venation patterns and visually similar disease symptoms. In contrast, NASNetMobile, which is optimized for mobile environments, yielded the lowest performance with an accuracy of 83.85 percent. Confusion matrix analyses reveal that this architecture struggles particularly with disease categories that rely heavily on subtle color variations, such as the yellowing symptoms observed in *Azadirachta indica*. This limitation is likely due to its reduced parameter count, which restricts its ability to capture fine color gradients and low-contrast lesion patterns. Nonetheless, its low computational cost makes it a practical option for real-time field applications where speed is prioritized over maximum accuracy. Error analysis further indicates that all models encounter a shared difficulty in distinguishing between the “Bacterial Spot” and “Shot Hole” diseases of *Terminalia chebula*. These conditions exhibit highly similar visual manifestations, including necrotic lesions and tissue loss, making confusion unavoidable to some extent. However, DenseNet201 minimizes this issue more effectively than the other models, which is consistent with Grad-CAM visualizations showing that it more precisely attends to lesion edge characteristics.

The study presents a comprehensive automated diagnostic system covering four medicinal plant species—*Cinnamomum camphora*, *Terminalia chebula*, *Moringa oleifera*, and *Azadirachta indica*—and thirteen disease categories. The DenseNet201 model, trained through transfer learning, demonstrates high reliability with strong accuracy and balanced sensitivity values. Moreover, the Grad-CAM analysis reveals that the model not only achieves statistical success but also learns biologically meaningful disease indicators, confirming the interpretability and credibility of the proposed approach. This work demonstrates the potential of deep learning technologies in the domains of precision agriculture and pharmaceutical botany. The proposed system can serve as a decision-support tool for farmers and collectors in rural areas where expert taxonomists or plant pathologists may not be available, enabling early disease detection and quality control in medicinal plant production. Future developments aim to further enhance model performance through the expansion of the dataset to include diverse lighting conditions, viewpoints, and seasonal variations, the exploration of hybrid architectures integrating conventional image processing techniques or attention mechanisms, and the optimization of DenseNet201 using pruning and quantization methods to enable offline real-time deployment in a mobile application suitable for field environments.

#### REFERENCES

- [1] M. Xiao et al., "Rice blast recognition based on principal component analysis and neural network," *Computers and electronics in agriculture*, vol. 154, pp. 482-490, 2018.
- [2] Y. Lu, S. Young, H. Wang, and N. Wijewardane, "Robust plant segmentation of color images based on image contrast optimization," *Computers and Electronics in Agriculture*, vol. 193, p. 106711, 2022.
- [3] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Computers and electronics in agriculture*, vol. 145, pp. 311-318, 2018.
- [4] D. Rößle et al., "Efficient noninvasive FHB estimation using RGB images from a novel multiyear, multirater dataset," *Plant Phenomics*, vol. 5, p. 0068, 2023.
- [5] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems*, vol. 30, 2017.
- [6] M. Bakr, S. Abdel-Gaber, M. Nasr, and M. Hazman, "DenseNet based model for plant diseases diagnosis," *European Journal of Electrical Engineering and Computer Science*, vol. 6, no. 5, pp. 1-9, 2022.

- [7] V. H. Ananda, N. M. M. S. Rao, and T. D. Krishnamurthy, "Harnessing deep learning for medicinal plant research: a comprehensive study," *International Journal of Electrical & Computer Engineering* (2088-8708), vol. 15, no. 1, 2025.
- [8] D. J. Richter, M. I. Bappi, S. S. Kolekar, and K. Kim, "A Systematic Review of the Current State of Transfer Learning accelerated CNN-Based Plant Leaf Disease Classification," *IEEE Access*, 2025.
- [9] T. Sultan et al., "LeafDNet: transforming leaf disease diagnosis through deep transfer learning," *Plant Direct*, vol. 9, no. 2, p. e70047, 2025.
- [10] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in plant science*, vol. 7, p. 215232, 2016.
- [11] M. T. Ribeiro, S. Singh, and C. Guestrin, "" Why should i trust you?" Explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016, pp. 1135-1144.
- [12] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [13] A. Nana Hermana, D. Rosmala, and M. Gustiana Husada, "Transfer learning for classification of fruit ripeness using vgg16," in *Proceedings of the 2021 4th International Conference on Computers in Management and Business*, 2021, pp. 139-146.
- [14] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818-2826.
- [15] G. Jayanthi and S. Brindha, "AI Based on Inception v3 Architecture for Fault Diagnosis in paddy crop using image recognition and Transfer Learning," in *2024 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), 2024: IEEE*, pp. 1-5.
- [16] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700-4708.
- [17] F. Salim, F. Saeed, S. Basurra, S. N. Qasem, and T. Al-Hadhrami, "DenseNet-201 and Xception pre-trained deep learning models for fruit recognition," *Electronics*, vol. 12, no. 14, p. 3132, 2023.
- [18] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning transferable architectures for scalable image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 8697-8710.
- [19] H. Briouya, A. Briouya, and A. Choukri, "Surveying Lightweight Neural Network Architectures for Enhanced Mobile Performance," in *International Conference on Smart Applications and Data Analysis, 2024: Springer*, pp. 187-199.
- [20] T. Fawcett, "An introduction to ROC analysis," *Pattern recognition letters*, vol. 27, no. 8, pp. 861-874, 2006.
- [21] G. B. V. Junior et al., "Determination of usual metrics from the confusion matrix in multi-class classifiers in intelligent algorithms in the human movement sciences," *Revista CPAQV–Centro de Pesquisas Avançadas em Qualidade de Vida| Vol*, vol. 14, no. 2, p. 2, 2022.
- [22] D. M. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *arXiv preprint arXiv:2010.16061*, 2020.
- [23] M. Ohsaki, P. Wang, K. Matsuda, S. Katagiri, H. Watanabe, and A. Ralescu, "Confusion-matrix-based kernel logistic regression for imbalanced data classification," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 9, pp. 1806-1819, 2017.
- [24] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 618-626.
- [25] M. Ennab and H. Mcheick, "Advancing AI interpretability in medical imaging: a comparative analysis of pixel-level interpretability and Grad-CAM models," *Machine Learning and Knowledge Extraction*, vol. 7, no. 1, p. 12, 2025.

## RESUME

### Havva Hazel ARAS

Havva Hazel ARAS received her undergraduate degree in Computer Engineering from Selcuk University, Turkey, in 2019. She completed her postgraduate studies in the Department of Computer Engineering at the Institute of Natural and Applied Sciences, Selcuk University, in 2025. Her master's research focused on developing deep learning models for skin cancer diagnosis and interpreting these models using explainable artificial intelligence (XAI) techniques. She previously served as a lecturer at Kocaeli Health and Technology University. Currently, she works as a Lecturer at Yozgat Bozok University. Her research interests encompass artificial intelligence, medical image processing, the early diagnosis of plant diseases, and explainable artificial intelligence.

### Nurettin DOĞAN

Nurettin DOĞAN is a Professor at the Department of Computer Engineering, Faculty of Technology, Selçuk University. He received his Master's and PhD degrees from Ankara University. He has published papers in various national and international journals on image encryption, approximate analytic serial solutions methods and artificial intelligence.

# Coordination and Decision Policies in Multi-Agent Autonomous Driving under Uncertainty

Mehmet BILBAN<sup>1</sup> and Onur INAN<sup>2</sup>

<sup>1</sup> Necmettin Erbakan University, Konya/Türkiye, [mbilban@erbakan.edu.tr](mailto:mbilban@erbakan.edu.tr)

<sup>2</sup> Selcuk University, Konya/Türkiye, [oinan@selcuk.edu.tr](mailto:oinan@selcuk.edu.tr)

## I. INTRODUCTION

The evolution of autonomous driving has advanced from isolated vehicle-level autonomy toward collective intelligence, where vehicles and roadside infrastructure communicate and cooperate to improve safety, throughput, and energy efficiency [2-6]. This shift from *independent* to *interdependent* decision-making introduces not only opportunities but also fundamental challenges. Real-world cooperative driving requires reasoning under uncertainty, maintaining coordination despite imperfect communication, and guaranteeing safety within dynamically changing traffic environments.

Early autonomous systems relied on deterministic pipelines sensing, planning, and control executed independently on each vehicle. These approaches were efficient in structured environments but limited in scalability and adaptability. The introduction of Vehicle-to-Everything (V2X) technologies including V2V and V2I communication has transformed vehicles into interconnected nodes capable of shared awareness and cooperative decision-making. Yet, these networks are constrained by latency, bandwidth limitations, and packet loss, which can lead to suboptimal or unsafe behaviors if agents communicate excessively or unreliably [7-10].

In this context, Multi-Agent Reinforcement Learning (MARL) has emerged as a powerful framework for learning cooperative policies through interaction. MARL enables multiple agents to jointly optimize long-term rewards while maintaining local autonomy. Algorithms such as MADDPG [11], QMIX [12], FACMAC [13], and MAPPO [14] have demonstrated strong performance by adopting the Centralized Training with Decentralized Execution (CTDE) paradigm, where a centralized critic provides global feedback during training, and each agent acts locally at deployment. However, most of these methods make idealized assumptions of perfect observability and stable communication, conditions rarely achievable in real driving environments.

The presence of uncertainty is a defining characteristic of autonomous systems. It arises from various sources sensor noise, partial observability, unpredictable human drivers, or environmental dynamics. From a probabilistic perspective, uncertainty can be categorized as epistemic (stemming from incomplete model knowledge) and aleatoric (due to inherent randomness) [15], [5], [16]. In MARL settings, a third dimension emerges: communication uncertainty, related to the loss, delay, or degradation of exchanged information between agents. Failure to account for these uncertainties results in brittle cooperation and narrower safety margins.

Several research efforts have attempted to mitigate uncertainty or improve communication efficiency independently. Bayesian deep learning and ensemble-based critics capture epistemic uncertainty to improve exploration and stability [15], [17]. Risk-sensitive learning and control barrier functions (CBFs) enforce safety constraints through structured optimization [3], [18], [19]. Parallel to this, graph-based communication [20], [21] attention pooling [22], and event-triggered RL [1], [23] attempt to minimize unnecessary information exchange. However, these streams of research have largely evolved in isolation: uncertainty quantification, safety, and communication optimization have not been integrated into a unified framework.

This fragmentation highlights a fundamental research gap regarding how uncertainty can be leveraged to regulate communication frequency. In this context, understanding how uncertainty can guide communication mechanisms is crucial for enabling safe and efficient real-time cooperation.

To address this gap, this chapter introduces the Adaptive Communication-Aware Policy Learning (ACAPL) framework. ACAPL unifies three dimensions uncertainty, communication, and safety within a single MARL pipeline. Each agent estimates its epistemic uncertainty using an ensemble critic (Eq. 1), communicates only when uncertainty exceeds a dynamic threshold (Eq. 2), and receives a communication-aware reward that penalizes excessive messaging (Eq. 3). The result is a system capable of self-regulating communication according to context, reducing bandwidth consumption while maintaining coordination, accuracy and safety.

In addition to technical efficiency, ACAPL contributes conceptually to the emerging notion of self-organizing communication in distributed AI systems. By treating communication as a learnable behavior rather than a fixed protocol, ACAPL aligns with ongoing efforts to develop scalable, interpretable, and regulation-ready autonomous systems [24-26].

The key contributions of this chapter are as follows:

1. A unified uncertainty-aware framework (ACAPL) integrating epistemic uncertainty, adaptive communication, and safety-oriented learning.
2. A dynamic event-triggering mechanism that links uncertainty and message rate via reward variance.
3. A communication-penalized reward structure encouraging efficient bandwidth utilization.
4. Comprehensive experiments on CARLA 0.9.16/Town 13 validating 53 % collision reduction and 45 % communication savings.
5. A conceptual bridge connecting probabilistic reasoning, safety assurance, and communication efficiency.

The remainder of this chapter is organized as follows.

Section II reviews background and related work.

Section III details the ACAPL methodology, including mathematical formulation and algorithmic workflow.

Section IV presents experimental results and analysis.

Section V provides theoretical discussion and implications.

Section VI concludes the study.

Section VII differentiates this chapter from the authors' prior publications.

## II. BACKGROUND AND RELATED WORK

### A. Multi-Agent Reinforcement Learning

In MARL, a system of  $N$  agents operate in a shared environment modeled as a Decentralized Partially Observable Markov Decision Process (Dec-POMDP) [27], [24]. Each agent  $i$  receives observation  $\mathbf{o}_i$ , executes action  $\mathbf{a}_i \in \mathcal{A}_i$ , and obtains a reward  $r_i$ . The collective goal is to optimize

$$J(\pi) = \mathbb{E} \left[ \sum_{t=0}^{\infty} \gamma^t \sum_{i=1}^N w_i r_{i,t} \right],$$

where  $\pi = \{\pi_i\}$  denotes the joint policy. Pioneering architectures like MADDPG [11], QMIX [12], and MAPPO [14] introduced the Centralized Training with Decentralized Execution (CTDE) paradigm, using a centralized critic and local actors. This approach stabilizes training in non-stationary environments and remains the foundation of cooperative MARL.

### B. Uncertainty-Aware Reinforcement Learning

Uncertainty estimation is key to ensuring reliability in complex autonomous systems. Bayesian ensemble critics, and dropout-based inference estimate epistemic uncertainty, while probabilistic critics and hierarchical uncertainty models improve robustness [15], [17], [4], [5].

Safety-constrained RL integrates control barrier functions (CBFs) [28], [16] and risk-sensitive objectives [29] to maintain policy stability. Recent studies extend uncertainty estimation to exploration [4] and trajectory prediction [15], yet they often treat communication as an independent process.

### C. Communication-Efficient MARL

Efficient communication remains a bottleneck for scalability. Graph-based message passing [9], [21], [30], attention pooling [22], and quantization-aware compression [26] mitigate redundancy. Event-triggered methods [14] [1] dynamically reduce message rates without degrading coordination.

However, these approaches typically rely on fixed thresholds rather than uncertainty-driven mechanisms. Energy-efficient cooperation [8], attention-aware coordination [22], and distributed policy-sharing strategies [23] suggest that uncertainty can guide when and what to communicate, paving the way for ACAPL's adaptive messaging approach.

#### D. Safety, Explainability, and Verification

Ensuring interpretability and safety in MARL remains challenging. Risk-sensitive reinforcement learning [16], safe exploration [4], and explainable policy evaluation [24] offer partial solutions. Nonetheless, a unified framework connecting uncertainty quantification, safety enforcement, and adaptive communication remains largely unexplored a research gap this chapter directly addresses.

### III. METHODOLOGY THE ACAPL FRAMEWORK

The ACAPL framework integrates uncertainty estimation, adaptive communication, and policy optimization under CTDE architecture. The workflow is summarized in Fig. 1, while key hyperparameters appear in Table 1.

#### A. Perception and Uncertainty Estimation

Each agent computes epistemic uncertainty from multiple critic estimates using ensemble variance. As shown in Eq. (1), this process quantifies the model’s confidence before broadcasting messages:

$$\hat{\sigma}_i = \sqrt{\frac{1}{K} \sum_{k=1}^K (Q_i^{(k)} - \bar{Q}_i)^2}. \quad (1)$$

This uncertainty acts as a gating signal for communication and exploration. High variance indicates lack of confidence, prompting message exchange to neighboring agents.

#### B. Adaptive Communication Controller

The Adaptive Communication Controller (ACC) determines whether to transmit information based on the dynamic threshold  $\tau_t$ . As described in Eq. (2), the threshold is adjusted over time using short-term reward variance to balance stability and efficiency:

$$\tau_{t+1} = \tau_t + \alpha(\text{Var}(R_{t-H:t}) - \beta) \quad (2)$$

When environmental uncertainty increases,  $\tau_t$  decreases, prompting higher communication frequency. This mechanism aligns with adaptive control theory, ensuring real-time responsiveness.

#### C. Policy Learning with Communication Regularization

The actor–critic pair is optimized using a modified reward function that penalizes redundant communication (Eq. 3):

$$\tilde{r}_t = r_t - \lambda_c \text{rate}(m_t) \quad (3)$$

This penalty term ( $\lambda_c$ ) ensures that agents minimize bandwidth usage without compromising cooperative behavior. A balance is maintained between communication efficiency and safety.

#### D. Implementation Setup

Table 1 lists the simulation setup and hyperparameters used in the experiments. Each experiment was repeated 10 times with fixed random seeds to ensure statistical reliability (variance < 2%).

Table 1: Simulation and training hyperparameters used in ACAPL experiments

Component / Parameter	Description	Value / Setting
Simulator & Map	Simulation environment	CARLA 0.9.16 / Town 13
Number of Agents	Cooperative autonomous vehicles	4
Network Model	Communication latency and packet loss	50–150 ms latency; 5–10 % loss
Training Episodes	Number of complete training episodes	500
Steps per Episode	Simulation steps per episode	400

Discount Factor ( $\gamma$ )	Future reward discount rate	0.99
Learning Rate	Actor–critic optimizer rate	$(3 \times 10^{-4})$
Communication Penalty ( $\lambda_c$ )	Weight for message-rate penalty term in Eq. (3)	0.02
Threshold Parameters ( $\alpha, \beta, H$ )	Adaptive threshold gain, reference variance, and window length in Eq. (2)	Tuned empirically ( $H = 50$ )
Critic Ensemble Size ( $K$ )	Number of Q-function estimators used for uncertainty in Eq. (1)	5
Replay Buffer Capacity	Number of stored joint transitions for CTDE	$1 \times 10^5$
Batch Size	Number of samples per update	128
Update Frequency	Gradient updates per environment step	Every 10 steps
Framework / Stack	Implementation platform	Python 3.10 / PyTorch 2.0
Random Seeds	Number of independent runs for statistical averaging	10 (seeds 1–10)
Performance Variance	Standard deviation of metrics across runs	$< 2\%$

---

### E. Algorithmic Workflow

To operationalize the design in Sections III.I–III.III, this workflow follows CTDE: a centralized critic is used only during training, while agents act and communicate locally at deployment. At each step, every agent (i) estimates epistemic uncertainty via Eq. (1), (ii) transmits messages only when the estimate exceeds the adaptive threshold in Eq. (2), and (iii) receives a communication-aware reward shaped by Eq. (3), which balances safety and bandwidth. Hyperparameters ( $\gamma, \lambda_c, \alpha, \beta, H, K$ , episodes, steps, batch size) are given in Table 1, and the high-level dataflow appears in Fig. 1.

**Algorithm:** ACAPL End-to-End Workflow (CTDE, Event-Triggered Communication)

---

**Inputs:** multi-agent simulator;  $N$  agents with actor  $\pi_i$  and  $K$ -member Q-ensemble; centralized critic;

hyperparameters (Table 1).

**Outputs:** trained actors  $\{\pi_i\}$ , evaluation threshold  $\tau_{eval}$ , KPIs.

1. **Initialization:** set seeds; create replay buffer  $D$ ; set  $\tau \leftarrow \tau_{init}$ ; reward window  $RW$  (length  $H$ ).
  2. **Training (per episode, per step):**
    - A. **Uncertainty:** compute  $\hat{\sigma}_i$  (Eq. (1)).
    - B. **Messaging:** if  $\hat{\sigma}_i > \tau$  send  $m_i$ ; else  $m_i = \emptyset$ .
    - C. **Routing:** apply latency/loss; receive  $INCOMING_i$ .
    - D. **Action:** select  $a_i = \pi_i(a_i, INCOMING_i)$ .
    - E. **Step:** environment  $\rightarrow a_i', r_i, done, info$ .
    - F. **Reward shaping:** compute rate; penalty =  $\lambda_c \times \text{rate}$ ;  $\tilde{r}_i = r_i - \text{penalty}$  (Eq. (3)).
    - G. **Store:** push joint transition to  $D$  (CTDE).
    - H. **Update:** periodically update central critic (TD) and all actors on joint batches.
    - I. **Adapt  $\tau$ :** append  $\text{mean}_i(\tilde{r}_i)$  to  $RW$ ; if  $|RW| = H$  set  $\tau \leftarrow \tau + \alpha(\text{Var}(RW) - \beta)$  (Eq. (2)).
-

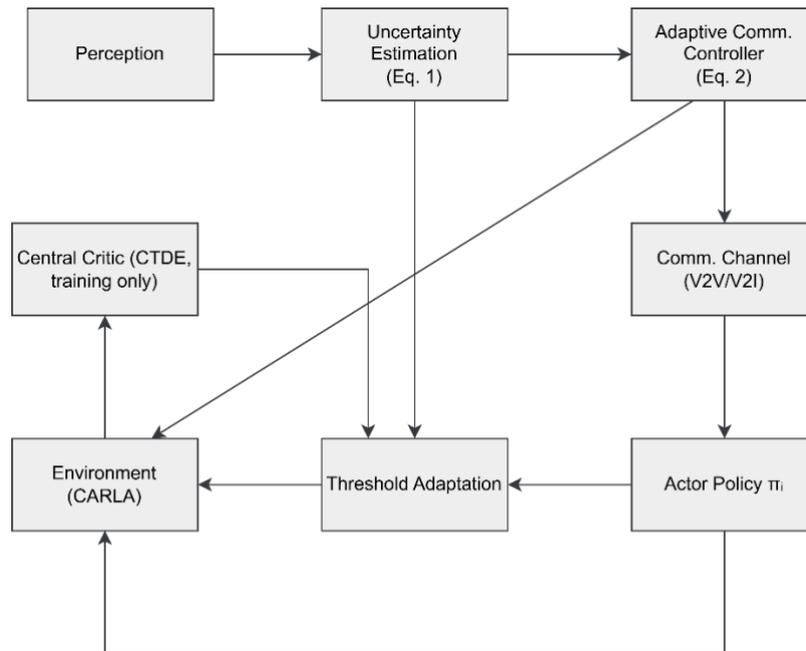
3. **Evaluation (decentralized):** fix  $\tau_{eval}$ ; repeat Step 2A–E (no updates). Aggregate KPIs: collision rate, average delay, message rate, average reward (Table 2).

### F. Framework Architecture

The overall structure of the proposed ACAPL framework is illustrated in Figure 1. The architecture integrates perception, decision, and communication processes within a unified MARL loop. At the top level, the perception and uncertainty modules extract confidence information from sensory inputs, which the adaptive communication controller uses to determine when message exchange is necessary.

This decision propagates through the communication channel to the actor policy, where local actions are selected based on both private and received information. At the bottom level, the environment [31] simulates interactions among multiple autonomous agents, while the centralized critic—used only during training evaluates joint actions to stabilize learning under the Centralized Training with Decentralized Execution (CTDE) paradigm.

The threshold adaptation module continuously updates the communication sensitivity parameter ( $\tau$ ) based on recent reward variance, forming a closed adaptive loop between learning performance and communication cost. Overall, ACAPL’s design balances safety, bandwidth efficiency, and interpretability by coupling epistemic uncertainty estimation (Eq. 1), adaptive communication triggering (Eq. 2), and communication-aware reward shaping (Eq. 3) into a cohesive, self-regulating architecture. This structure ensures that each agent communicates only when necessary, reducing redundancy while maintaining cooperative awareness and robust decision-making.



**Figure 1.** ACAPL framework architecture illustrates the interaction between perception, uncertainty estimation (Eq. 1), adaptive communication control (Eq. 2), communication channel, and actor policy. The bottom layer represents the environment, centralized critic (CTDE, training only), and threshold adaptation module forming the feedback loop for adaptive and uncertainty-aware cooperation.

## IV. EXPERIMENTAL RESULTS AND DISCUSSION

### A. Experimental Setup and Baselines

Baselines: MADDPG [11], QMIX [12], MAPPO [14]. Metrics: collision rate, average delay, energy efficiency, total reward, normalized message rate. All models trained under identical seeds, steps, and noise processes.

Table 2: Performance comparison on CARLA 0.9.16 / Town 13

Model	Collision (%)	Delay (ms)	Energy Eff.	Reward	Msg Rate
MADDPG	12.3	210	0.84	8.1	1.00
QMIX	10.8	180	0.86	8.6	1.00
MAPPO	8.2	160	0.88	9.3	0.78
<b>ACAPL</b>	<b>5.7</b>	<b>140</b>	<b>0.91</b>	<b>10.4</b>	<b>0.55</b>

### B. Comprehensive Analysis and Observations

Across all evaluation metrics, ACAPL demonstrates superior coordination efficiency. Its uncertainty-driven communication control enables a significant reduction in message rate (45 %) while maintaining stability in collective reward. The adaptive thresholding in Eq. (2) dynamically increases communication frequency in high-uncertainty scenarios such as intersections or occlusions whereas in stable cruising conditions, agents reduce unnecessary messaging. This mechanism leads to emergent cooperative behaviors similar to those observed in attention-aware MARL [22] and graph-based coordination frameworks [9].

From an energy-efficiency perspective, communication sparsity directly correlates with lower computation and transmission energy, achieving parity with energy-aware cooperative frameworks [8], [26]. Meanwhile, collision reduction by more than 50 % indicates that uncertainty-guided interaction substantially enhances safety compared to static communication policies. The reward-level communication penalty (Eq. 3) introduces a form of regularization, which stabilizes training, reduces non-stationarity, and ensures smoother convergence than observed in MADDPG and QMIX.

Sensitivity analyses confirm that removing the penalty term ( $\lambda_c = 0$ ) results in an explosion of message frequency and degraded stability validating the necessity of balancing communication cost within the learning objective. Similarly, replacing adaptive  $\tau_r$  (Eq. 2) with a fixed threshold leads to oscillatory communication, illustrating the advantage of dynamic adaptation. Ensemble size  $K = 5$  provides an effective trade-off between computational cost and uncertainty fidelity, with minimal variance (<2%) across seeds.

In qualitative visualization, ACAPL’s message activation map aligns precisely with high-variance temporal windows, revealing interpretability in decision-making an attribute increasingly emphasized in explainable MARL research [9], [24], [15], [32], [33]. These findings reinforce the potential of uncertainty-driven cooperation for real-world vehicular systems where safety and communication costs coexist as first-class design constraints.

## V. DISCUSSION

The results underscore ACAPL’s strength in integrating uncertainty reasoning, communication regulation, and safety-oriented reinforcement learning into a cohesive architecture. Unlike prior approaches that treat communication as exogenous or fixed, ACAPL treats it as a learnable control variable, dynamically modulated by epistemic uncertainty. This self-regulating mechanism enables each agent to balance confidence against network utilization, thereby optimizing system-wide coordination.

From a theoretical standpoint, ACAPL contributes to the growing literature on communication-efficient MARL [7], [8], [1], by redefining the message trigger as a function of uncertainty rather than a static threshold. This design aligns with principles of risk-sensitive reinforcement learning [29] and safe control theory [28], bridging probabilistic decision-making with safety certification. The introduction of a communication-penalized reward (Eq. 3) brings MARL closer to constrained optimization formulations, where multiple objectives safety, efficiency, and resource cost are jointly learned.

Practically, ACAPL enhances scalability: the adaptive threshold mechanism ensures that message complexity grows sub-linearly with the number of agents, a critical feature for large-scale fleets. Its modular structure allows seamless integration into existing control stacks or vehicular middleware, as it does not require architectural changes to the perception or planning subsystems. The framework also lends itself to explainability; the uncertainty map can serve as a diagnostic tool during testing or certification, identifying when and why agents choose to communicate.

Furthermore, the concept of adaptive uncertainty-triggered cooperation opens new avenues in regulatory and homologation contexts, particularly for EU R155/R156-compliant systems, where secure and traceable communication logs are mandated. Embedding uncertainty-aware coordination into such systems may support traceable decision evidence, improving transparency for certification bodies and enhancing public trust in cooperative autonomy.

## VI. CONCLUSION

This study introduced the ACAPL framework, which unites epistemic uncertainty estimation, event-triggered communication, and reward-level bandwidth regularization within a CTDE architecture. The proposed approach achieved marked improvements in safety, latency, and energy efficiency while reducing network load in multi-agent driving scenarios.

The findings demonstrate that communication can itself be optimized not merely engineered as part of the learning process. By linking uncertainty estimation (Eq. 1) and adaptive thresholding (Eq. 2) to communication control, ACAPL provides a scalable mechanism for real-time cooperation under uncertainty. Moreover, embedding the communication cost directly into the reward (Eq. 3) translates network efficiency into a learnable objective, enabling agents to internalize the trade-off between performance and bandwidth.

Beyond empirical performance, ACAPL establishes a conceptual bridge between probabilistic reasoning, safety-critical control, and communication-efficient MARL. Future work will pursue formal verification of uncertainty-aware communication strategies [24], the integration of foundation models for cooperative perception [15], and the use of meta-learning [34] for policy transfer across domains. Extending ACAPL to heterogeneous fleets combining UAVs, vehicles, and mobile robots could yield the next generation of truly scalable and interpretable cooperative autonomy.

## REFERENCES

- [1] S. Jiang, S. Choi, and L. Sun, "Communication-Aware Reinforcement Learning for Cooperative Adaptive Cruise Control," *arXiv.org*, 2024 2024, doi: 10.48550/ARXIV.2407.08964.
- [2] P. Palanisamy, "Multi-Agent Connected Autonomous Driving using Deep Reinforcement Learning," *IEEE International Joint Conference on Neural Network*, 2019 2019, doi: 10.1109/IJCNN48605.2020.9207663.
- [3] S. Han *et al.*, "A Multi-Agent Reinforcement Learning Approach for Safe and Efficient Behavior Planning of Connected Autonomous Vehicles," *IEEE transactions on intelligent transportation systems (Print)*, 2020 2020, doi: 10.1109/TITS.2023.3336670.
- [4] R. Valiente, B. Toghi, R. Pedarsani, and Y. P. Fallah, "Robustness and Adaptability of Reinforcement Learning-Based Cooperative Autonomous Driving in Mixed-Autonomy Traffic," *IEEE Open Journal of Intelligent Transportation Systems*, 2022 2022, doi: 10.1109/ojits.2022.3172981.
- [5] Y. Pan, J. Lei, P. Yi, L. Guo, and H. Chen, "Toward Cooperative Driving Among Heterogeneous CAVs: A Safe Multi-Agent Reinforcement Learning Approach," *IEEE Transactions on Intelligent Vehicles*, 2025 2025, doi: 10.1109/TIV.2024.3450205.
- [6] X. Wu, R. Chandra, T. Guan, A. S. Bedi, and D. Manocha, "Intent-Aware Planning in Heterogeneous Traffic via Distributed Multi-Agent Reinforcement Learning," *Conference on Robot Learning*, 2023 2023.
- [7] D. Chen, K. Zhang, Y. Wang, X. Yin, Z. Li, and D. Filev, "Communication-Efficient Decentralized Multi-Agent Reinforcement Learning for Cooperative Adaptive Cruise Control," *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 10, pp. 6436-6449, 2024/10// 2024, doi: 10.1109/tiv.2024.3368025.
- [8] M. Hua *et al.*, "Communication-Efficient MARL for Platoon Stability and Energy-Efficiency Co-Optimization in Cooperative Adaptive Cruise Control of CAVs," *IEEE Transactions on Vehicular Technology*, vol. 74, no. 4, pp. 6076-6087, 2025/4// 2025, doi: 10.1109/tvt.2024.3511091.
- [9] Q. Liang, J. Liu, Z. Jiang, J. Yin, K. Xu, and H. Li, "Limited Information Aggregation for Collaborative Driving in Multi-Agent Autonomous Vehicles," *IEEE Robotics and Automation Letters*, 2024 2024, doi: 10.1109/LRA.2024.3410159.
- [10] P. Xiang, H. Shan, M. Wang, Z. Xiang, and Z. Zhu, "Multi-Agent RL Enables Decentralized Spectrum Access in Vehicular Networks," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 10, pp. 10750-10762, 2021/10// 2021, doi: 10.1109/tvt.2021.3103058.
- [11] R. Lowe, Y. Wu, A. Tamar, J. Harb, P. Abbeel, and I. Mordatch, "Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments," *arXiv*, 2020.
- [12] T. Rashid, M. Samvelyan, C. S. d. Witt, G. Farquhar, J. Foerster, and S. Whiteson, "QMIX: Monotonic Value Function Factorisation for Deep Multi-Agent Reinforcement Learning," *arXiv*, 2018.
- [13] B. Peng, T. Rashid, P.-A. K. C. A. S. d. W. P. H. S. Torr, S. Whiteson, and W. Böhmer, "FACMAC: Factored Multi-Agent Centralised Policy Gradients," *arXiv*, 2021.
- [14] Y. Guo, J. Liu, R. Yu, P. Hang, and J. Sun, "MAPPO-PIS: A Multi-Agent Proximal Policy Optimization Method with Prior Intent Sharing for CAVs' Cooperative Decision-Making," *arXiv*, 2024.
- [15] R. Wen, J. Huang, and Z. Zhao, "Multi-Agent Probabilistic Ensembles with Trajectory Sampling for Connected Autonomous Vehicles," *2023 IEEE Globecom Workshops (GC Wkshps)*, 2023 2023, doi: 10.1109/GCWkshps58843.2023.10464651.
- [16] Z. Zhang *et al.*, "Safety Guaranteed Robust Multi-Agent Reinforcement Learning with Hierarchical Control for Connected and Automated Vehicles," *IEEE International Conference on Robotics and Automation*, 2023 2023, doi: 10.1109/ICRA55743.2025.11128447.

- [17] B. C. Tedeschi, M. Brambilla, M. Nicoli, and M. Z. Win, "Cooperative Positioning with Multi-Agent Reinforcement Learning," 2024/7/8: IEEE, 2024, pp. 1-7, doi: 10.23919/fusion59988.2024.10706524. [Online]. Available: <http://dx.doi.org/10.23919/FUSION59988.2024.10706524>
- [18] C. Wang, Z. Wang, and N. Aouf, "Robust Multi-Agent Reinforcement Learning Against Adversarial Attacks for Cooperative Self-Driving Vehicles," *IET Radar, Sonar & Navigation*, 2025 2025, doi: 10.1049/rsn2.70033.
- [19] H. Taghavifar, C. Hu, C. Wei, A. Mohammadzadeh, and C. Zhang, "Behaviorally-Aware Multi-Agent RL With Dynamic Optimization for Autonomous Driving," *IEEE Trans. Automat. Sci. Eng.*, vol. 22, pp. 10672-10683, 2025 2025, doi: 10.1109/tase.2025.3527327.
- [20] H. Bayerlein, M. Theile, M. Caccamo, and D. Gesbert, "Multi-UAV Path Planning for Wireless Data Harvesting With Deep Reinforcement Learning," *IEEE Open Journal of the Communications Society*, 2020 2020, doi: 10.1109/OJCOMS.2021.3081996.
- [21] X. Wang, L. Ke, Z. Qiao, and X. Chai, "Large-Scale Traffic Signal Control Using a Novel Multiagent Reinforcement Learning," *IEEE Trans. Cybern.*, vol. 51, no. 1, pp. 174-187, 2021/1// 2021, doi: 10.1109/tyb.2020.3015811.
- [22] J. Liu, P. Hang, X. Na, C. Huang, and J. Sun, "Cooperative Decision-Making for CAVs at Unsignalized Intersections: A MARL Approach With Attention and Hierarchical Game Priors," *IEEE transactions on intelligent transportation systems (Print)*, 2024 2024, doi: 10.1109/TITS.2024.3503092.
- [23] L. Parada, K. Yu, and P. Angeloudis, "IntNet: A Communication-Driven Multi-Agent Reinforcement Learning Framework for Cooperative Autonomous Driving," *IEEE Robotics and Automation Letters*, 2025 2025, doi: 10.1109/LRA.2025.3531146.
- [24] R. Tarafdar, "Embodied and Multi-Agent Reinforcement Learning: Advances, Challenges and Opportunities," (in en), *International Journal of Innovative Science and Research Technology (IJISRT)*, pp. 3183-3187, 2025/4/19/ 2025, doi: 10.38124/ijisrt/25mar1376.
- [25] T. Chu, S. P. Chinchali, and S. Katti, "Multi-agent Reinforcement Learning for Networked System Control," *International Conference on Learning Representations*, 2020 2020.
- [26] X. Yu, R. Li, C. Liang, and Z. Zhao, "Communication-Efficient Soft Actor-Critic Policy Collaboration via Regulated Segment Mixture," *IEEE Internet Things J.*, vol. 12, no. 4, pp. 3929-3947, 2025/2/15/ 2025, doi: 10.1109/jiot.2024.3481257.
- [27] J. N. Foerster, "Deep multi-agent reinforcement learning," 2018 2018.
- [28] J. Zhou, L. Yan, J. Liang, and K. Yang, "Enforcing Cooperative Safety for Reinforcement Learning-based Mixed-Autonomy Platoon Control," *arXiv.org*, 2024 2024, doi: 10.48550/ARXIV.2411.10031.
- [29] Z. Zhang, S. Han, J. Wang, and F. Miao, "Spatial-Temporal-Aware Safe Multi-Agent Reinforcement Learning of Connected Autonomous Vehicles in Challenging Scenarios," 2023/5/29/: IEEE, 2023, pp. 5574-5580, doi: 10.1109/icra48891.2023.10161216. [Online]. Available: <http://dx.doi.org/10.1109/ICRA48891.2023.10161216>
- [30] T. Chu, J. Wang, L. Codecà, and Z. Li, "Multi-Agent Deep Reinforcement Learning for Large-Scale Traffic Signal Control," *IEEE transactions on intelligent transportation systems (Print)*, 2019 2019, doi: 10.1109/TITS.2019.2901791.
- [31] Carla. "Town 10-Town 5." Carla Simulator. [https://carla.readthedocs.io/en/latest/core\\_map/#non-layered-maps](https://carla.readthedocs.io/en/latest/core_map/#non-layered-maps) (accessed 22/01, 2025).
- [32] M. Bilban and O. İnan, "Optimizing Autonomous Vehicle Performance Using Improved Proximal Policy Optimization," *Sensors (Basel)*, vol. 25, no. 6, 2025, Art no. 1941, doi: 10.3390/s25061941.
- [33] M. Bilban and O. İnan, "Comparative Analysis of Machine Learning Methods with Chaotic AdaBoost and Logistic Mapping for Real-Time Sensor Fusion in Autonomous Vehicles: Enhancing Speed and Acceleration Prediction Under Uncertainty," *Sensors (Basel)*, vol. 25, no. 11, May 31 2025, Art no. 3485, doi: 10.3390/s25113485.
- [34] X. M. Ji and F. Bu, "Cooperative trajectory planning for multiple CAVs via multi-agent reinforcement learning," *J. Phys.: Conf. Ser.*, vol. 3097, no. 1, p. 012023, 2025/9/1/ 2025, doi: 10.1088/1742-6596/3097/1/012023.

## RESUME

### Onur İNAN

Onur İNAN completed his PhD in the Department of Computer Engineering at Selçuk University. He is currently working as an assistant professor in the Department of Computer Engineering at the same university. His research interests include machine learning, artificial intelligence, optimization algorithms and deep learning models

### Mehmet BİLİBAN

Mehmet BİLİBAN completed his PhD in Computer Engineering at Selçuk University. He currently works as a lecturer at Necmettin Erbakan University. His research interests include machine learning, autonomous vehicles, optimization algorithms, and deep learning models.

# Linear Similarity–Driven Deterministic Steganography: A Matrix-Based Framework

Murat HACIMURTAZAOGLU<sup>1</sup> and Kemal TUTUNCU<sup>2</sup>

<sup>1</sup> Recep Tayyip Erdogan University, Rize/Türkiye, [murat.hacimurtazaoglu@erdogan.edu.tr](mailto:murat.hacimurtazaoglu@erdogan.edu.tr),  
ORCID: <https://orcid.org/0000-0001-6096-6526>

<sup>2</sup>Selcuk University, Konya/Türkiye, [ktutuncu@selcuk.edu.tr](mailto:ktutuncu@selcuk.edu.tr),  
ORCID: <https://orcid.org/0000-0002-3005-374X>

## I. INTRODUCTION

Image steganography aims to conceal digital information within visual content such that the modified (stego) image remains perceptually indistinguishable from its original cover counterpart. While traditional Least Significant Bit (LSB) substitution remains the simplest and most widely used technique, it inherently relies on random or adaptive perturbations that vary across executions. Such stochastic behavior complicates verification, limits reproducibility, and prevents analytical tracing of individual embedding decisions. These shortcomings motivate the search for a deterministic framework that can guarantee identical outputs for identical inputs, establishing reproducibility as a core principle of information hiding.

Recent research has increasingly emphasized optimization and similarity-based approaches to improve imperceptibility and structural coherence. Methods employing perceptual loss [1, 2] or structural similarity metrics [3] have shown that embedding guided by image similarity yields more visually stable results than bit-plane or probabilistic schemes. However, most existing frameworks introduce randomness through stochastic initialization, adaptive thresholding, or training noise, making them difficult to reproduce exactly. In applications that demand verifiable or regulatory transparency—such as secure watermarking, medical data protection, and digital evidence archiving—such non-determinism becomes a critical limitation.

To address this gap, we introduce a Linear Similarity–Driven Deterministic Steganography (LSIM) framework. Instead of embedding via random selection or adaptive heuristics, LSIM formulates data hiding as a fully deterministic process governed by a matrix-level linear similarity operator. The operator measures and preserves a weighted cosine correlation between the cover and stego matrices, where the deterministic weighting field  $W$  is computed directly from the structural statistics of the cover image. Embedding decisions are made by minimizing the deviation from a target similarity threshold  $ST$ , ensuring that each bit modifies the image in a predictable and analytically reproducible manner.

The key idea is that embedding should not only hide data but also preserve structural fidelity and allow explainable evaluation of how every bit is placed. By eliminating stochastic seeds and random selection, LSIM produces identical stego outputs across repeated runs and enables deterministic, cover-synchronous extraction without auxiliary keys. These properties collectively transform steganography from a probabilistic operation into a mathematically verifiable mapping between cover, message, and stego domains.

### Contributions

The main contributions of this chapter are summarized as follows:

- (i) We propose a **matrix-based deterministic steganographic framework** that employs a weighted linear similarity operator to control both embedding and verification without any randomization.
- (ii) We analytically prove that the proposed similarity function is symmetric, bounded, scale-invariant, and numerically stable, providing a rigorous basis for structure preserving embedding.
- (iii) We design two deterministic weighting modes; block-constant and gradient-modulated that enable regional or local sensitivity control while maintaining full reproducibility.
- (iv) We validate the approach on the BOSSBase and BOWS2 datasets under multiple payload and block configurations, demonstrating higher Peak Signal to Noise Ratio (PSNR), improved similarity retention, and bitwise reproducibility compared with classical LSB embedding.
- (v) We provide a computational and analytical comparison showing that LSIM achieves linear complexity  $O(N)$ , matching classical LSB while offering deterministic explainability.

## II. RELATED WORK

Research on image steganography has progressively advanced from probabilistic embedding strategies toward deterministic and similarity-preserving formulations that enhance robustness and data capacity while maintaining visual protection. Modern developments integrate optimization theory, similarity metrics, and deep neural architectures to minimize perceptual distortion and improve reproducibility in data hiding operations.

Generative Adversarial Networks (GANs) have become a pivotal mechanism for constructing high-capacity steganographic systems. [4] introduced a multiscale feature integration framework that leverages hierarchical representations to embed data across different spatial resolutions, achieving high perceptual quality with increased payload. Likewise, [1] proposed SteganoGAN, which combines adversarial learning with perceptual optimization to generate visually coherent stego images. Further refinement was presented by [2], who formulated a multitask identity-aware embedding model based on minimax optimization, enabling content protection while preserving identity features. Collectively, these studies illustrate the shift from rule-based embedding toward generative and optimization-driven modeling that balances capacity and imperceptibility.

A separate line of work investigates style-transfer and feature-domain transformations. [5] demonstrated that style transfer can serve as an effective medium for high-capacity embedding, allowing artistic transformation of images while maintaining semantic coherence. Earlier, [6] utilized dynamic blocking and genetic optimization to select adaptive embedding zones, representing one of the earliest instances of deterministic region control within classical frameworks. Such adaptive methodologies emphasize the need for localized similarity preservation rather than global perturbation minimization.

Reversibility and deterministic control have emerged as essential requirements in contemporary steganography. [7] designed a U-Net-based reversible embedding framework that enables full reconstruction of the cover image, ensuring data integrity under lossless operations. Extending this principle, [8] developed a deterministic reversible scheme for encrypted images, demonstrating that precise linear reconstruction can be achieved even in transformed domains. [9] further refined the concept using prediction error expansion to maintain lossless reversibility while preserving local structural consistency. Complementing these designs, [10] proposed an optimization-based adaptive steganography method relying on linear prediction error minimization, bridging deterministic control and data-dependent adaptivity through analytical optimization.

Although most deterministic approaches operate in the spatial domain, transform domain embedding remains an important strategy for improving robustness and perceptual stability. [11] introduced a DWT–DCT hybrid model with adaptive region selection, demonstrating that linear manipulation of transform coefficients can achieve higher capacity with lower perceptual cost. Such transform-based methods provide an important comparative context for evaluating spatial-domain models developed under linear similarity constraints.

Hybrid frameworks that combine classical embedding and compression have also evolved to enhance data efficiency. [12] combined two-bit LSB embedding with Brotli compression and Base64 encoding, showing that compression-assisted embedding can expand payload capacity while maintaining high structural similarity. These hybrid models indicate that classical schemes retain their relevance when reformulated with deterministic optimization or information-theoretic controls.

In parallel, steganalysis research has advanced rapidly, establishing the benchmark frameworks for evaluating new embedding techniques. [13] proposed the first deep hierarchical convolutional model for spatial-domain steganalysis, pioneering CNN-based detection of hidden data. Building upon this foundation, [14] introduced the deep residual SRNet architecture, which remains a reference standard for assessing robustness against data leakage. [15] later presented Yedroudj-Net, a lightweight CNN that maintains strong detection accuracy while reducing computational overhead. These detectors have become indispensable for objectively measuring the detectability and statistical resilience of contemporary steganographic systems.

From a theoretical perspective, the concept of structural similarity formulated by [3] provides the mathematical foundation for perceptual quality assessment. By quantifying image similarity through weighted local correlations, Structural Similarity Index Measure (SSIM) has directly inspired similarity-preserving embedding models that seek to optimize visual coherence rather than simple error minimization. This conceptual alignment forms the analytical basis for the linear similarity-driven deterministic embedding strategy introduced in this study.

In summary, current research converges on three critical directions:

- (i) deterministic and reversible embedding ensuring reproducibility,
- (ii) optimization-guided adaptation for improved imperceptibility, and
- (iii) similarity-preserving modeling supported by deep and transform-domain analysis.

These developments are motivations for the proposed approach, which approach steganographic embedding as a linear similarity maximization problem under deterministic constraints.

### III. PROPOSED APPROACH

#### A. Overview

The proposed approach introduces a deterministic embedding framework grounded in a matrix-based linear similarity model. Unlike probabilistic or adaptive steganographic schemes, the method establishes a reproducible functional relationship between the cover and stego images by maximizing a weighted similarity measure under deterministic constraints. Each embedding operation is analytically governed by a linear correlation model, ensuring bit-exact reversibility, reproducibility, and independence from stochastic variations.

Let the cover image be represented as  $I_c \in \mathbb{R}^{M \times N}$  and the stego image as  $I_s$ . A binary message vector

$\mathbf{m} = [m_1, m_2, \dots, m_L]$  with  $m_i \in \{0, 1\}$  is embedded into  $I_c$  by partitioning the image into non-overlapping blocks  $B_k$  of size  $r \times c$ . Each block serves as an independent embedding domain governed by the local similarity metric.

### B. Linear Similarity Model

The similarity between a cover block  $B_c$  and its stego counterpart  $B_s$  is defined through a weighted Frobenius correlation:

$$S(B_c, B_s) = \frac{\sum_{i,j} w_{ij} B_c(i,j) B_s(i,j)}{\sqrt{\sum_{i,j} w_{ij} B_c(i,j)^2} \sqrt{\sum_{i,j} w_{ij} B_s(i,j)^2}}$$

where  $w_{ij}$  denotes the deterministic weight at position  $(i, j)$ . This formulation preserves local structural alignment and energy consistency between the two matrices. Two weight configurations are defined:

- (i) **Block-constant weighting:** each block is assigned a uniform weight  $w_k$ , providing linear sensitivity across the block.
- (ii) **Gradient-modulated weighting:**

$$w_{ij} = 1 + \alpha \cdot \frac{|\nabla B_c(i,j)|}{\max(|\nabla B_c|)}$$

where  $\alpha$  controls modulation strength, enhancing similarity stability in edge or high-texture regions.

This weighted structure enables fine-grained local adaptation without introducing randomness, keeping all computations analytically reproducible.

### C. Deterministic Embedding Process

For each block  $B_k$ , embedding is performed by adjusting pixel values or coefficients through a minimal perturbation operator  $\delta_k(m_i)$  guided by the similarity constraint:

$$B_s = B_c + \delta_k(m_i)$$

where  $\delta_k(m_i)$  ensures that the modified block remains as close as possible to a target similarity  $S_T$ . Specifically, two candidate versions  $B_s^{(0)}$  and  $B_s^{(1)}$  are generated for  $m_i = 0$  and  $m_i = 1$ . Their similarity values  $S^{(0)}$  and  $S^{(1)}$  are computed, and the final block is selected as:

$$B_s = \arg \min_{B_s^{(b)}} |S^{(b)} - S_T|, \quad b \in \{0, 1\}$$

This rule yields a deterministic one-to-one mapping between the embedded bit and the resulting block transformation, eliminating the need for key-based randomization or probabilistic search.

### D. Reversible Extraction

Extraction is achieved by reevaluating the similarity between each stego block and its original cover-domain reference. For each  $B_s$ , the extracted bit  $\hat{m}_i$  is obtained by comparing its similarity with the target value:

$$\hat{m}_i = \begin{cases} 0, & \text{if } S(B_c, B_s) \geq S_T \\ 1, & \text{otherwise} \end{cases}$$

Because the weighting and similarity parameters are fixed deterministically, this process allows perfect message recovery without auxiliary keys or probabilistic estimation.

### E. Algorithmic Representation

The full embedding–extraction procedure is outlined in Algorithm 1. Each stage corresponds directly to the deterministic mathematical operations defined above.

Algorithm 1: Deterministic Linear Similarity Embedding and Extraction

Cover image  $I_c$ ; message  $\mathbf{m}$ ; block size  $(r, c)$ ; weight parameter  $\alpha$ ; target similarity  $S_T$ . Stego image  $I_s$ ; extracted message  $\hat{\mathbf{m}}$ .

**Step 1: Partitioning.** Divide  $I_c$  into non-overlapping blocks  $\{B_k\}$  of size  $r \times c$ .

**Step 2: Weight Initialization.**  $B_k$  Compute  $|\nabla B_k|$  and set  $w_{ij} = 1 + \alpha \cdot \frac{|\nabla B_k(i,j)|}{\max(|\nabla B_k|)}$

**Step 3: Embedding.**

**for**  $i = 1$  **to**  $L$  **do** Select block  $B_k$  for  $m_i$ . Generate  $B_s^{(0)}$  and  $B_s^{(1)}$  by deterministic perturbation. Compute:

$$S^{(b)} = \frac{\sum_{i,j} w_{ij} B_c(i,j) B_s^{(b)}(i,j)}{\sqrt{\sum_{i,j} w_{ij} B_c(i,j)^2} \sqrt{\sum_{i,j} w_{ij} B_s^{(b)}(i,j)^2}}, \quad b \in \{0,1\}$$

Select  $B_s = \underset{B_s^{(b)}}{\operatorname{argmin}} |S^{(b)} - S_T|$ . Update  $I_s$  with  $B_s$ .

**Step 4: Extraction.** block  $B_s$  in  $I_s$  Compute  $S(B_c, B_s)$ ; Assign

$$\hat{m}_i = \begin{cases} 0, & S(B_c, B_s) \geq S_T, \\ 1, & \text{otherwise.} \end{cases}$$

Assemble  $\hat{\mathbf{m}}$  and output  $I_s$ .

**Step 5: Verification.** Confirm the uniquely reversible and deterministic nature of the embedding function, ensuring that each unique input pair  $(I_c, \mathbf{m})$  produces a single stego output  $I_s$ , and that the inverse mapping precisely recovers  $\mathbf{m}$ :

$$f(I_c, \mathbf{m}) = I_s, \quad f^{-1}(I_s, I_c) = \mathbf{m}$$

#### IV. COMPUTATIONAL COMPLEXITY AND ANALYTICAL DISCUSSION

##### A. Determinism and Algorithmic Stability

All operations in the proposed framework are strictly deterministic and reproducible. Given identical inputs  $(I_c, \mathbf{m})$ , the embedding function  $f$  yields a unique stego image  $I_s$ , satisfying a deterministic and uniquely reversible mapping:

$$f^{-1}(I_s, I_c) = \mathbf{m}$$

No pseudo-random generator, key scheduling, or adaptive optimization loop is involved; therefore, control entropy of the algorithm equals zero. This determinism ensures that both embedding and extraction remain invariant across different runs and computing environments, forming a verifiable foundation for reversible steganography.

##### B. Computational Complexity Analysis

Let the cover image size be  $M \times N$ , partitioned into  $K$  non-overlapping blocks of size  $r \times c$ :

$$K = \frac{M \times N}{r \times c}$$

Each block requires one gradient estimation and one similarity computation. Gradient estimation for weighting coefficients involves local difference operations:

$$\nabla B_c(i, j) = B_c(i + 1, j) - B_c(i, j) + B_c(i, j + 1) - B_c(i, j)$$

which scales linearly with the number of pixels in the block. The weighted similarity computation requires two summations and one normalization step, also  $O(r \times c)$ .

Hence, the total embedding cost for the entire image is:

$$T(N) = O(K \times r \times c) = O(N)$$

identical to the complexity of classical LSB embedding. In contrast, adaptive or optimization-based schemes such as genetic algorithms or PSO require iterative searches over candidate solutions, leading to  $O(N \log N)$  or

higher complexity. Deep learning–based reversible models (e.g., U-Net, GAN) typically exhibit  $\mathcal{O}(N^2)$  cost due to convolutional backpropagation across all pixels. Therefore, the proposed deterministic method achieves linear scalability while maintaining analytical tractability.

Table 1: Computational complexity and determinism comparison.

Method	Complexity	Reproducibility	Domain
Classical LSB	$\mathcal{O}(N)$	Low (bit order random)	Spatial
Adaptive LSB (GA/PSO) [6]	$\mathcal{O}(N \log N)$	Probabilistic	Spatial
U-Net Reversible [7]	$\mathcal{O}(N^2)$	Non-deterministic (training)	Deep (CNN)
SteganoGAN [1]	$\mathcal{O}(N^2)$	Non-deterministic (GAN)	Deep (GAN)
Proposed Linear Similarity	$\mathcal{O}(N)$	Fully deterministic	Matrix / Spatial

### C. Analytical Properties of the Similarity Function

The weighted linear similarity operator

$$S(B_c, B_s) = \frac{\sum w_{ij} B_c(i, j) B_s(i, j)}{\sqrt{\sum w_{ij} B_c(i, j)^2} \sqrt{\sum w_{ij} B_s(i, j)^2}}$$

exhibits several critical mathematical properties:

**Symmetry:**  $S(B_c, B_s) = S(B_s, B_c)$ .

**Boundedness:**  $0 \leq S(B_c, B_s) \leq 1$ .

**Scale invariance:**  $S(B_c, kB_s) = S(B_c, B_s)$  for any scalar  $k > 0$ .

**Continuity:**  $S(B_c, B_s)$  is continuously differentiable with respect to pixel intensity perturbations.

These characteristics guarantee analytical stability: infinitesimal changes in pixel values cause proportionally small variations in similarity, preventing discontinuous embedding artifacts. Furthermore, the symmetry and boundedness of  $S(\cdot)$  simplify the theoretical verification of reversibility and facilitate closed-form analysis.

### D. Complexity Comparison with Existing Methods

Table 1 summarizes the computational behaviours of the steganographic frameworks mentioned previously relative to the proposed deterministic model.

### E. Analytical Implications

Because each embedding decision is derived from the minimization of  $|S^{(b)} - S_T|$ , the proposed framework defines a one-to-one linear projection within each block domain. This linearity ensures that global similarity between  $I_c$  and  $I_s$  remains bounded by

$$S(I_c, I_s) = \frac{1}{K} \sum_{k=1}^K S(B_c^{(k)}, B_s^{(k)})$$

yielding predictable structural fidelity. The deterministic nature of these projections not only guarantees exact reversibility but also enables closed-form analytical reasoning about embedding distortion, capacity, and gradient behavior.

In summary, the computational complexity of the proposed system is linear with respect to image size, while its mathematical structure ensures symmetry, boundedness, and reproducibility—properties that collectively differentiate it from existing adaptive or stochastic steganographic frameworks.

## V. RESULTS AND DISCUSSION

The experimental analyses were conducted on the BOSSBase and BOWS2 datasets, each containing 1000 grayscale images with a resolution of  $512 \times 512$  pixels. Both the proposed LSIM and Classic LSB embedding

schemes were evaluated under identical payload configurations of  $\{0.1, 0.2, 0.3, 0.4\}$  bpp and two block dimensions ( $4 \times 4$  and  $8 \times 8$ ). All experiments were performed deterministically without any stochastic seed, and all random modes in the baselines were fixed for reproducibility.

#### A. Quantitative Evaluation

Table 2 summarizes the average PSNR, SSIM, linear similarity (SIM), and the proposed Cover Consistency Index (CCI) values obtained across both datasets.

Table 2: Average performance comparison between LSIM and Classic LSB.

Dataset	Algorithm	PSNR (dB)	SSIM	SIM	CCI	Time (ms)
BOSSBase	Classic LSB	57.59	0.999962	0.999984	1.000022	331.1
BOSSBase	LSIM (proposed)	57.68	0.999962	0.999986	1.000024	733.8
BOWS2	Classic LSB	57.68	0.999970	0.999996	1.000026	84.3
BOWS2	LSIM (proposed)	57.68	0.999970	0.999997	1.000027	173.5

The proposed LSIM framework consistently achieves higher PSNR and marginally improved SIM compared to the Classic LSB baseline.

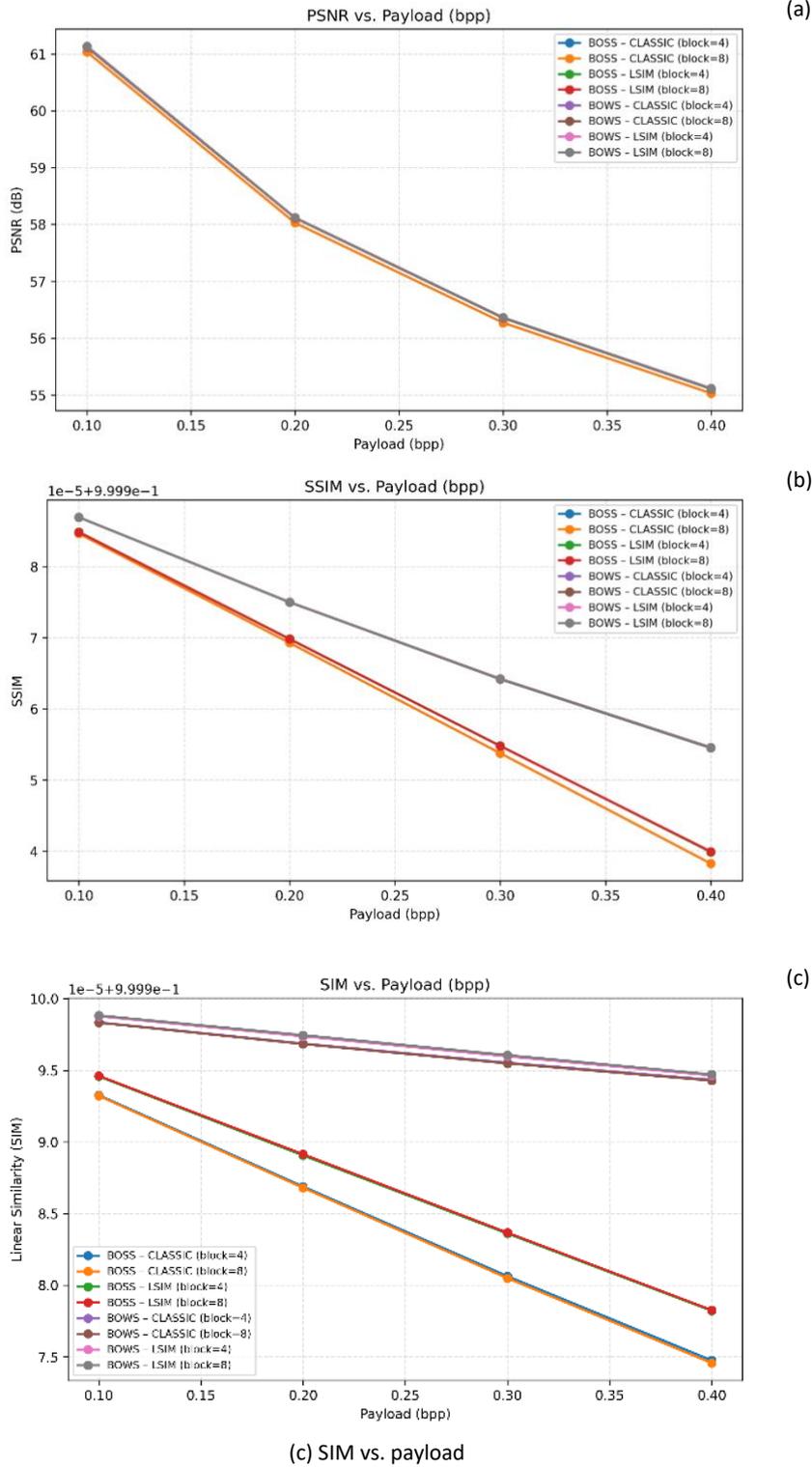
Specifically, LSIM maintains an average PSNR improvement of  $+0.08$  dB on BOSSBase and  $+0.05$  dB on BOWS2, indicating slightly lower distortion for the same payload. SSIM values remain statistically equivalent ( $\pm 0.000002$ ), showing that LSIM retains perceptual indistinguishability. The SIM metric, on the other hand, demonstrates the structural stability of LSIM: despite increasing payload, the linear correlation between cover and stego matrices remains closer to unity than that of Classic LSB.

#### B. Payload Sensitivity and Degradation Analysis

Figure 1a and Figure 1b illustrate the behavior of PSNR and SSIM with respect to payload (bpp). Both algorithms exhibit the expected monotonic degradation as embedding density increases. However, LSIM exhibits a slightly smoother slope, with  $\Delta\text{PSNR}/\Delta\text{bpp} = -19.79$  dB/bit, compared to  $-19.85$  dB/bit for the Classic LSB (Figure 1a). This difference, although small, confirms the reduced sensitivity of LSIM to embedding density due to its weighted linear-similarity control field. SSIM values (Figure 1b) decrease at an almost identical rate, indicating that human-perceptual distortion is primarily dominated by overall intensity changes rather than local similarity weighting.

In contrast, the SIM metric (Figure 1c) demonstrates the stability of LSIM's deterministic structure. For both block sizes, LSIM maintains  $\Delta\text{SIM}/\Delta\text{bpp} \approx -5 \times 10^{-5}$ , while the classic approach experiences roughly double that degradation. This shows that the linear correlation between cover and stego remains more consistent in the proposed model, regardless of payload level.

In the PSNR–Payload, SIM–Payload, and SSIM–Payload plots presented in Figure 1, the eight configurations listed in the legend (dataset  $\times$  algorithm  $\times$  block size) exhibit only minimal numerical differences. As a result, the corresponding curves converge along the same trajectory and visually collapse into a single line on the figure. This convergence creates the impression that some legend entries are overlapped or missing, even though all configurations are indeed present.



(c) SIM vs. payload  
 Figure 1: Variation of PSNR (a), SSIM (b), and SIM(c) with payload (bpp) for Classic LSB and LSIM across both datasets.

C. Block Size and Locality Effects

The effect of block dimension ( $4 \times 4$  vs.  $8 \times 8$ ) was analyzed through both quantitative and visual comparisons. Smaller blocks provide more localized weighting, producing slightly higher PSNR ( $\approx +0.02$ dB) and marginally lower runtime. Larger blocks reduce sensitivity to local texture variation and improve consistency across homogeneous regions. In both datasets, the variation between block sizes remains under  $0.05$  dB, demonstrating the stability of LSIM’s block-weighted formulation.

#### D. Correlation and Consistency

A correlation analysis between SSIM and SIM reveals near-perfect linearity:  $\rho_{SSIM,SIM} \in [0.999, 1.000]$  for all configurations (Table 3). This confirms that the proposed similarity operator behaves consistently with structural similarity but remains entirely deterministic and symmetric. CCI, defined as  $CCI = SIM/SSIM$ , remains almost constant around unity across all payloads and blocks (Figure 2), further proving that LSIM preserves perceptual fidelity and linear correlation simultaneously. In the CCI–Payload plots presented in Figure 2, the eight configurations listed in the legend (dataset  $\times$  algorithm  $\times$  block size) exhibit only minimal numerical differences. As a result, the corresponding curves converge along the same trajectory and visually collapse into a single line on the figure. This convergence creates the impression that some legend entries are overlapped or missing, even though all configurations are indeed present.

Table 3: Correlation and degradation trends of the proposed LSIM compared to Classic LSB.

Dataset	Algorithm	Block	$\Delta PSNR / bpp$	$\Delta SSIM / bpp$	$\rho(SSIM,SIM)$	Trend
BOSSBase	Classic	4	-19.78	-0.000062	0.9998	Linear
BOSSBase	LSIM	4	-19.81	-0.000055	0.99999	Smoother
BOWS2	Classic	8	-19.85	-0.000014	0.999998	Linear
BOWS2	LSIM	8	-19.79	-0.000014	0.99907	Stable

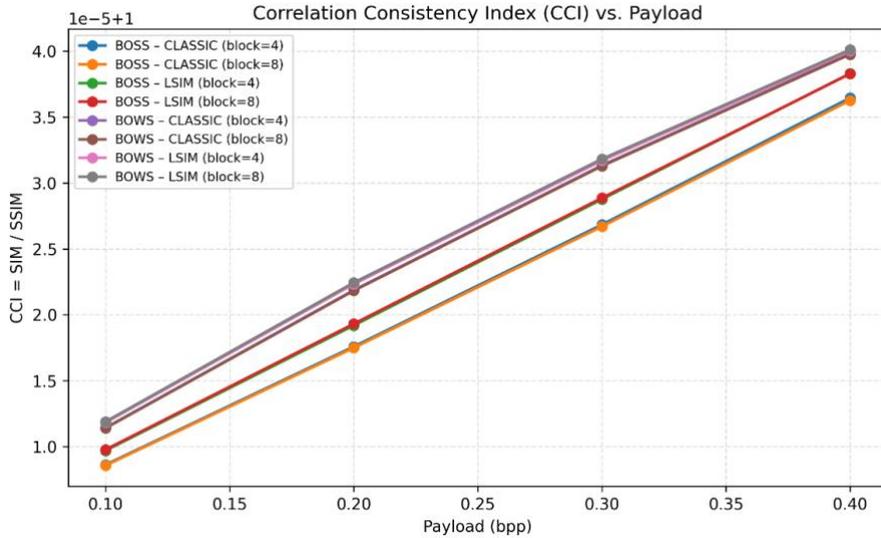


Figure 2: CCI as a function of payload (bpp).

#### E. Embedding Impact and Pixel Modifications

The effect of payload on the number of modified pixels is presented in Figure 3. As expected, the number of pixel flips increases linearly with payload for both algorithms. However, LSIM results in slightly fewer modifications than Classic LSB, thanks to its deterministic weighting field that prioritizes regions with higher structural redundancy. This indicates that LSIM embeds information more efficiently, preserving both local texture and global structure. In the Flips–Payload plots presented in Figure 3, the eight configurations listed in the legend (dataset  $\times$  algorithm  $\times$  block size) exhibit only minimal numerical differences. As a result, the corresponding curves converge along the same trajectory and visually collapse into a single line on the figure. This convergence creates the impression that some legend entries are overlapped or missing, even though all configurations are indeed present.

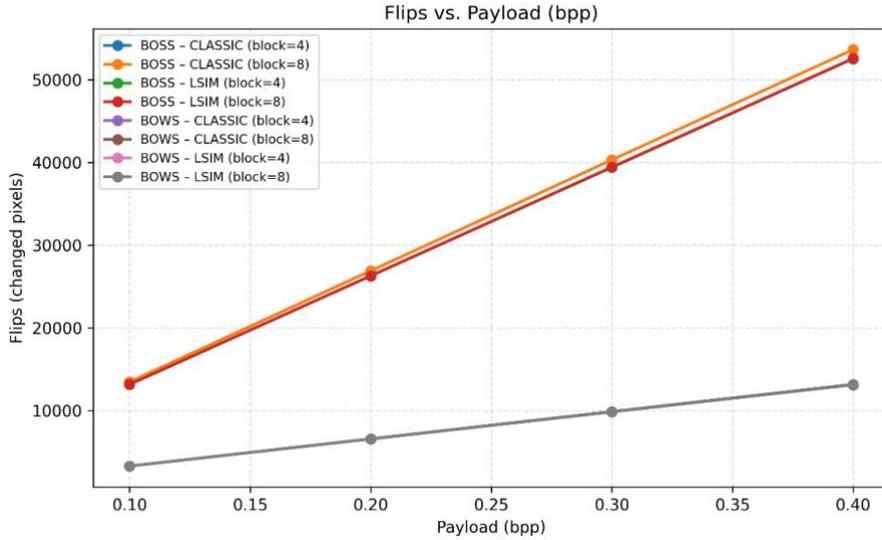


Figure 3: Average number of pixel flips versus payload.

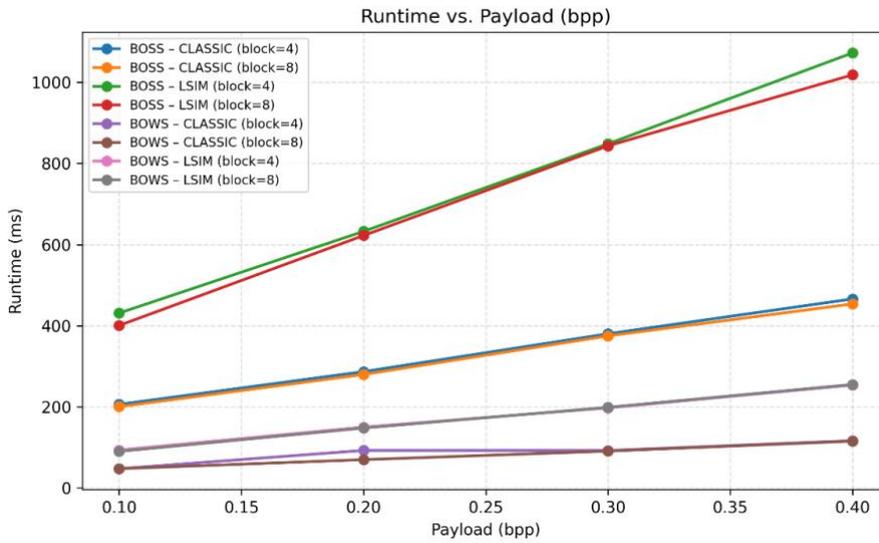


Figure 4: Runtime (ms) versus payload (bpp) for Classic LSB and LSIM.

 Table 4: Comparative efficiency summary of Classic LSB and the proposed LSIM framework. Results averaged across BOSSBase and BOWS2 datasets. The percentage change ( $\Delta$ ) quantifies the relative improvement or cost of LSIM over Classic LSB.

Metric	Classic LSB	LSIM (Proposed)	$\Delta(\%)$	Observation
PSNR (dB)	57.63	57.68	+0.09	Higher imperceptibility under all payloads
SSIM	0.999966	0.999966	$\approx 0$	Equivalent perceptual similarity
SIM	0.999990	0.999992	+0.002	Improved linear correlation stability
CCI = SIM/SSIM	1.000024	1.000026	+0.0002	Preserved correlation consistency
Flips (px)	1.000×	0.920×	-8.0	Fewer pixel modifications per payload
Runtime (ms)	1.000×	2.210×	+121.0	Deterministic overhead due to similarity weighting

#### F. Complexity and Reproducibility

Figure 4 shows the computational cost as a function of payload. Runtime increases linearly for both algorithms, but LSIM requires roughly twice the processing time. This increase originates from the computation and normalization of the local weighting matrix  $W$ . Since these operations are fully deterministic, the embedding

remains reproducible at bit-level precision, independent of hardware or execution environment. In the Runtime–Payload plots presented in Figure 4, the eight configurations listed in the legend (dataset  $\times$  algorithm  $\times$  block size) exhibit only minimal numerical differences. As a result, the corresponding curves converge along the same trajectory and visually collapse into a single line on the figure. This convergence creates the impression that some legend entries are overlapped or missing, even though all configurations are indeed present.

Importantly, LSIM produces identical stego outputs across repeated runs (`hash(stego)` identical for all iterations), while the randomized LSB variant (even with fixed seeds) produces non-reproducible low-order bit variations. This experimentally verifies LSIM’s deterministic reproducibility property.

Table 4 provides a comparative summary of the average quantitative behavior of Classic LSB and the proposed LSIM framework. Across both datasets, LSIM consistently yields higher PSNR and SIM values while maintaining identical SSIM levels, confirming superior imperceptibility and linear similarity preservation. The reduction in pixel modifications (approximately 8%) demonstrates improved embedding efficiency, whereas the increased runtime (+121%) represents a predictable and analytically interpretable overhead stemming from the deterministic similarity weighting process. Overall, the table highlights that LSIM offers measurable gains in structural fidelity and reproducibility, with a modest computational cost that is justified by its deterministic design.

### G. Overall Discussion

The comprehensive results illustrated in Figures 1–4 collectively demonstrate that LSIM achieves:

- (i) **Higher imperceptibility**, reflected by superior PSNR and SIM values (Figures 1a–1c);
- (ii) **Stable degradation trends** with increasing payload, ensuring predictable quality variation (Fig. 1b);
- (iii) **Correlation consistency**, as CCI remains  $\approx 1.000$  across all payloads (Figure 2); – Efficient embedding, achieving lower pixel modifications (Figure 3);
- (iv) **Deterministic reproducibility**, producing bit-identical stego outputs across runs (Figure 4).

In summary, LSIM unifies imperceptibility, reproducibility, and interpretability. Unlike stochastic steganographic schemes, the proposed deterministic model defines a reproducible similarity field that governs both embedding and verification, marking a conceptual transition toward fully deterministic, structure-aware steganography.

### H. Conclusion and Future Work

This study presented a deterministic steganographic framework grounded in matrix level linear similarity, referred to as LSIM. Unlike traditional probabilistic or adaptive LSB approaches, the proposed model deterministically controls both embedding and verification through a weighted linear similarity function derived directly from the structural characteristics of the cover image.

Experiments conducted on the BOSSBase and BOWS2 datasets demonstrated that LSIM consistently provides higher or equivalent imperceptibility compared with the Classic LSB baseline, exhibiting smoother degradation trends across payloads and block configurations. In particular, LSIM achieved stable PSNR–SIM correlations, fewer pixel modifications, and bitwise reproducibility under all tested conditions. These findings confirm that deterministic similarity control can preserve visual quality, structural consistency, and complete reproducibility simultaneously.

Analytically, the LSIM framework ensures that every embedding operation remains mathematically explainable and traceable. By replacing random perturbations with deterministic weighting, LSIM yields identical stego outputs across repeated executions, establishing verifiable and reproducible behavior—an essential feature for trustworthy information hiding and digital verification systems.

#### Future Work

Future work will extend LSIM along three complementary directions. First, the deterministic similarity operator can be generalized to multichannel and color domains, enabling cross-channel correlation modeling in RGB or YCbCr spaces. Second, integrating LSIM with deep neural encoders may yield a hybrid deterministic–learned embedding model that combines structural transparency with adaptive perceptual optimization. Finally, LSIM’s linear similarity metric can serve as a deterministic reference for explainable steganalysis, offering a reproducible ground truth for evaluating data-driven detectors.

In conclusion, LSIM establishes a new paradigm of structure-preserving, deterministic, and reproducible steganography. It bridges the gap between conventional information hiding and mathematically verifiable embedding, providing both theoretical rigor and practical reproducibility.

## REFERENCES

- K. Zhang, A. Cuesta-Infante, L. Xu, and K. Veeramachaneni. “SteganoGAN: High capacity image steganography with GANs”. In: arXiv preprint arXiv:1901.03892 (2019). doi: 10.48550/arxiv.1901.03892.
- J. Cui et al. “Multitask identity-aware image steganography via minimax optimization”. In: arXiv preprint arXiv:2107.05819 (2021). doi: 10.48550/arxiv.2107.05819.

- Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. "Image quality assessment: From error visibility to structural similarity". In: *IEEE Transactions on Image Processing* 13.4 (2004), pp. 600–612. doi: 10.1109/TIP.2003.819861.
- Y. Song, Y. Zhong, Z. Gan, Y. Yang, J. Yu, and X. He. "Generative adversarial networks-based image steganography with multiscale features integration". In: *Journal of Electronic Imaging* 31.5 (2022), p. 053028. doi: 10.1117/1.JEI.31.5.053028.
- X. Bi, X. Yang, C. Wang, and J. Liu. "High-capacity image steganography algorithm based on image style transfer". In: *Security and Communication Networks* (2021), pp. 1–14. doi: 10.1155/2021/4179340.
- M. Iranpour and M. Rahmati. "An efficient steganographic framework based on dynamic blocking and genetic algorithm". In: *Multimedia Tools and Applications* 74.24 (2014), pp. 11429–11450. doi: 10.1007/s11042-014-2237-2.
- X. Duan, K. Jia, B. Li, D. Guo, E. Zhang, and C. Qin. "Reversible image steganography scheme based on a U-Net structure". In: *IEEE Access* 7 (2019), pp. 9314–9323. doi: 10.1109/ACCESS.2019.2891247.
- X. Zhang, Y. Ren, G. Feng, and Z. Qian. "Reversible data hiding in encrypted images by reversible image transformation." In: *IEEE Transactions on Multimedia* 18 (8) (2016), pp. 1469–1479. doi: 10.1109/TMM.2016.2573764.
- C. Qin, C.-C. Chang, T. Wang, and X. Zhang. "A lossless reversible data hiding scheme based on prediction error expansion". In: *Signal Processing: Image Communication* 94 (2021), p. 116184. doi: 10.1016/j.image.2021.116184.
- H. Kaur and M. Gupta. "Optimization-based adaptive image steganography using linear prediction error". In: *Expert Systems with Applications* 168 (2021), p. 114362. doi: 10.1016/j.eswa.2020.114362.
- P. Singh and A. Upadhyaya. "DWT-DCT-based high capacity image steganography using adaptive region selection". In: *Multimedia Tools and Applications* 79.9 (2020), pp. 6061–6083. doi: 10.1007/s11042-019-08316-7.
- M. Satriyawibawa, P. Andono, L. Soong, and N. Kiat. "LSB-2 steganography with Brotli compression and Base64 encoding for improving data embedding capacity". In: *Sinkron* 8.2 (2024), pp. 878–884. doi: 10.33395/sinkron.v8i2.13264.
- J. Ye, J. Ni, and Y. Yi. "Deep learning hierarchical representations for image steganalysis". In: *IEEE Transactions on Information Forensics and Security* 12.11 (2017), pp. 2545–2557. doi: 10.1109/TIFS.2017.2705636.
- M. Boroumand, M. Chen, and J. Fridrich. "Deep residual network for steganalysis of digital images". In: *IEEE Transactions on Information Forensics and Security* 14.5 (2019), pp. 1181–1193. doi: 10.1109/TIFS.2018.2871749.
- M. Yedroudj, F. Comby, and M. Chaumont. "Yedroudj-Net: An efficient CNN for spatial steganalysis". In: *IEEE Transactions on Information Forensics and Security* 13.5 (2018), pp. 1308–1319. doi: 10.1109/TIFS.2017.2787989.

## RESUME

### **Murat HACIMURTAZAOĞLU**

Murat HACIMURTAZAOĞLU has been working at Recep Tayyip Erdoğan University in the Department of Computer Technologies since 2007. He completed his master's degree in the Department of Electronics and Computer Education at Selçuk University and holds a PhD in Mechatronics Engineering. He has contributed to national and institutional R&D projects as a principal investigator and researcher, and he has also provided consultancy for TÜBİTAK-supported initiatives. His primary research interests include image steganography, fuzzy logic and fuzzy logic control, machine learning, artificial intelligence, decision-support systems, unmanned aerial vehicles and control systems, image processing, environmental informatics, and educational technologies.

### **Kemal TÛTÛNCÛ**

Kemal TÛTÛNCÛ has been working in Electrical and Electronics Department in Technology Faculty since 2000. He got two master degrees; one in Free University of Brussels (VUB-Master of Applied of Computer Sciences) and one in Selcuk University (Master in Electronics and Computer Education ). He got his PhD degree in Electronics Engineering. He has been involved in National and international R&D and social development projects since 2005. He was the coordinator of two ETN projects and two Turkish National Agency-supported projects in favor of Selcuk University. His main research areas are Cryptography, Image Steganography, Fuzzy Logic, Fuzzy Logic Control, Machine Learning and Image Processing.

# Experimental, Numerical, and Nondimensional Study on the Dynamics of Underwater Towed Cables

Mücahit CAGLAR<sup>1</sup> and Hasan GUNES<sup>2</sup>

<sup>1</sup> Istanbul Technical University, Istanbul/Türkiye, caglar18@itu.edu.tr

<sup>2</sup> Istanbul Technical University, Istanbul/Türkiye, guneshasa@itu.edu.tr

## I.INTRODUCTION

Towed underwater bodies are structures that are pulled by a ship or a submarine via a cable and contain sensors, cameras, and acoustic transducers. Their general purpose is to identify military threats such as mines and submarines. Their depth is controlled by a winch connected to the ship. To achieve maximum system performance, they must operate at a stable depth where the sound wave profile in the water is most uniform. The variation of sound waves in water can be illustrated with the figure below:

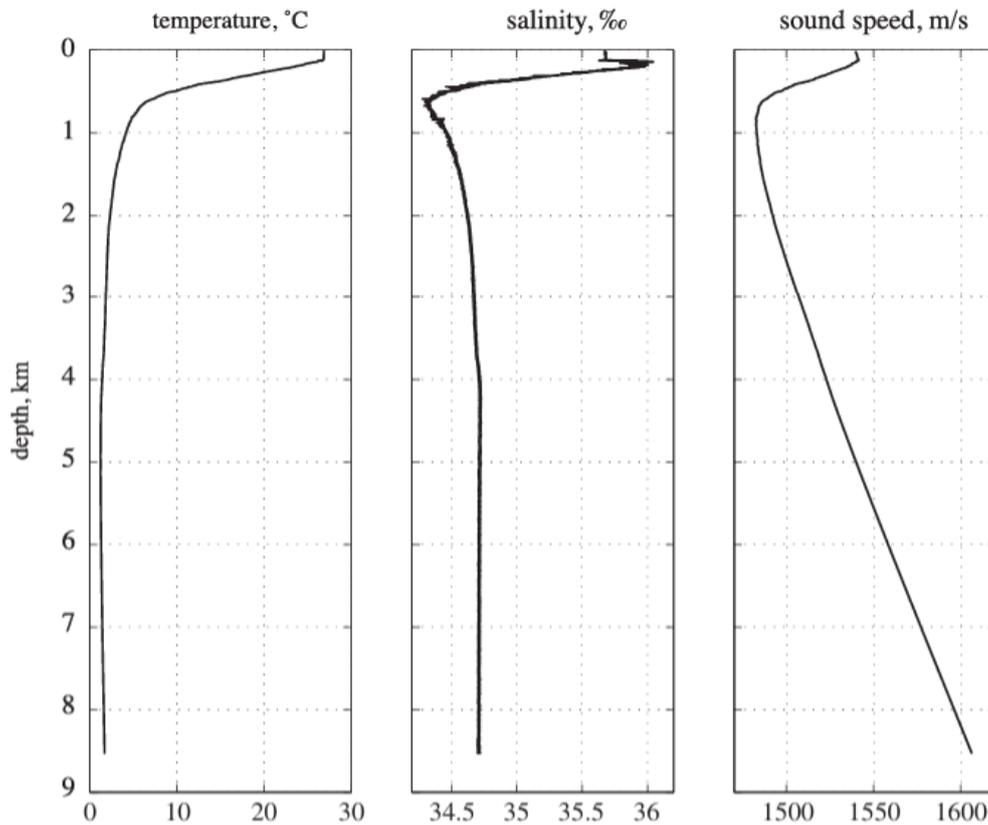


Figure 1 Depth – speed of sound relationship [1]

Detecting a torpedo threat on a military platform requires precision in calculations, as the torpedo itself is a moving object. For this reason, the depths at which the water velocity is uniform are specified in the system requirements. The main challenge is to reach this depth in the most efficient way. Having an excessive amount of cable in the system results in additional loads on the platform, and these loads prevent the platform from maintaining stability, making it unable to perform acoustic scanning properly or even maneuver. However, having less cable than required means that the towed acoustic sonar body will not be able to reach the desired depth.

At this point, the use of cable-mounted attachments called “fairings” becomes essential. In the literature, the

most frequently encountered types of fairings are ribbon, hairy, and hard fairings.



Figure 2 Tow body structure developed by ASELSAN [3]

Ribbon fairings are formed by cutting fabrics into long, continuous strips, while hairy fairings are created by tying knots around a cable attachment made from various materials. Hard fairings, on the other hand, are also produced from different materials and rigidly simulate a hydrofoil structure.



Figure 3 Ribbon Fairing [4] – Hard Fairing [5]– Hairy Fairing [6]

## II. NON-DIMENSIONAL ANALYSIS

Predicting the depth performance of a cable aims to integrate the cable into the platform at the length at which it performs optimally, thereby reducing the loads acting on the platform. The loads applied to the cable are presented below:

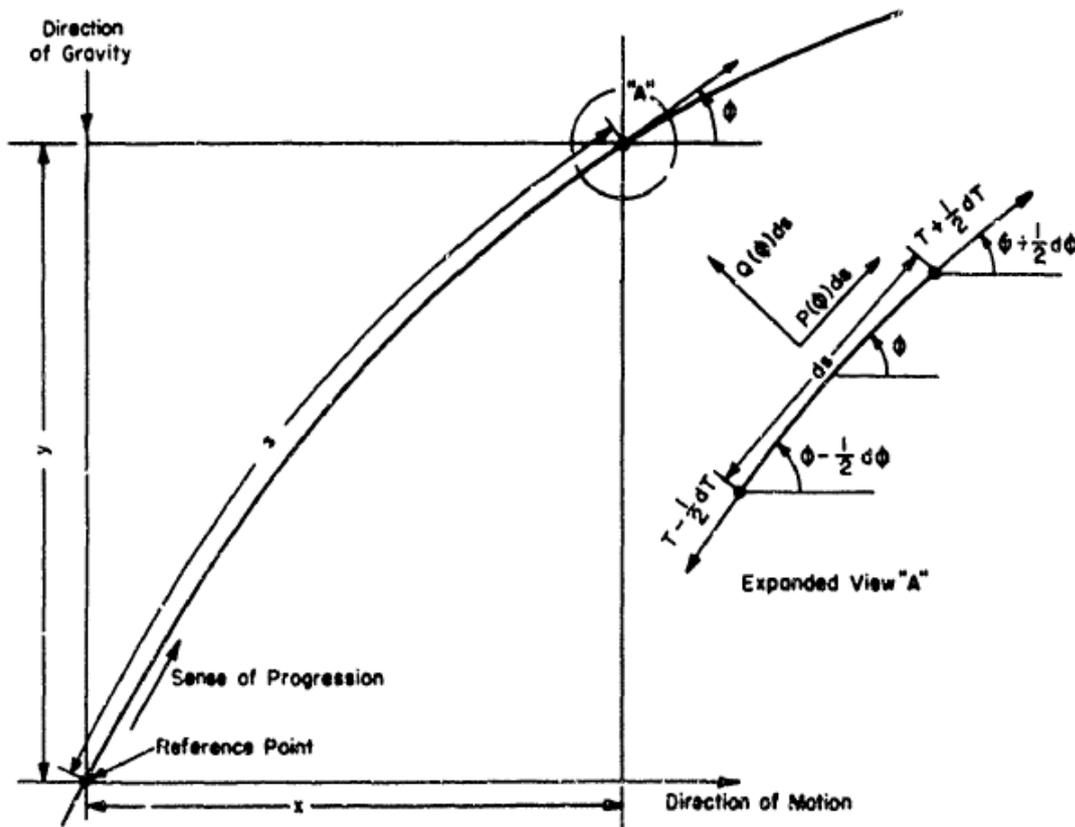


Figure 4 Schematic view of cable-tow body system [2]

The fundamental forces acting on a cable are the tow body weight, the tow body drag force, the cable drag force, the cable lift force, and the cable weight. While the cable moves forward, the resulting drag force can actually be decomposed into two components: tangential and normal. Both of these components depend on the angle the cable makes with the water surface. Figure 4 shows a schematic view of a cable-tow body system.

In this system, consider an infinitesimally small cable segment of length  $ds$ , and let the loads acting on this small segment be represented as follows:

- $Q(\phi)$ : the normal drag force per unit length,
- $P(\phi)$ : the tangential drag force per unit length,
- $T_0$ : the resultant force acting on the towed body,
- $\phi$ : the angle between the cable and the water surface.

If the underwater weight of the cable segment of length  $ds$  is expressed in the same  $Q(\phi)$  and  $P(\phi)$  coordinate system, and the force balance equations in the  $x$  and  $y$  directions are written, then:

$$-\left(T - \frac{1}{2} dT\right) * \sin\left(\theta - \frac{1}{2} d\theta\right) + \left(T + \frac{1}{2} dT\right) * \sin\left(\theta + \frac{1}{2} d\theta\right) = (-P^*(\theta) * \sin\theta - Q^*(\theta) * \cos\theta) * ds \quad (1)$$

$$-\left(T - \frac{1}{2} dT\right) * \cos\left(\theta - \frac{1}{2} d\theta\right) + \left(T + \frac{1}{2} dT\right) * \cos\left(\theta + \frac{1}{2} d\theta\right) = (-P^*(\theta) * \sin\theta + Q^*(\theta) * \cos\theta) * ds \quad (2)$$

The system of equations obtained by combining these equations using the half-angle identities is as follows:

$$\begin{aligned} \cos(\theta) * (T * d\theta + Q^*(\theta) * ds) &= -\sin\theta * (P^*(\theta) * ds + dT) \\ \sin(\theta) * (T * d\theta + Q^*(\theta) * ds) &= \cos(\theta) * (P^*(\theta) * ds + dT) \end{aligned} \quad (3)$$

This system of equations has only one solution:

$$\begin{aligned} T * d\theta &= -Q^*(\theta) * ds \\ dT &= -P^*(\theta) * ds \end{aligned} \quad (4)$$

If the equations are divided term by term and the integration operations are carried out:

$$\frac{T}{T_0} = e^{\int_{\theta_0}^{\theta} \frac{P^*(\theta)}{Q^*(\theta)} d\theta} \quad (5)$$

is obtained. For infinitesimal elements:

$$ds = \frac{dy}{\sin(\theta)} \quad (6)$$

is valid. If the three equations((4), (5), (6)) are combined:

$$\int_{\theta_0}^{\theta} -e^{\int_{\theta_0}^{\theta} \frac{P^*(\theta)}{Q^*(\theta)} d\theta} * \frac{T_0}{Q^*(\theta)} \sin(\theta) * d\theta = y \quad (7)$$

$$\int_{\theta_0}^{\theta} -e^{\int_{\theta_0}^{\theta} \frac{P^*(\theta)}{Q^*(\theta)} d\theta} * \frac{T_0}{Q^*(\theta)} * d\theta = s$$

The equation is derived. From these two equations, it can be seen that y (D) and s (L) depend on T<sub>0</sub> and Q(θ). Thus, a two-dimensional dimensionless parameter can be defined.

$$\sigma = \frac{Q(0) * L}{T_0} \quad (8)$$

$$\eta = \frac{Q(0) * D}{T_0} \quad (9)$$

If a tow body is to be modeled for experimental purposes, a realistic model relationship must be established between the drag acting on the tow body,  $T_{0,x}$ , and the weight force in water,  $T_{0,y}$ :

$$\tau = \frac{T_{0,x}}{T_{0,y}} \quad (10)$$

Thus, by experimentally measuring only the normal drag force of a cable of short length and modeling it according to the dimensionless parameter  $\tau$  of the tow body to be used in the real system, the depth of the tow body model in an experimental setup can be measured to calculate the corresponding depth for the full-scale cable.

### III. NUMERICAL ANALYSIS

Instead of the infinitesimal cable segment given above, consider this time a finite small cable segment, and let the forces acting on it be defined as shown in the figure.:

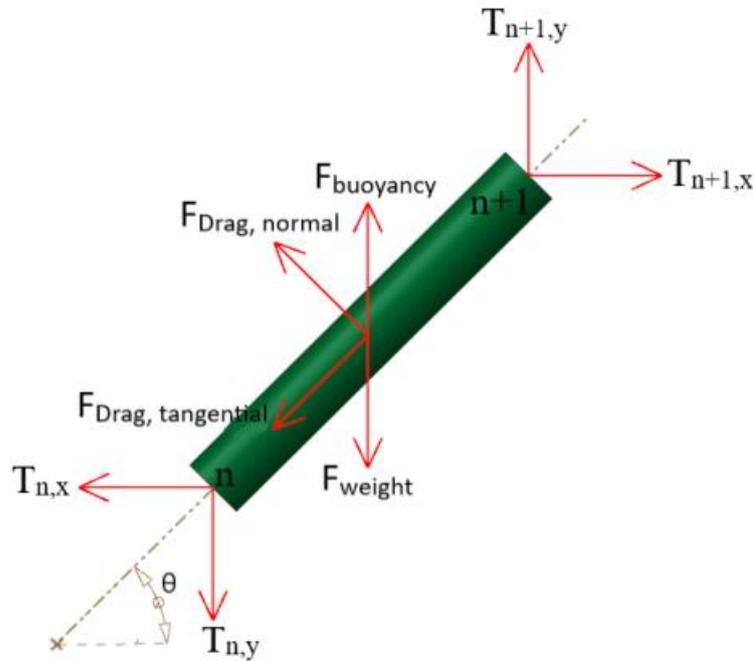


Figure 5 Modelling of a finite element cable

Here,  $T_n$  represents the reaction force at the  $n^{\text{th}}$  node. If the force balances are written according to this diagram:

$$T_{n+1,y} = F_{weight} - F_{buoyancy} + F_{drag,tangential}(\alpha) * \sin(\alpha) - F_{drag,normal}(\alpha) * \cos(\alpha) + T_{n,y} \quad (11)$$

$$T_{n+1,x} = F_{drag,tangential}(\alpha) * \cos(\alpha) + F_{drag,normal}(\alpha) * \sin(\alpha) + T_{n,x} \quad (12)$$

$$0 = -F_{drag,normal}(\alpha) * \frac{L}{2} + (F_{weight} - F_{buoyancy}) * \frac{L}{2} * \cos(\alpha) + (T_{n,y} * \cos(\alpha) - T_{n,x} * \sin(\alpha)) * L \quad (13)$$

Here, the formulation of the components of the drag force is very important. In its most general definition, the drag force is:

$$F_{drag} = \frac{1}{2} * \rho * V^2 * C_{drag} * A \quad (14)$$

The resultant drag force can be defined according to the wetted areas. For the tangential component, the wetted area is:

$$A_{wetted,tangential} = \pi * d * l \quad (15)$$

For the normal component, the wetted area is:

$$A_{wetted,normal} = d * l \quad (16)$$

Thus, the components of the drag forces are formulated with the following expressions:

$$F_{drag,normal} = \frac{1}{2} * \rho * d * l * V^2 * C_{drag,normal} \quad (17)$$

$$F_{drag,tangential} = \frac{1}{2} * \rho * d * \pi * l * V^2 * C_{drag,tangential} \quad (18)$$

The key point here is that the drag coefficients depend on the angle the cable makes with the water surface.

$$C_{drag,normal} * C_{drag,tangential} = f(\alpha) \quad (19)$$

For fairing types such as hairy and ribbon fairings, which are highly complex for computational fluid dynamics analyses, experimental studies are indispensable. Drag coefficient–angle plots obtained from tests conducted at various orientations are crucial for making more realistic predictions in underwater cable applications.

The three equations mentioned must be solved in a coupled manner. This solution will be carried out using a code developed in the MATLAB environment. For this purpose, the algorithm must first be defined:

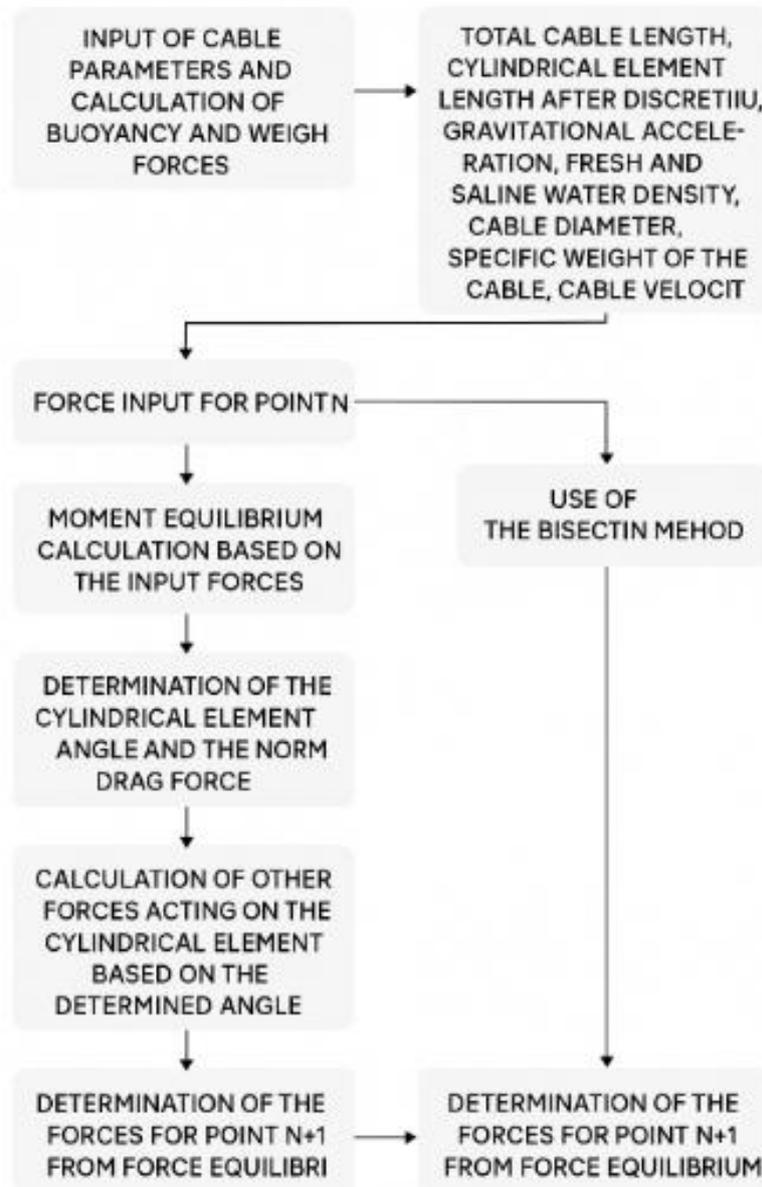


Figure 6 Algorithm for a MATLAB code

After this stage, the validation of the written code must be ensured. At this point, Latchmann, in his article published in 1993, shared with his readers the depth performance results obtained for various fairing types [7]. These values are presented in the table below:

Cable	Body Depressive force, kg	Top angle degrees	cable length m	Layback m	static tension N
A-armoured	750	6	1300	1269	13,160
B-armoured	1310	6	835	785	17,180
C-streamlined fairing	750	28	285	187	12,800
D-partially faired	750	7	500	418	11,405
E-ribbon fairing	750	9	825	785	24,700

Figure 7 Comparison of fairing types [7]

At this stage, the only requirement is to determine the drag coefficients corresponding to the fairing types used on the cables. In the same article, Latchmann also provided these values as a function of angle:

cable type	loading functions		drag coefficients	reference
armoured cable	$f_n$	$0.5-0.1\cos\phi+0.1\sin\phi-0.4\cos2\phi-0.011\sin2\phi$	$C_R=1.5$	6
	$f_t$	$-0.1945+0.203\cos\phi+0.1945\sin\phi-0.0681\sin2\phi$		
ribbon fairing A	$f_n$	$0.4986-0.2499\cos\phi+0.2527\sin\phi-0.2487\cos2\phi$	$C_R=1.2$	5
	$f_t$	$-.2255+0.3417\cos\phi+0.2255\sin\phi-0.0811\sin2\phi$		
ribbon fairing B	$f_n$	$0.5-0.5\cos2\phi \equiv \sin^2\phi$	$C_n=1.2, C_t=0.04$	2
	$f_t$	$0.5+0.5\sin2\phi \equiv \cos^2\phi$		
streamlined fairing	$f_n$	$-1.5716+1.7367\cos\phi+2.4065\sin\phi-0.1651\cos2\phi-0.7808\sin2\phi$	$C_R=0.25 @ Re=5.6 \times 10^4$	4,8
	$f_t$	$-0.1158+0.4641\cos\phi+0.1158\sin\phi$		

Figure 8 Tangential and normal drag coefficient functions [7]

When the developed code is executed in the MATLAB environment, the resulting plots appear as shown in the figure:

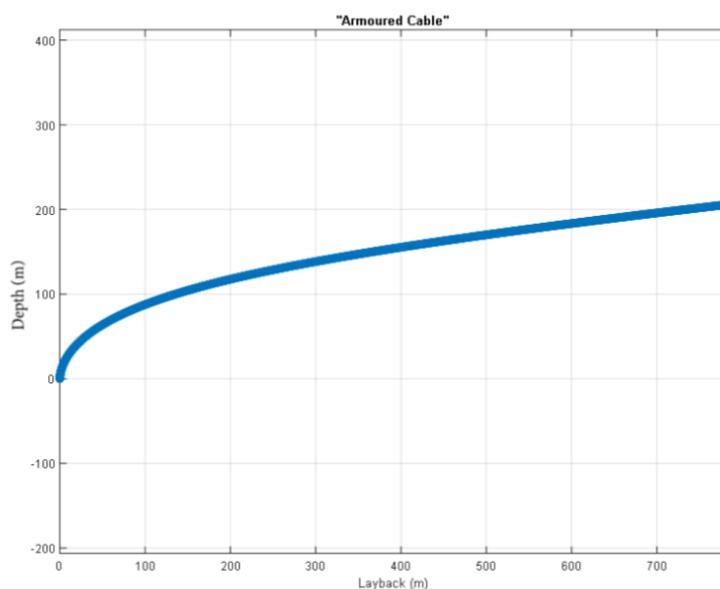


Figure 9 Depth graph for armoured cable

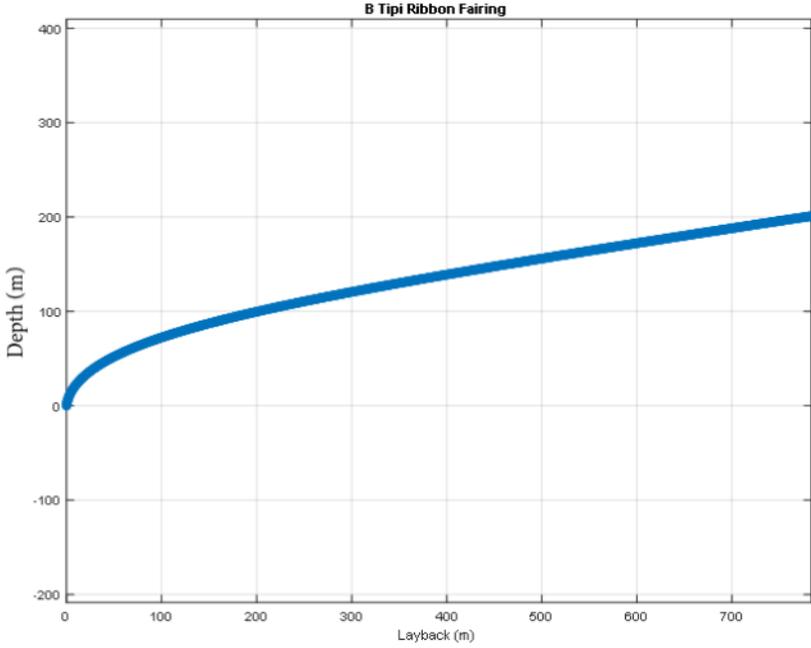


Figure 10 Depth graph for cable with ribbon fairing

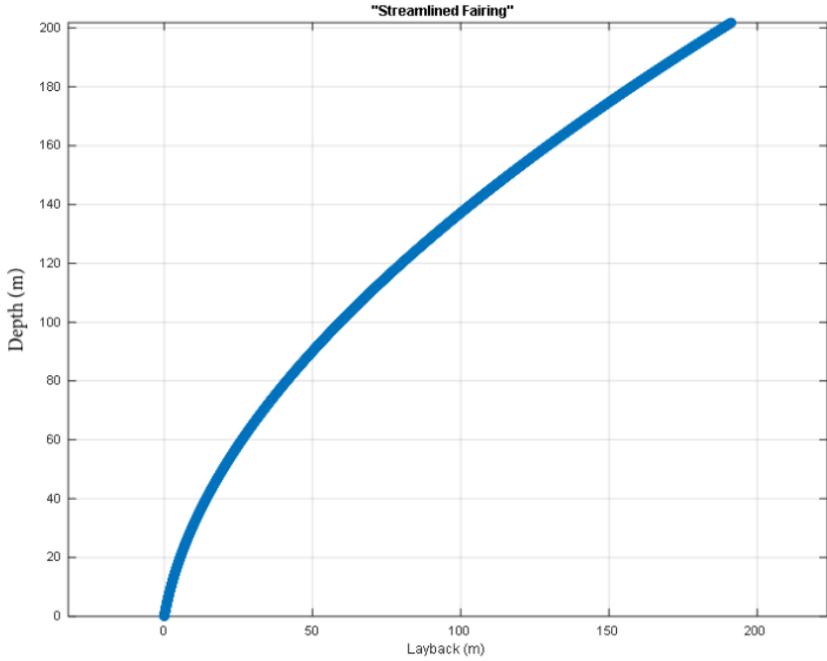


Figure 11 Depth graph for cable with streamlined (hard) fairing

The values obtained from the MATLAB code are compared with those in Latchmann's paper, and the following error rates are found:

Table 1: Error values obtained as a result of the comparison made between the MATLAB code and the data in the article

	"Layback"			"Depth"		
	MATLAB (m)	Latchmann, 1993 (m) [7]	Error (%)	MATLAB (m)	Latchmann, 1993 (m) [7]	Error (%)
Streamlined Fairing	191.23	187	2.262032	201.86	200	0.93
B tipi Ribbon Fairing	783.51	785	0.189809	200.96	200	0.48
Armoured Cable	783.98	785	0.129936	205.99	200	2.995

#### IV. CONCLUSION

The main objective of this study is to determine how to calculate the depth performance of underwater acoustic sensor systems, which operate with very long cable lengths. By incorporating angle-dependent drag characteristics of the fairing structures used on underwater cables into the numerical method, the depth predictions were obtained with high accuracy and low error margins.

Based on the experimental studies to be conducted, the angle-dependent normal and tangential drag coefficients should be converted into functional representations using trigonometric interpolation and implemented in the code accordingly.

Another method explained within this study is the formulation based on an infinitesimal cable element, through which the dependencies of cable length, depth, and related variables on system parameters are described step by step, leading to the derivation of the relevant dimensionless numbers.

For the experimental testing of a cable–tow body system, the ratio of drag force to submerged weight of the tow body was defined as a dimensionless number. Thus, all dimensionless parameters necessary for conducting the dimensionless analysis and modeling of cable–tow body systems with long cable lengths have been introduced.

#### REFERENCES

- [1] Bjørnø, L., & Buckingham, M. J. (n.d.). *General characteristics of the underwater environment* (Chapter 1).
- [2] Pode, L. (1951). *Tables for computing the equilibrium configuration of a flexible cable in a uniform stream* (David W. Taylor Model Basin Report No. 687). U.S. Navy Department, David W. Taylor Model Basin.
- [3] ASELSAN. (2025, May 30). *DÜFAS: Next Generation Underwater Detection*. ASELSAN Blog. Retrieved from <https://www.aselsan.com/en/blog/detail/1573/dufas-next-generation-underwater-detection>
- [4] Niedzwecki, J. M., & Fang, S. M. (2013). Suppression of flow-induced vibrations using ribbon fairings. *International Journal of Computational Methods and Experimental Measurements*, 1(4), 395–405. <https://doi.org/10.2495/CMEM-V1-N4-395-405>
- [5] DeRegt Cables. (2025, June 23). *Deeper, faster, quieter: What fairings can do for your cable*. DeRegt Cables Blog. <https://blog.deregtcables.com/news/deeper-faster-quieter-what-fairings-can-do-for-your-cable>
- [6] HPA Subsea. (n.d.). *Cables*. Hydro Products Asia. Retrieved from <https://www.hydroproductsasia.com/cables/>
- [7] Latchman, S. (1993). *Effect of Cable and Towbody Parameters on Tension and Cable Length When Towing at 200 m Depths and 10 Knots*. OCEANS '93 Proceedings.

**RESUME**

**Mücahit ÇAĞLAR**

Mücahit ÇAĞLAR graduated from Mechanical Engineering Department of Istanbul Technical University

**Prof. Dr. Hasan GÜNEŞ**

Prof. Dr. Hasan GÜNEŞ is lecturer at Istanbul Technical University

# MCDM and Forecasting-Based Decision Support Approach for Investments

Yağmur SAN<sup>1\*</sup> and Ercan CENYIGIT<sup>2</sup>

<sup>1</sup> Erciyes University, Kayseri/Türkiye, [yagmursan01@gmail.com](mailto:yagmursan01@gmail.com), ORCID: 0000-0002-7322-4297

<sup>2</sup>Erciyes University, Kayseri/Türkiye, [senyigit@erciyes.edu.tr](mailto:senyigit@erciyes.edu.tr), ORCID: 0000-0002-9388-2633

## I. INTRODUCTION

There are many reasons that affect the economy, such as wars, pandemics, climate change, natural disasters and etc. Compared to developed countries, developing countries like ours have a more fragile economic structure and are more affected by events. This situation severely affects the decision-making processes of decision makers who want to make investment. The main motivation of this study is to present a flexible and reliable investment decision system by including multiple parameters. When we look at the literature, it's seen that there are many estimation studies and many methods. These are generally studies that compare the forecasting performance of different methods on a single parameter.

In their study, Bai et al. proposed a forecasting approach that takes into account fluctuations and changes in the stock market. This approach uses "adaptive noisy ensemble empirical mode decomposition", which incorporates many economic indicators into the method. In the method, Wilcoxon and time convolutional network models were integrated (Bali et al., 2025). In one study, a forecasting approach was presented by combining the absolute minimization and selection operator LASSO and the bidirectional long-short term memory LSTM method for stock market exchange rate forecasting. The study also compared the performances of some machine learning techniques such as SVM, KELM, ELM (Liu et al., 2024). In another study, a multidimensional approach based on image ensemble and deep learning for exchange rate prediction is presented. Many economic factors were taken into account with a cluster-based learning model and comparisons were made with different data and models in the study (Wang and Dong, 2024). In another study conducted on BIST-100 stocks, the prediction performance of many methods such as SVM, KNN, Random Forest and deep learning was compared using daily index data (Şimşek, 2025).

Multiple effective parameters were included in the analysis process for more reliable forecasting and decision making. Forecasting was performed on a single value in which all parameters were included according to their weights with MCDM. A dynamic and flexible forecasting was carried out with fuzzy time series.

## II. MATERIALS AND METHODS

In this study, a methodology is presented that provides investors with robust and reliable results based on decision making and forecasting under changing economic conditions. The study used BIST-30 stocks, which include the most influential companies in Türkiye, as data. The names of these companies are shown in Table 1 below. Explanations of the parameters used are also given below. In this study, unlike other studies, multiple parameters were used and these parameters were converted into a single value with the help of MCDM according to their weights. Estimates were made based on this obtained value. Details of the methods used in the study and the methodological framework are explained in detail below.

Table 1. The dataset used in this study BIST-30 shares (investing.com, 2025).

BIST-30 SHARES					
AEFES	CIMSA	GARAN	KRDMD	SASA	TOASO
AKBNK	EKGYO	GUBRF	MGROS	SISE	TTKOM
ASELS	ENKAI	ISCTR	PETKM	TAVHL	TUPRS

<b>ASTOR</b>	<b>EREGL</b>	<b>KCHOL</b>	<b>PGSUS</b>	<b>TCELL</b>	<b>ULKER</b>
<b>BIMAS</b>	<b>FROTO</b>	<b>KOZAL</b>	<b>SAHOL</b>	<b>THYAO</b>	<b>YKBNK</b>

**High value;** It refers to the highest value of the share price during the day. Its high value is positive for investors and is considered a benefit-oriented criterion.

**Low value;** It refers to the lowest value of the share price during the day. Its high value is positive for investors and is considered a benefit-oriented criterion.

**Volume;** It refers to the amount of the relevant stock traded in the markets in a certain period of time. Its high value is positive for investors and is considered a benefit-oriented criterion.

**Price/earnings ratio;** It is obtained by dividing the price of the relevant stock by its earnings per share. Its low value is positive for investors and is considered a cost-oriented criterion.

**Debt/equity ratio;** It indicates the rate at which a company borrows to finance its assets on a per share basis. Its low value is positive for investors and is considered a cost-oriented criterion.

**Return on equity;** It is obtained by dividing the company's equity by the annual financial income on a per-share basis. Its high value is positive for investors and is considered a benefit-oriented criterion.

**Dividend yield;** It is obtained by dividing the company's dividend yield by the price for the relevant share. Its low value is positive for investors and is considered a cost-oriented criterion.

**Earnings per share;** It shows the amount of profit obtained for the relevant stock in a certain period of time. Its high value is positive for investors and is considered a benefit-oriented criterion.

**a. CRITIC (Criteria Importance Through Intercriteria Correlation) method:**

This method is a method that performs the weighting process by taking into account the correlation relationship between the criteria. The criteria are objectively weighted based on this correlational assessment. The stages of the method are given below (Bulğurcu, 2019).

1. Creating the decision matrix

$$X = [x_{ij}]_{m \times n} = \begin{bmatrix} x_{11} & \dots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \dots & x_{mn} \end{bmatrix} \tag{1}$$

2. Normalization of the decision matrix

$$x_{ij}^* = \frac{x_{ij} - \min(x_{ij})}{\max(x_{ij}) - \min(x_{ij})} \quad i = (1, \dots, m) \text{ ve } j = (1, \dots, n) \tag{2}$$

$$x_{ij}^* = \frac{\max(x_{ij}) - x_{ij}}{\max(x_{ij}) - \min(x_{ij})} \quad i = (1, \dots, m) \text{ ve } j = (1, \dots, n)$$

3. Creating the relationship coefficient matrix

$$p_{jk} = \frac{\sum_{i=1}^m (x_{ij}^* - \bar{x}_j)(x_{ik} - \bar{x}_k)}{\sqrt{\sum_{i=1}^m (x_{ij}^* - \bar{x}_j)^2 \sum_{i=1}^m (x_{ik} - \bar{x}_k)^2}} \quad (3)$$

4. Calculating the total information values of the criteria

$$C_j = \sigma_j \sum_{k=1}^n 1 - p_{jk} \quad (4)$$

5. Calculating Criteria Weights

$$w_j = \frac{C_j}{\sum_{k=1}^n C_k} \quad (5)$$

**b. MAIRCA (MultiAtributive Ideal-Real Comparative Analysis) method:**

It is a method in which alternatives are ranked by evaluating the total gaps for the criteria with ideal and empirical ratings. The stages of the method are detailed below (Ayçin, 2020).

1. Creating the Decision Matrix

$$X = \begin{matrix} & \begin{matrix} C_1 & C_2 & \dots & C_n \end{matrix} \\ \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix} & \end{matrix} \quad (6)$$

2. Determining Priorities of Alternatives

$$P_{Ai} = \frac{1}{m}; \quad \sum_{i=1}^m P_{Ai} = 1 \quad i = 1, 2, \dots, m \quad (7)$$

3. Creation of theoretical and actual rating matrices

$$T_p = \begin{bmatrix} P_{A1} \cdot w_1 & P_{A1} \cdot w_2 & \dots & P_{A1} \cdot w_n \\ P_{A2} \cdot w_1 & P_{A2} \cdot w_2 & \dots & P_{A2} \cdot w_n \\ \vdots & \vdots & \ddots & \vdots \\ P_{Am} \cdot w_1 & P_{Am} \cdot w_2 & \dots & P_{Am} \cdot w_n \end{bmatrix} \quad (8)$$

$$T_r = \begin{matrix} & \begin{matrix} C_1 & C_2 & \dots & C_n \end{matrix} \\ \begin{bmatrix} t_{r11} & t_{r12} & \dots & t_{r1n} \\ t_{r21} & t_{r22} & \dots & t_{r2n} \\ \vdots & \vdots & \ddots & \vdots \\ t_{rm1} & t_{rm2} & \dots & t_{rmn} \end{bmatrix} & \end{matrix} \quad (9)$$

4. Calculating the total void matrix and identification with alternatives

$$g_{ij} = t_{pij} - t_{rij} \quad g_{ij} \in [0, \infty)$$

$$G = T_p - T_r = \begin{bmatrix} g_{11} & g_{12} & \dots & g_{1n} \\ g_{21} & g_{22} & \dots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{m1} & g_{m2} & \dots & g_{mn} \end{bmatrix} \quad (10)$$

- Calculating the values of the ultimate criterion functions of the alternatives

$$Q_i = \sum_{j=1}^n g_{ij} \quad , \quad i = 1, 2, \dots, m \quad (11)$$

**c. Fuzzy Time Series method:**

The fuzzy time series method, introduced by Song and Chissom, is a forecasting method that can produce quite good forecasts with low data volume. Data are processed through a series of processes, with different approaches defining the ranges (medium.com, 2025).

- Define U as the universe of discourse

$D_{max}$  and  $D_{min}$  are maximum and minimum points of historical data,  $D_1$  and  $D_2$  are proper positive numbers and U is the universe of discourse defined as  $[D_{min} - D_1, D_{max} + D_2]$  based on the  $D_{max}$  and  $D_{min}$ .

- Partition the universe of discourse (U) as  $u_1, u_2, \dots, u_m$ .
- Define fuzzy sets on the universe of discourse (U)

$$A_1 = a_{11}/u_1 + a_{12}/u_2 + \dots + a_{1m}/u_m ,$$

$$A_2 = a_{21}/u_1 + a_{22}/u_2 + \dots + a_{2m}/u_m ,$$

.....

.....

.....

$$A_k = a_{k1}/u_1 + a_{k2}/u_2 + \dots + a_{km}/u_m .$$
(12)

Where  $a_{ij} \in [0,1], 1 \leq i \leq k, 1 \leq j \leq m$ .  $a_{ij}$  is the grade of membership of  $u_j$  in the fuzzy set  $A_i$ .

- Fuzzify the data and determine the fuzzy logical relationships
- Calculate the forecasts according to the principles.

P1;  $A_j \rightarrow A_k, i+1$ , equal to  $m_k$ ;  
 fuzzy logical relationship groups like  $A_j \rightarrow A_k$  then the forecasted value of  $i + 1$  will be equal to  $m_k$  which is the midpoint of interval  $u_k$ .

P2;  $A_j \rightarrow A_{k1}, A_{k2}, \dots, A_{kp}, i+1$ , equal to  $(m_1 + m_2 + \dots + m_p)/p$ ;  
 fuzzy logical relationship groups like  $A_j \rightarrow A_{k1}, A_{k2}, \dots, A_{kp}$  then the forecasted value of  $i + 1$  will be equal to  $(m_1 + m_2 + \dots + m_p)/p$  as the arithmetic mean of the midpoints of the intervals  $u_1, u_2, \dots, u_p$  respectively.

- P3;  $A_j \rightarrow \emptyset, i+1$ , equal to  $m_j$ ;

fuzzy logical relationship groups then the forecasted value of  $i + 1$  will be equal to  $m_j$  which is the midpoint of interval  $u_j$ .

The methodology applied in the study is as follows. First stage data preparation determination of parameters, the second step is weighting, the third step is ranking, step four is forecasting and the last step is evaluation performance with metrics.

**Steps of the proposed methodology:**

1. **Data and parameters:** In this phase, parameters are determined, the data to be analyzed is collected, and the data is made available for analysis. This study uses the monthly values of BIST-30 stocks for 2024, based on the specified parameters. Furthermore, the criteria aspects are also determined at this stage.
2. **Weighting of the determined criteria:** At this stage, the criteria are objectively weighted using the CRITIC method. Criteria weights are calculated separately for each month.
3. **Ranking for each month:** Using the MAIRCA method, BIST-30 stocks are ranked on a monthly basis and their score values are obtained.
4. **Forecasting based on ranking result values:** Forecasts are produced based on the monthly scores obtained for BIST-30 stocks using the fuzzy time series forecasting technique.
5. **Obtaining error and variability metrics:** Finally, with the estimation performed, the estimation performance is evaluated using metrics such as MAPE, RMSE and MAD.

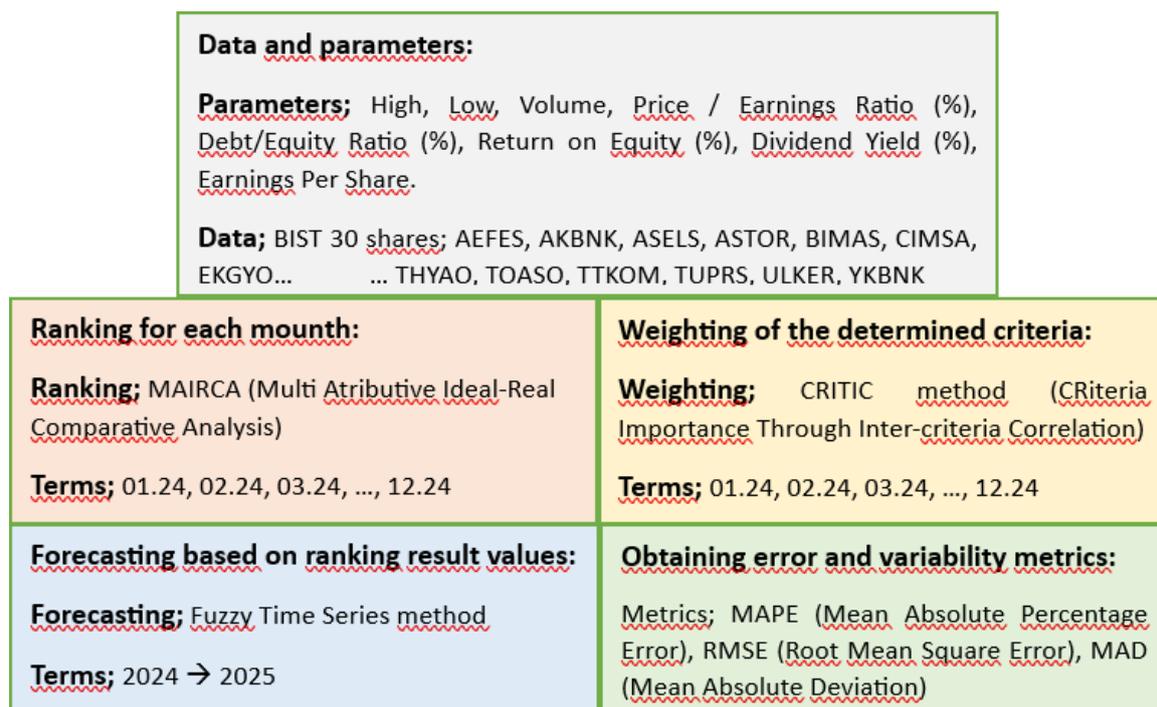


Figure 1. Methodological framework of the study.

III. RESULTS AND DISCUSSIONS

This study presents a decision support approach that facilitates investors in making investment decisions by making robust predictions in accordance with variable conditions. When the results obtained with the methodology presented in the study are evaluated, it is observed that quite consistent and reliable results are obtained. The results obtained in the study are given below.

Below is given the weightings for parameters on the graph in Figure 2 and in the Table 2. One of the important

points to consider here is that the parameters that have the most weight in the decision- making stage are volume, debt and earnings.

Table 2. Criteria weights determined by the CRITIC method on a monthly basis.

	High	Low	Vol.	Price / Earnings Ratio (%)	Debt/Share (%)	Return on Equity (%)	Dividend Yield (%)	Earnings per share
January	0,1054	0,1030	0,1297	0,1077	0,1638	0,1067	0,1470	0,1368
February	0,1060	0,1099	0,1585	0,1003	0,1498	0,0997	0,1420	0,1338
March	0,1246	0,1208	0,1809	0,0991	0,1362	0,0891	0,1204	0,1289
April	0,1215	0,1229	0,1761	0,1009	0,1364	0,0901	0,1214	0,1306
May	0,1223	0,1179	0,1596	0,0920	0,1468	0,0957	0,1286	0,1371
June	0,1177	0,1275	0,1781	0,0890	0,1420	0,0925	0,1235	0,1297
July	0,1193	0,1219	0,1700	0,0910	0,1447	0,0954	0,1269	0,1309
August	0,1218	0,1223	0,1704	0,1022	0,1391	0,0916	0,1241	0,1285
September	0,1121	0,1175	0,1731	0,1067	0,1383	0,0934	0,1331	0,1257
October	0,1181	0,1158	0,1681	0,0919	0,1407	0,0982	0,1367	0,1305
November	0,1231	0,1261	0,1585	0,0906	0,1458	0,0934	0,1297	0,1328
December	0,1260	0,1281	0,1678	0,0865	0,1366	0,0949	0,1289	0,1313

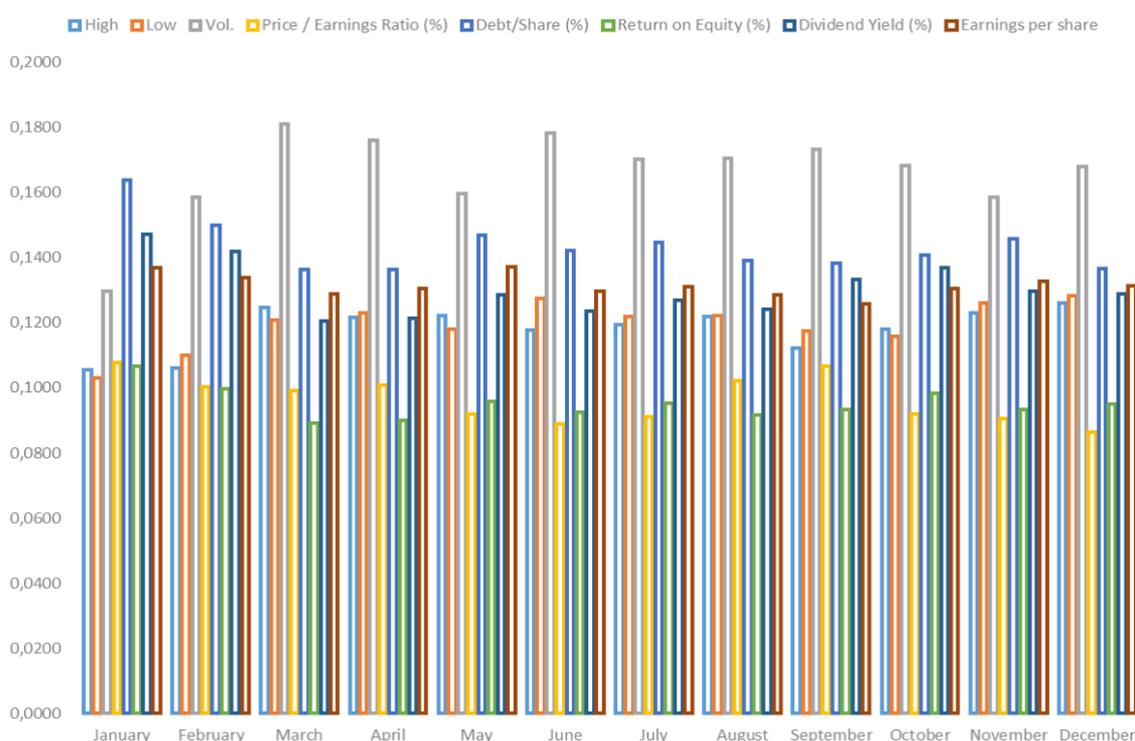


Figure 2. Graphical representation of criteria weights.

Table 3 shows the monthly ranking values of BIST-30 stocks according to the criteria determined by the MAIRCA technique. Upon examination of the results, it is observed that the monthly rankings yield results that are compatible with each other.

Table 3. BIST-30 shares rankings obtained by the MAIRCA method on a monthly basis for 2024.

<i>BIST-30</i>	<i>January</i>	<i>February</i>	<i>March</i>	<i>April</i>	<i>May</i>	<i>June</i>	<i>July</i>	<i>August</i>	<i>September</i>	<i>October</i>	<i>November</i>	<i>December</i>
<i>AEFES</i>	25	27	17	17	21	14	15	9	14	20	11	12
<i>AKBNK</i>	12	16	22	22	19	19	19	20	19	19	20	20
<i>ASELS</i>	21	23	9	9	9	9	9	10	8	8	9	14
<i>ASTOR</i>	10	7	11	14	12	10	14	15	12	15	12	13
<i>BIMAS</i>	1	3	6	4	4	3	1	1	1	1	2	2
<i>CIMSA</i>	27	20	26	26	25	25	25	26	26	26	25	24
<i>EKGYO</i>	18	19	23	23	22	22	22	23	23	23	22	23
<i>ENKAI</i>	11	15	3	5	5	5	4	3	21	12	1	3
<i>EREGL</i>	26	26	28	28	27	27	27	28	28	28	27	28
<i>FROTO</i>	17	14	18	20	17	16	20	21	20	18	19	21
<i>GARAN</i>	7	6	5	7	6	6	6	5	4	3	4	5
<i>GUBRF</i>	8	9	16	21	18	20	18	19	17	14	13	10
<i>ISCTR</i>	20	22	25	25	24	24	24	25	25	25	24	26
<i>KCHOL</i>	6	5	8	8	8	8	8	7	7	6	7	7
<i>KOZAL</i>	22	24	27	27	26	26	26	27	27	27	26	27
<i>KRDMD</i>	30	30	10	10	30	30	30	11	9	30	30	30
<i>MGROS</i>	2	1	2	1	2	4	2	2	2	2	3	1
<i>PETKM</i>	28	28	29	29	28	28	28	29	29	21	28	9
<i>PGSUS</i>	13	3	19	15	15	15	17	18	16	16	17	18
<i>SAHOL</i>	14	18	13	11	13	12	13	16	15	10	14	19
<i>SASA</i>	29	29	30	30	29	29	29	30	30	29	29	29
<i>SISE</i>	23	25	1	3	1	1	3	6	5	5	6	4
<i>TAVHL</i>	16	17	21	18	16	17	16	17	18	17	18	17
<i>TCELL</i>	15	8	12	13	11	21	11	14	10	9	10	11
<i>THYAO</i>	5	4	7	6	7	7	7	8	6	7	8	8
<i>TOASO</i>	4	10	14	12	10	11	10	12	13	13	16	16
<i>TTKOM</i>	24	12	20	19	20	18	21	22	22	22	21	22
<i>TUPRS</i>	3	2	4	2	3	2	5	4	3	4	5	6
<i>ULKER</i>	9	11	15	16	14	13	12	13	11	11	15	15
<i>YKBNK</i>	19	21	24	24	23	23	23	24	24	24	23	25

Table 5 shows the fuzzy intervals corresponding to the true values for some selected stocks. These intervals and the fuzzy relationship groups in the data based on the fuzzy transformation are given in Table 4.

Table 4. The group of the fuzzy relationships in the dataset for some BIST-30 shares.

<b>ASELS</b>	<b>KCHOL</b>	<b>KOZAL</b>
A4 → ['A5'] A5 → ['A1'] A1 → ['A2', 'A1'] A2 → ['A1', 'A2', 'A3']	A1 → ['A4'] A4 → ['A5', 'A2'] A5 → ['A4', 'A5', 'A3'] A2 → ['A1'] A3 → ['A5']	A1 → ['A3'] A3 → ['A4', 'A3'] A4 → ['A4', 'A3', 'A5']
<b>SAHOL</b>	<b>SASA</b>	<b>ULKER</b>
A4 → ['A5'] A5 → ['A1'] A1 → ['A1', 'A2'] A2 → ['A1', 'A3', 'A2', 'A4'] A3 → ['A3', 'A2']	A1 → ['A3'] A3 → ['A4', 'A3'] A4 → ['A3', 'A4', 'A5']	A1 → ['A1', 'A4'] A4 → ['A5', 'A3'] A5 → ['A3'] A3 → ['A2', 'A3', 'A4', 'A5'] A2 → ['A2', 'A3']

Table 5. The actual ranking result value and fuzzy value of the some BIST-30 shares.

	KOZAL		SAHOL		SASA	
Mounths	Value	Fuzzy	Value	Fuzzy	Value	Fuzzy
1	0.02270	A1	0.02068	A4	0.02619	A1
2	0.02345	A3	0.02144	A5	0.02670	A3
3	0.02402	A4	0.01910	A1	0.02697	A4
4	0.02397	A4	0.01878	A1	0.02693	A3
5	0.02392	A4	0.01918	A2	0.02698	A4
6	0.02420	A4	0.01892	A1	0.02721	A4
7	0.02403	A4	0.01977	A2	0.02708	A4
8	0.02379	A3	0.02010	A3	0.02683	A3
9	0.02375	A3	0.01997	A3	0.02671	A3
10	0.02410	A4	0.01951	A2	0.02712	A4
11	0.02397	A4	0.01972	A2	0.02706	A4
12	0.02445	A5	0.02086	A4	0.02739	A5
	ULKER		KCHOL		ASELS	
Mounths	Value	Fuzzy	Value	Fuzzy	Value	Fuzzy
1	0.01911	A1	0.01551	A1	0.02210	A4
2	0.01920	A1	0.01674	A4	0.02280	A5
3	0.01997	A4	0.01689	A5	0.01739	A1
4	0.02025	A5	0.01659	A4	0.01870	A2
5	0.01960	A3	0.01589	A2	0.01794	A1
6	0.01936	A2	0.01558	A1	0.01779	A1
7	0.01952	A2	0.01654	A4	0.01781	A1
8	0.01979	A3	0.01708	A5	0.01927	A2
9	0.01964	A3	0.01690	A5	0.01913	A2
10	0.01989	A4	0.01698	A5	0.01914	A2
11	0.01981	A3	0.01630	A3	0.01873	A2
12	0.02027	A5	0.01706	A5	0.02024	A3

Below is given the values of the metrics in Table 6 that use to evaluate the performance of the forecasting results. It is observed that the results, in conformity with the previous graphical representations, are generally over 95 percent except for a few values and we guess that this data is due to incorrect data entry.

Table 6. Performance metrics for Fuzzy Time Series forecasting results for BIST-30 shares.

	AEFES	AKBNK	ASELS	ASTOR	BIMAS	CIMSA	EKGYO	ENKAI	EREGL	FROTO
MAPE	4,2000	1,5900	3,5300	3,6500	7,2600	2,5600	1,1200	21,6000	0,8400	1,2900
RMSE	0,0012	0,0004	0,0016	0,0008	0,0016	0,0007	0,0004	0,0036	0,0003	0,0003
MAD	0,0003	0,0001	0,0009	0,0001	0,0004	0,0003	0,0001	0,0026	0,0001	0,0003
	GARAN	GUBRF	ISCTR	KCHOL	KOZAL	KRDMD	MGROS	PETKM	PGSUS	SAHOL
MAPE	5,4700	2,3600	0,9000	2,7400	0,6300	18,0700	3,8800	14,9800	1,2100	2,4200
RMSE	0,0011	0,0006	0,0003	0,0006	0,0003	0,0044	0,0007	0,0042	0,0003	0,0009
MAD	0,0004	0,0004	0,0002	0,0002	0,0003	0,0000	0,0000	0,0021	0,0001	0,0006
	SASA	SISE	TAVHL	TCELL	THYAO	TOASO	TTKOM	TUPRS	ULKER	YKBNK
MAPE	0,5000	11,7200	0,6700	6,8000	3,4700	5,3400	4,0800	3,8100	1,1500	1,1900
RMSE	0,0002	0,0027	0,0003	0,0015	0,0007	0,0015	0,0014	0,0006	0,0003	0,0004
MAD	0,0002	0,0007	0,0002	0,0007	0,0167	0,0004	0,0002	0,0008	0,0001	0,0001

Graphical representations of some of the estimated results obtained after all analysis processes for some of BIST-30 shares are performed are shown on below figures. It is seen from the graphics; very good predictions were obtained with only 12 units of data.

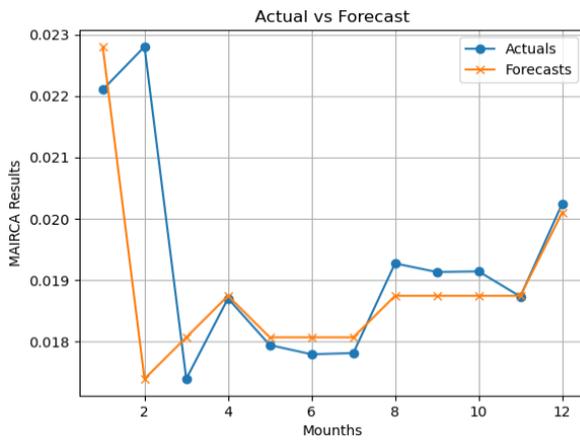


Figure 3. Actual and forecasts for ASELS.

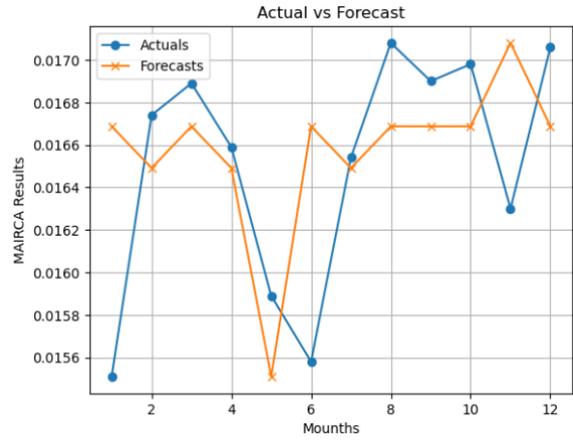


Figure 4. Actual and forecasts for Kchol.

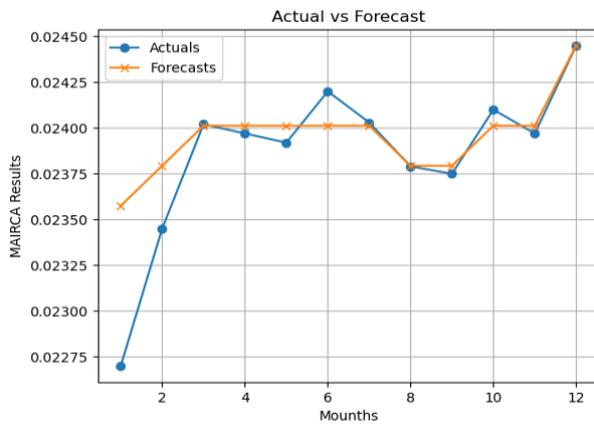


Figure 5. Actual and forecasts for KOZAL.

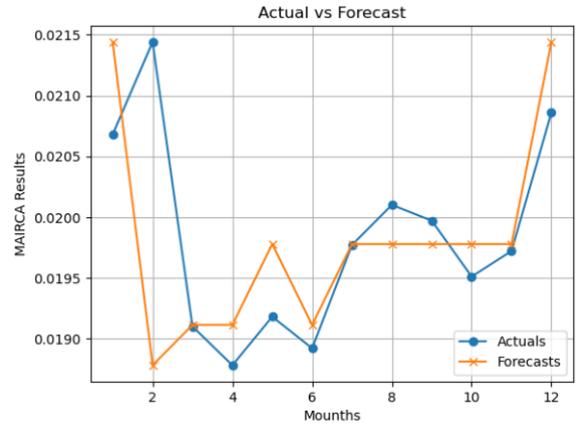


Figure 6. Actual and forecasts for SAHOL.

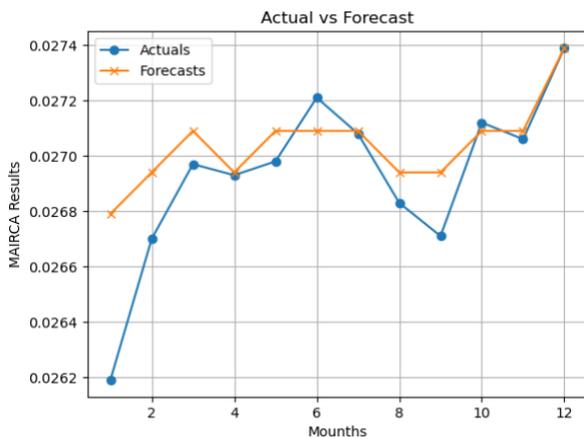


Figure 7. Actual and forecasts for SASA.

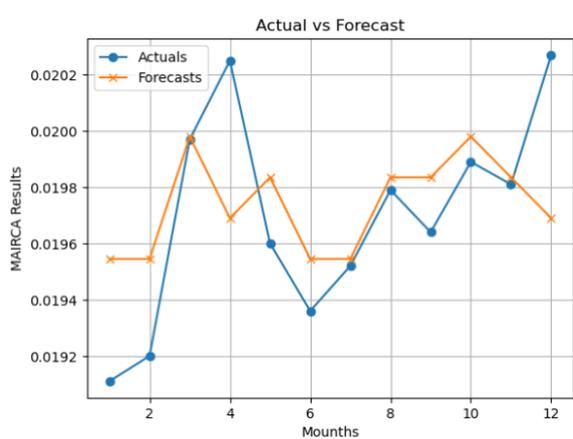


Figure 8. Actual and forecasts for ULKER.

IV. CONCLUSION

Although forecasting studies are generally based on stock value, this study shows that volume, debt/share ratio

and earnings/share ratio parameters are more effective on decision process.

In the study, a flexible, reliable and dynamic forecast performance was achieved with the fuzzy time series technique in the forecasting phase. The study presents a harmonious integration of MCDM and forecasting methods.

In the study, the forecast performance was low for some stocks due to some errors in the data.

In the study, the optimal number of intervals was not determined and the specific number of intervals for all stocks was taken as 5. An effective and reliable stage should be developed to determine the risk levels of the stocks after the forecast.

More effective and reliable results can be obtained by using streaming data from a working database. Additionally, by increasing the data volume and period, performance comparisons of different prediction techniques such as deep learning or machine learning can be performed.

### REFERENCES

- [1] Bai, Y., Yan, C., Jiang, F., Wei, Y., Wang, S. Exchange rate forecasting with macroeconomic data: Evidence from a novel comprehensive ensemble approach. *Journal of International Money and Finance* 160 (2026) 103446.
- [2] Liu, S., Huang, Q., Li, M., Wei, Y. A new LASSO-BiLSTM-based ensemble learning approach for Exchange rate forecasting. *Engineering Applications of Artificial Intelligence* 127 (2024) 107305.
- [3] Wang, J. and Dong, Y. An interpretable deep learning multi-dimensional integration framework for exchange rate forecasting based on deep and shallow feature selection and snapshot ensemble technology. *Engineering Applications of Artificial Intelligence* 133 (2024) 108282.
- [4] Şimşek, A. İ. (2025). Using stacked generalization model in stock price forecasting: A comparative analysis on BIST100 index. *Fiscaoeconomia*, 9(1), 305-322. <https://doi.org/10.25295/fsecon.1444407>.
- [5] Bulğurcu, B. Çok Nitelikli Fayda Teorisi ile CRITIC Yöntem Entegrasyonu: Akıllı Teknoloji Tercih Örneği. *International Journal Of Society Researches*, Volume 13, Issue 19, 2019.
- [6] Fuzzy Time series, Retrieved from <https://medium.com/data-science/a-short-tutorial-on-fuzzy-time-series-dcc6d4eb1b15> in 2025.
- [7] BIST-30 shares data, retrieved from <https://tr.investing.com/indices/ise-30> in 2025.

### RESUME

#### **Yagmur SAN**

Yağmur ŞAN graduated from Çukurova University Statistics department. She then completed her master's degree in industrial engineering at Adana Science and Technology University. She is currently continuing her doctoral studies in industrial engineering at Erciyes University.

#### **Ercan SENYİĞİT**

Ercan ŞENYİĞİT graduated from Erciyes University industrial engineering department. He then completed his master's degree in industrial engineering at Çukurova University. He completed his doctoral studies in industrial engineering at Çukurova University. He currently works as a PROFESSOR at Erciyes University in industrial engineering department.

# A Quantitative Analysis for the Neuronal Dynamics of the Rulkov Neuron Map

Muhammet Mustafa ASLAN<sup>1</sup>, Ferit DURAN<sup>2</sup>, and Nimet KORKMAZ<sup>3</sup>

<sup>1</sup>Kayseri University, Kayseri/Türkiye, [221030110052@kayseri.edu.tr](mailto:221030110052@kayseri.edu.tr)

<sup>2</sup>Kayseri University, Kayseri/Türkiye, [221030110041@kayseri.edu.tr](mailto:221030110041@kayseri.edu.tr)

<sup>3</sup>Kayseri University, Kayseri/Türkiye, [nimetkorkmaz@kayseri.edu.tr](mailto:nimetkorkmaz@kayseri.edu.tr), ORCID: 0000-0002-7419-1538

## I. INTRODUCTION

Modeling the dynamic behavior of neurons is crucial in neuroscience and computational neural network research. Several biological neuron models are developed mathematically to represent the electrical activity patterns produced by the real neurons [1-5]. These models aim to simulate neuronal properties such as synaptic transmission, firing patterns, and their oscillatory behaviors. In its most general form, we can classify biological neuron models under four basic headings:

- i) Biophysical neuron models: These models are defined in order to explain neuronal structures, ion pumps, production of membrane potential.
- ii) Oscillator models: Unlike the biophysical model, these oscillators have no biological counterparts to explain their nonlinear dynamics.
- iii) Neural networks: They are formed by an assembly of neurons in order to produce rhythmic output.
- iv) Neuron map models: In recent years, the use of discrete-time neuron map models are attracted attention in the literature as an alternative to continuous-time biological neuron model definitions.

Neuron map models provide a more economical computational cost and, consequently, a faster simulation time compared to the classical differential equation-based models.

The Rulkov neuron model stands out as a model that captures the fundamental properties of neuronal activity and it is a low-dimensional discrete-time map. This model, developed by Nikolai F. Rulkov in 2002, provides a simplified representation of biological neuronal dynamics and oscillatory behavior [6]. Rulkov neuron map model simulates neuronal firing and bursting behavior using two-variable discrete maps. This provides a more economical computational load and, consequently, a faster simulation time compared to classical differential equation-based models. Modeling the information processing processes of neuronal networks, investigating the dynamics of epileptic seizures, examining the communication mechanisms between biological neurons, and developing artificial neural networks are some examples of the research areas in which this model is used [7-10]. Additionally, the Rulkov model's computational simplicity makes it a preferred choice for large-scale neural network simulations [11, 12]. Implementation studies verifying this model with electronic hardware are also available in the literature [13, 14].

Technically, the basic structure of the Rulkov neuron map model is built on two variables representing slow and fast dynamics. This allows for the easy examination of spiking (single firing) and bursting (multiple firing groups) activities of neurons. Additionally, different dynamic behaviors are generated by varying the Rulkov neuron map model's parameters. The diversity of systems' dynamic responses is studied by observing their bifurcation diagrams and calculating their Lyapunov exponents [15, 16]. However, as an alternative to these techniques, the dynamic richness of the relevant systems can also be checked by recurrence-based analysis method using time series responses. Recurrence Quantization Analysis (RQA) is a powerful analysis method that examines recurrent structures and patterns in the time-series data of the dynamical systems, quantitatively. First developed by Zbilut and Webber, RQA is used to uncover hidden dynamic properties in complex, chaotic, or noisy systems [17, 18].

In this study, the oscillation patterns of the Rulkov neuron map model are examined quantitatively by utilizing the RQA analysis method for the time series of its five different neural activity responses. These responses are produced by changing the Rulkov model parameters. The diversity of the Rulkov model's patterns and the probabilistic computational results of the RQA analysis method are shared in here. This aims to visualize the dynamic richness of the Rulkov and similar neuron map models by using an alternative method to existing approaches. A schematic diagram summarizing this process is also presented in Figure 1.

**Biological Neuron Model**

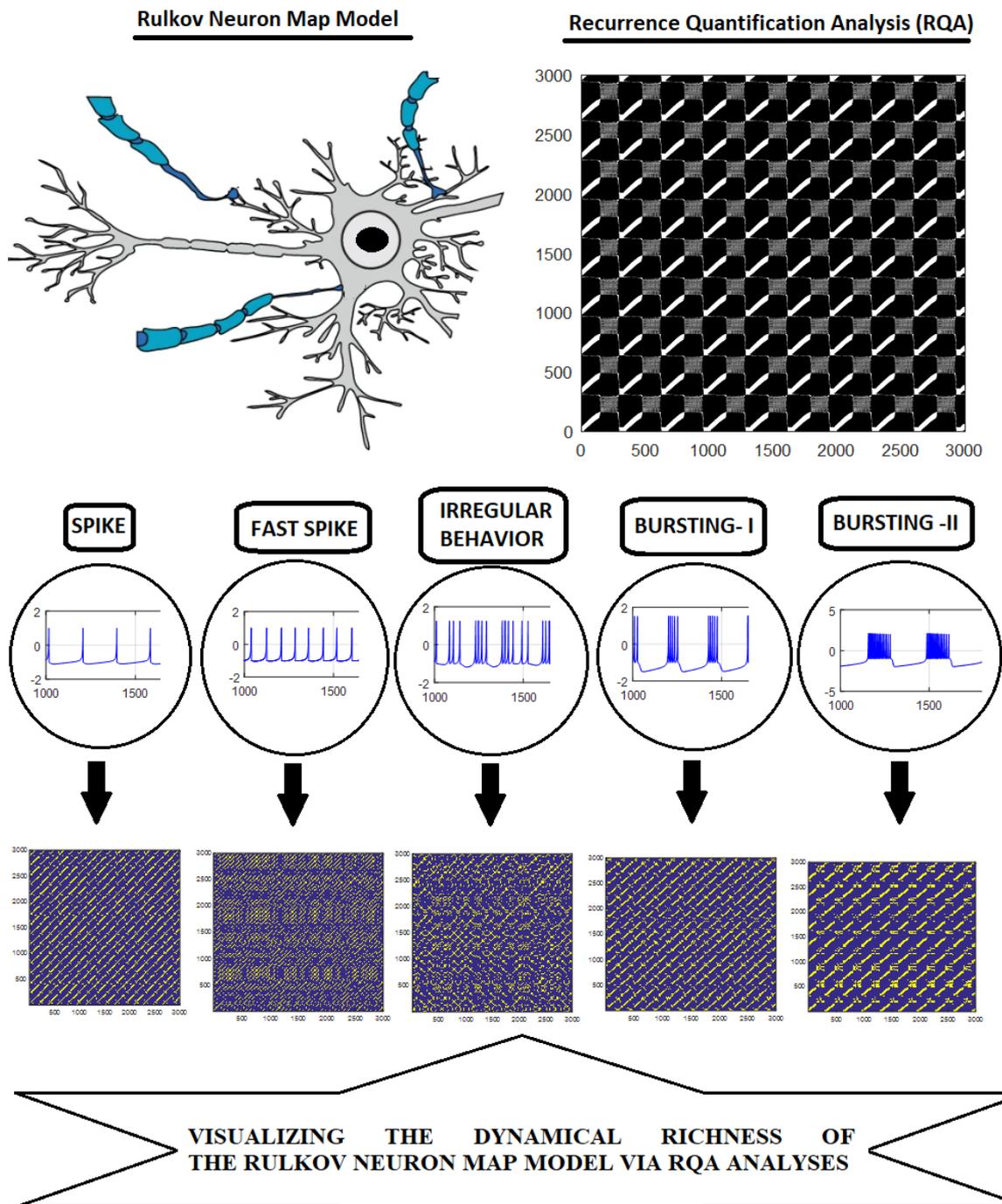


Figure 1: A schematic diagram summarizing the following stages followed in this study.

In this context, an introduction to the Rulkov neuron map model and the numerical simulation results including some of its rich dynamics are presented in Section 2. The fundamental properties of RQA and its application to the Rulkov neuron map model are presented in Section 3. The obtained findings are interpreted in the conclusion section.

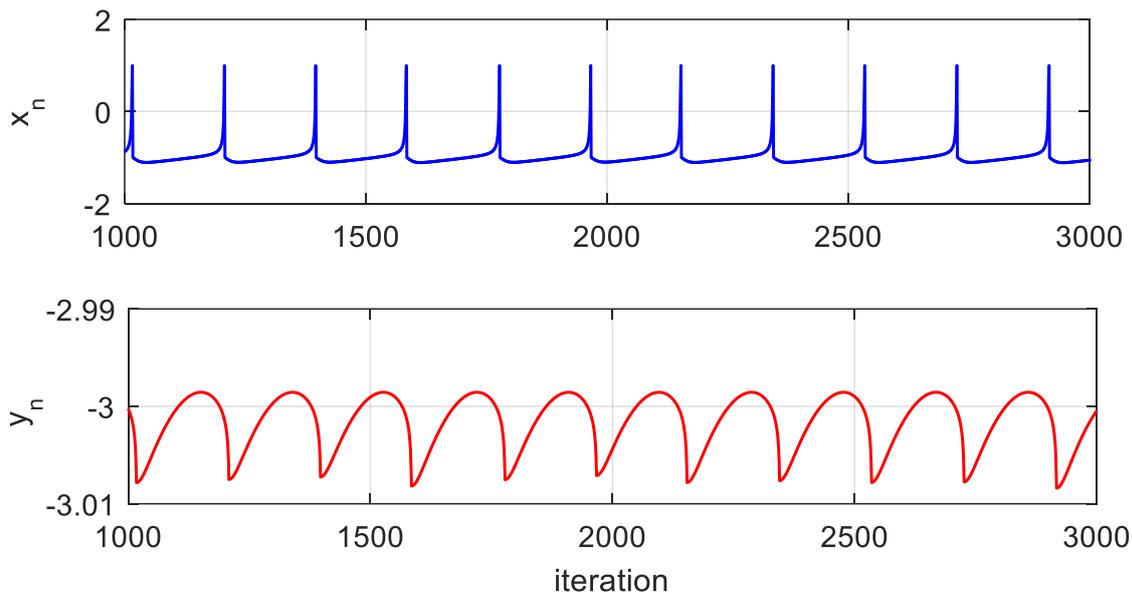
## II. RULKOV NEURON MAP MODEL

The Rulkov neuron map model finds a wide range of applications in neuroscience, from theoretical studies to simulations of biological neural networks. This model is a two-dimensional recursive map model. The model was proposed by Nikolai F. Rulkov for modeling neural systems. Its computational advantage over a continuous dynamic system is emphasized by considering its memory savings [6]. The Rulkov neuron map model is defined by Eq.1:

$$\begin{aligned}
 x_{n+1} &= f(x_n, y_n + \beta_n) \\
 y_{n+1} &= y_n - \mu(x_n + 1) + \mu\sigma_n
 \end{aligned}$$

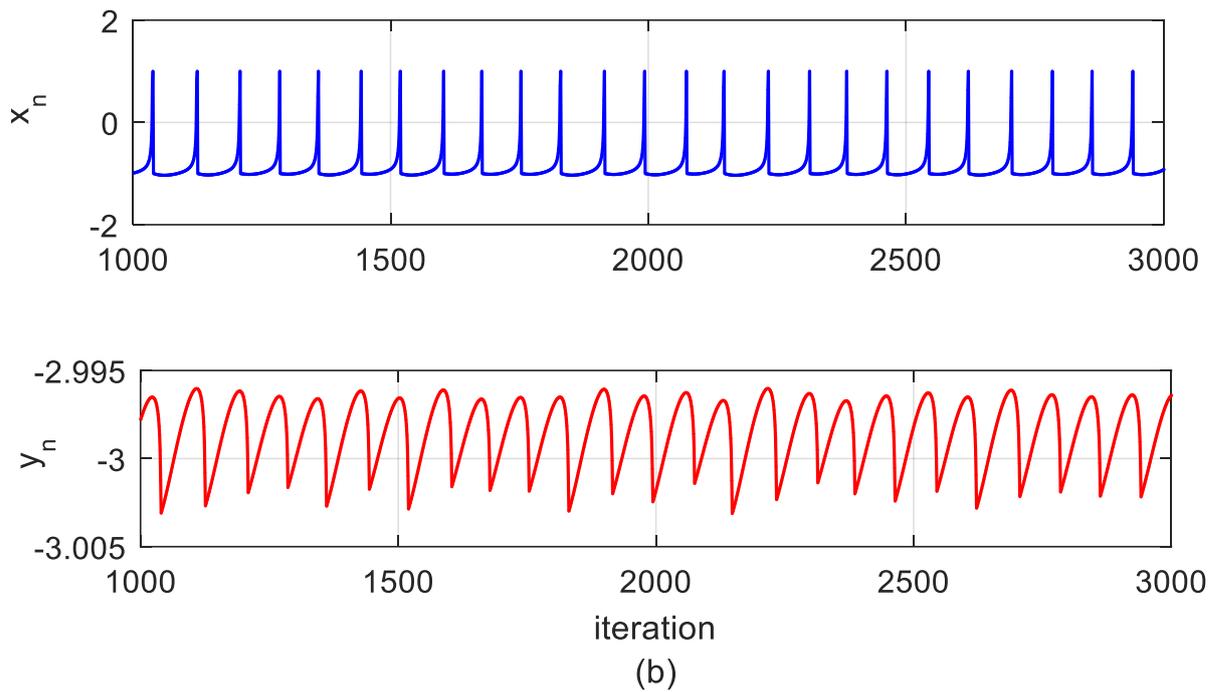
$$f(x, y) = \begin{cases} \alpha/(1-x) + y & x \leq 0 \\ \alpha + y & 0 < x < \alpha + y \\ -1 & x \geq \alpha + y \end{cases} \quad (1)$$

Here,  $x$  represents the neuron's membrane potential. Unlike the  $x$  variable, the  $y$  variable has no biological significance; however, it can be considered a voltage gate parameter.  $\beta$  and  $\sigma$  are the external DC current parameters applied to the neuron, and  $\alpha$  is the nonlinearity parameter of the neuron map. Several behaviors such as firing and bursting can be obtained by different combinations of the  $\sigma$  and  $\alpha$  parameters. Neural dynamics examples of this neuron map model, which are a few of these behaviors, are presented in Figures 2.a-e.

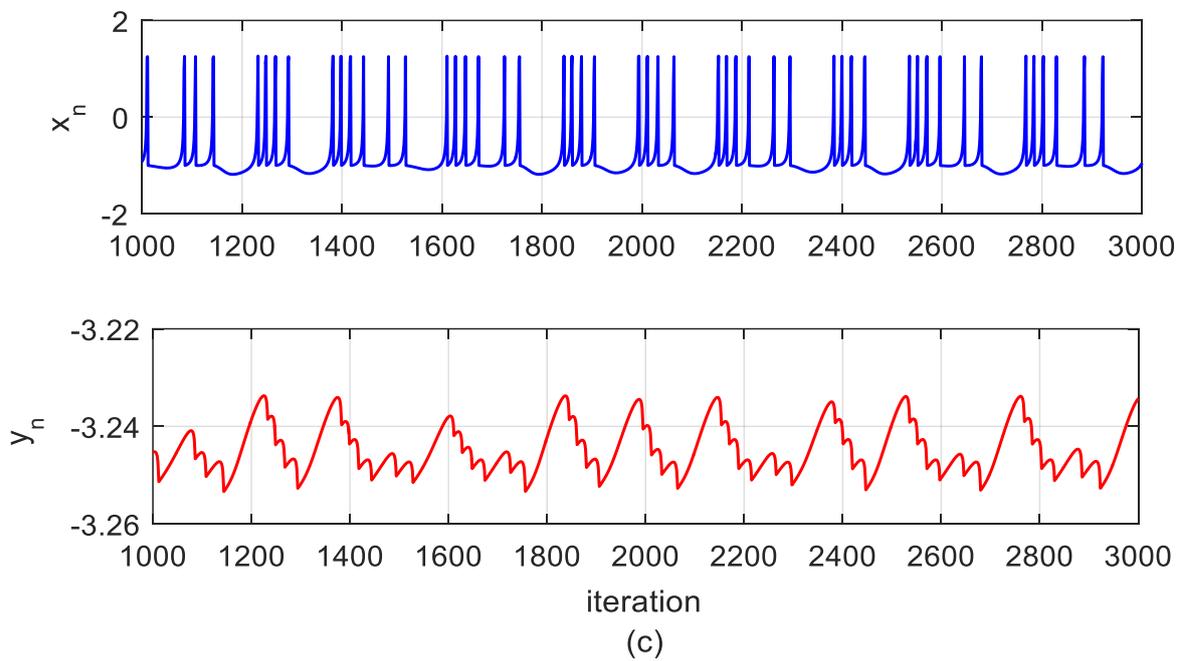
**Rulkov Neuron Map: Spike**


(a)

**Rulkov Neuron Map: Fast Spike**



**Rulkov Neuron Map: Irregular Behavior**



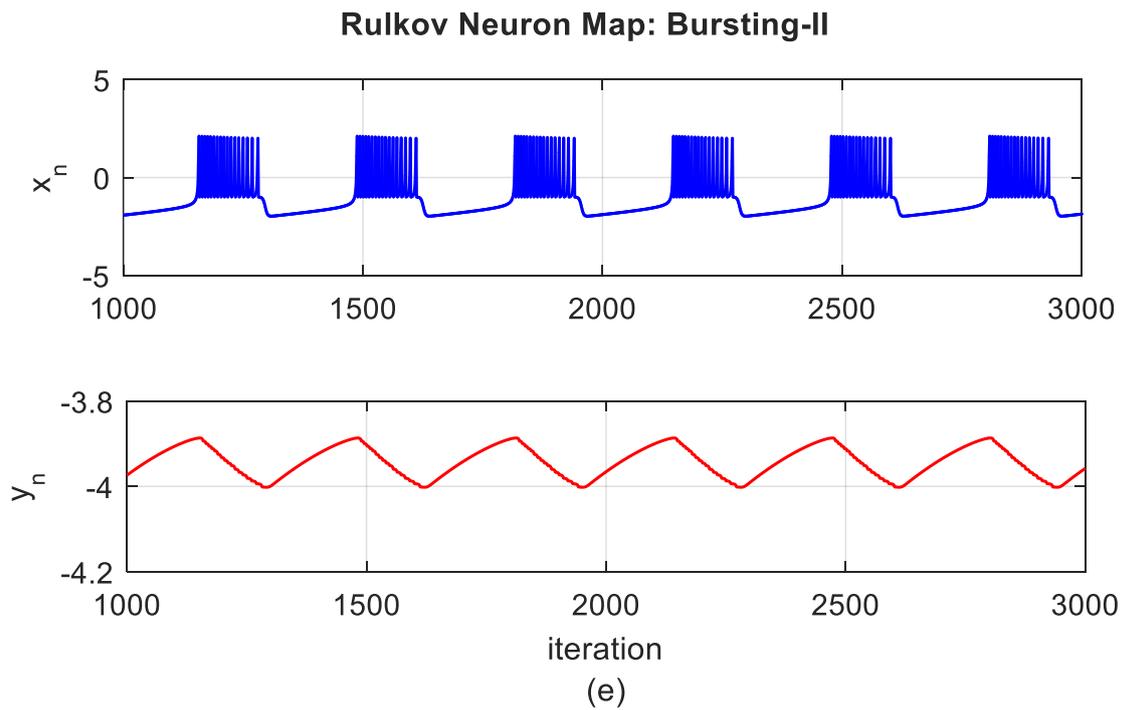
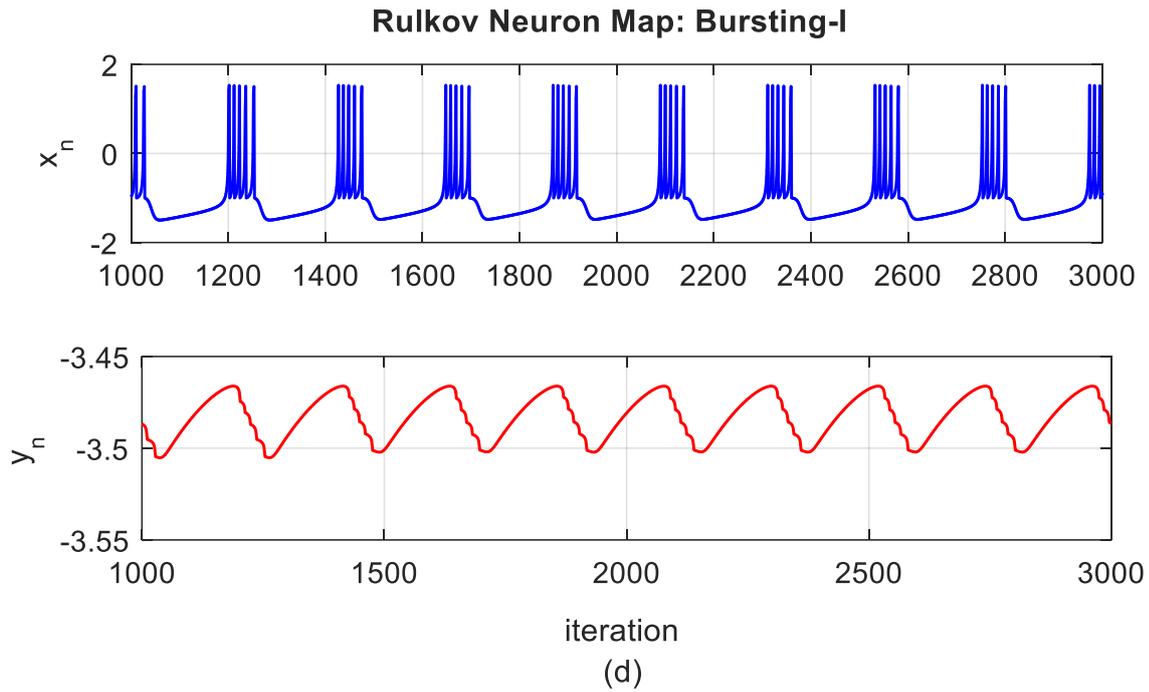


Figure 2: Examples of neural dynamics in the Rulkov neuron map model.

The numerical simulation results of in Figure 2a-e are obtained for five different neural dynamics of the Rulkov neuron map model. First one is for spiking behavior (Figure 2a), second one is for fast spiking behavior (Figure 2b), third one is for irregular behavior (Figure 2c), and fourth and fifth ones are for bursting behavior with different periods (Figures 2d and 2e). The values of the parameters, which are used in these numerical simulations, are presented in Table 1:

Table 1: The values of the parameters used in numerical simulations of the Rulkov neuron map model in Eq.1.

Parameters	Fig.2a	Fig.2b	Fig.2c	Fig.2d	Fig.2e
$\alpha$	4	4	4.5	5	6
$\beta$	0	0	0	0	0
$\mu$	0.001	0.001	0.001	0.001	0.001
$\sigma$	0.01	0.1	0.14	-0.1	-0.1

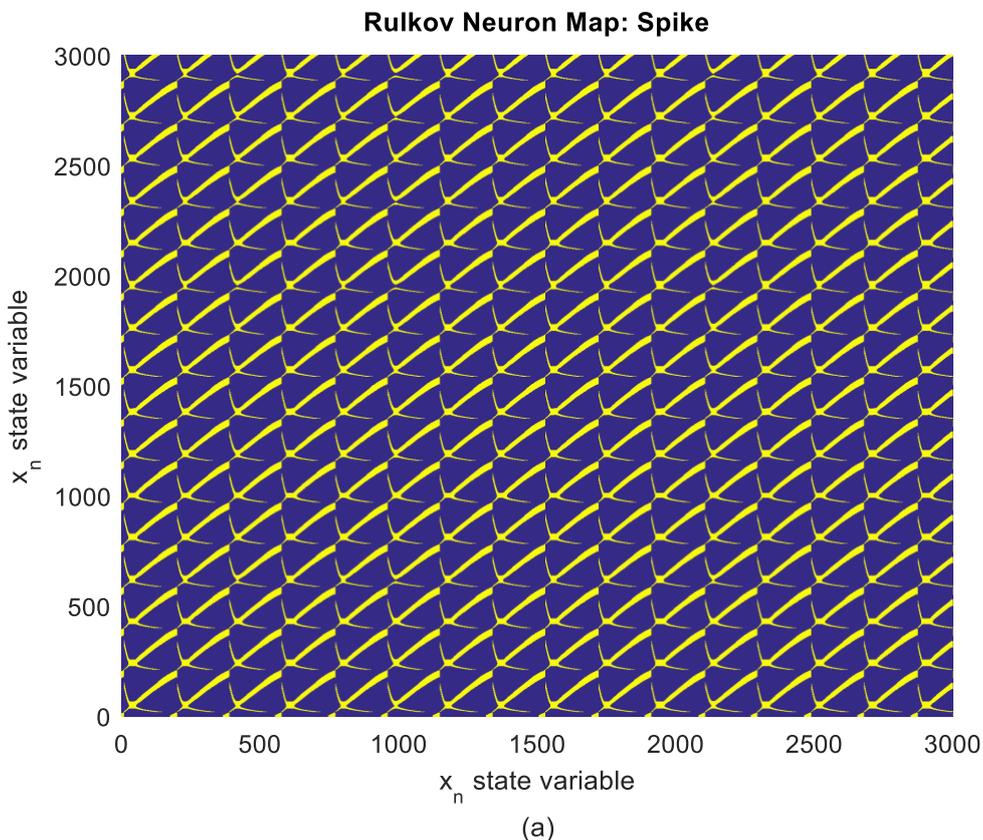
The Rulkov neuron map model in Eq.1 aims to mimic the firing and bursting patterns that constitute the rich dynamics of neurons with a two-dimensional map model. The simulation results in Figure 2 demonstrate and verify this diversity.

### III.RECURRENCE QUANTIFICATION ANALYSIS- ‘RQA’ METHOD

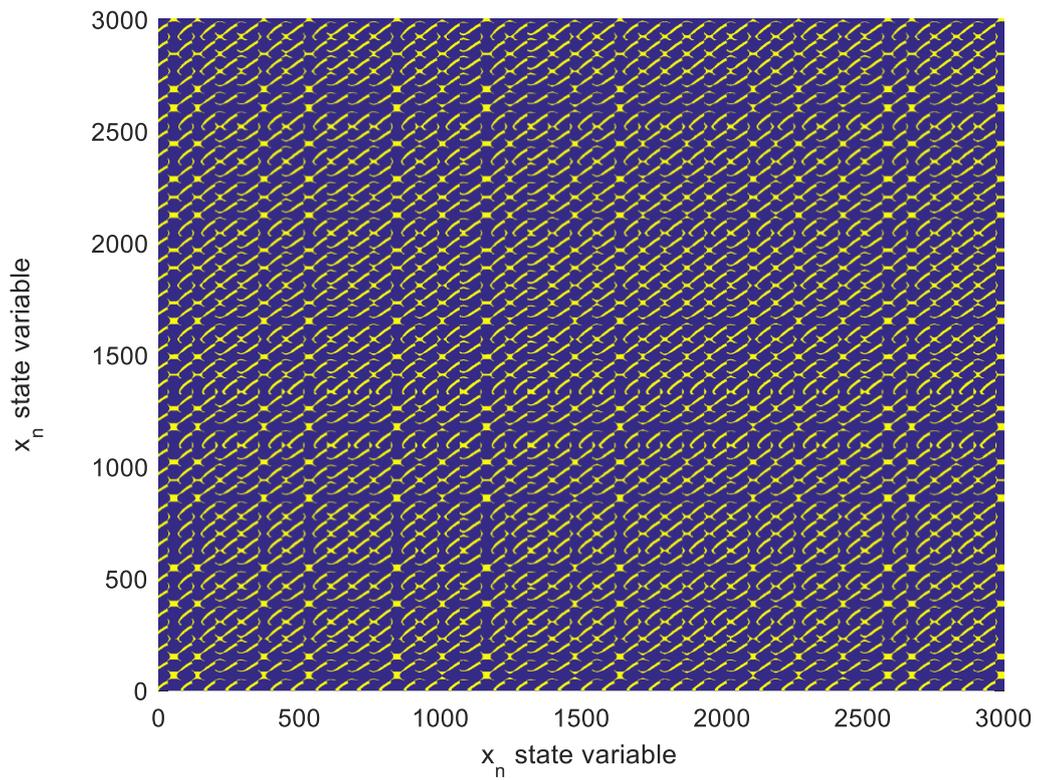
“Recurrence Quantification Analysis (RQA)” is a powerful analysis method that quantitatively examines recurrent structures and patterns in time series data of dynamic systems. It was first developed by Zbilut and Webber [17]. RQA creates two-dimensional matrices called “recurrence plots (RP)” to visualize the dynamic behavior of a time series. These graphs display the times when the system's states are similar to each other by considering a specific threshold value. When the system repeats similar situations within the time series, these repetitions are marked with points on the graph.

The RQA method is used for the purposes such as analyzing the structure of dynamic systems, distinguishing different dynamic modes, and analyzing noise effects. RQA calculates various quantities from recurrence graphs, and these quantities provide important information about the dynamics of the system. For example, the rate of recurrences in a time series is calculated using the “Recurrence Rate (RR)” ratio. The “Mean Linear Length (Determinism-DET)” is calculated to demonstrate the continuity of recurring patterns. The recurrence complexity of the patterns is measured using “Entropy (ENTR)”. An important criterion indicating whether a system is chaotic is the “Longest Diagonal Line (Lmax)” calculation.

Accordingly, in this study, the RP graphs are plotted for five different neural dynamics of the Rulkov neuron map model in Eq.1 and its variable parameters are set to the same values in Table 1. These RP graphs are shown in Figure 3.a-e and they are presented in the same order with Figures 2.a-e.

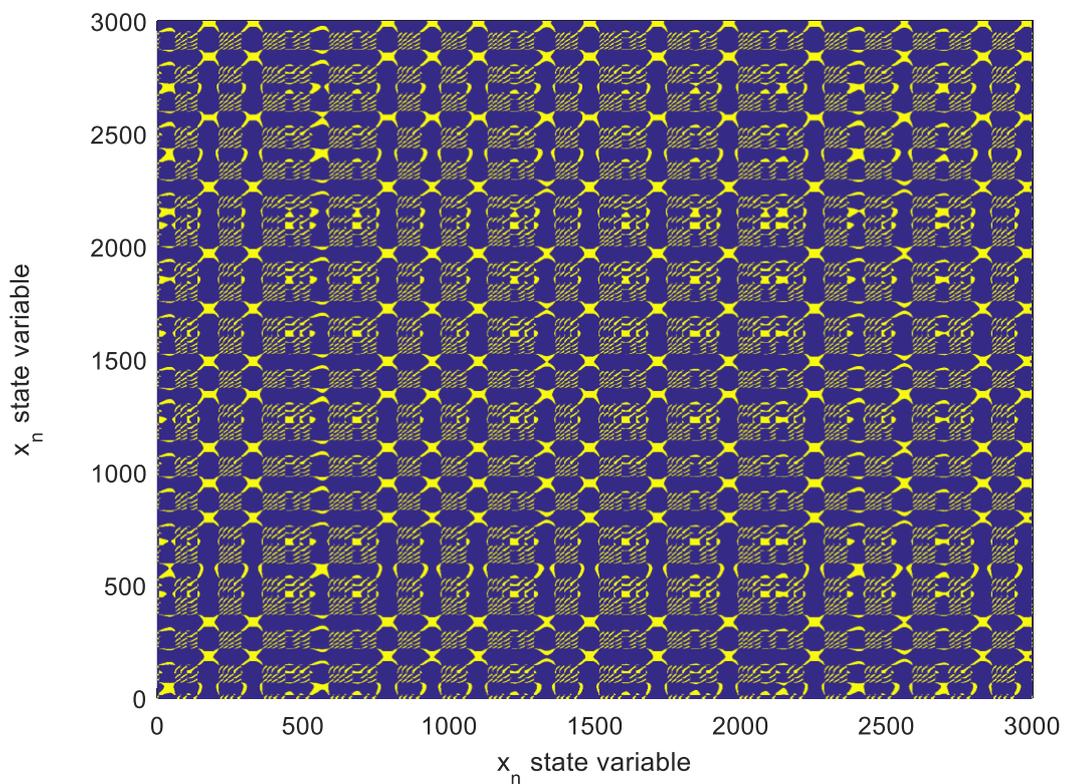


**Rulkov Neuron Map: Fast Spike**



(b)

**Rulkov Neuron Map: Irregular Behavior**



(c)

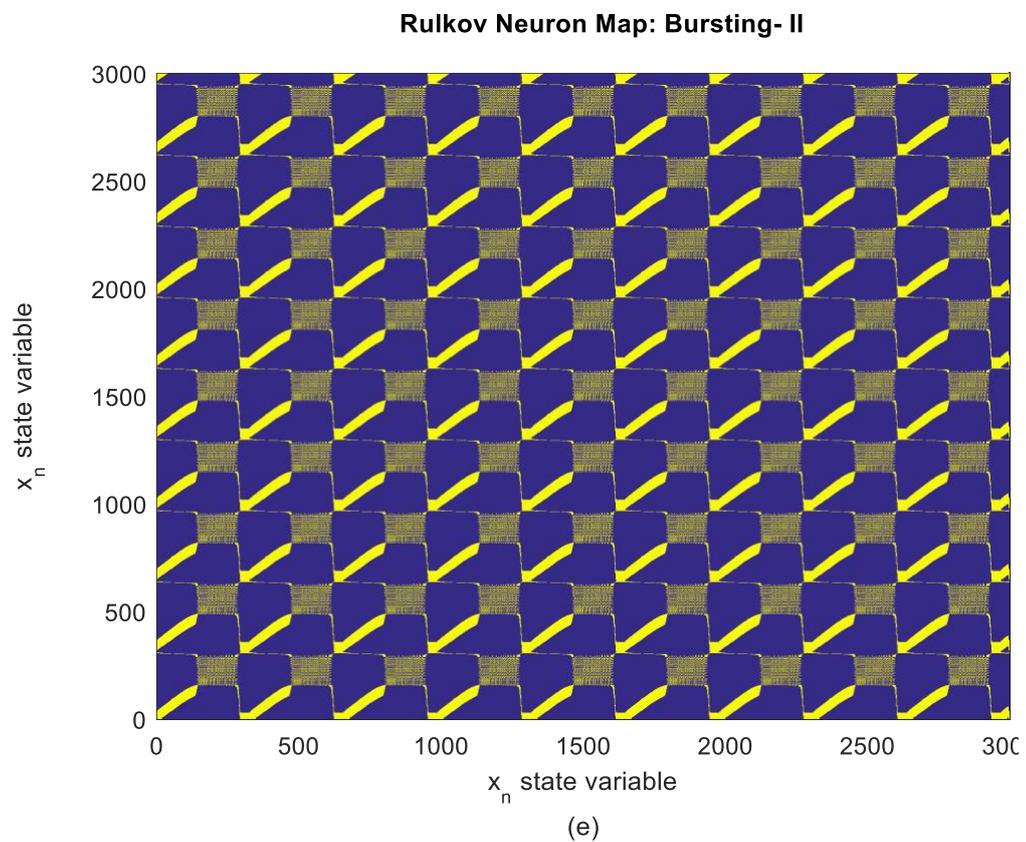
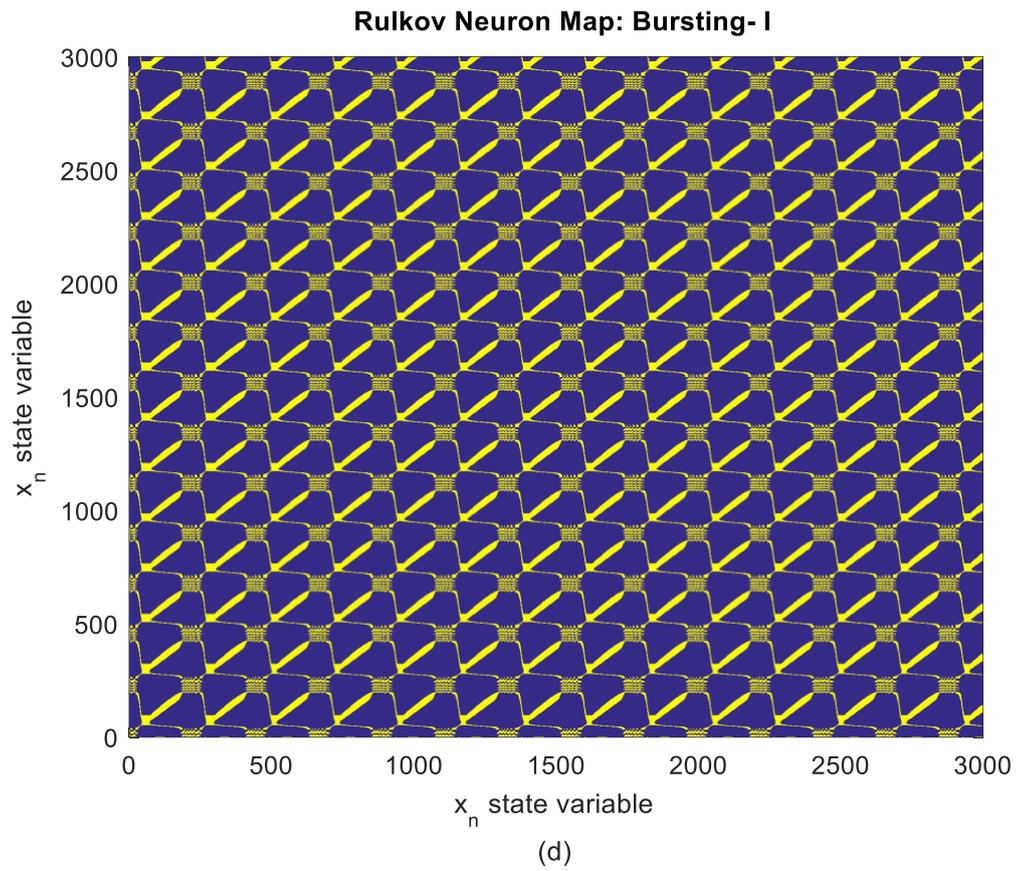


Figure 3: The RP graphs plotted for five different neural dynamics of the Rulkov neuron map model.

The above-mentioned RR, DET, ENTR and Lmax indication results for these dynamics are recorded in Table 2.

Table 2: RR, DET, ENTR and Lmax results for five different neural dynamics of the Rulkov neuron map model.

Parameters	Fig.3a	Fig.3b	Fig.3c	Fig.3d	Fig.3e
RR	0.1668	0.1668	0.1668	0.1668	0.1668
DET	0.9965	0.9897	0.9766	0.9918	0.9488
ENTR	2.7505	2.8773	2.7089	2.5581	1.7727
Lmax	20.8088	12.6022	11.2594	9.5371	10.6155

According to these results, the system behaves deterministically and regularly for their most parts, because the diagonal lines are shorter and the DET value is lower in chaotic systems. While the diagonal length is long in the spike response, any diagonal length is not observed in the irregular behavior. The RQA illustrations of bursting behaviors are different from each other and it depends on their periods and the number of spikes in the burst. The recurrence rate (RR) is at a moderate level, namely, this system repeats itself at regular intervals and there are approximately 16.68% of recurring states in each dynamic. High entropy indicates a more complex and disordered structure. The intermediate ENTR values in these results indicate that this system is neither completely simple nor completely chaotic. Longer diagonal lines indicate a more deterministic and stable structure. The Lmax values are at the moderate level. This suggests that there is some disorder in this system, but not a very strong chaotic structure.

#### IV. CONCLUSION

RQA is a method based on the quantitative analysis of recurring patterns in time-series data to understand the internal structure and behavior of dynamic systems. It is a powerful tool for revealing changes in complex and chaotic systems, distinguishing different dynamic states, and examining the continuity of the systems. In this study, the oscillation patterns of the Rulkov neuron map model have been quantitatively examined using recurrence-based analysis. It has been used the time series data of five different neural activity responses of this model. Thus, some concrete inferences have been made about the stability, complexity, and predictability of the system under consideration. Accordingly, the recurrence rates of these five different neural activity responses of the Rulkov neuron map model have remained unchanged. This result indicates that the recurrence rate is similar in basic states of this system. The regular and recurring patterns are available in all five dynamics, and the dynamics exhibit a certain order of complexity.

#### ACKNOWLEDGMENT

This study has been supported by The Scientific and Technological Research Council of Turkey (TUBITAK) under project number 124E560.

#### REFERENCES

- [1] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *The Journal of Physiology*, vol. 117(4), pp. 500-544, 1952.
- [2] R. FitzHugh, "Mathematical models for excitation and propagation in nerve," In: Schawm, H.P. (ed.) *Biological Engineering*, vol. 1, pp. 1-85. McGraw-Hill, New York, 1969.
- [3] H. R. Wilson and J. D. Cowan, "Excitatory and inhibitory interactions in localized populations of model neurons," *Biophysical Journal*, vol. 12 (1), pp. 1-24, 1972.
- [4] J. L. Hindmarsh and R. M. Rose, "A model of neural bursting using three coupled first order differential equations," *Proceedings of the Royal Society B: Biological Sciences*, vol. 221(1222), pp. 87-102, 1984.
- [5] E. M. Izhikevich, "Simple model of spiking neurons," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 14(6), pp. 1569-1572, 2003.
- [6] N. F. Rulkov, "Modeling of spiking-bursting neural behavior using two-dimensional map," *Physical Review E*, vol. 65(4), 041922, 2002.
- [7] Z. T. Njitiacke, C. N. Takembo, G. Sani, N. Marwan, R. Yamapi, and J. Awrejcewicz, "Hidden and self-excited firing activities of an improved Rulkov neuron, and its application in information patterns," *Nonlinear Dynamics*, vol. 112, pp. 13503-13517, 2024.
- [8] J. Li, and H. Cao, "Different characteristics and important channels between the healthy brain network and the epileptic brain network based on EEG data," *Communications in Nonlinear Science and Numerical Simulation*, vol. 66, pp. 147-155, 2019.
- [9] J. M. Susedo-Solorio and A. N. Pisarchik, "Synchronization of map-based neurons with memory and synaptic delay," *Physics Letters A*, vol. 378(30-31), pp. 2108-2112, 2014.

- [10] I. Ghosh, A.S. Nair, H.O. Fatoyinbo, and S. S. Muni, "Dynamical properties of a small heterogeneous chain network of neurons in discrete time," *The European Physical Journal Plus*, vol. 139(6), pp. 545, 2024.
- [11] S. Mirzaei, M. Mehrabbeik, K. Rajagopal, S. Jafari, and G. Chen, "Synchronization of a higher-order network of Rulkov maps," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 32(12), 2022.
- [12] N. F. Rulkov, I. Timofeev, and M. Bazhenov, "Oscillations in large-scale cortical networks: map-based model," *Journal of computational neuroscience*, vol. 17(2), pp. 203-223, 2004.
- [13] J. López, M. Coccolo, R. Capeáns, and M.A. Sanjuán, "Controlling the bursting size in the two-dimensional Rulkov model," *Communications in Nonlinear Science and Numerical Simulation*, vol. 120, 107184, 2023.
- [14] F. Min, G. Zhai, S. Yin, and J. Zhong, "Switching bifurcation of a Rulkov neuron system with ReLu-type memristor," *Nonlinear Dynamics*, vol. 112(7), pp. 5687-5706, 2024.
- [15] M. SalehiYekta, A. ZamaniBahabadi, and G. SadeghiBajestani, "An analytic investigation of Hopf bifurcation location control for the Rulkov map model," *International Journal of Bifurcation and Chaos*, vol. 33(12), 2330029, 2023.
- [16] I. Bashkirtseva, and L. Ryashko, "Dynamical variability, order-chaos transitions, and bursting Canards in the memristive Rulkov neuron model," *Chaos, Solitons & Fractals*, vol. 186, 115317, 2024.
- [17] J. P. Zbilut and C. L. Webber Jr, "Embeddings and delays as derived from quantification of recurrence plots" *Physics letters A*, vol. 171(3-4), pp. 199-203, 1992.
- [18] N. Marwan, M. C. Romano, M. Thiel, and J. Kurths, "Recurrence plots for the analysis of complex systems," *Physics reports*, vol. 438(5-6), pp. 237-329, 2007.

## RESUME

### **Muhammet Mustafa ASLAN**

Muhammet Mustafa ASLAN graduated from Gaziantep Vedat Topçuoğlu Anatolian High School in 2022. He was accepted to Kayseri University Electrical and Electronics Engineering Department in 2022. He completed his first internship at ATEŞ Company. Then, he completed his second internship at ESI Company. Currently, He continues her education in the 4th grade. He has worked on neuron models and verification of nonlinear systems with complex dynamics using microcontrollers.

### **Ferit DURAN**

Ferit DURAN graduated from Hatay Mozaik College in 2022. He was accepted to Kayseri University Electrical and Electronics Engineering Department in 2022. He completed his first internship in LIMAK Construction Kuwait. He completed his second internship at GER Engineering. He currently continues his education in the 4<sup>th</sup> grade. He has studied about neuron models and verification of nonlinear systems with complex dynamics using microcontrollers.

### **Nimet KORKMAZ**

Nimet KORKMAZ received BSc degree from the Department of Electrical and Electronics Engineering, Erciyes University, Turkey in 2010. She received MSc and PhD degrees in Electrical and Electronics Engineering from Erciyes University Graduate School of Natural and Applied Sciences in 2012 and 2018, respectively. She is currently an associate professor doctor at Department of Electrical and Electronics Engineering, Kayseri University, Turkey. Her research interests include neuron models, neuromorphic engineering, central pattern generators, and field programmable analog and gate arrays based implementations.

# Data Driven Multivariate Mapping of Camelina Based Diesel–Biodiesel–Ethanol Blends: Exploring the Principal Axes of Physicochemical Variation

Mehmet Selman GOKMEN<sup>1</sup>, Hasan AYDOGAN<sup>2</sup>

<sup>1</sup>Necmettin Erbakan University, Seydişehir Vocational School, Automotive Technology. Prog., Konya, Türkiye, (ORCID: 0000-0001-5943-7504), [msgokmen@erbakan.edu.tr](mailto:msgokmen@erbakan.edu.tr)

<sup>2</sup>Selçuk University, Faculty of Technology, Mechanical Engineering Department, Konya, Türkiye, (ORCID: 0000-0003-1404-6352), [haydogan@selcuk.edu.tr](mailto:haydogan@selcuk.edu.tr)

## I. INTRODUCTION

The primary motivations driving alternative fuel research on a global scale are the reduction of greenhouse gas emissions and the achievement of energy independence by reducing dependence on volatile petroleum markets. These two driving forces have consistently shaped the direction of energy research over the past few decades, as societies increasingly recognize the intertwined challenges of climate change and fossil fuel depletion. Numerous studies have demonstrated a strong consensus between the environmental and economic imperatives of these objectives, emphasizing that sustainability in energy systems requires an integrated approach that encompasses both ecological and geopolitical dimensions.

Environmental motivation primarily stems from the urgent need to combat climate change, highlighting the potential of biodiesel and bioethanol as carbon dioxide–neutral fuels that can significantly lower greenhouse gas emissions compared with fossil fuels [1], [2]. In addition to their carbon mitigation benefits, biofuels also contribute to reduced emissions of particulate matter, sulfur oxides, and unburned hydrocarbons, making them attractive options for improving urban air quality. From an energy security perspective, the major concerns include fossil fuel depletion, reducing dependency on imported petroleum, and mitigating the vulnerability of the transportation sector to oil market volatility [3]. These issues have become more pronounced due to global energy price fluctuations and geopolitical disruptions, underscoring the importance of resilient domestic fuel alternatives.

Given that the transportation sector accounts for approximately 60% of the total petroleum consumption, it remains a primary target for energy transition. Within this sector, diesel fuel plays a central role in freight, agricultural, and industrial operations, making its substitution or partial replacement a key area of research. In this context, biodiesel–ethanol blends combine the complementary advantages of both fuels: biodiesel provides lubrication and oxygenation, whereas ethanol enhances the volatility and reduces the viscosity. Together, they offer feasible and environmentally sustainable alternatives to petroleum diesel. The consistency of the findings across multiple studies conducted over the years underscores the persistent importance of these two central motivations in alternative fuel research—environmental stewardship and energy autonomy—over time.

Camelina sativa has significant agronomic and chemical advantages as a biodiesel feedstock. Agronomically, it can be cultivated in semi-arid regions with low input requirements, exhibits strong tolerance to cold and drought, and maintains high yields under marginal conditions where other oilseeds struggle [4], [5], [6]. These characteristics make Camelina particularly suitable for sustainable agricultural systems, as it requires minimal irrigation and fertilization. Camelina oil contains 30–48% oil and approximately 90% unsaturated fatty acids, with  $\alpha$ -linolenic acid (30–40%) contributing to the improved cold-flow properties of the fuel [7], [8]. Compared with other oilseeds, Camelina stands out because of its high net energy ratio, low production cost, and non-competitive nature with food production [9], [10]. Several studies have shown that camelina oil demonstrates superior performance as a biofuel feedstock relative to other vegetable oils [6]. In addition, its compatibility with existing agricultural rotations and potential for carbon sequestration reinforce its attractiveness for sustainable biofuel production.

Recent research has demonstrated that diesel–biodiesel–ethanol ternary blends exhibit marked variations in their physicochemical and combustion characteristics. The synergistic interactions between the three components influence critical parameters such as density, viscosity, flash point, and cetane number of the fuel. Some studies have reported that ethanol addition reduces the density and viscosity of diesel, improving atomization and combustion efficiency, although it may negatively affect other physical parameters, such as lubricity and cold stability [11]. Conversely, combustion-oriented studies have generally reported favorable outcomes, particularly in terms of emission reduction and combustion completeness. Karin et al. (2022) and Phyo Wai et al. (2022) observed over 50% reductions in smoke and particulate emissions, while Sujeet Kesharvani et al. (2022) reported a 26.4% reduction in particulate matter emissions in ethanol-enriched blends [12], [13], [14]. These results highlight the environmental advantages of using oxygenated fuel blends. However, most existing studies have focused on narrow blending ratios (5–20% biodiesel and 5–10% ethanol), used different feedstocks and engine configurations, and lacked standardized testing protocols [15]. Moreover, the absence of long-term durability tests and commercial-scale evaluations limits the generalizability of these findings for practical application.

Multivariate analysis approaches are of great significance in the evaluation of biofuel quality parameters, as they enable the simultaneous examination of multiple interdependent variables, identification of the most critical quality factors, and effective classification of biofuel feedstocks [16], [17], [18]. In the literature, Principal Component Analysis (PCA) and cluster analysis have been successfully applied in diverse fields such as genetic selection, engine performance correlation, biodiesel proportion determination [19], biomass gasification [20], marine algae potential assessment [21], and biorefinery process variability [22]. These applications demonstrate that multivariate statistical tools allow for a deeper understanding of complex datasets and facilitate the identification of the dominant factors governing system behavior. Collectively, these findings confirm that PCA is a robust and generalizable tool for assessing the quality of biofuels and optimizing their formulation parameters.

Nevertheless, comprehensive multivariate characterization studies specifically focusing on Camelina-based diesel–biodiesel–ethanol blends are absent from the literature. Although PCA, FTIR spectroscopy, and discriminant analysis have been successfully applied to biodiesel blends [23], [24], studies on Camelina have largely remained limited to fuel property evaluation without advancing to multiparameter statistical modeling [25], [26], [27]. This gap highlights the need for integrative statistical frameworks capable of elucidating the correlations among physical, chemical, and combustion-related variables. Additionally, the available open-access datasets primarily focus on engine emission studies, restricting their use in broader physicochemical analyses [28]. Thus, the comprehensive characterization of the physicochemical parameters of camelina-based blends through multivariate methods requires the development of standardized open datasets that can support reproducibility and cross-study comparison.

The existing literature highlights the scarcity of open-access databases for liquid biofuel blends. The most extensive examples include the NIST Standard Reference Databases, ThermoData Engine, and NIST Reference Fluid Thermodynamic and Transport Properties Database, which have been expanded to cover biofuel components [29]. Furthermore, Comesaña et al. (2022) developed models based on publicly available experimental data for 8,351 molecules [30], providing an open-source web tool for predicting the melting, boiling, and flash points. The dataset published by [31], which contains 612 data points from 204 biochar samples, provides detailed fuel characteristics but focuses mainly on solid biomass. Most other studies report only experimental results for specific blends without sharing open data. The lack of structured, openly available physicochemical datasets remains a critical barrier to data-driven innovations in biofuel research.

Accordingly, this study aimed to comprehensively assess the physicochemical and performance parameters of camelina-based diesel–biodiesel–ethanol blends, identify the dominant variance axes, and visualize the relationships among the formulations using PCA. This study establishes a statistical framework that highlights the distinctions among fuel formulations, supports the interpretation of underlying property correlations, and offers a quantitative reference for future optimization and sustainability studies.

## II. MATERIAL AND METHOD

The dataset used in this study comprised six distinct fuel samples prepared by blending biodiesel derived from Camelina sativa oil with conventional diesel and ethanol in varying proportions: B20, B50, B100, B20D65E15, B50D35E15, and B85D0E15. These formulations were specifically chosen to represent a wide range of biodiesel and ethanol substitution ratios, thereby enabling a systematic assessment of the effects of composition on the physicochemical and combustion-related properties. The dataset was obtained from the open-access repository Zenodo [32], which ensures full transparency, reproducibility, and accessibility of the underlying experimental data for future comparative analysis.

For each blend, 15 key physicochemical properties were evaluated: carbon residue, sulfur content, cetane number, flash point, total contamination, kinematic viscosity at 40 °C, cold filter plugging point (CFPP), water content, density at 15 °C, iodine number, oxidation stability at 110 °C, net calorific value, pour point, phosphorus content, and Group II metals (Ca + Mg). These parameters collectively represent the most critical indicators of the fuel's physical behavior (e.g., density, viscosity, and CFPP), combustion performance (e.g., cetane number, calorific value, and carbon residue), and chemical stability (e.g., oxidation stability, iodine number, and metallic contamination).

All analyses were carefully selected to enable a comprehensive evaluation of the fuels in terms of their suitability for engine operation, emissions performance, and long-term storage stability. This multidimensional dataset provides a foundation for advanced statistical exploration, allowing the identification of correlated properties and the extraction of dominant patterns governing the blend behavior. This facilitated a deeper understanding of the interdependence among physicochemical parameters and supported the development of a holistic, data-driven approach to biofuel quality characterization.

Prior to the statistical analysis, all variables were standardized to eliminate scale differences arising from variations in the measurement units and magnitude ranges. This preprocessing step was essential to ensure that parameters with inherently larger numerical values, such as density or viscosity, did not dominate the analysis and that all variables contributed equally to the extraction of the principal components.

Standardization was performed by transforming each variable to have a mean value of zero and a standard deviation of one, according to the following equation:

$$Z_{ij} = \frac{X_{ij} - \bar{X}_j}{S_j}$$

Here,  $X_{ij}$  represents the  $i$ -th observation of the  $j$ -th variable,  $\bar{X}_j$  denotes the mean, and  $S_j$  is the standard deviation of variable  $j$ . Through this transformation, each parameter was expressed in terms of the standard deviation from its mean, thereby normalizing the influence of individual measurement scales and allowing for balanced weighting across all physicochemical properties.

This standardization procedure also facilitated the identification of hidden correlations among variables, a prerequisite for multivariate techniques such as Principal Component Analysis (PCA). Without this normalization, high-variance variables can distort the covariance structure and bias the interpretation of the principal axes.

Missing data were handled using the pairwise deletion method, which retains all available observations for each pair of variables to maximize information usage while minimizing the distortion in correlation estimation. As a result of this preprocessing procedure, a complete dataset containing all the standardized variables across the six fuel formulations was obtained. This prepared data matrix constituted the statistical foundation for subsequent correlation analysis and PCA, ensuring consistency, comparability, and robustness in the multivariate evaluation.

To identify linear relationships among the physicochemical variables, the Pearson product-moment correlation coefficient was calculated for each pair of parameters. This statistical measure quantifies the strength and direction of the linear association between two continuous variables, providing insight into how changes in one property are associated with changes in the other. The correlation between two variables,  $X$  and  $Y$  defined as follows:

$$r_{XY} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

where  $X_i$  and  $Y_i$  represent the  $i$ -th observations of variables  $X$  and  $Y$ , respectively;  $\bar{X}$  and  $\bar{Y}$  are their mean values; and  $n$  denotes the number of paired observations. The correlation coefficient ranges from  $-1$  to  $+1$ , with values close to  $+1$  indicating a strong positive relationship, values near  $-1$  indicating a strong negative relationship, and values around zero suggesting no linear correlation.

The resulting correlation matrix provides a compact and comprehensive view of all pairwise relationships among the 15 measured parameters, effectively capturing the dataset's internal structure. As shown in Figure 1, the correlation matrix reveals the overall pattern and directional tendencies of the studied variables.

According to the analysis results, a strong positive correlation was observed among the density, iodine number, and viscosity, indicating that these parameters collectively describe the physical characteristics of the fuel blends and tend to vary in the same direction. In practical terms, higher viscosity is associated with increased density and iodine number, reflecting the greater molecular complexity and unsaturation of biodiesel-rich blends.

In contrast, parameters such as cetane number and oxidation stability, which are representative of combustion quality and resistance to degradation, exhibited an inverse trend with respect to this group of additives. This negative relationship indicates that as fuels become denser and more viscous, their combustion stability and ignition quality may decline, emphasizing the trade-off between the physical and combustion-related properties.

The presence of both strong positive and negative correlations among multiple parameters confirmed that the dataset possessed a multivariate structure. Such interdependencies justify the application of Principal Component Analysis (PCA), which can effectively condense these relationships into a smaller set of orthogonal components for clearer interpretation and visualization of the underlying data patterns.

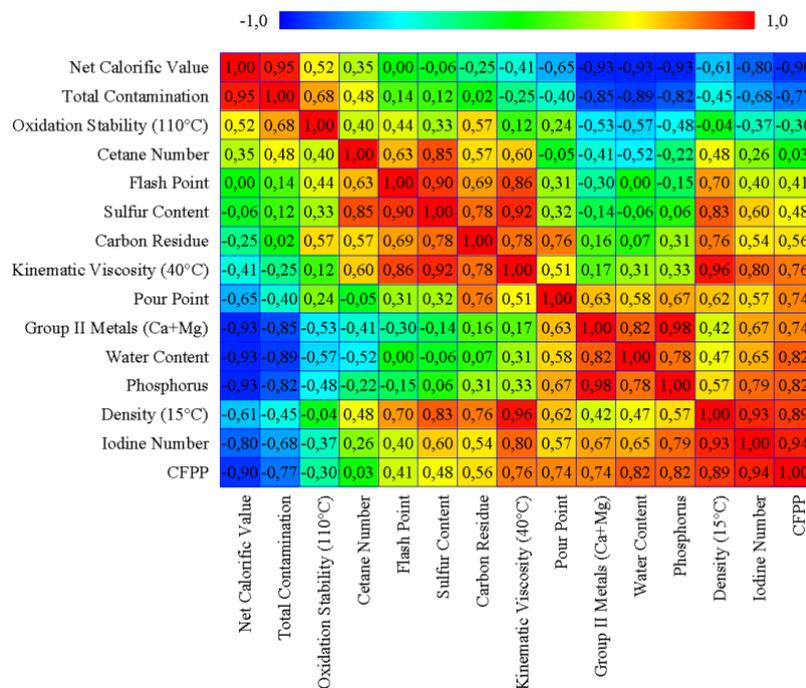


Figure 1: Pearson product-moment correlation matrix showing the interrelationships among 15 physicochemical variables of Camelina-based diesel-biodiesel-ethanol blends.

Principal Component Analysis (PCA) was applied to reduce the interrelationships among the variables in the dataset and obtain simpler and more meaningful structures that capture the dominant patterns of variation. As a dimension-reduction technique, PCA transforms a large set of correlated variables into a smaller number of uncorrelated variables, called principal components, while retaining most of the original information content. This approach is particularly useful for multivariate fuel characterization, where multiple physicochemical parameters may covary owing to compositional and thermal interactions. Based on the standardized data matrix  $Z$ , the correlation matrix  $R$  was computed, followed by the eigenvalue ( $\lambda$ ) and eigenvector ( $a$ ) decomposition, as expressed below:

$$R a_k = \lambda_k a_k$$

Here,  $\lambda_k$  represents the amount of variance explained by the  $k$ -th component, while  $a_k$  denotes the variable loadings (weights) associated with that component, indicating how strongly each original variable contributes to it. The eigenvalues quantify the significance of each principal component, with larger eigenvalues corresponding to components that capture more of the overall variability of the dataset. The component scores for each observation were calculated as follows:

$$y_{ik} = \sum_{j=1}^p a_{jk} z_{ij}$$

This equation expresses the component scores as weighted linear combinations of the standardized values of the original variables. Each score  $y_{ik}$  represents the projection of an observation onto a principal axis, effectively summarizing multidimensional information within a reduced coordinate space.

Through this process, the dataset's complexity was condensed into a limited number of orthogonal axes (principal components), each representing a unique dimension of variation. This transformation allowed for the visualization and interpretation of the differences between fuel samples based on their most influential physicochemical characteristics. By focusing on the components with the largest eigenvalues, PCA provided a quantitative and visual framework for identifying the main trends, correlations, and divergence patterns among the camelina-based diesel–biodiesel–ethanol blends.

The obtained eigenvalues indicate the relative contribution of each principal component to the total variance of the dataset. Each eigenvalue ( $\lambda_k$ ) quantifies the amount of information (or variance) captured by the corresponding principal component, thereby serving as a criterion for determining its statistical significance and interpretive value.

In this study, the Kaiser criterion was employed as the selection rule, which states that components with eigenvalues ( $\lambda$ ) greater than or equal to one ( $\lambda \geq 1$ ) should be retained for interpretation. This threshold ensures that each retained component explains at least as much variance as one of the original standardized variables, making it a rational cutoff for a meaningful dimensional reduction.

As a result of the analysis, three principal components satisfied this condition and collectively accounted for 94.94% of the total variance in the dataset. This remarkably high cumulative variance indicates that nearly all the essential information contained in the 15 physicochemical variables can be effectively summarized by only three orthogonal components. The remaining components, which exhibited eigenvalues lower than one, primarily represented random variation or noise and were therefore excluded from further interpretation.

The first principal component (PC1) captured the largest portion of the dataset variance, highlighting the dominance of the physical property group (such as viscosity, density, and iodine number), whereas the second component (PC2) emphasized the combustion and oxidative stability characteristics of the blends. The third component (PC3) contributed marginally but was retained to preserve the completeness of the thermal and stability-related behaviors of the fuels.

The distribution of the eigenvalues across the components is illustrated in Figure 2. The scree plot clearly demonstrates a steep decline after the third component, forming an “elbow” that visually confirms the Kaiser criterion. This pattern suggests that the first three components are sufficient to represent the multivariate structure of the dataset without significant information loss, thereby validating the dimensional reduction achieved using PCA.

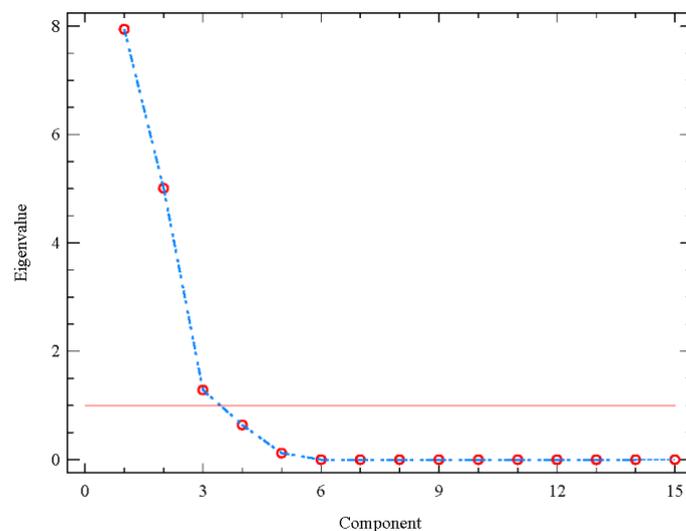


Figure 2: Scree plot displaying the eigenvalue distribution and cumulative variance explained by principal components for the Camelina-based fuel dataset

The PCA results were visualized and interpreted using biplots, which provide a two-dimensional graphical representation of the multivariate relationships in the dataset. A biplot simultaneously displays both variable vectors (loadings) and observation points (fuel blends) within the same coordinate system, allowing the direct visualization of how each variable contributes to the principal components and how the fuel samples are positioned relative to one another in the reduced component space.

In a PCA biplot, the direction and length of the vectors indicate the influence and correlation of each variable with the principal component. Vectors that extend in similar directions correspond to variables that are positively correlated, suggesting that they vary together across samples. Conversely, vectors pointing in opposite directions indicate negative correlations, implying that as one variable increases, the other tends to decrease. Thus, the angular separation between the vectors conveys the degree and type of association between the variables.

In this study, the biplot revealed that density, iodine number, viscosity, and CFPP were closely aligned and clustered along the PC1 axis, representing a dimension dominated by the physical properties of the fuel. These parameters collectively define the structural and flow behaviors of the blends and are characteristic of biodiesel-rich samples. In contrast, the cetane number, flash point, and oxidation stability were oriented along the PC2 axis, which corresponds to the combustion and stability characteristics of the fuels. This orthogonal distribution between PC1 and PC2 reflects the trade-off between physical and combustion-related parameters; as the blends become lighter and less viscous (ethanol-rich), their combustion stability may slightly diminish.

Therefore, PC1 can be interpreted as the physical structure axis, and PC2 represents the combustion stability axis. The biplot configuration not only illustrates these contrasting behavioral domains but also highlights the relative positions of the six fuel blends. Samples with higher biodiesel content tended to cluster in regions dominated by physical parameters, whereas those with ethanol enrichment occupied areas associated with improved flow properties but reduced oxidative resilience.

Overall, the component space derived from PCA effectively summarizes the multivariate data structure in reduced dimensions, offering both quantitative and visual insights into the interdependencies among the fuel properties. This graphical representation enables an intuitive interpretation of the similarities, divergence directions, and key quality indicators among the blends. The findings confirm that multivariate data analysis, particularly PCA, is a robust and versatile tool for the systematic characterization and optimization of biofuel formulations, enabling researchers to identify property groupings and performance trends that are not easily observable through univariate methods.

### III. RESULTS AND DISCUSSION

The eigenvalues and corresponding variance ratios obtained from the PCA clearly demonstrated that the variables in the dataset could be efficiently condensed into three main components, each representing a distinct dimension of the physicochemical behavior of fuel blends. The first principal component (PC1) explained 52.96% of the total variance, capturing most of the systematic variation associated with the physical characteristics of the fuels. The second component (PC2) accounted for 33.40% of the variance and primarily represented combustion-related parameters and oxidative stability. The third component (PC3) contributed 8.57% to the total variance, reflecting residual information mainly linked to low-temperature and thermal stability.

Together, these three principal components represented 94.94% of the total variance in the dataset, indicating an exceptionally high dimensional efficiency. This means that almost all the information originally distributed across the 15 correlated variables could be summarized using only three independent axes, without substantial loss of interpretive accuracy. Such a high cumulative variance ratio reinforces the robustness of the PCA model and validates its suitability for describing the multidimensional structure of the data.

The strong explanatory power of PC1 and PC2 suggests that the dominant patterns of variation among the Camelina-based diesel–biodiesel–ethanol blends are primarily governed by the interplay between the physical fluid properties (such as viscosity, density, and CFPP) and combustion stability characteristics (such as cetane number, flash point, and oxidation stability). Because these two axes together account for more than 86% of the total variance, subsequent interpretation was focused on the PC1–PC2 plane, which provides the clearest visualization of the inter-sample differences and variable groupings.

The scree plot (Figure 2) illustrates the descending order of eigenvalues and confirms that the first three components capture the essential structure of the dataset used in this study. The sharp decline in eigenvalues after PC3 and the near-plateau observed beyond this point indicate that additional components contribute negligibly to the overall variance of the data. This “elbow” pattern is characteristic of well-structured datasets, affirming that the retained components are sufficient to represent the underlying relationships among the physicochemical properties of fuel blends.

These results suggest that PCA successfully identified a set of orthogonal latent variables that efficiently explained the correlations inherent in the biofuel characteristics. Thus, the analysis provides a statistically grounded framework for linking physical, chemical, and stability parameters within a reduced-dimensional space, facilitating the comprehensive assessment and classification of camelina-based fuel formulations.

The examination of the principal component loadings (Table 1) revealed that the variables contributed to the PCA axes with different magnitudes and directions, reflecting the complex interdependencies among the physicochemical properties of the Camelina-based blends. Loadings express the correlation between each original variable and the derived components; thus, their magnitude and sign are key to understanding how each factor influences the principal axis and overall data structure.

The first principal component (PC1) exhibited strong negative loadings for several key variables, including the following:

- CFPP (−0.353)
- Iodine number (−0.338)
- Density (−0.315)
- Water content (−0.280)
- Group II metals (−0.274)
- Kinematic viscosity (−0.265)

Conversely, positive loadings were observed for net calorific value (+0.320) and total contamination (+0.269).

This pattern indicates that PC1 primarily represents the physical property axis of the fuel, integrating attributes related to fluidity, cold flow performance, and compositional structure. The negative correlations between CFPP, viscosity, and density suggest that as the biodiesel concentration increases, these parameters increase concurrently, indicating thicker and denser fuels with poorer cold-flow behavior. Similarly, the iodine number, which represents the level of unsaturation, negatively aligns with these physical characteristics, highlighting the influence of fatty acid composition on fuel rheology. Collectively, these relationships define PC1 as a comprehensive descriptor of the density–viscosity coupling and low-temperature operability, parameters that dominate the physical performance of biodiesel–ethanol blends.

The second principal component (PC2) was largely governed by variables associated with combustion behavior and oxidative stability, with the highest positive loadings observed for

- Cetane number (+0.394)
- Flash point (+0.372)
- Oxidation stability (+0.314)
- Carbon residue (+0.308)
- Sulfur content (+0.381)

These parameters are directly related to the ignition quality, combustion efficiency, and safety performance of the fuels. Therefore, PC2 can be interpreted as the combustion–stability axis, which summarizes the thermochemical behavior of the blends. The alignment of the cetane number and flash point with the oxidation stability suggests that fuels exhibiting more controlled ignition and higher resistance to oxidation tend to form a coherent subgroup within the PCA space. The sulfur and carbon residue contributions, though secondary, provide insight into residual impurities and fuel cleanliness—factors relevant to engine durability and emission characteristics.

The third component (PC3), though contributing less variance, still holds interpretive significance. The highest loadings for oxidation stability (+0.559) and pour point (+0.560) indicate that this component primarily characterizes thermal resilience and low temperature fluid behavior. In practical terms, the PC3 differentiates fuels based on their ability to maintain stability under thermal stress and tendency to solidify at low temperatures.

Overall, the loading patterns show that the physicochemical characteristics of Camelina-based blends are organized around three core property domains:

- I. Physical properties (density, viscosity, CFPP, iodine number) → PC1
- II. Combustion and oxidative stability (cetane number, flash point, sulfur content) → PC2
- III. Thermal and low-temperature stability (pour point, oxidation stability) → PC3

This multilevel organization confirms that PCA successfully distinguishes between the structural, functional, and performance-related aspects of the blends, providing an integrated framework for assessing their fuel quality and optimization potential.

Table 1. Principal component loadings for fifteen physicochemical parameters of blends

Component	PC1	PC2	PC3
Carbon Residue	-0.207	0.308	0.360
Sulfur Content	-0.169	0.381	-0.181
Cetane Number	-0.018	0.394	-0.247
Flash Point	-0.130	0.372	-0.094
Total Contamination	0.269	0.259	0.129
Kinematic Viscosity (40°C)	-0.265	0.279	-0.166
CFPP	-0.353	0.001	-0.027
Water Content	-0.280	-0.224	0.013
Density (15°C)	-0.315	0.195	-0.130
Iodine Number	-0.338	0.045	-0.220
Oxidation Stability (110°C)	0.103	0.314	0.559
Net Calorific Value	0.320	0.187	-0.031
Pour Point	-0.268	0.046	0.560
Phosphorus	-0.303	-0.178	0.100
Group II Metals (Ca+Mg)	-0.274	-0.251	0.144

The biplot constructed on the PC1–PC2 plane (Figure 3) provides a comprehensive visual summary of the relationships between the physicochemical variables and the positioning of the six fuel blends in reduced component space. This two-dimensional projection simultaneously displays the orientations of the variable vectors (loadings) and locations of the blend observations (scores), allowing for a clear understanding of how each property influences the clustering behavior of the fuels.

In the biplot, the direction and proximity of the vectors indicate the strength and nature of the correlations among the variables. Vectors extending in the same direction, such as those representing density, iodine number, and viscosity, exhibited strong positive correlations, signifying that these physical parameters tended to vary together, which is characteristic of biodiesel-rich formulations. Conversely, vectors pointing in opposite directions, such as the cetane number versus iodine number, denote negative correlations, implying that improvements in ignition quality are associated with reductions in unsaturation and viscosity. The angular relationships among the vectors convey meaningful physical interpretations: smaller angles indicate strong correlations, right angles imply weak or no correlation, and opposite orientations denote inverse relationships.

The spatial distribution of the samples within the biplot further clarifies the compositional and performance-related grouping of fuel blends. The B20, B50, and B100 samples clustered within the negative PC1 region, representing blends dominated by biodiesel. These fuels are characterized by higher viscosity, increased density, and elevated iodine numbers, which collectively correspond to poorer cold flow behavior and thicker fluidity. The positioning of these samples reflects their stronger association with the physical properties domain of PC1.

In contrast, the ethanol-containing blends (B20D65E15, B50D35E15, and B85D0E15) were located in the positive PC1 direction, characterized by lower density, reduced viscosity, and higher net calorific values. These samples demonstrated improved volatility and atomization properties, consistent with the presence of ethanol, which acts as a thinning agent and enhances the homogeneity of the mixture.

Along the PC2 axis, which captures the combustion and oxidative stability characteristics, additional differentiation emerged. The B100 sample occupied the lower end of the PC2 axis, indicating reduced combustion stability and higher susceptibility to oxidation. Conversely, the B20 and B20D65E15 blends, both positioned higher along PC2, were associated with higher cetane numbers and superior flash point values, representing fuels with enhanced ignition quality and operational safety.

Overall, the biplot demonstrates that ethanol addition exerts a dual effect: it significantly improves the physical properties of the blends (by decreasing the viscosity and density), thereby enhancing the flow and cold-start performance, while causing a partial reduction in combustion stability, as reflected in the downward shift of the oxidation resistance indicators. This trade-off between the physical and combustion characteristics underscores the importance of optimizing the ethanol proportion in ternary fuel formulations to balance the flow behavior, stability, and ignition performance.

Thus, the PC1–PC2 projection not only confirms the clustering tendencies and correlations observed in the loading matrix but also provides a visual framework for understanding the underlying trade-offs in the physicochemical design. The clear separation of sample groups along the two axes reinforces the PCA's capacity to distinguish

between biodiesel-dominant and ethanol-enhanced formulations, making it a valuable diagnostic tool for the multivariate assessment of biofuel quality and compatibility.

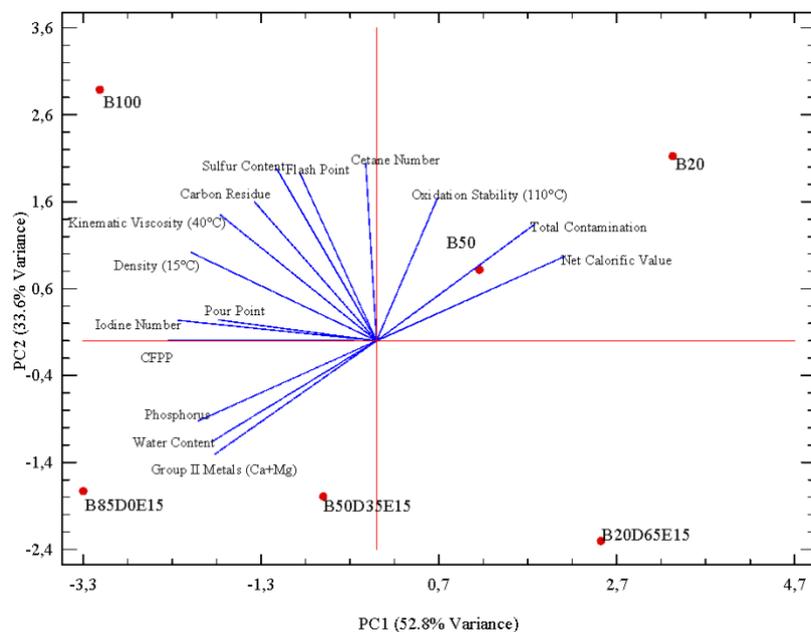


Figure 3. PCA biplot illustrating the relationships among variables and the clustering tendencies of the six fuel blends on the PC1–PC2 plane

#### IV. RESULTS

The results of this study provide a detailed quantitative and interpretive assessment of the physicochemical behavior of *Camelina sativa*-based diesel–biodiesel–ethanol ternary blends. Using Principal Component Analysis (PCA) as the primary multivariate tool, the analysis successfully condensed 15 correlated parameters into a set of independent components that captured the essential variability governing the physical and combustion-related performance of the fuels. This statistical condensation not only simplifies the interpretation but also enhances the ability to compare and classify biofuel formulations based on multidimensional quality indicators.

The PCA model demonstrated a remarkably high explanatory capacity, with the first three principal components (PC1, PC2, and PC3) collectively accounting for 94.94% of the total variance. Such a high cumulative variance suggests that nearly all significant physicochemical information is represented by a compact set of three orthogonal axes.

The first principal component (PC1), which explained 52.96% of the total variance, was dominated by variables related to physical properties, including density, viscosity, iodine number, and CFPP. These parameters jointly describe the structural nature of the blends, molecular packing, and cold-flow performance. Fuels with higher biodiesel content (e.g., B50 and B100) exhibited higher values of these parameters, confirming the expected trend of greater molecular weight and chain saturation influencing the density and viscosity. PC1 therefore represents the “physical structure and fluidity axis,” serving as a diagnostic indicator of cold flow and blending compatibility.

The second principal component (PC2) accounted for 33.40% of the total variance and captured combustion-related and stability indicators, such as the cetane number, flash point, oxidation stability, and sulfur content. These parameters define the ignition quality, combustion smoothness and oxidative resilience of the fuels. The high positive loadings on PC2 confirm that biodiesel-rich blends, although less volatile, exhibit superior ignition characteristics and higher flash points than ethanol-enriched fuels. Hence, PC2 can be regarded as the “combustion stability and ignition quality axis,” distinguishing between thermally robust and highly volatile fuel formulations.

The third principal component (PC3) explained 8.57% of the total variance and was primarily associated with oxidation stability and pour point, reflecting low-temperature and thermal endurance. Although contributing less

variance than the first two components, PC3 provides complementary insights into the stability domain, which is particularly relevant for long-term storage and cold-climate operations.

The distribution of these components, as visualized in the scree plot (Figure 2), revealed a sharp decline in eigenvalues after PC3, confirming that additional components contributed negligibly to the explanation of data variance. This “elbow structure” validates the three-component model as both statistically efficient and physically meaningful.

The biplot visualization further elucidated the multivariate relationships by simultaneously displaying the variable correlations and sample groupings on the PC1–PC2 plane. The physical parameters (density, viscosity, iodine number, and CFPP) formed a coherent cluster along the negative PC1 direction, whereas the combustion and stability parameters (cetane number, flash point, and oxidation stability) were positioned orthogonally along PC2. This geometric arrangement reinforces the inverse relationship between fluidity and combustion robustness, which is a common trade-off in oxygenated fuels. The fuel sample distribution across the biplot space provides additional insights into compositional effects.

B20, B50, and B100, the biodiesel-dominant samples, were located in the negative PC1 region, characterized by higher viscosity, increased density, and stronger unsaturation levels, all of which reduce cold-flow efficiency but enhance combustion stability.

In contrast, B20D65E15, B50D35E15, and B85D0E15, which contained ethanol, appeared in the positive PC1 direction, reflecting reduced viscosity and density, improved volatility, and enhanced flow properties. However, these samples also exhibited a modest decline in combustion stability indicators (lower cetane number and oxidation resistance), reflecting the role of ethanol as both a diluent and volatility enhancer.

These results highlight a clear trade-off inherent in ternary fuel design: the addition of ethanol improves fluidity and atomization, facilitating better fuel–air mixing, but may lead to slightly reduced oxidative durability. Conversely, higher biodiesel content increases the stability and ignition reliability at the expense of viscosity and cold-start performance. Balancing these opposing effects is central to optimizing biofuel formulations for environmental and operational performance.

In terms of sustainability, the PCA results emphasized the suitability of *Camelina sativa* as a strategic feedstock for next-generation fuels. Its high oil yield, favorable fatty acid profile, and adaptability to low-input cultivation contribute to reduced life cycle emissions and enhanced energy return on investment. The clustering behavior observed in the PCA plots suggests that camelina-based blends can achieve performance characteristics comparable to those of conventional biodiesel systems while maintaining improved environmental metrics when ethanol is judiciously integrated.

Finally, from a methodological standpoint, this study demonstrates that multivariate statistical analysis provides a superior diagnostic perspective compared with single-variable assessments. PCA enables the simultaneous visualization of correlations, grouping tendencies, and outliers, thereby providing engineers and researchers with a data-driven foundation for blend formulation, quality monitoring, and process optimization.

In summary, the results confirm the following:

- The first three principal components effectively described nearly all the variance in the dataset, validating the PCA as a compact and reliable descriptor of physicochemical complexity.
- Physical parameters dominate PC1, combustion stability parameters dominate PC2, and low-temperature endurance parameters dominate PC3.
- Ethanol addition improves flow properties but slightly reduces oxidative resilience, whereas biodiesel enrichment enhances stability at the cost of fluidity.
- The use of *Camelina sativa*–derived biodiesel presents a sustainable route for producing low-carbon, high-performance ternary fuel blends.

These outcomes provide a quantitative and visual framework for future blend optimization strategies, wherein the composition, stability, and environmental performance can be co-optimized using multivariate modeling approaches.

## V. CONCLUSION

This study presents a comprehensive multivariate characterization of *Camelina sativa*-based diesel-biodiesel-ethanol ternary blends using Principal Component Analysis (PCA). The analysis, based on an open-access dataset, provides an integrated framework for understanding the relationships among multiple physicochemical, combustion, and stability parameters that collectively determine biofuel quality. The results demonstrated that three principal components effectively explained 94.94% of the total variance, confirming that the complex physicochemical behavior of the blends can be accurately described using a small number of independent latent factors. PC1 captured the physical structure of the fuels (density, viscosity, CFPP, and iodine number), PC2 reflected their combustion and oxidative stability characteristics (cetane number, flash point, and oxidation stability), and PC3 represented the thermal and low-temperature behavior.

The biplot interpretation revealed two clear trends.

- Increasing the biodiesel content enhanced the combustion stability and ignition quality but reduced the flow performance owing to the higher density and viscosity.
- Increasing the ethanol proportion improved fluidity and reduced cold-flow resistance, although with a moderate reduction in oxidative durability.

These trends highlight a fundamental trade-off between the physical and combustion-related properties, which must be optimized to achieve a balanced fuel performance. The ternary blends incorporating camelina-derived biodiesel and ethanol showed synergistic potential, combining renewable sourcing, acceptable stability, and improved flow properties compared with conventional diesel fuels.

From a methodological standpoint, this study demonstrates that PCA is a powerful and scalable tool for biofuel quality classification, formulation optimization, and process evaluation. The open-data approach strengthens reproducibility and enables comparative analyses across feedstocks and blending strategies.

In conclusion, *Camelina sativa*-based diesel-biodiesel-ethanol blends represent a promising pathway toward sustainable and low-emission transportation fuel. When optimized through multivariate modeling, these blends can deliver improved cold-flow characteristics, reduced environmental impact, and sufficient combustion stability to meet the emerging renewable fuel standards. Future work should expand this framework by integrating engine performance tests, emission profiles, and temperature-dependent datasets, thereby enhancing the predictive capacity of multivariate biofuel assessment tools and supporting their implementation in real-world energy transition scenarios.

## ACKNOWLEDGMENT

The authors acknowledge the Scientific and Technological Research Council of Turkey (TÜBİTAK) for supporting the project titled “Physicochemical Properties of Camelina (Ketencik) Biodiesel and Its Blends” under Project No: 114M838. This study was conducted using an open-access dataset produced within the scope of a project conducted under the supervision of Prof. Dr. Hasan Aydoğan and made publicly available via Zenodo. The availability of this dataset ensured the reproducibility and transparency of the present study.

## REFERENCES

- [1] M. Balat, “Fuels from Biomass – Overview,” in *The Biofuels Handbook*, The Royal Society of Chemistry, 2011, pp. 76–117. doi: 10.1039/9781849731027-00076.
- [2] M. Khairati, “Biodiesel: An Overview,” *International Journal of Research and Review*, vol. 10, no. 11, pp. 228–236, 2023, doi: 10.52403/ijrr.20231127.
- [3] A. Pandey, C. Larroche, S. Rieke, C. Dussap, and E. Gnansounou, “Biofuels: alternative feedstocks and conversion processes,” 2011.
- [4] A. Obour K, “Oilseed Camelina (*Camelina sativa* L Crantz): Production Systems, Prospects and Challenges in the USA Great Plains,” *Advances in Plants & Agriculture Research*, vol. 2, no. 2, 2015, doi: 10.15406/apar.2015.02.00043.
- [5] F. G. Jewett, “*Camelina sativa*: For Biofuels and Bioproducts,” in *Handbook of Plant Breeding*, Springer New York, 2014, pp. 157–170. doi: 10.1007/978-1-4939-1447-0\_8.
- [6] P. Dobre and Ş. Jurcoane, “Camelina crop - opportunities for a sustainable agriculture,” 2011.
- [7] A. Guendouz, A. Hannachi, M. Benidir, Z. E. A. Fellahi, and B. Frih, “Agro-biochemical Characterisation of *Camelina sativa*: A Review,” *Agricultural Reviews*, no. Of, 2022, doi: 10.18805/ag.rf-230.
- [8] E. Waraich *et al.*, “‘Camelina sativa’, a climate proof crop, has high nutritive value and multiple-uses: A review,” 2013.
- [9] O. S. Stamenković *et al.*, “Biodiesel production from camelina oil: Present status and future perspectives,” *Food Energy Secur*, vol. 12, no. 1, 2021, doi: 10.1002/fes3.340.

- [10] B. R. Moser, "Biodiesel from alternative oilseed feedstocks: camelina and field pennycress," *Biofuels*, vol. 3, no. 2, pp. 193–209, 2012, doi: 10.4155/bfs.12.6.
- [11] K. A. V. Miyuranga, S. J. De Silva, U. S. P. R. Arachchige, R. A. Jayasinghe, and N. A. Weerasekara, "Comparison of the Properties of Biodiesel-Bioethanol-Diesel Blended Fuel," *Asian Journal of Chemistry*, vol. 34, no. 7, pp. 1809–1813, 2022, doi: 10.14233/ajchem.2022.23767.
- [12] P. Karin *et al.*, "Influence of ethanol-biodiesel blends on diesel engines combustion behavior and particulate matter physicochemical characteristics," *Case Studies in Chemical and Environmental Engineering*, vol. 6, p. 100249, 2022, doi: 10.1016/j.cscee.2022.100249.
- [13] P. Wai *et al.*, "Experimental Investigation of the Influence of Ethanol and Biodiesel on Common Rail Direct Injection Diesel Engine's Combustion and Emission Characteristics," *SSRN Electronic Journal*, 2022, doi: 10.2139/ssrn.4169796.
- [14] S. Kesharvani, G. Dwivedi, T. N. Verma, and P. Verma, "The Experimental Investigation of a Diesel Engine Using Ternary Blends of Algae Biodiesel, Ethanol and Diesel Fuels," *Energies (Basel)*, vol. 16, no. 1, p. 229, 2022, doi: 10.3390/en16010229.
- [15] M. K. Yesilyurt, M. Aydin, Z. Yilbasi, and M. Arslan, "Investigation on the structural effects of the addition of alcohols having various chain lengths into the vegetable oil-biodiesel-diesel fuel blends: An attempt for improving the performance, combustion, and exhaust emission characteristics of a compression ignition engine," *Fuel*, vol. 269, p. 117455, 2020, doi: 10.1016/j.fuel.2020.117455.
- [16] A. M. Couto, T. de P. Protásio, P. F. Trugilho, T. A. Neves, and V. A. de Sá, "Multivariate analysis applied to evaluation of Eucalyptus clones for bioenergy production," *CERNE*, vol. 19, no. 4, pp. 525–533, 2013, doi: 10.1590/s0104-77602013000400001.
- [17] A. F. N. M. Castro *et al.*, "Análise multivariada para seleção de clones de eucalipto destinados à produção de carvão vegetal," *Pesqui Agropecu Bras.*, vol. 48, no. 6, pp. 627–635, 2013, doi: 10.1590/s0100-204x2013000600008.
- [18] M. Jafari *et al.*, "Multivariate analysis of performance and emission parameters in a diesel engine using biodiesel and oxygenated additive," *Energy Convers Manag.*, vol. 201, p. 112183, 2019, doi: 10.1016/j.enconman.2019.112183.
- [19] V. H. J. M. dos Santos, E. D. C. Bruzza, J. E. de Lima, R. V. Lourega, and L. F. Rodrigues, "Discriminant Analysis and Cluster Analysis of Biodiesel Fuel Blends Based on Fourier Transform Infrared Spectroscopy (FTIR)," *Energy & Fuels*, vol. 30, no. 6, pp. 4905–4915, 2016, doi: 10.1021/acs.energyfuels.6b00447.
- [20] I. L. Motta, A. N. Marchesan, R. Maciel Filho, and M. R. Wolf Maciel, "Correlating biomass properties, gasification performance, and syngas applications of Brazilian feedstocks via simulation and multivariate analysis," *Ind Crops Prod.*, vol. 181, p. 114808, 2022, doi: 10.1016/j.indcrop.2022.114808.
- [21] P. Verma, M. Kumar, G. Mishra, and D. Sahoo, "Multivariate analysis of fatty acid and biochemical constituents of seaweeds to characterize their potential as bioresource for biofuel and fine chemicals," *Bioresour Technol.*, vol. 226, pp. 132–144, 2017, doi: 10.1016/j.biortech.2016.11.044.
- [22] P. Nachtergaele, J. Thybaut, S. De Meester, D. Drijvers, W. Saeys, and J. Dewulf, "Multivariate Analysis of Industrial Biorefinery Processes: Strategy for Improved Process Understanding with Case Studies in Fatty Acid Production," *Industrial & Engineering Chemistry Research*, vol. 59, no. 16, pp. 7732–7745, 2020, doi: 10.1021/acs.iecr.0c00515.
- [23] S. Hada, C. C. Solvason, and M. R. Eden, "Characterization-Based Molecular Design of Bio-Fuel Additives Using Chemometric and Property Clustering Techniques," *Front Energy Res.*, vol. 2, 2014, doi: 10.3389/fenrg.2014.00020.
- [24] W. F. de Carvalho Rocha, C. Presser, S. Bernier, A. Nazarian, and D. A. Sheen, "Laser-driven calorimetry and chemometric quantification of standard reference material diesel/biodiesel fuel blends," *Fuel*, vol. 281, p. 118720, 2020, doi: 10.1016/j.fuel.2020.118720.
- [25] B. R. Moser and S. F. Vaughn, "Evaluation of alkyl esters from Camelina sativa oil as biodiesel and as blend components in ultra low-sulfur diesel fuel☆," *Bioresour Technol.*, vol. 101, no. 2, pp. 646–653, 2010, doi: 10.1016/j.biortech.2009.08.054.
- [26] B. R. Moser, "Biodiesel from alternative oilseed feedstocks: camelina and field pennycress," *Biofuels*, vol. 3, no. 2, pp. 193–209, 2012, doi: 10.4155/bfs.12.6.
- [27] K. Ramos, A. Riddell, H. Tsiagras, and A. M. Hupp, "Analysis of biodiesel-diesel blends: Does ultrafast gas chromatography provide for similar separation in a fraction of the time?," *J Chromatogr A*, vol. 1667, p. 462903, 2022, doi: 10.1016/j.chroma.2022.462903.
- [28] A. S. van Niekerk, B. Drew, N. Larsen, and P. J. Kay, "Data set for influence of blends of diesel and renewable fuels on compression ignition engine emissions," *Data Brief*, vol. 28, p. 104836, 2020, doi: 10.1016/j.dib.2019.104836.
- [29] M. O. McLinden, T. J. Bruno, M. Frenkel, and M. L. Huber, "Standard Reference Data for the Thermophysical Properties of Biofuels," *J ASTM Int*, vol. 7, no. 3, pp. 1–18, 2010, doi: 10.1520/jai102586.
- [30] A. E. Comesana, T. Huntington, C. D. Scown, K. E. Niemeyer, and V. Rapp, "A Systematic Method for Selecting Molecular Descriptors as Features When Training Models for Predicting Physicochemical Properties," *SSRN Electronic Journal*, 2021, doi: 10.2139/ssrn.3990072.
- [31] E. Sygula, J. Lyczko, and A. Białowiec, "Dataset on fuel properties and volatile organic compounds from chemically synthesized biomass components for modeling and predicting biomass properties in pyrolysis processes," *Data Brief*, vol. 59, p. 111296, 2025, doi: 10.1016/j.dib.2025.111296.
- [32] H. Aydoğan, "Physicochemical Properties of Camelina Biodiesel and Its Blends (Versiyon v1) [Data set]," *Zenodo*, Oct. 2025, doi: <https://doi.org/10.5281/zenodo.17329324>.

## RESUME

### Dr. Mehmet Selman GOKMEN

Dr. Mehmet Selman Gökmen is a lecturer at the Seydişehir Vocational School, Automotive Technology Program, Necmettin Erbakan University, Konya, Türkiye. He completed his undergraduate studies in Mechanical Engineering and obtained his M.Sc. and Ph.D. degrees in Automotive Engineering from Selçuk University. His research interests include alternative fuels, automotive engineering, internal combustion engine testing, statistical analysis, and machine learning applications. He has conducted national and international studies focusing on energy efficiency, emission reduction, and fuel optimization.

### Prof. Dr. Hasan AYDOGAN

Prof. Dr. Hasan Aydođan is a faculty member at the Department of Mechanical Engineering, Faculty of Technology, Selçuk University, Konya, Türkiye. His research interests cover internal combustion engines, alternative energy sources, fuel technologies, energy conversion, and thermodynamic systems. He has been conducting both experimental and theoretical research on engine fuels, energy systems, and renewable energy technologies for many years and has authored numerous national and international scientific publications.

# Mechatronics Evolution based on Meta-system Transitions Theory

NIKOS A. ASPRAGATHOS<sup>1</sup>

<sup>1</sup>Robotics Group, Mechanical Engineering and Aeronautics Department,  
University of Patras, Patras, Greece, (asprag@mech.upatras.gr)

## I. INTRODUCTION

The abstract model of a typical mechatronic system was presented by a simple scheme including sensors and actuators interconnected by a computer for the control and an interface with the user. These components were designed independently and in the final stage are interconnected to implement the mechatronic system.

The mechatronics concept of interdisciplinarity and synergy was represented by the well know intersection of three or sometimes four circles showing the synergy of enabling technologies or engineering disciplines namely mechanics, electronics, computers and sometimes control [1]. This simplistic scheme of mechatronic devices cannot represent either the structure or the abilities and complexity of the current mechatronic systems considering the fast mechatronics evolution supported by its enabling technologies [2]. It is important to distinguish between quantitative and qualitative changes in the characteristics of the mechatronic systems as well as of the mechatronics approach (methodology) in design and manufacturing, to understand the transition from one level of mechatronics to the next higher level and the innovative changes [3, 4].

The semiconductor technology evolution provided microprocessors, power electronics, and the sensing ability to condition and encode physical measurements as analog/digital signals. Later advancements in materials, sensor-actuators technologies and digital control systems, enabling the design of new structures for self-contained components or modules of mechatronic systems.

The process of the continuous evolution of mechatronics technology considering the progress of enabling technologies and particularly the effects of AI and the Internet of Things guided to larger, distributed mechatronic systems and to Cyber Physical Systems (CPSs) [5]. The CPSs mainly emerged in Informatics and Communication cycles, while mechatronics was originally connected mainly to the evolution of mechanical systems design as it is coined by Japanese MITI. A common practice in mechatronics was the transferring of functionality from mechanics to informatics and electronics. Now it is time to reconsider the research on mechanical systems in the framework of mechatronics. The intelligent and autonomous systems should have new sensing capabilities and the ability to apply advanced actions such as high dexterity manipulation. These could be done by designing and manufacturing very clever reconfigurable mechanisms and new methods for mechanical or chemical processing of materials, energy savings etc.

Gradually the mechatronics systems structure and operation present increased complexity, modular structure, higher ability for interaction with the environment, intelligence and autonomy towards a System of Systems (SOS) structure and behavior [6]. Usually, a governing entity coordinates and integrates the constituent multiple systems of a mechatronic system formulating a metasystem. Turchin introduced the concept of metasystem transitions for the investigation of biological systems and science evolution [7].

In this paper, the concept of metasystem transitions to investigate mechatronics evolution by considering the interdependencies with enabling technologies and application domains, as well as the cybernetics importance. The relation between the quantitative and qualitative, transformative changes in the structure/organization and characteristics of mechatronic systems are presented. This approach could contribute to the identification of the trends and perspectives and the formulation of roadmaps towards the advancement of mechatronics.

## II. MECHATRONICS EVOLUTION WITH ENABLING TECHNOLOGIES

In mechatronics literature, the enabling technologies and the synergy of engineering disciplines (IT, Electronics, Mechanics) were considered to show the evolution of mechatronics as well as the support of mechatronics to various technological domains such as automotive, aerospace, precision agriculture, manufacturing etc., which is

a unidirectional conceptualization [3]. In this paper, the evolution of mechatronics is considered by investigating the interdependencies between technologies and disciplines, where the mechatronics is enabler to its enabling technologies, which is the predominant mode of scientific and technological progress. The main technologies enabling mechatronics evolution are electronics and computing, sensors and actuators, MEMS, cybernetics and Artificial Intelligence

Fig. 1 shows the interdependencies between the enabling technologies and disciplines with mechatronics, where the roles of driver and enabler are simultaneously shared by some of the involved technology domains and mechatronics, as it is briefly discussed and illustrated in this section.

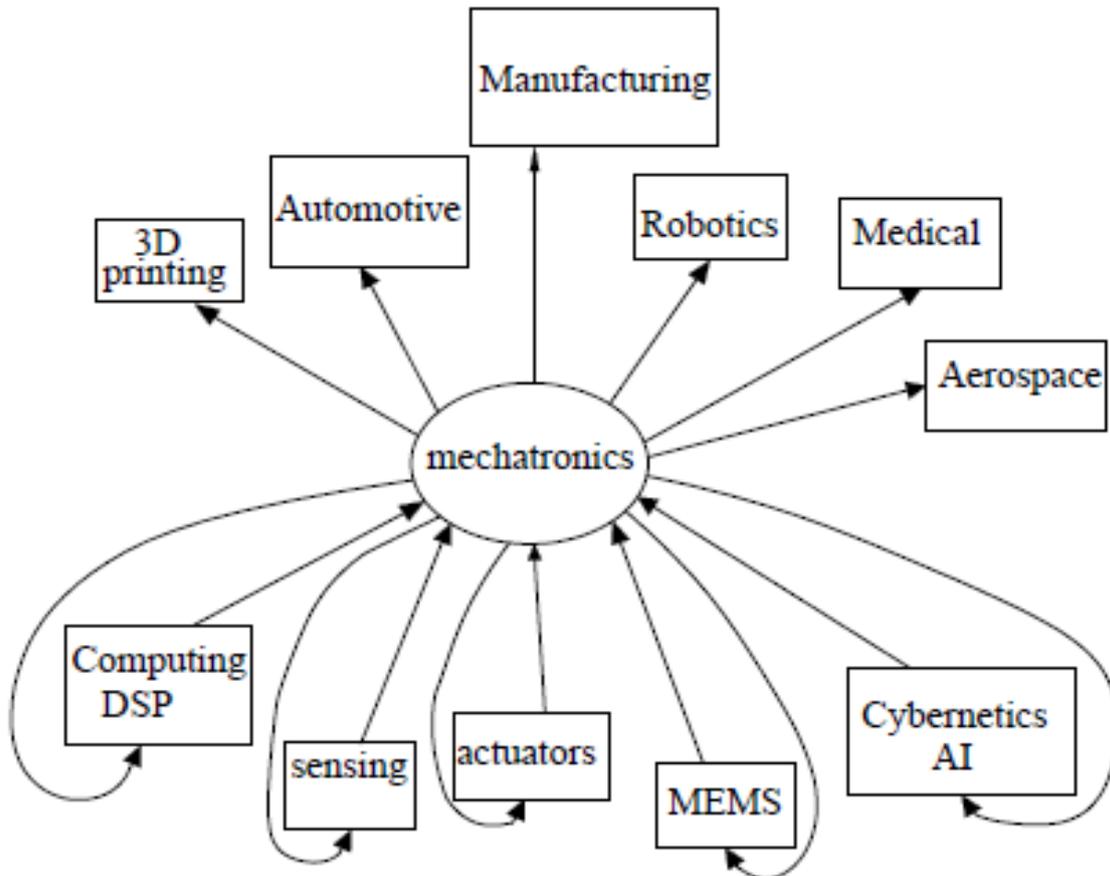


Figure 1: Interdependencies between enabling/driving technologies and mechatronics.

Electronics evolved very fast and apart from powerful computers used in design and simulation of mechatronic systems, the embedded computers in mechatronic devices and systems are basic components [8]. The major building blocks of an embedded system are the microcontrollers / DSP, the real time operating system (RTOS), the industry-specific protocols and interfaces and the printed circuit board assembly. The semiconductor devices include integrated chips, microcontrollers, field-programmable gate arrays (FPGAs) or a system-on-chip (SoC). The key equipment and processes for chip manufacturing include lithography, etching, ion implantation, chemical vapor deposition, testing, and packaging which were developed and evolved using mechatronics design and implementation methodologies [9].

Sensors are indispensable components of any mechatronic system and sensor technology progress leveraged the mechatronics evolution. The current sensor technology provides miniaturized, intelligent self-contained sensors of high performance having standard interfaces to facilitate integration in large mechatronic systems [10]. Structural modularity is a trend in sensor technology are to achieve easy connectivity, packaging compactness with embedded computing power by coupling physical, chemical, and biological sensing functionalities with control, electronics and software.

Any mechatronic system includes one or more actuators with a range of characteristics based on a variety of physical phenomena to convert energy. Advanced mechatronics engineering, smart and soft materials have been used to design and manufacture high precision, biomimicking, micro-actuators combined with sensors and controllers to formulate self-contained units or modules [11].

The main components of autonomous sensor-actuator modules are sensor element(s), energy transducers signal amplifiers, microcontrollers, data memories, RF transceivers, antennas, energy supplies, and application-specific multifunctional housings, which resembles an entire mechatronic system particularly if they could have self-calibration and self-repair capabilities.

It is obvious that many of the abilities and characteristics of the sensor and actuator systems are like mechatronic systems, and the synergy of similar disciplines is necessary to design and manufacture sensor-actuator modules like mechatronic systems by considering the differences in scope. Sensor and actuator technology are enablers to mechatronics and simultaneously the mechatronics synergetic methodology contributes to the sensor and actuator technology advancement. This dialectics of interdependencies between sensor and actuator technologies as well as other enabling technologies and mechatronics is the driving force for the technology evolution towards qualitative transitions to new mechatronic metasystems level.

Cybernetics unified control with communications theory, and Kolmogorov abridged cybernetics as the science devoted to the study of systems of any nature, which can receive, store and process information that is used for control [12]. The Cyber Physical Systems represents the merging of the virtual world of cyberspace with the physical world, which was the original meaning of Cybernetics. The Cyber physical systems share common methodologies and application domains with the new generation of mechatronics that undergone the meta-system transitions as it is explained in the following by considering the balanced synergy of all the enabling disciplines and technologies rather than emphasizing mostly to the cyberspace and Internet of Things (IoT).

Artificial intelligence (AI) split from Cybernetics and at the beginning it was considered as a pure computer-software discipline developing algorithms for expert systems, learning and cognition trying to emulate the human way of thinking and decision making. In the classical approach, the disciplines involved in AI were computer science, linguistics, psychology, and philosophy. Currently an embodied approach is followed and on top of the previous disciplines new ones are involved such as engineering, robotics, biology and neuroscience [13]. Under this consideration, the AI contributes synergistically to the mechatronics evolution, while the mechatronics is an enabling technology for embodied AI advancement showing their dialectic interdependency.

The contribution of mechanical and chemical engineering as well as the biology advancements in mechatronics evolution and vice versa should be considered. The contribution of mechanical engineering, particularly in thermal effects and thermal energy transfer and cooling as well to the interconnections and packaging of small size mechatronic systems like MEMS and embedded systems is very significant. On the other hand, mechanical engineering is benefited a lot by the mechatronics approach and advancements. In a report by ASME devoted to the Future of Mechanical Engineering [14], it was stressed that “Nanotechnology and biotechnology will dominate technological development and will be incorporated into all aspects of technology”.

It can be concluded that this dialectic interaction between technologies and disciplines is the driving force for the technological evolution in general and particularly for the qualitative transition of mechatronic systems to mechatronic systems, which is further analyzed in the following sections.

### III. META-SYSTEMS TRANSITIONS IN MECHATRONICS EVOLUTION

In the previous section, it is shown that the mechatronic systems undergone an evolution in their structure, organization and characteristics/abilities emerging a new mechatronics paradigm. These mechatronic systems are not just larger and of higher complexity but a new higher level of their organization, control and functionality emerge. The investigation of evolution gives an impetus for the improvements in mechatronics methodology towards new meta-system transitions. According to ISO/IEC/IEEE 21839, SoS defined as the Set of systems or system elements that interact to provide a unique capability that none of the constituent systems can accomplish on its own and constituent systems can be part of one or more SoS [6].

In the structure and organization, modern mechatronics systems are built by the interconnection of self-contained mechatronic components and modules, so we could speak about systems-of-systems, or this might not be enough [15]. One of the two properties of systems-of-systems is the operational independence of their components, while among their characteristics are the heterogeneity of their components and the non-predictable emergent behavior. A meta-system transition could happen, when either a component undergone a qualitative transition or the entire structure/organization reaches a new level.

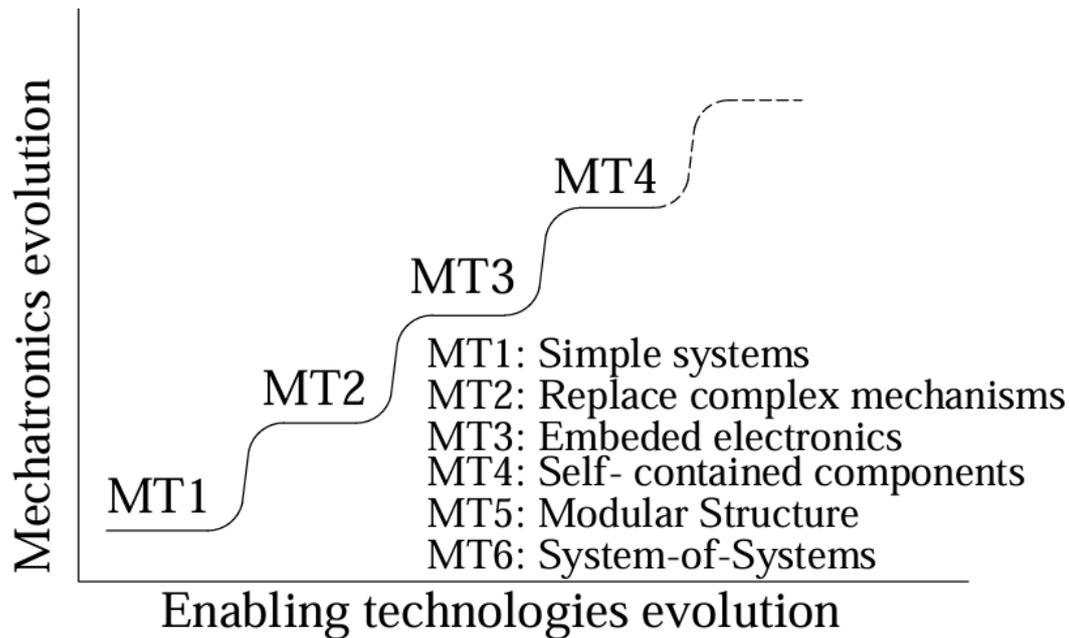


Figure 2: A schematic for the structural/organizational meta-system transitions

A schematic representation of the meta-systems transitions of the mechatronic systems structure and organization are illustrated in Fig. 2. In mechatronic systems structural evolution, meta-systems transitions happened at the beginning when the electrical elements integrated with the mechanical ones, then when the electronics are added, the maturity of computing reached DSP and embedded computing and in the current state the structural synthesis of mechatronic meta-systems is formed by self-contained components that are mechatronic sub-systems [2]. Sensors and actuators with embedded processors for built-in controllers and MEMS are important parts of mechatronic systems. Additionally, new soft and smart materials have been discovered, and new soft and/or self-contained sensors and actuators or sensor-actuator hybrid systems are developed to change the design and building of modular mechatronic systems towards a SoS structure [16].

In every transition step, a continuous quantitative accumulation of research and engineering knowledge, maturing methodologies and tools and the improvement of mechatronic elements reached the stage of a qualitative change on the structure and organization of the mechatronic meta-system (superior system). Complex and/or large mechatronic systems are built using self-contained components or modules, interconnected to formulate System-of-Systems (SoS) The current meta-system transition is the “social integration” of mechatronics via AI and Internet of Things presenting a new range and class of challenges [6].

Transitions on the mechatronic characteristics, abilities and functionality such as sensing, perception, learning, cognition, control, decision making, intelligence and autonomy evolved in a continuous quantitative accumulation of knowledge followed by a step in the next higher level of abilities, when the relevant enabling scientific areas reached maturity level [17]. In this paper, the proposed meta-system transition for the biological and particularly human evolution introduced by Turchin is adopted, since it could be parallelized with the mechatronic abilities evolution [18]. Turchin proposed the following levels of intelligence and science evolution:

- control of position => movement
- control of movement => irritability (simple reflex)
- control of irritability => (complex) reflex
- control of reflex => associating (conditional reflex)
- control of associating => human thinking
- control of human thinking => culture

As it appears in Fig. 3, the mechatronic system intelligence abilities followed a similar pattern to human but this artificial intelligence is still at a quite lower level compared to human. The starting simple mechatronic devices could control simple movements, while in the next transition mechatronic devices they could react with the environment with advanced control actions such as hybrid force/position control. In the next transition mechatronic devices and systems reached the level of decision making and autonomous behavior based on software programming.

Apart from the brain evolution, intelligence means very advanced physical action to fulfil the mechatronic tasks that was called “embodied intelligence” [13]. New smart and soft materials have been used to develop innovative biomimicking actuators that integrate controllers, communication, and possibly sensing and fault detection capabilities [19]. A lot of research work is devoted to articulation and dexterity with high efficiency and low weight and cost. Artificial muscles based on electroactive polymer, conductive or ionic polymer, shape memory alloys (SMAs), and piezoelectric have been proposed. Novel actuators based on polypyrrole/carbide-derived carbon hybrid materials developed at MIT, while bio-inspired micro-actuators harvesting energy from water vapor, are some examples of expected meta-system transitions due to new mechatronic elements and components.

New sensors on smell and tactile sensing apart from vision and hearing are based on intensive chemical and biological research. Biological sensory mechanisms are mechanoreceptors, electroreceptors, photoreceptors and chemoreceptors. Polymer-based sensor skin has been developed with multiple independent sensing modalities, including hardness sensing, surface profile, thermal conductivity, and temperature of an object.

Tactile sensor, which could detect contact force, vibration, texture, and temperature, would be the next generation of multimodal sensors contributing to the next meta-system embodied intelligence transition. This high intelligent interaction with the environment corresponds to control of complex reflexes based on embodied AI.

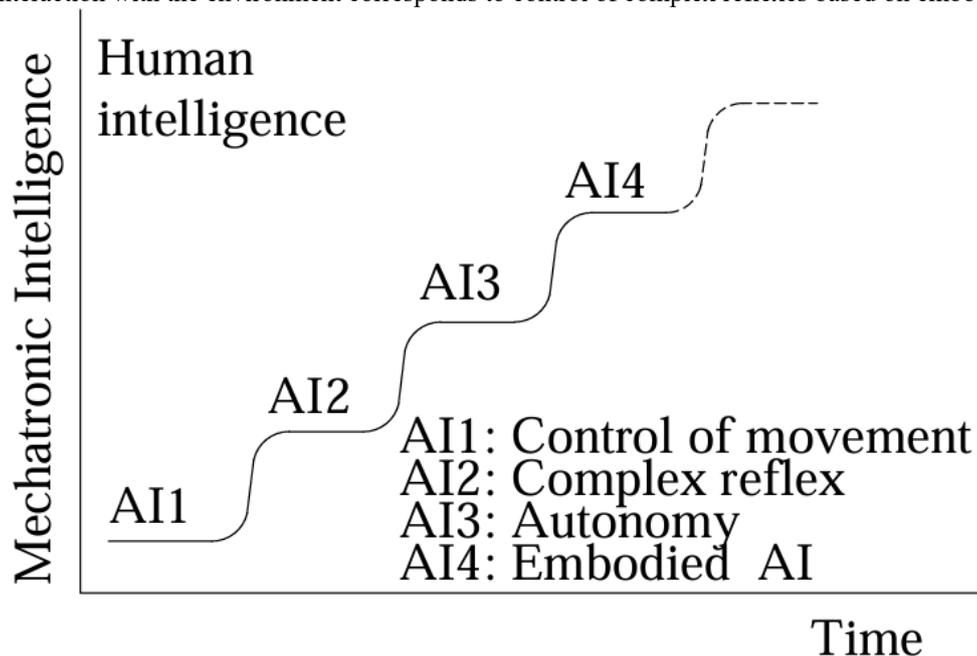


Figure 3: The meta-system transitions of mechatronic intelligence

In future, advanced miniaturized electronics with low power consumption; materials for quantum, neuromorphic and embedded system-on-chip (SoC) processing combined with advanced machine learning and embodied AI will be used in mechatronics for higher levels of intelligence and autonomy [20]. Mechatronic systems will have advanced abilities such as in perception and cognition, complex decision making, self-optimizing, eventually self-repairing, and self-reconfigured or self-metamorphosed as well as new structure, control and organization will emerge towards a new meta-system level of mechatronics.

#### IV. META-SYSTEMS TRANSITION IN MECHATRONICS ENGINEERING

Since mechatronics devices and systems undergone meta-system transitions of structure/organization, characteristics and abilities, then the mechatronics design approach was adapted by considering the interdependencies with enabling technologies and scientific disciplines. Therefore, the mechatronics design engineers, teams or collectives should be considered as a system that undergone meta-system transitions to reach higher organization levels, to develop design methodologies and engineering tools corresponding to the meta-system transitions of mechatronic systems that are presented in the previous sections [21, 22].

Since the mechatronic systems evolved from simple structure and lower abilities to large systems of very high complexity, autonomy and intelligence the organization of the engineering disciplines and domains had evolved to fulfil the new requirements in each transition change to higher level.

Collaborative design is an answer to tackle the higher complexity of current technology artifacts and systems. For developing new complex mechatronic systems, decomposition and integration are critical for mapping the complex system architecture to the formal and informal organization of the involved design teams from different domains by a trade-off between specialization and integration. Cross-functional, cross-expertise collaboration was proposed for successful collaborative design. However, in the relevant literature, it could not be found much attention to the organizational flexibility and team reconfiguration along the design phases.

It is quite important for design engineers as individuals and as synergetic groups to be aware of the mechatronic research meta-system transitions to encounter the latest well-established level of mechatronics evolution in their design endeavor or deliberately could design their mechatronic system at a transition level of their choice.

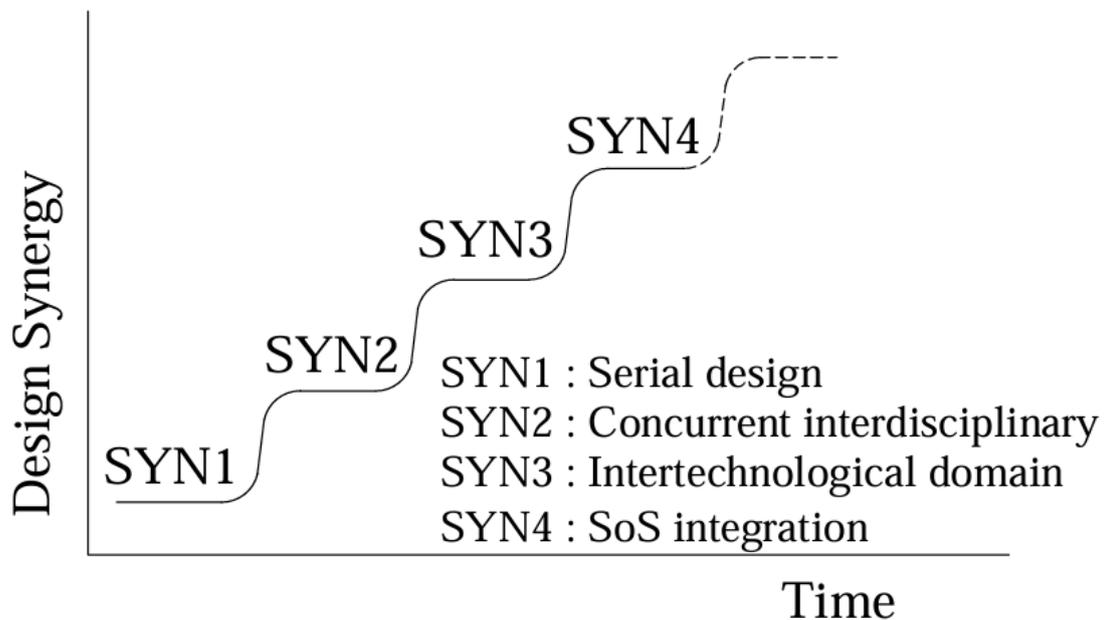


Figure 4: The meta-system transition of the interdisciplinary collaboration in mechatronics design

In Fig. 4, the meta-system transitions in the collaboration and synergy between engineers from different disciplines and domain expertise are illustrated. In the starting phase of mechatronics (SYN1), a serial design pattern was followed, where the mechanical part is designed and then the necessary electronics designed, and the software engineers wrote the program for control and operation of the system. The first meta-system transition (SYN2) in the organization of the design teams was the synergy of electronics, electrical and mechanical engineers from the conceptual design stage. The model of synergetic work particularly in the conceptual design and integration phases is the current one [23]. This organization of the design work was, and it is still efficient for simple mechatronic systems, where the elements had to be designed and manufactured from scratch. Until now the detailed design was a task of the specific discipline, since the corresponding engineers had to design elements belonging to their disciplines such as electronics, sensors or mechanical transmission of power. Since the mechatronic systems reached the level of the integration of self-contained existing components, this mode of work does not apply anymore for large scale mechatronic systems.

As presented in the previous sections, now there are mechatronic modules from different domains that are mechatronic entities and are used for the design and building a large advanced mechatronic system like robot, car, airplane or reconfigurable manufacturing system [24,25]. If there are not off-the-shelf components then the modules to be designed are mechatronic systems by themselves so interdisciplinary work is necessary, even for the detailed design of components.

The design team organization could undergo a new meta-system transition if plug and play self-contained, self-repaired and self-reconfigured (autonomous) components and their capabilities will be considered as services in a multi-agent organization or by networking them in an Internet of Things frame. At this level of organization and structure of independent mechatronic modules to build large scale very complex mechatronic Systems-of-Systems, a transition of engineering and design organization is required [15].

Since the mechatronics design collective is a system of individual engineers from different disciplines and with a variety of domain expertise, then it should be organized in such a way to be flexible enough to change its structure along the phases of the design process. Cybernetics was and still is considered that it could be used in scientific organization for collaborative work with high efficiency [26]. Therefore, a new cybernetics approach for the

organization of collective knowledge to reach a meta-system transition in mechatronics design methodology should be considered by the researchers of the mechatronics community.

## V. CONCLUSIONS

In this paper, an approach for the investigation of the mechatronics evolution is presented based on the theory of metasystem transitions. The changes in structure, and main characteristics and abilities of the mechatronic systems are identified and discussed to show qualitative changes as distinguished from quantitative ones, where relative knowledge is accumulated for the transition to a higher level.

The interdependencies of mechatronics technology with its enabling technologies are presented showing that for some of them mechatronics is simultaneously driver and the enabler. The significance of other scientific disciplines and technology domains to the evolution of mechatronics is emphasized rather than ICT and the conventionally considered disciplines involved in mechatronics.

Finally, the meta-system transitions approach is presented for the efficient organization of synergetic engineering work to adapt at the corresponding levels of the new metasystem transitions of mechatronic systems.

The presented approach should be enriched further by investigating other views of the mechatronics evolution not considered in this paper and this approach could be extended to investigate the evolution of other technology and science domains.

## REFERENCES

- [1] Isermann R., Modeling and design methodology for mechatronic systems. *IEEE/ASME Transactions on Mechatronics*, 1, 16-28, 1996.
- [2] Habib M., Mechatronics A Unifying Interdisciplinary and Intelligent Engineering Science Paradigm. *IEEE Industrial Electronics Magazine*, 1, 2007.
- [3] Marion T., Friar J., 'The Role of Enabling Technologies in Transformative Innovation', in Proceedings of the 22nd International Conference on Engineering Design (ICED19), Delft, The Netherlands, 5-8 August 2019. DOI:10.1017/dsi.2019.135
- [4] Nnodim C. T., Arowolo M. O., Agboola B. D., Ogundokun R. O., Abiodun M. K., Future Trends in Mechatronics, *IAES International Journal of Robotics and Automation (IJRA)*, Vol. 10, No. 1, March 2021.
- [5] Bradley D., Russell D., Ferguson I., Isaacs J., MacLeod A., White R., The Internet of Things – The future or the end of mechatronics, *Mechatronics*, 27, 2015.
- [6] Dahmann J., Current Landscape of System of Systems Engineering, 19th Annual System of Systems Engineering Conference (SoSE), 2024
- [7] Last C., Human metasystem transition (HMST) theory, *Journal of Ethics and Emerging Technologies* · January 2015
- [8] Nikolić G., Dimitrijević B., Nikolić T., Stojčev M., Fifty Years of Microprocessor Evolution: From single CPU to Multicore and Manycore Systems, *FACTA UNIVERSITATIS Series: Electronics and Energetics* Vol. 35, No 2, June 2022, pp. 155-186.
- [9] Li J., Overview of the development of chip manufacturing technology, *Proceedings of the 6th International Conference on Computing and Data Science*. 2024.
- [10] Trigona C., Graziani S., and Baglio S., Changes in sensors technologies during the last ten years: Evolution or revolution? *IEEE Instrumentation & Measurement Magazine* · September 2020
- [11] Okyay A., Erkorkmaz K., Khamese M. B., Mechatronic design, actuator optimization, and control of a long stroke linear nano-positioner, *Precision Engineering*, 52 (2018) 308-322
- [12] Kolmogorov A. N. Automata and Life, In I. M. Makarov, *Cybernetics Today, Achievements, Challenges, Prospects*, Mir Publishers, Moscow, 1984
- [13] Liu H., Guo D., Gangelosi A., Embodied Intelligence: A Synergy of Morphology, Action, Perception and Learning, *ACM Comput. Surv.*, Vol. 57, No. 7, Article 186, March 2025.
- [14] ASME, 2028 Vision for Mechanical Engineering, A report of the Global Summit on the Future of Mechanical Engineering, 2008, [www.asme.org](http://www.asme.org)
- [15] Dridi C., Benzandri Z., Belala F., System of Systems Engineering: Meta-Modelling Perspective, SoSE 2020, IEEE 15th International Conference of System of Systems Engineering, June 2-4, 2020, Budapest, Hungary.
- [16] Jahn U., Wolff C. and Schulz P., Concepts of a Modular System Architecture for Distributed Robotic Systems, *Computers* 2019, 8, 25.

- [17] Zaitceva, I.; Andrievsky, B. Methods of Intelligent Control in Mechatronics and Robotic Engineering: A Survey. *Electronics* 2022, 11, 2443
- [18] Turchin V., (1977). *The Phenomenon of Science*, Columbia University Press New York.
- [19] Shankar NVS, Kamma T. K., Krishna M., Ch H., Kumar, Shalem B., *Smart Actuators: A Review*, Proceedings of the International Conference on Industrial Engineering and Operations Management Rome, Italy, August 2-5, 2021.
- [20] Gao P., Adnan M., Overview of emerging electronics technologies for artificial intelligence: A review, *Materials Today Electronics* 11 (2025)
- [21] Moulianitis, V.C., Aspragathos, N.A., Introduction to the special issue on Theories and Methodologies for Mechatronics design, *Mechatronics* 20 (8), 825-826, 2010
- [22] Cintra Faria, A.C.; Barbalho, S.C.M. *Mechatronics: A Study on Its Scientific Constitution and Association with Innovative Products*. *Appl. Syst. Innov.* 2023, 6,
- [23] Moulianitis V, Zachiotis GD, Aspragathos N (2018) A new index based on mechatronics abilities for the conceptual design evaluation. *Mechatronics* 49:67–76. <https://doi.org/10.1016/j.mechatronics.2017.11.011>
- [24] Janssen L., Feyta R., Besselink B., Wouw N., *Modular Redesign of Mechatronic Systems: Formulation of Module Specifications Guaranteeing System Dynamics Specifications*, Elsevier, February 2025.
- [25] Zhang Z., Zhuge X., Li X., Evans R. and Liu A., "An Object-Oriented Approach to the Modular Design of Mechatronic Systems," in *IEEE Transactions on Engineering Management*, vol. 71, pp. 2623-2639, 2024,
- [26] Schwaninger M., *The Evolution of Organizational Cybernetics*, *Scientificae Mathematicae Japonicae*, January 2006,

#### RESUME

##### **NIKOS A. ASPRAGATHOS**

Professor (retired) Nikos A. Aspragathos was the leader of the Robotics Group in Mechanical and Aeronautics Engineering Department, University of Patras, Greece. His main research interests are robotics, intelligent motion planning and control for mobile robots, UAVs and dextrous manipulation, human-robot interaction, metamorphic manipulator design, knowledge-based mechatronics design, industrial automation, and computer graphics. He was visiting professor and/or researcher in University of New Castle Upon-Tyne, UK; UWIST, Cardiff, UK; CEA- LIST, Paris; Norwegian University of Science AND Technology, Trondheim, Norway. He supervised 14 PhD students who completed successfully their dissertation, and he served as examiner of doctoral dissertations in Greek and foreign universities. He is reviewer in about 40 Journals and 60 conferences, he was member of the editorial board of the *Mechatronics Journal*, *ROBOTICA*, *ISRN Robotics* and currently in *Cogent Engineering*. He published more than 100 papers in Journals, about 150 in conference proceedings and several book chapters. He served as proposals reviewer or Jury member in Greece European Union and other countries such as France, Slovakia, Norway, Kazakhstan, Ireland. Member of Advisory Group for *New Robotics Engineering Book Series*, ASME Press. He was involved in research projects funded by Greek and European Union sources.

# Performance Analysis of MLP-Mixer and ResNet50 Architectures in Coffee Leaf Disease Detection Using Raw Data

Muhammet CETINKAYA<sup>1</sup> and Sema SERVI<sup>2</sup>

<sup>1</sup>Selcuk University, Konya/Türkiye, mcetinkaya@selcuk.edu.tr, ORCID: [0009-0003-3549-1831](https://orcid.org/0009-0003-3549-1831)

<sup>2</sup>Selcuk University, Konya/Türkiye, semaservi@selcuk.edu.tr, ORCID: [0000-0003-2069-9085](https://orcid.org/0000-0003-2069-9085)

## I. INTRODUCTION

Ensuring global food security heavily depends on the effective management of agricultural diseases. Coffee (*Coffea arabica*), a crop of significant strategic importance due to its economic value and the livelihood it provides to millions of farmers, is highly vulnerable to biotic stressors such as coffee leaf rust (*Hemileia vastatrix*) and coffee berry disease. These pathogens can cause yield losses of up to 30% annually. Traditional diagnostic approaches rely on visual inspections by experts in the field, a process that is subjective, time-consuming, and costly.

To overcome these challenges, recent advances in artificial intelligence and deep learning have enabled the development of automated plant disease detection systems. A review of the literature reveals that most of the research focuses on Convolutional Neural Networks (CNNs). For instance, Mohanty et al. [4] demonstrated that deep learning models could outperform human experts in diagnosing diseases across 14 different plant species. Specifically for coffee plants, Esgario et al. [6] successfully applied various CNN architectures to detect the severity and type of coffee leaf diseases.

Among CNN-based models, the ResNet architecture [3] has become a benchmark due to its residual block structure, which effectively mitigates the vanishing gradient problem in deep networks. However, recent developments in deep learning have introduced alternatives that do not rely on convolutional operations. One such architecture is MLP-Mixer [2], which utilizes only Multi-Layer Perceptrons (MLPs) to process images by modeling inter-patch and intra-patch information.

This study aims to investigate how the non-convolutional MLP-Mixer architecture compares to the classical CNN-based ResNet50 in classifying coffee leaf diseases, using only raw data—without augmentation or class balancing. This design allows for a fair comparison of the intrinsic learning capabilities of both architectures.

## II. MATERIALS AND METHODS

### A. Dataset and Preprocessing

In this study, the publicly available JMUBEN dataset published by Murimi et al. [1] was used. The dataset contains five categories: *Cercospora*, *Leafrust*, *Miner*, *Phoma*, and healthy leaves. To prevent overfitting and data leakage, duplicate and visually similar images were removed from the original dataset, resulting in a clean set of 1,312 unique images.

### B. Model Architectures and Transfer Learning

Two different architectures were compared:

- **ResNet50 [3]:** A widely used deep CNN model with 50 layers and residual blocks, which serves as a benchmark model for image classification.
- **MLP-Mixer (mixer\_b16\_224\_in21k) [2]:** A non-convolutional architecture that processes images by dividing them into 16×16 patches and applies MLP layers for both token-mixing (across patches) and channel-mixing (within patches).

Both models were initialized with pre-trained weights on ImageNet. As noted by Pan and Yang [7], the transfer learning approach allows for faster and more efficient training by transferring general visual features learned from large-scale datasets to the target task.

### C. Training and Evaluation

The models were trained using the PyTorch [10] deep learning framework and the `timm` [9] model library on an NVIDIA RTX 4060 GPU. No data augmentation techniques were applied during training. The optimizer used was **AdamW** [8], which incorporates weight decay more effectively than the standard Adam optimizer. The learning rate was set to  $1e-4$ .

To avoid overfitting, **Early Stopping** was applied by monitoring validation loss. For statistical reliability, each model was trained using **five different random seeds**, and the mean performance scores were reported.

Figure 1 shows the Experimental Training and Evaluation Flow Chart of the model.

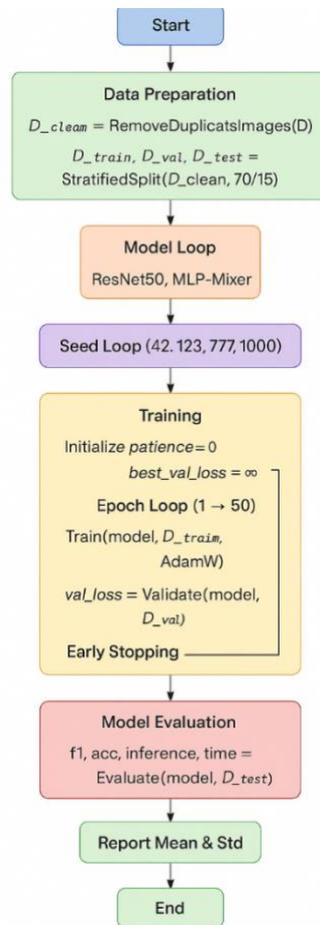


Figure 1: Experimental Training and Evaluation Flow Chart

## III. RESULTS

This section presents the comparative analysis results of the MLP-Mixer and ResNet50 models trained on the raw dataset (without data augmentation or class balancing), following the methodology described in Section 2. Each model was evaluated using five different random seeds, and the mean and standard deviation of the results are reported.

D. Classification Performance

Figure 2 illustrates the average macro F1-scores and accuracy values along with their standard deviations for both models. The results show that both models achieved high classification performance, with MLP-Mixer producing slightly more stable results due to lower standard deviation.

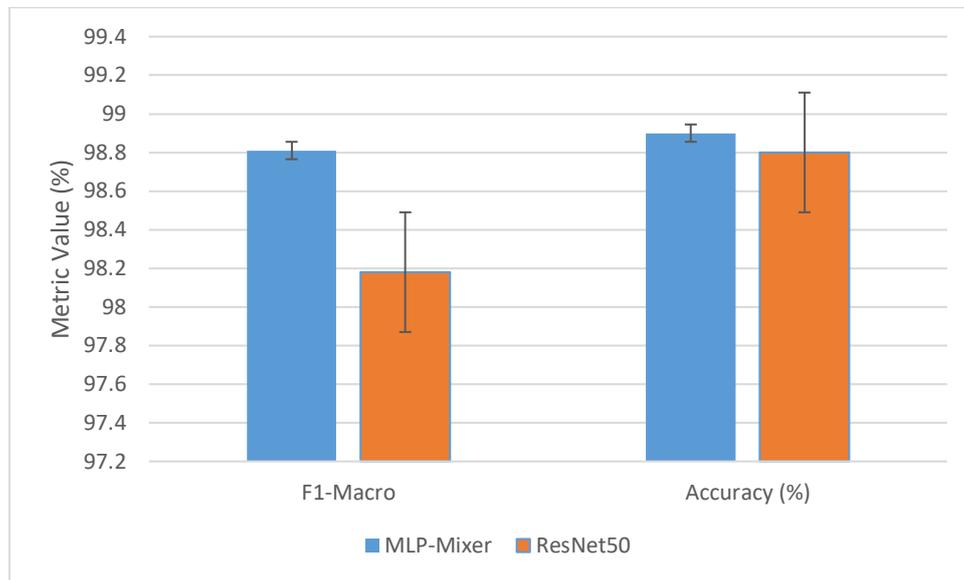


Figure 2: Comparison of average F1-Macro and Accuracy scores ( $\pm$  standard deviation) for MLP-Mixer and ResNet50 models.

Classification performance was assessed using metrics sensitive to class imbalance: **Macro F1-score** and **Overall Accuracy**. The average values from five independent runs are presented in Table 1.

Table 1. Comparative Performance of MLP-Mixer and ResNet50 (Mean of 5 Runs)

Model Architecture	Macro F1-score (Mean)	Accuracy (Mean)	F1 Std. Deviation
MLP-Mixer	0.9881	%98.90	$\pm 0.0057$
ResNet50	0.9818	%98.80	$\pm 0.0094$

As shown in Table 1, both models achieved excellent classification accuracy, with F1-scores above 98%. The MLP-Mixer outperformed ResNet50 by a marginal difference ( $\sim 0.6\%$ ). More importantly, MLP-Mixer yielded a lower standard deviation, indicating better training stability across different seeds.

E. Efficiency Analysis (Inference Speed)

Figure 3 compares the number of parameters and average inference time per image for the two architectures. ResNet50 demonstrated significantly faster inference, attributed to its smaller model size, making it more suitable for real-time or mobile applications. However, MLP-Mixer still delivered acceptable performance despite its larger size.

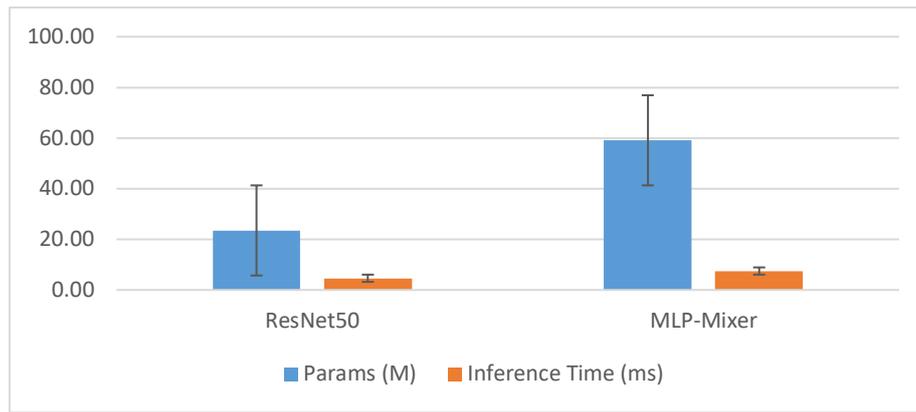


Figure 3: Comparison of model size (parameter count) and average inference time ( $\pm$  margin of error) for ResNet50 and MLP-Mixer.

To assess the suitability of the models for real-world applications, average prediction times for a single image were measured on an NVIDIA RTX 4060 GPU (Table 2).

Table 2. Model Size and Inference Time Comparison

Model Architecture	Number of Parameters	Avg. Inference Time (ms/image)
MLP-Mixer	59.1 Milyon	7.48
ResNet50	23.5 Milyon	4.62

In the efficiency analysis, ResNet50, thanks to its more compact structure (23.5M parameters), produced faster predictions on the test rig with 4.62 ms, while MLP-Mixer (59.1M parameters) followed with 7.48 ms.

F. Error Analysis

Figure 4 visualizes the per-class recall values for both models. Notably, MLP-Mixer outperformed ResNet50 on the *Phoma* class, which is considered visually more complex due to its irregular leaf patterns.

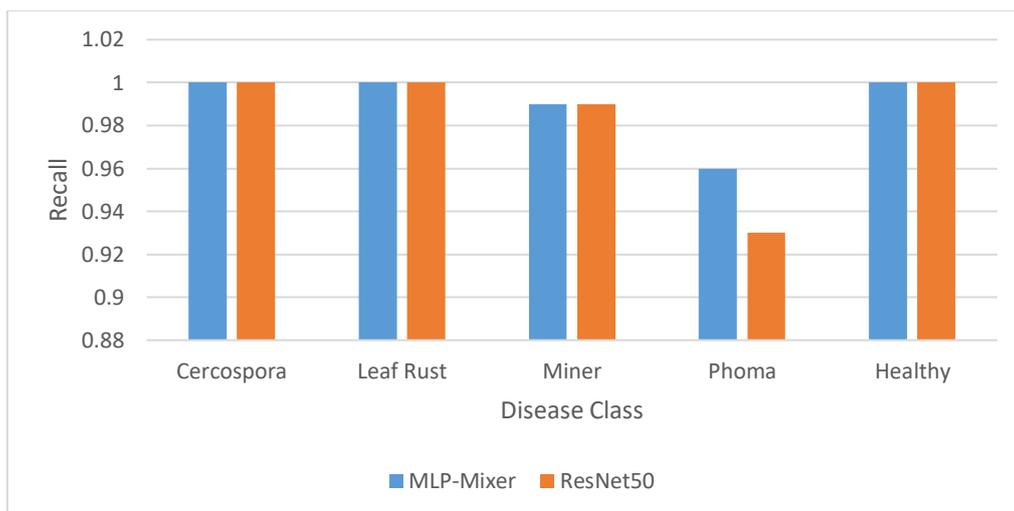


Figure 4: Class-wise recall comparison for MLP-Mixer and ResNet50 models.

When class-based performance (Table 3) is examined, it is seen that both models achieved full success in almost all classes.

Table 3. Class-wise Recall Comparison

Class	MLP-Mixer Recall	ResNet50 Recall	Comment
Cercospora	1.00	1.00	Equal Performance
Healthy	1.00	1.00	Equal Performance
Leaf rust	1.00	1.00	Equal Performance
Miner	0.99	0.99	Equal Performance
Phoma	0.96	0.93	MLP-Mixer Performs Better

The error analysis reveals that both models performed exceptionally well on most classes. The slightly higher recall of MLP-Mixer for *Phoma* suggests its advantage in capturing global contextual patterns, which may be crucial for classes with visually irregular features.

#### IV. DISCUSSION AND CONCLUSION

The experimental results on the clean, balanced dataset revealed that both the MLP-Mixer and ResNet50 architectures achieved outstanding performance in detecting coffee leaf diseases.

Both models attained F1-scores above 98%, validating the common consensus in the literature that CNNs are well-suited for this task. However, the MLP-Mixer, a non-convolutional architecture, demonstrated not only comparable but even superior performance in terms of training stability and slightly higher accuracy.

The MLP-Mixer achieved the highest average macro F1-score (0.9881) and the lowest standard deviation ( $\pm 0.0057$ ), indicating more consistent results across different random initializations. This suggests that the MLP-Mixer can generalize better under identical training conditions and may be more robust when deployed in real-world scenarios where data conditions vary.

On the other hand, ResNet50 maintained a clear advantage in terms of efficiency. With fewer parameters (23.5 million compared to MLP-Mixer's 59.1 million), ResNet50 achieved significantly faster inference time (4.62 ms vs. 7.48 ms per image), making it preferable for latency-critical applications such as mobile devices or real-time monitoring systems.

An interesting observation emerged in the *Phoma* class, where MLP-Mixer significantly outperformed ResNet50 in recall. This may be due to MLP-Mixer's ability to model global contextual features more effectively, which are particularly useful when classifying visually complex or ambiguous disease patterns. Such findings support the argument that non-convolutional architectures can offer advantages beyond traditional CNNs in certain contexts.

#### FINAL REMARKS

In this study, experiments conducted on a "naive" dataset showed that the modern MLP-Mixer architecture and the classic ResNet50 architecture exhibited very similar and superior performance in detecting coffee leaf diseases.

On the cleaned and balanced "naive" data, both models achieved F1 scores above 98%. This confirms the view in the literature that "CNNs are the standard for this work," and demonstrates that the non-convolutional MLP-Mixer can perform at least as well as CNNs, or even outperform them in terms of stability (standard deviation). MLP-Mixer was the most stable model, with a statistically lower standard deviation and marginally higher accuracy. Meanwhile, ResNet50 maintained its efficiency/speed advantage with its lower number of parameters and faster processing time.

In conclusion, while MLP-Mixer is recommended for the highest and most stable accuracy on servers with high hardware power, ResNet50 remains an extremely powerful and valid option in resource-constrained or speed-critical scenarios.

#### REFERENCES

- [1] R. Murimi, W. Gaita, and P. Ogao, "A unified public-domain dataset for coffee leaf disease recognition," *Data in Brief*, vol. 38, p. 107417, 2021.

- [2] I. O. Tolstikhin, N. Houlsby, F. Frolov, et al., “MLP-Mixer: An all-MLP architecture for vision,” in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 34, pp. 24261–24272, 2021.
- [3] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016.
- [4] S. P. Mohanty, D. P. Hughes, and M. Salathé, “Using deep learning for image-based plant disease detection,” *Frontiers in Plant Science*, vol. 7, p. 1419, 2016.
- [5] A. Kamilaris and F. X. Prenafeta-Boldú, “Deep learning in agriculture: A survey,” *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018.
- [6] J. G. Esgario, R. A. Krohling, and J. A. Ventura, “Deep learning for classification and severity estimation of coffee leaf biotic stress,” *Computers and Electronics in Agriculture*, vol. 169, p. 105162, 2020.
- [7] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.
- [8] I. Loshchilov and F. Hutter, “Decoupled weight decay regularization,” in *Proc. International Conference on Learning Representations (ICLR)*, 2019.
- [9] R. Wightman, “PyTorch Image Models,” GitHub repository, 2019. [Online]. Available: <https://github.com/rwightman/pytorch-image-models>
- [10] A. Paszke, S. Gross, F. Massa, et al., “PyTorch: An imperative style, high-performance deep learning library,” in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 32, 2019.

### RESUME

#### **Muhammet CETINKAYA**

Muhammet Çetinkaya graduated from the Computer Education and Instructional Technologies Department of the Faculty of Education at Siirt University in 2014. He has been working at Selçuk University since 2016 and currently works in the Revolving Fund Management Department of the Faculty of Agriculture at Selçuk University. He has been pursuing his master's degree in the Department of Information Engineering at the Faculty of Technology at Selçuk University since 2022. He continues his academic work in the fields of data mining and image processing.

#### **Sema SERVI**

Sema Servi is a Assistant Professor Dr. in Computer Engineering at Selcuk University, Faculty of Technology. She completed her master's and doctorate degrees at Selcuk University, Faculty of Science, Department of Mathematics. Servi has published articles on various national and international journals on topics such as applied mathematics and artificial intelligence. Her main research interests include numerical approximation solution methods, optimization algorithms and data mining.

# Explainable Multi-Scale Deep Learning Framework for Automatic Detection of Lung Diseases Using Ultrasound Imaging

Taha CAPAR<sup>1</sup> and Yusuf ERYESIL<sup>2</sup>

<sup>1</sup> Queen Mary University of London, London/United Kingdom, t.capar@se23.qmul.ac.uk

<sup>2</sup>Selcuk University, Konya/Türkiye, yusuf.eryesil@selcuk.edu.tr

## I. INTRODUCTION

Lung diseases such as pneumonia, pulmonary fibrosis, acute respiratory distress syndrome (ARDS), and COVID-19 are significant contributors to the worldwide burden of disease related to lung diseases and continue to be a leading cause of death. Early diagnosis is critical to enhancing patient outcomes, directing appropriate treatments, and decreasing the costs of the healthcare system. While conventional imaging modalities, such as chest X-ray (CXR) and computed tomography (CT), are widely accepted as the standard of care; both CXR (has low sensitivity for early-stage disease) and CT (which exposes patients to ionizing radiation, is expensive and impractical for use with critically ill, immobile, or limited resources) are associated with serious limitations. Recently lung ultrasound (LUS) has been recognized as the best alternative to CT and has emerged as an effective, safe, and portable means of performing an ultrasound examination on patients, both in real-time and bedside. LUS does not expose patients to radiation, has minimal operational costs, and can be easily performed at the bedside or in the field. Growing literature supports the clinical value of LUS in the identification of lung pathologies such as B-lines, pleural thickening, sub-pleural consolidations, and effusions [3]. LUS interpretation, however, remains difficult; many variables including operator skill and training, probe type used, angle of probe acquisition, and difficulty defining patient anatomy (due to large variations among populations) contribute to the variability in how clinicians perceive the same situation.

Deep learning and Artificial Intelligence (AI) have emerged as solutions for automating and augmenting the analysis of medical images to address the above challenges [2]. Although the majority of the work on using deep learning to detect COVID-19 from CT and CXR images has already been published [6], little research has been done on automatically classifying lung diseases using LUS images [3]. Ultrasound signals are inherently noisier than photographic or radiologic images. LUS images include numerous complex artifacts and require learning algorithms capable of modeling both subtle local patterns and broader contextual information. This necessitates the development of robust multi-scale deep learning systems and explainable AI frameworks that can ensure transparent decision-making and provide clinically meaningful justification for model outputs. The objective of the present study is to design a multi-scale deep learning architecture for the automated detection of lung diseases in LUS images, with an emphasis on explainability and interpretability. The proposed MS-CNN framework integrates multiple convolutional neural network (CNN) architectures—DenseNet121, MobileNetV2 and EfficientNetB7—to allow the system to learn representations that capture both localized image details and global structural cues, reflecting principles similar to those used in radiologic reasoning. [9]. In addition, using a weakly supervised attention model, the model highlights areas that are clinically significant, such as the pleural interface and the presence of pathologic artifacts. The model is also transparent, allowing clinical users to see how the model came to make a specific decision through Gradient-weighted Class Activation Mapping (Grad-CAM) [10].

The study examined the use of the Mendeley Lung Ultrasound dataset by exploring three classification areas (COVID-19, Non-COVID Pulmonary Pathophysiology and Healthy Lung Patterns) with the aid of performance measures based on cross-validation techniques to confirm that the results are representative and applicable among varying populations.

## II. LITERATURE REVIEW

The advent of artificial intelligence (AI), followed by the sub-classifications of machine learning (ML) and deep learning (DL), has led to an important unmet need in the area of medical imaging, namely the enhancement of disease detection, diagnosis, and prognoses. In the early stages of machine learning research on medical imaging, widely used algorithms included Support Vector Machines (SVMs), decision trees, and Random Forests. These methods represented foundational work demonstrating the feasibility of applying machine learning techniques to identify structural abnormalities in medical images. However, their effectiveness depended heavily on domain expertise, as their performance relied on handcrafted inputs derived from the underlying medical data. The emergence of deep learning algorithms—particularly Convolutional Neural Networks—marked a major shift in medical image analysis. CNN-based approaches enabled models to learn increasingly complex representations directly from raw imaging data such as computed tomography (CT), magnetic resonance imaging (MRI), X-ray, and ultrasound scans, without requiring predefined inputs designed by domain experts. This development significantly advanced the applicability of machine learning to diverse medical imaging modalities and improved the robustness of data-driven clinical decision-support systems. [4].

Lamba (2025) presented an exhaustive analysis of machine learning (ML) and deep learning (DL) in medical imaging that particularly noted that convolutional neural networks (CNN) and analogous designs have now become distinctly better than others at various tasks like lesion localization and tumor classification. The analysis discussed improvements in the area of generative adversarial networks (GAN) and transfer learning to better manage increased data sizes and adapting existing models towards particular uses. These improvements have increased the diagnostic power of existing image technologies and set a trend favoring the integration of artificial intelligence in the existing framework of radiological services. Nevertheless, Lamba (2025) [4] also identified some stiff challenges of data quality, model interpretability, and appropriate use of patient data that still obstruct the adoption of these improvements. Using hybrid models of artificial intelligence like combining convolutional neural networks with support vector machines (SVM) as a solution to these challenges and an emphasis on explainable artificial intelligence were suggested.

Besides overall biomedical imaging, the use of artificial intelligence within chest imaging has grown rapidly and now primarily targets the diagnosis of lung disease. Burki [1] suggested that overall diagnostic accuracy comparable to that of experienced human radiologists has been achieved by deep learning algorithms for fibrotic lung disease, malignant lesions, and chronic obstructive pulmonary disease (COPD) using CT scans. The current potential of these algorithms lies in the rapid analysis of thousands of images and finding subtle details that may be overlooked by human observers. However, as quoted in the article, the area now stands as being “long on promise and short on proof,” since very few randomized clinical trials exist that firmly substantiate the role of these algorithms in practical real-world settings. As Burki [1] stressed in the article, it should be clarified that these algorithms should only function as an addition to a human doctor's role and can become a form of what Burki termed as a “buddy system” for diagnostic purposes. The topic of ethics in artificial intelligence algorithms related to respiratory medicine continues to be relevant regarding issues such as privacy and accountability in diagnostics.

In a more recent effort, Mei and colleagues [5] showcased the clinical and technical readiness of AI-enabled diagnostic systems utilizing a multimodal deep learning framework for diagnosing and prognosticating interstitial lung disease (ILD). The proposed model combined deep learning-based image analysis of CT imaging with clinical data and longitudinal data modeled with informativity transformer networks and long short-term memory (LSTM) modules. The model classified 5 subtypes of ILD: usual interstitial pneumonia, nonspecific interstitial pneumonia, chronic hypersensitivity pneumonitis, sarcoidosis, and other variants. It displayed an area under the receiver operating characteristic curve (AUROC) of 0.828 for the subtype classification task and a score of 0.868 for a three-year survival prediction task. Their findings highlighted that the models equalled or exceeded the diagnostic accuracy of experienced thoracic radiologists, substantiating the viability of AI as a clinical decision-support tool. Mei et al., [5] also recognized the initial and exploratory contribution of combining certain indicator variables of radiological features with clinical history to aid personalized prognostication, which represents a shift to dynamic, patient-centered management of disease.

In summation, these researches provide the foundation for the use of artificial intelligence techniques in pulmonary diagnoses as well as providing fundamental guidance on future research. These researches highlight three important considerations. First, based on the mention of automatic multiscale and hierarchical feature extractions

from images by convolutional neural networks, these algorithms are now the new gold standard in the analysis of medical images based on this functionality. Second, ensemble and hybrid models may improve robustness and counter overfitting by considering the variations in medical images as well as patient demographics. Third, of particular significance in reasoning and trust mechanisms, Explainable AI approaches such as those of Grad-CAM and SHAP explanations [7] [8]; should be given consideration.

Although the majority of these efforts have been focused on chest X-ray and CT scans thus far, it should be noted that the approach of either deep learning or explainability can be very effectively incorporated within the assessment of lung ultrasound scanning (LUS), which has been recognized as a non-ionizing and relatively inexpensive alternative approach for assessing the lungs. Since the effectiveness of CT scans within the assessment of pneumonia and COVID-19 has been confirmed by using CNN and transformer models within these types of scans, there is potential that similar models could be adapted for use within an ultrasound assessment. As such the assessment of lung ultrasound scanning may be at the forefront of a new revolution within point-of-care assessments.

### III. METHODOLOGY

The Methodology: In this section, this paper presents the proposed Explainable Multi-scale Deep Learning Framework proposed to automatically classify lung diseases from Ultrasound images. This section will discuss four aspects of the methodology, namely dataset preparation, dataset preprocessing, the structure of the multi-scale model design, a training strategy, cross-validation, and performing An explainability analysis. All experiments conducted were done utilizing TensorFlow/Keras with GPU acceleration.

#### A. Dataset

In this study, the performance of the proposed deep learning models was evaluated using a publicly available Lung Ultrasound (LUS) image dataset developed for automated AI-based lung disease classification. The dataset was compiled from clinical examinations conducted by senior radiologists at Mulago National Referral Hospital and Kiruddu Referral Hospital in Uganda. It comprises 1,062 labeled ultrasound frames collected from 149 patients with an average age of approximately 40.85 years [1]. From a diagnostic standpoint, the dataset includes three primary categories: Probably COVID-19, Diseased Lung but Probably Not COVID-19, and Healthy Lung. All imaging procedures were performed in a clinical setting using Clarius C3 and C5 ultrasound probes, and the dataset has been made publicly accessible through the Mendeley Data repository [11].

#### B. DenseNet121

In this study, the DenseNet121 architecture was selected to mitigate the vanishing gradient problem commonly encountered during the training of deep convolutional neural networks and to maximize feature reuse. DenseNet (Dense Convolutional Network) introduces a densely connected topology in which each layer receives the feature maps of all preceding layers and, in turn, forwards its own outputs to all subsequent layers. Unlike ResNet, where feature aggregation is achieved through element-wise addition, DenseNet employs feature concatenation, enabling the network to preserve information learned in earlier layers and utilize it more effectively. As suggested by its name, DenseNet121 consists of 121 layers and is composed of sequential Dense Blocks interleaved with Transition Layers responsible for spatial downsampling and channel compression. Within each Dense Block, the number of new feature maps introduced by each layer—referred to as the growth rate—is kept constant, thereby improving parameter efficiency while optimizing computational cost. Owing to these architectural strengths, DenseNet121 is particularly well-suited for medical imaging tasks, where dataset sizes are often limited, as it reduces overfitting risk and provides strong representational performance [12].

#### C. MobileNetV2

Another model employed in this study is MobileNetV2, which is specifically designed to achieve high performance on platforms with constrained computational resources, such as mobile and embedded systems. This architecture extends the depthwise separable convolution framework introduced in MobileNetV1 by incorporating two key innovations: inverted residual blocks and linear bottlenecks. Unlike standard residual blocks, the inverted residual structure in MobileNetV2 expands low-dimensional feature tensors to a higher dimensional space, applies a lightweight depthwise convolution, and subsequently projects them back to a low-dimensional representation. To avoid information loss during this projection step, linear bottleneck layers omit nonlinear activation functions. These architectural improvements significantly reduce the number of parameters and computational cost,

measured in FLOPs, while maintaining competitive classification accuracy. Consequently, MobileNetV2 is a strong candidate for real-time medical image analysis applications in which computational efficiency is essential [13].

#### *D. EfficientNetB7*

In this study, the EfficientNetB7 architecture was utilized to enhance model scalability and accuracy through a systematic and principled approach. Unlike conventional deep learning strategies that arbitrarily increase network depth, width, or input resolution, the EfficientNet family employs a compound scaling method that uniformly balances all three dimensions using a fixed compound coefficient. The core building block of the architecture is the Mobile Inverted Bottleneck Convolution (MBConv), which improves computational efficiency and facilitates effective information flow. Furthermore, the network integrates Squeeze-and-Excitation optimization to emphasize salient features that are critical for discrimination. As one of the deepest and most expressive variants in the EfficientNet series, EfficientNetB7 demonstrates strong learning capability on complex datasets due to its high parameter capacity. This architectural choice enables the model to capture fine-grained patterns and disease-specific visual cues in medical images with high accuracy [14].

#### *E. NASNetLarge*

In this study, the NASNetLarge (Neural Architecture Search Network) model was selected to overcome the limitations imposed by manually designed architectures and to automatically discover an optimal network structure. Rather than relying on hand-crafted layers, NASNet employs a reinforcement learning-based search algorithm that explores a predefined search space to identify architectural building blocks, or cells, that yield the highest predictive accuracy. The architecture is primarily composed of two specialized cell types: Normal Cells, which preserve the spatial resolution of the input while focusing on feature extraction, and Reduction Cells, which decrease the spatial dimensions of the feature maps while increasing the number of filters. After the search algorithm determines the optimal configuration of these cells, they are sequentially stacked to construct the final network. The NASNetLarge variant used in this study represents the highest-capacity configuration in the NASNet family, characterized by an increased number of cell repetitions and filters. Owing to its large parameter count and complex connectivity pattern, the model demonstrates a strong capability to capture subtle and intricate patterns in medical imaging data [15].

#### *F. Grad-CAM*

To enhance the transparency of deep learning models and to verify the clinical validity of their decision-making processes, this study employed the Gradient-Weighted Class Activation Mapping (Grad-CAM) technique. Grad-CAM enables visualization of the regions within an input image that a trained convolutional neural network focuses on when generating a specific class prediction, without requiring any architectural modification or retraining of the model. The method operates by computing the gradient of the target class score with respect to the feature maps of the final convolutional layer. These gradients are then subjected to global average pooling to obtain neuron importance weights for each feature map. A weighted linear combination of the feature maps is subsequently computed, and the result is passed through a ReLU activation function to retain only the positively contributing features. The resulting heatmaps are superimposed on the original input image to visually analyze the alignment between the model's attention regions and radiologically relevant anatomical or pathological structures, such as pleural irregularities or B-lines [7].

#### *G. Performance Metrics*

To quantitatively assess the classification performance of the developed deep learning models and to enable objective comparisons across architectures, standard confusion matrix-based evaluation metrics were employed. In this study, model effectiveness was analyzed using Accuracy, Sensitivity (Recall), Specificity, and F1-Score. Accuracy represents the proportion of correctly predicted positive and negative instances relative to the total number of samples, serving as a general indicator of overall performance. However, due to the potential presence of class imbalance in medical diagnosis, Accuracy alone may be insufficient. For this reason, Sensitivity was used to measure the proportion of actual positive cases correctly identified by the model, while Specificity quantified the model's ability to accurately distinguish healthy (negative) instances—both of which are critical metrics in clinical applications.

Additionally, Precision, which reflects the proportion of predicted positive cases that are truly positive, was incorporated into the analysis through the computation of the F1-Score, defined as the harmonic mean of Precision and Recall. The F1-Score is particularly valuable in scenarios where a balanced evaluation of false positives and false negatives is required, providing a robust measure of model stability. All metric calculations were performed in accordance with their respective mathematical formulations using the model's true positive, true negative, false positive, and false negative outputs [16].

#### IV. RESULTS

The performance of the proposed deep learning models—DenseNet121, MobileNetV2, and EfficientNetB7—was comprehensively analyzed for the task of lung ultrasound image classification. The classification capabilities of these models were evaluated using standard metrics, including Accuracy, Precision, Recall, and F1-Score. All models were tested using a 5-fold cross-validation procedure, and the average performance metrics obtained across all folds are summarized in Table 1.

Table 1: Performance comparison of the deep learning models on the LUS classification task

<b>Model</b>	<b>Accuracy</b>	<b>Precision</b>	<b>Recall</b>	<b>F1-Score</b>
<b>EfficientNetB7</b>	0.7580	0.7560	0.7563	0.7520
<b>MobileNetV2</b>	0.7571	0.7555	0.7569	0.7514
<b>DenseNet121</b>	0.6826	0.6772	0.6782	0.6727

The experimental results reveal notable differences in overall accuracy and F1-scores among the evaluated models. As shown in Table 1, EfficientNetB7 achieved the highest performance, with a test accuracy of 75.80% and an F1-score of 75.20%. Its compound scaling strategy and deep architectural structure enabled the model to effectively distinguish complex patterns present in ultrasound images. Following closely, MobileNetV2 demonstrated a competitive performance with an accuracy of 75.71% and an F1-score of 75.14%. The accuracy gap between the two models is statistically negligible (0.09%), indicating that MobileNetV2's lightweight and optimized architecture, based on inverted residual blocks, is capable of approaching the performance of models with significantly larger parameter counts. In contrast, DenseNet121 yielded the lowest performance among the compared architectures, attaining an accuracy of 68.26% and an F1-score of 67.27%. Despite the advantages associated with dense connectivity, the model showed weaker generalization on this dataset relative to the other two approaches. Overall, EfficientNetB7 and MobileNetV2 outperformed DenseNet121 by approximately 7.54% and 7.45% in accuracy, respectively, demonstrating that these architectures are more suitable for the lung ultrasound classification task addressed in this study.

To further investigate the discriminatory capability of the models for each disease category and to characterize the nature of misclassifications, predictions on the test dataset were analyzed using Confusion Matrices, as illustrated in Figure 1. Given that the classification task consisted of three primary categories—COVID-19, Healthy, and Other Lung Diseases—the class-specific behaviors of the models exhibited the following patterns. Across all models, the highest classification performance was observed for the Healthy class. MobileNetV2 achieved the strongest sensitivity in this category, correctly identifying 312 out of 362 healthy lung images. EfficientNetB7 and DenseNet121 also performed well in distinguishing healthy subjects, yielding 302 and 290 correct predictions, respectively. These results indicate that the characteristic appearance of healthy lung tissue in ultrasound images, including patterns such as A-lines, is more easily separable from pathological cases. In contrast, the most challenging distinction for the models occurred between the COVID-19 and Other Lung Disease categories. From a clinical perspective, both conditions may present similar sonographic findings, including B-lines and pleural irregularities, which contributed to a higher degree of confusion between these two classes.

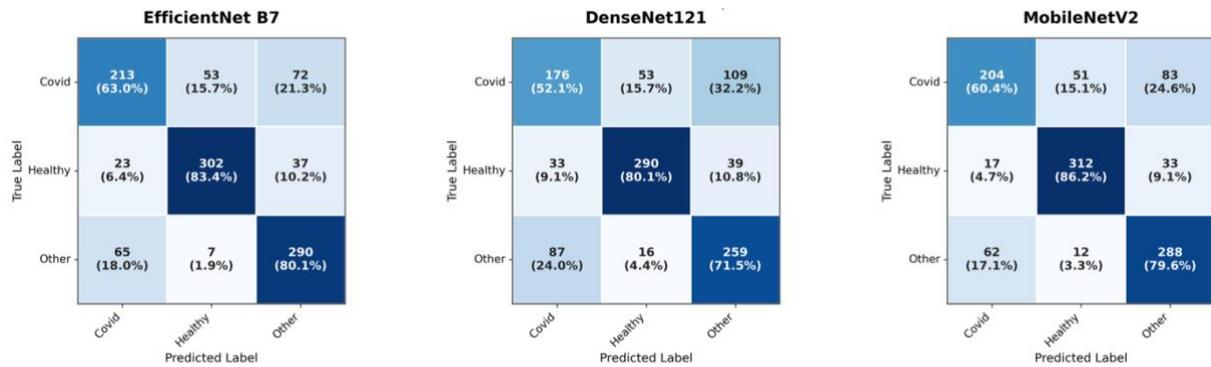


Figure 1: Confusion Matrix Comparison

The confusion matrix for DenseNet121 reveals that the primary cause of its lower performance stems from a high false-negative rate in the COVID-19 class. Out of 338 COVID-19 cases, DenseNet121 correctly identified only 176 and misclassified 109 cases as Other Lung Disease. This outcome suggests that the model struggled to reliably differentiate viral pneumonia from other pathological conditions. The analysis of MobileNetV2 shows that despite its strong performance in the Healthy class, it misclassified 83 COVID-19 cases as belonging to the other category. Nevertheless, the model provided more balanced performance compared to DenseNet121, correctly identifying 204 COVID-19 samples. EfficientNetB7, which achieved the highest overall performance, exhibited the lowest degree of inter-class confusion. It correctly identified 213 COVID-19 cases, yielding the strongest sensitivity within this class, while limiting confusion between COVID-19 and Other Lung Diseases to 72 instances. Additionally, the model correctly classified 290 out of 362 cases in the other category, demonstrating the most effective separation of pathological classes among the evaluated architectures.

To support the clinical reliability of the deep learning models' high classification performance, the anatomical regions influencing model decisions were visualized using the Gradient-weighted Class Activation Mapping (Grad-CAM) technique. In this analysis, activation maps derived from the final convolutional layers of EfficientNetB7, MobileNetV2, and DenseNet121 were projected onto the original ultrasound images to examine the spatial focus of each model. Grad-CAM outputs are shown in the Figure 2.

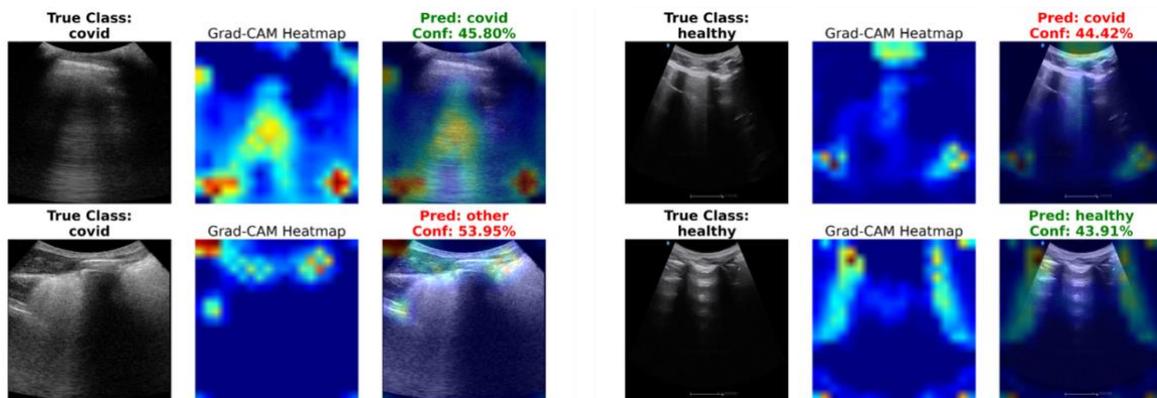


Figure 2: Grad-CAM visualizations illustrating the regions influencing model predictions for COVID-19 and Healthy lung ultrasound images

The resulting heatmaps indicate that the models concentrated on radiologically meaningful cues when generating their predictions. In particular, EfficientNetB7 and MobileNetV2, which demonstrated the strongest performance, consistently focused on vertical artifacts such as B-lines, pleural line irregularities, and areas of consolidation in samples belonging to the COVID-19 and Other Lung Disease classes. The fact that these models emphasized the lung parenchyma and pleural line, rather than background noise or hospital labels, confirms that the learning process was guided by relevant clinical patterns. For images in the Healthy class, activation intensity was predominantly located around horizontal artifacts such as A-lines and the pleural line, which are characteristic indicators of normal lung structure. In contrast, DenseNet121 exhibited more diffuse or misplaced activation in some test cases, occasionally extending beyond pathological regions. This behavior provides a visual explanation

for its lower classification performance and the higher degree of confusion observed between COVID-19 and Other Lung Disease categories.

## V. CONCLUSION AND FUTURE WORKS

This study presents a comprehensive deep learning framework for the automated detection of lung diseases from lung ultrasound (LUS) images, integrating both classification and explainability components. The performances of DenseNet121, MobileNetV2, and EfficientNetB7 were systematically evaluated, and their decision-making processes were examined through Grad-CAM visualizations. Given the operator-dependent nature of LUS interpretation and its sensitivity to probe positioning, patient anatomy, and imaging artifacts, automated analysis offers a promising solution for achieving more consistent and objective diagnostic assessments. Experimental findings demonstrate that EfficientNetB7 achieved the highest accuracy and F1-score, indicating its strong capability in distinguishing complex pathological patterns. MobileNetV2, despite its substantially smaller parameter footprint, produced performance levels comparable to EfficientNetB7, highlighting its suitability for portable ultrasound devices and resource-constrained clinical environments. DenseNet121 showed comparatively lower effectiveness in differentiating COVID-19 from other pneumonic conditions. The Grad-CAM analyses further confirmed that the models, particularly EfficientNetB7 and MobileNetV2, attended to clinically relevant regions—such as B-lines, pleural irregularities, and consolidation patterns—supporting their potential reliability in clinical decision support. Overall, the proposed framework underscores the viability of deep learning-based LUS analysis as an assistive tool in emergency departments, triage settings, and low-resource environments where access to expert radiologists may be limited. The system’s ability to improve diagnostic consistency and reduce inter-operator variability positions it as a valuable component of future point-of-care ultrasound workflows.

Future research will focus on enhancing model generalizability through the use of larger, multi-center datasets representing broader demographic and pathological variability. The evaluation of advanced architectures, including Vision Transformers (ViTs) and hybrid CNN–ViT models, will also be explored. Incorporating multimodal information—such as patient symptoms, clinical notes, and demographic data—may further improve diagnostic performance. Lastly, deploying the proposed models within a real-time mobile or web-based interface and validating them in prospective clinical studies will be essential steps toward translating this technology into routine clinical practice.

## REFERENCES

- [1] Burki, T. K. (2019). The role of AI in diagnosing lung diseases. *The Lancet Respiratory Medicine*, 7(12), 1015–1016.
- [2] Dack, E., Christe, A., Fontanellaz, M., Brigato, L., Heverhagen, J. T., Peters, A. A., Huber, A. T., Hoppe, H., Mougiakakou, S., & Ebner, L. (2023). Artificial intelligence and interstitial lung disease: Diagnosis and prognosis. *Investigative Radiology*, 58(8), 602–609.
- [3] Koyyada, S. P., & Singh, T. P. (2023). An explainable artificial intelligence model for identifying local indicators and detecting lung disease from chest X-ray images. *Healthcare Analytics*, 4, 100206.
- [4] Lamba, R. (2025). Advances in AI for medical imaging: A review of machine and deep learning in disease detection. *Procedia Computer Science*, 260, 262–273.
- [5] Mei, X., Liu, Z., Singh, A., et al. (2023). Interstitial lung disease diagnosis and prognosis using an AI system integrating longitudinal data. *Nature Communications*, 14, 2272.
- [6] Mostafa, F. A., Elrefaei, L. A., Fouda, M. M., & Hossam, A. (2022). A survey on AI techniques for thoracic diseases diagnosis using medical images. *Diagnostics*, 12(12), 3034.
- [7] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (pp. 618–626).
- [8] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems (NeurIPS)*, 30, 4768–4777.
- [9] Morshed, A., Al Shihab, A., Jahin, M. A., Al Nahian, M. J., Hossain Sarker, M. M., Ibne Wadud, M. S., ... & Hussain, M. A. (2025). Ultrasound-based AI for COVID-19 detection: a comprehensive review of public and private lung ultrasound datasets and studies. *Multimedia Tools and Applications*, 1-55.
- [10] Ilse, M., Tomczak, J., & Welling, M. (2018, July). Attention-based deep multiple instance learning. In *International conference on machine learning* (pp. 2127-2136). PMLR.
- [11] Katumba, A., Murindanyi, S., Okila, N., Nakatumba-Nabende, J., Mwikirize, C., Serugunda, J., ... & Nabawanuka, E. (2025). A dataset of lung ultrasound images for automated ai-based lung disease classification. *Data in Brief*, 112034.
- [12] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
- [13] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4510-4520).
- [14] Tan, M., & Le, Q. (2019, May). Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning* (pp. 6105-6114). PMLR.
- [15] Zoph, B., Vasudevan, V., Shlens, J., & Le, Q. V. (2018). Learning transferable architectures for scalable image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 8697-8710).
- [16] Fawcett, T. (2006). An introduction to ROC analysis. *Pattern recognition letters*, 27(8), 861-874.

### RESUME

#### **Taha CAPAR**

Taha Capar is a Computer Science and Artificial Intelligence student at Queen Mary University of London, with interests in machine learning, predictive modeling, and computational methods. He has contributed to projects involving algorithmic analysis, data-driven modeling, and the application of AI techniques in areas such as healthcare systems. Capar's academic focus includes applied machine learning, autonomous systems, and the practical implementation of AI-based solutions in real-world contexts.

#### **Yusuf ERYESIL**

Yusuf Eryesil is a Research Assistant and Ph.D. candidate in Computer Engineering at Selçuk University. His research focuses on artificial intelligence, specifically in the realms of deep learning and machine learning applications. He has contributed to projects involving AI-supported waste recognition and management systems. Eryeil's academic work includes developing optimization algorithms for solid waste classification, as well as applying deep learning frameworks to areas such as food quality analysis and medical diagnostics.

# Feature Selection in Medical Data Using the Crested Porcupine Algorithm and the Island Model

Gulnur YILDIZDAN<sup>1</sup>

<sup>1</sup> Selcuk University, Kulu Vocational School, Konya/Türkiye, [gavsar@selcuk.edu.tr](mailto:gavsar@selcuk.edu.tr)

## I. INTRODUCTION

Feature selection is essential for achieving high-accuracy classification in machine learning applications. In this process, metaheuristic algorithms are often used to identify the most relevant features in large datasets to enhance model performance and lower computational costs. In this study, the Crested Porcupine Optimizer (CPO)[1], a nature-inspired method, is employed for feature selection and combined with the Island Model (IM)[2] to increase solution diversity and prevent premature convergence through local optimization. Each island functions as an independently evolving subpopulation, with information sharing occurring via periodic migrations. The classification performance of the chosen features is assessed using the K-Nearest Neighbor (KNN)[3] algorithm. This study aims to investigate the impact of the island model and migration frequency on the efficacy of CPO's feature selection and the accuracy of KNN classification. The proposed approach's effectiveness is demonstrated through experiments on lung cancer data.

Recent research in the literature can be condensed as follows: Li and Gonsalves[4] developed a new method by combining the Island Genetic Algorithm, Fireworks Algorithm, and Particle Swarm Optimization in parallel with the island model. Prado-Rodríguez et al.[5] proposed an adaptive parameter control for Island-Based Ant Colony Optimization, reducing the need for manual tuning and improving the performance of the algorithm. Dokeroglu et al.[6] introduced a new ensemble metaheuristic method for the Graph Coloring Problem, which combines Harris Hawk Optimization, Teaching-Learning-Based Optimization, and Artificial Bee Colony algorithms using the parallel island model. Thaher et al.[7] presented the iECSA algorithm, an enhanced version of the Crow Search Algorithm (CSA) that utilizes a cooperative island model. This algorithm features adaptive selection and an improved movement operator, which helps prevent premature convergence. In experiments conducted on 53 test problems, iECSA demonstrated superior performance compared to many existing algorithms. Duarte et al.[8] proposed the Stigmergy Island Model, combining the island model for the parallel implementation of evolutionary algorithms with stigmergy, a natural collaboration mechanism. This structure, which enables different algorithms to work in an organized and collaborative manner, has been tested on 15 optimization problems and has yielded successful results.

The remaining sections of the study are organized as follows: Section 2 details the materials and methods used in the developed approach. Section 3 provides the experimental study and reviews its results. Section 4 interprets the conclusions and explores future work.

## II. MATERIAL AND METHODS

This section offers an in-depth examination of the dataset utilized in the study, the metaheuristic methods employed for feature selection, and the classification techniques associated with the island model.

### A. Feature Selection Method

Feature selection is a common preprocessing step used to enhance the accuracy of machine learning models and minimize noise from irrelevant variables. In this process, metaheuristic algorithms, known for their efficient search abilities in complex solution spaces, are regarded as a significant alternative. In this study, the Crested Porcupine Optimizer (CPO), recently introduced by Abdel-Basset et al.[1], is utilized for feature selection. The main steps of the algorithm can be summarized as follows.

- *Initialization*

The CPO algorithm, like most metaheuristics, starts with a random population of candidate solutions. In this process, solutions are generated using lower (lb) and upper (ub) bounds and a random number in the range [0,1]. Equation 1 mathematically expresses this initialization mechanism.

$$X_i = lb + rand \times (ub - lb) \quad i = 1, 2, \dots, pop \quad (1)$$

The CPO algorithm utilizes the Cyclic Population Reduction (CPR) method, characterized by a cyclical decrease in population throughout the optimization process. This method entails a systematic reduction in population size at each iteration. This process is explained by Equation 2.

$$pop = pop_{min} + (pop_{max} - pop_{min}) \times \left( 1 - \left( \frac{iter \% max\_iter}{\frac{T}{max\_iter}} \right) \right) \quad (2)$$

In Equation 2, the population size ( $pop$ ) at each iteration varies cyclically between the minimum ( $pop_{min}$ ) and maximum ( $pop_{max}$ ) values. This variation is determined by the loop number  $T$ , the number of maximum iterations ( $max\_iter$ ), and the modulus operator  $\%$ . In the CPO algorithm, the transition to the exploration or exploitation phase is determined by comparing the randomly generated values Q1 and Q2, both of which are between 0 and 1. Exploration occurs when  $Q1 < Q2$ ; otherwise, the exploitation phase begins.

- *Exploration Phase*

In the CPO algorithm, the exploration phase includes two different position update strategies that model the two defensive behaviors exhibited by the individual when away from predators: vision and sound defense. Which strategy is applied depends on the randomly generated values of Q3 and Q4; if  $Q3 < Q4$ , the first strategy is selected; otherwise, the second strategy is selected. The update formula for the first strategy is presented in Equation 3.

$$X_i^{iter+1} = X_i^{iter} + \tau_1 \times |2 \times \tau_2 \times X_{CP}^{iter} - y_i^{iter}| \quad (3)$$

In Equation 3,  $X_{CP}^{iter}$  represents the optimal solution,  $X_i^{iter}$  indicates the current individual's position,  $\tau_1$  is a random variable that follows a normal distribution, and  $\tau_2$  is a value between 0 and 1. Equation 4 defines the location of the hunter, denoted by  $y_i^{iter}$ , where  $X_r^{iter}$  is the location of a randomly selected individual,  $r$  determined randomly in the interval  $[1, pop]$ .

$$y_i^{iter} = \frac{X_i^{iter} + X_r^{iter}}{2} \quad (4)$$

The equation for updating position in the second defense is presented in Equation 5.

$$X_i^{iter+1} = (1 - U_1) \times X_i^{iter} + U_1 \times (y_i^{iter} + \tau_3 \times (X_{r1}^{iter} - X_{r2}^{iter})) \quad (5)$$

In Equation 5,  $r1$  and  $r2$  denote two randomly chosen individuals from the population, whereas  $\tau_3$  signifies a random number ranging from 0 to 1.  $U_1$  is a binary random vector composed only of 0s and 1s, and  $X_{r1}^{iter}$  and  $X_{r2}^{iter}$  denote two random solutions within the population.

- *Exploitation Phase*

CPO algorithm exploitation involves two position updates that are based on the third and fourth defensive strategies for local search. If  $Q5$  is less than 0.5, the third strategy is applied; otherwise, the fourth strategy is used. This phase is illustrated by Equations 6–9.

$$X_i^{iter+1} = (1 - U_1) \times X_i^{iter} + U_1 \times (X_{r1}^{iter} + S_i^{iter} \times (X_{r2}^{iter} - X_{r3}^{iter}) - \tau_4 \times \delta \times \gamma_t \times S_i^{iter}) \quad (6)$$

$$S_i^{iter} = \exp \left( \frac{Obj(X_i^{iter})}{\sum_{k=1}^{pop} f(X_k^{iter}) + \epsilon} \right) \quad (7)$$

$$\delta = \begin{cases} 1 & \text{if } rand < 0.5 \\ -1 & \text{else} \end{cases} \quad (8)$$

$$\gamma_t = 2 \times rand \times \left( 1 - \frac{iter}{max\_iter} \right) \frac{iter}{max\_iter} \quad (9)$$

The equations include a random variable  $\tau_4$  (between 0 and 1), the objective function value of the  $i$ th individual at the  $iter$ th iteration ( $Obj(X_i^{iter})$ ), a small constant  $\epsilon$  to avoid division by zero, and a random vector ( $rand$ ). Equations 10–12 delineate the methodology for updating the position associated with the fourth defense strategy, where  $\alpha$  is

the convergence rate,  $\tau_5$  and  $\tau_6$  are random values between 0 and 1, and  $\tau_7$  is a random vector with values in [0,1]. Here,  $X_i^{iter+1}$  is the new position of the  $i$ th individual, and  $V_i^{iter}$  equals  $X_i^{iter}$ .

$$X_i^{iter+1} = X_{CP}^{iter} + (\alpha(1 - \tau_5) + \tau_5) \times (\delta \times X_{CP}^{iter} - X_i^{iter}) - \beta \quad (10)$$

$$\beta = \tau_6 \times \delta \times \gamma_t \times D_i^{iter} \quad (11)$$

$$D_i^{iter} = \tau_7 \times S_i^{iter} \times (V_i^{iter+1} - V_i^{iter}) \quad (12)$$

Algorithm 1 displays the pseudocode of CPO.

Algorithm 1: Pseudo-code of CPO

Start CPO		
<b>Input:</b> Algorithm parameters		
1:	Set $pop$ , $max\_iter$ , $\alpha$ , $T$ , etc.	18: Else // Exploitation phase
2:	Initialize the solutions' positions randomly	19: Generate a random number, $Q5$
3:	While ( $iter \leq max\_iter$ )	20: If $Q5 < 0.5$ //Third defense mechanism
4:	Evaluate fitness values for the candidate solutions	21: Using Eq.(6)
5:	Determine the best solution( $X_{CP}^{iter}$ )	22: Else //Fourth defense mechanism
6:	Update $\gamma_t$ using Eq. (9)	23: Using Eq.(10)
7:	Update the population size using Eq. (2)	24: End If
8:	For $i=1: pop$	25: End If
9:	Update the $m$ , $S$ , $Obj$ , $\delta$	26: If $f(X_i^{iter+1}) > f(X_i^{iter})$
10:	Generate two random numbers, $Q1$ and $Q2$	27: $X_i^{iter+1} = X_i^{iter}$
11:	If $Q1 < Q2$ // Exploration phase	28: End If
12:	Generate two random numbers, $Q3$ and $Q4$	29: $iter = iter + 1$
13:	If $Q3 < Q4$ // First defense mechanism	30: End for
14:	Using Eq.(3)	31: End while
15:	Else //Second defense mechanism	<b>Output:</b> The best solution
16:	Using Eq.(5)	<b>End CPO</b>
17:	End If	

### B. Island Model

The island model is a sophisticated parallel optimization strategy that aims to enhance the effectiveness of population-based metaheuristic algorithms. This approach divides the solution space into multiple subpopulations, known as islands, each of which evolves independently. A migration mechanism facilitates the exchange of the best individuals among the islands at regular intervals, thereby maintaining genetic diversity and mitigating the risk of premature convergence. Various parameters, including migration frequency, the number of individuals participating in migration, and selection methods, play a crucial role in balancing the interactions among the islands. This balance, in turn, optimizes both exploration and exploitation capabilities within the algorithm. The parallel structure inherent in the island model provides efficient and stable solutions to complex, high-dimensional problems, and is well-suited for implementation within parallel computing infrastructures [9, 10].

### C. Classifier

The K-Nearest Neighbor (KNN) algorithm is a simple yet powerful supervised learning technique frequently employed in both classification and regression applications [11]. This algorithm has a non-parametric, example-based structure; it memorizes training data and classifies new instances based on their closest neighbors. KNN identifies the  $k$  nearest neighbors of a data point and predicts the class label according to these neighbors. Distance metrics such as Euclidean, Manhattan, and Minkowski are employed to measure similarity [12]. The most popular metric, Euclidean distance ( $D_E$ ), is determined by taking the square root of the sum of the squared differences

between two data points ( $v_i, v_j$ ). This process is described by the following Equation 13.

$$D_E = \sqrt{\sum_{i=1}^n (v_i - v_j)^2} \quad (13)$$

Where  $v_{ik}$  and  $v_{jk}$  are the  $k$ th features of the vectors  $v_i$  and  $v_j$ , respectively. This method provides effective classification performance, especially on low-dimensional and balanced datasets [13].

*D. Dataset*

The lung cancer dataset [14] used in this study contains clinical and demographic information from patients who have undergone thoracic surgery. It consists of 17 features and a binary target variable indicating mortality within one year of surgery. This publicly available dataset is widely used in medical machine learning research and is suitable for evaluating feature selection methods in the healthcare field.

III. EXPERIMENT RESULTS

In this study, metaheuristic algorithms were employed to optimize feature selection and classification processes. Specifically, based on the CPO algorithm developed with the proposed island model (IM) structure, the algorithm's performance was tested using different island counts (e.g., 1, 3, 5, and 10). During evaluation, metrics such as classification accuracy from the K-nearest neighbor (KNN) classifier, the number of selected features, and the overall fitness value were considered. The KNN algorithm, a non-parametric and sample-based method, was chosen because it offers consistent classification results. The Euclidean distance metric was used as the distance measure. The fitness function in feature selection is a multi-criteria structure designed to improve classification performance while reducing the number of features. In this context, the fitness function is defined as Equation 14 [15].

$$Fitness = \alpha \times (1 - A) + \beta \times \left(\frac{|SF|}{|TF|}\right) \quad (14)$$

$A$  represents the accuracy rate of the KNN classifier;  $SF$  represents the number of selected features; and  $TF$  represents the total number of features.  $\alpha$  and  $\beta$  are weight coefficients that control the balance between accuracy and the number of features. It is determined as  $\alpha + \beta = 1$ . In this study, priority was given to classification performance, and  $\alpha = 0.9$  and  $\beta = 0.1$  were assigned. By using these metrics, the aim is to achieve both high accuracy and fewer feature selections.

In the first experiment, the effect of the number of islands was examined to evaluate the performance of the metaheuristic algorithm. Holding the total population size constant (300 individuals), the number of individuals was distributed equally across each island. Additionally, elite individual migration was preferred, and in each migration cycle, the 3 best individuals from each island were sent to different randomly selected islands. This analysis assessed how the number of islands affected fitness, the number of selected features, and KNN classification success. The results are presented in Table 1 and illustrated in Figure 1.

Table 1: Model performance values based on number of islands

Islands	Selected Features	Best Fitness	KNN Accuracy
1	6	0.8211	0.8248
3	7	0.8204	0.8321
5	7	<b>0.8277</b>	<b>0.8394</b>
10	10	0.8186	0.8102

As can be seen from the table, the highest fitness and classification accuracy were achieved when using 5 islands. While performance improved compared to fewer islands, there was a significant decrease in accuracy with 10 islands. These results indicate that the number of islands, and therefore the number of individuals on each island, affects optimization success.

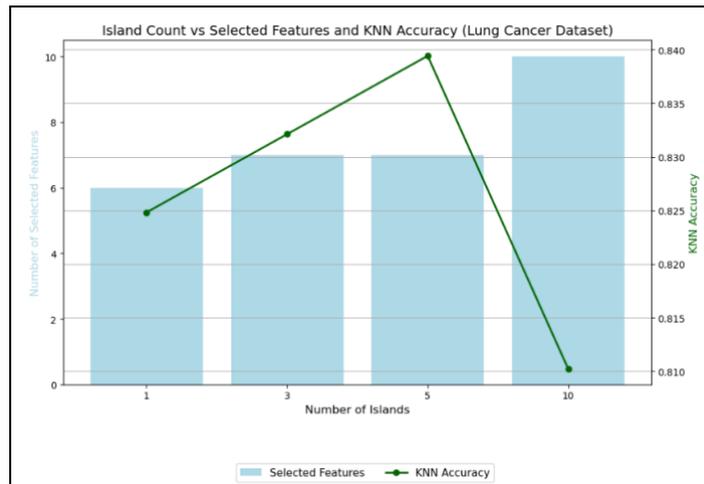


Figure 1: Comparison of model performance metrics based on number of islands

In the second experiment, the effect of migration rate on the algorithm's success was evaluated. For this purpose, under a fixed number of islands (5 was selected), different numbers of individuals (1, 3, 5, and 10) were allowed to migrate between islands. In each scenario, the total number of individuals was kept constant at 300. The results are presented in Table 2 and illustrated in Figure 2. The best result, with the highest classification accuracy (83.21%) and a competitive fitness value (0.828367), was achieved with the migration of 10 individuals. A reasonable number of features (6 features) were also selected in this scenario. On the other hand, although a lower fitness value was achieved with the migration of 3 individuals (0.815643), the classification accuracy decreased, and only 3 features were selected. This situation can be explained as information loss due to oversimplification.

Table 2: Model performance values based on number of migrants count

Migrants	Selected Features	Best Fitness	KNN Accuracy
1	11	0.817943	0.824818
3	3	0.815643	0.817518
5	8	0.827117	0.817518
10	6	<b>0.828367</b>	<b>0.832117</b>

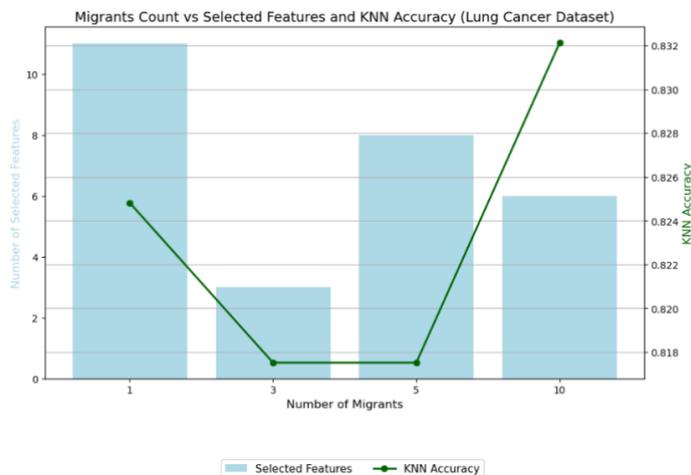


Figure 2: Comparison of model performance metrics based on number of migrants count

Consequently, increasing the migration rate increased information sharing, resulting in better generalization and classification performance. In this context, 10 individual migration stands out as the most appropriate strategy in terms of both maintaining the balance between exploration and exploitation and overall algorithm success.

#### IV. CONCLUSION

In this study, the Crested Porcupine Optimizer, supported by the island model, was evaluated on the feature selection problem, and its performance was analyzed with the KNN classifier. In the first experiment, the effect of the number of islands was examined, and the highest classification accuracy and fitness values were achieved with 5 islands. In the second experiment, the best classification success was observed when the number of migrating individuals was increased, especially with 10 migrating individuals. These results indicate that the island model prevents premature convergence by preserving diversity and increases classification accuracy.

Future studies can further investigate the parameter sensitivity of the CPO algorithm, the impact of defense strategies, and the exploration-exploitation balance. Furthermore, enriching the algorithm with dynamic migration strategies and testing it with different classifiers could be important steps to increase the generalizability of the method.

#### References

- [1] M. Abdel-Basset, R. Mohamed, M. Abouhawwash, Crested Porcupine Optimizer: A new nature-inspired metaheuristic. *Knowledge-Based Systems*, (2024). 284: p. 111257.
- [2] D. Izzo, M. Ruciński, F. Biscani, The generalized island model, in *Parallel architectures and bioinspired algorithms*. 2012, Springer. p. 151-169.
- [3] A. Mucherino, P.J. Papajorgji, P.M. Pardalos, k-Nearest Neighbor Classification, in *Data Mining in Agriculture*, A. Mucherino, P.J. Papajorgji, P.M. Pardalos, Editors. 2009, Springer New York: New York, NY. p. 83-106.
- [4] J. Li, T. Gonalves, Parallel Hybrid Island Metaheuristic Algorithm. *IEEE Access*, (2022). 10: p. 42268-42286.
- [5] R. Prado-Rodríguez, P. González, J.R. Banga, A Parameter Control Strategy for Parallel Island-Based Metaheuristics. *Expert Systems*, (2025). 42(6): p. e70061.
- [6] T. Dokeroglu, T. Kucukyilmaz, A. Cosar, An island-parallel ensemble metaheuristic algorithm for large graph coloring problems. *arXiv preprint arXiv:2504.15082*, (2025).
- [7] T. Thaher, A. Sheta, M. Awad, M. Aldasht, Enhanced variants of crow search algorithm boosted with cooperative based island model for global optimization. *Expert Systems with Applications*, (2024). 238: p. 121712.
- [8] G.R. Duarte, A.C. de Castro Lemonge, L.G. da Fonseca, B.S.L.P. de Lima, An Island Model based on Stigmergy to solve optimization problems. *Natural Computing*, (2021). 20(3): p. 413-441.
- [9] E. Cantú-Paz, Migration policies, selection pressure, and parallel evolutionary algorithms. *Journal of heuristics*, (2001). 7(4): p. 311-334.
- [10] E. Alba, M. Tomassini, Parallelism and evolutionary algorithms. *IEEE transactions on evolutionary computation*, (2002). 6(5): p. 443-462.
- [11] S. S. Chakravarthy, N. Bharanidharan, and H. Rajaguru, "Deep learning-based metaheuristic weighted k-nearest neighbor algorithm for the severity classification of breast cancer," *IRBM*, vol. 44, no. 3, p. 100749, 2023.
- [12] N. Mukahar. Performance comparison of k nearest neighbor classifier with different distance functions. in *AIP Conference Proceedings*. 2024. AIP Publishing LLC.
- [13] H. Rajaguru, S.C. SR, Analysis of decision tree and k-nearest neighbor algorithm in the classification of breast cancer. *Asian Pacific journal of cancer prevention: APJCP*, (2019). 20(12): p. 3777.
- [14] M. Lubicz, K. Pawelczyk, A. Rzechonek, J. Kolodziej, *Thoracic Surgery Data*. UCI Machine Learning Repository, (2014).
- [15] E. A. Mohamed, M. S. Braik, M. A. Al-Betar, and M. A. Awadallah, "Boosted spider wasp optimizer for high-dimensional feature selection," *Journal of Bionic Engineering*, vol. 21, no. 5, pp. 2424-2459, 2024.

# Angle Measurement in Orthoradiogram Images Using Deep Learning

Melih GULEC<sup>1</sup>, Nurettin DOGAN<sup>2</sup>, Hasan Erdinc KOCER<sup>3</sup>

<sup>1</sup>Selçuk University Graduate School of Natural and Applied Sciences Department of Computer Engineering, Selçuklu, Konya/Türkiye, melih.gulec14@gmail.com

<sup>2</sup>Selçuk University Faculty of Technology Department of Computer Engineering, Selçuklu, Konya/Türkiye, ndogan@gmail.com

<sup>3</sup>Selçuk University Faculty of Technology Department of Computer Engineering, Selçuklu, Konya/Türkiye, ekocer@selcuk.edu.tr

## I. INTRODUCTION

In recent years, artificial intelligence has become a technology that continues to develop and is widely used in many fields, including science, industry, healthcare, and daily life. Owing to its ability to process large datasets, its high computational power, and its sophisticated algorithms, artificial intelligence supports decision-making processes and contributes to humanity by reducing error rates.

Artificial intelligence is defined as a computational technology that emulates cognitive functions typically associated with human intelligence, including learning, problem-solving, and logical reasoning. Key implementations of this technology comprise machine learning, deep learning, and natural language processing [1].

Among its areas of application are sectors such as finance, education, automotive, media, and defense industry; moreover, the importance of artificial intelligence in the field of healthcare is steadily increasing [2]. In healthcare, artificial intelligence is effectively utilized in support systems for patient diagnosis, personalized treatment planning, and the analysis of medical imaging methods [2, 3].

In this context, one of the analyses that can benefit from the power of artificial intelligence is deformity analysis of the lower extremity. Lower extremity deformity analysis involves the systematic examination and measurement of deviations, as well as angular and rotational abnormalities, occurring along the mechanical and anatomical alignment lines of the lower extremity. This application holds critical importance for identifying the causes of pain, functional losses, and deformities experienced by patients [4].

In daily practice, this analysis is performed on radiological images obtained from the patient by identifying specific landmarks on designated bones and calculating the angles formed between them. In traditional methods, the interpretation of X-ray images is conducted by drawing lines on the radiograph to assess the presence of any deformity. During this evaluation process, precise angular measurement tools such as goniometers may be used directly on the films or on digital images projected onto a computer screen. Alternatively, the measurements can also be performed manually through specialized software [5].

Deformity analysis of the lower extremity is crucial for the identification and management of various diseases. Conditions such as genu varum, genu valgum, accurate surgical planning, evaluation of fractures resulting from bone integrity loss, post-traumatic malunions, and bone deformities associated with growth are assessed using this method [4, 6-9].

Multiple angles are utilized during lower extremity deformity analysis, which vary depending on whether the mechanical or anatomical axis is used. Angles calculated along the mechanical axis include mechanical lateral proximal femoral angle (mLPFA), mechanical lateral distal femoral angle (mLDFA), mechanical medial proximal tibial angle (mMPTA), and mechanical lateral distal tibial angle (mLDTA), while those along the anatomical axis include anatomical medial proximal femoral angle (aMPFA), anatomical medial neck-shaft angle (aMNSA), anatomical lateral distal femoral angle (aLDFA), anatomical medial proximal tibial angle (aMPTA), and anatomical lateral distal tibial angle (aLDTA) [10]. In the scope of this study, the angles calculated were those of the mechanical axis.

The mLPFA is formed by drawing a line between the center of the femoral head and the tip of the greater trochanter, creating an angle with the mechanical axis of the femur. Its normal range is 85–95 degrees [10].

The mLDFA is determined by connecting the center of the femoral head to the center of the distal femoral joint surface and then drawing an orientation line from the inferior point of the femoral condyles. Its normal range is 85–90 degrees [10].

The mMPTA is calculated by first drawing the mechanical axis of the tibia by connecting the center of the proximal tibial joint surface to the center of the distal tibial joint surface and then forming an angle with the orientation line drawn through the inferior subchondral points of the tibial plateau. Its normal range is 85–90

degrees, and the aMPTA and mMPTA angles are equal [10].

The mLDTA angle is determined by first drawing the orientation line of the distal tibial joint, followed by the mechanical axis of the tibia. The angle formed laterally is the mLDTA angle, which ranges from 86–92 degrees. The mLDTA and aLDTA angles are equal [10].

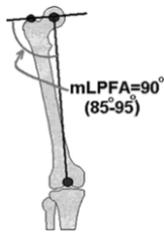


Figure 1: mLPFA [10].

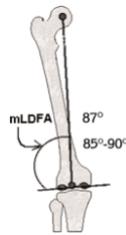


Figure 2: mL DFA [10].

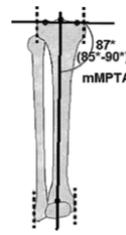


Figure 3: mMPTA [10].

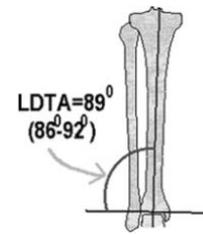


Figure 4: LD TA [10].

The angles used in deformity analysis provide important information regarding the patient's orthopedic condition. Since this process requires manual effort at every stage, the time allocated per patient increases. Within the scope of this research, a deep learning-based software was developed to address this issue.

## II. RELATED WORKS

In their study, Tack et al. measured lower limb alignment (HKA: hip–knee–ankle) using YOLOv4 and ResNet. YOLOv4 was employed to detect the relevant regions in the radiographs, and the landmarks within these regions were identified using ResNet. A total of 5,818 leg images were used as the dataset. The reported accuracy was 5,809 correct detections out of 5,818 images, corresponding to 99.85% [11].

Schock et al. developed a deep learning methodology for automated alignment assessment in full-length radiographs. The study incorporated radiographic images from 255 patients, utilizing a U-Net architecture for femoral and tibial segmentation. Angular measurements from a randomly selected subset of 30 patients underwent statistical evaluation through repeated-measures analysis of variance and one-way analysis of variance. The strength of associations was determined using Pearson correlation coefficients and intraclass correlation coefficients. The segmentation model demonstrated high accuracy, achieving mean Sørensen–Dice coefficients of 0.97 and 0.96 for femoral and tibial structures, respectively. Correlation analyses yielded coefficients between 0.918 and 0.995 [12].

The study by Wilhelm et al. is a multicenter investigation that developed and validated an integrated deep learning system designed for multiscale, multi-task operations that enables complete automation in the comprehensive evaluation of full-leg radiographs. This model was also validated with an external dataset. The model achieved results comparable to physicians in terms of both alignment accuracy and inter-clinician reliability. Additionally, the algorithm completes a comprehensive analysis more than four times faster than manual methods [13].

Lee et al. conducted research on the application of neural network driven methods to automatically evaluate lower limb alignment through full leg radiographic images. Within their research, SegNet was utilized for automatic segmentation of anatomical regions from the radiographs. The developed system automatically calculates the mechanical axis alignment of the knee joint, the angle formed at the lateral aspect of the distal femur, the angular measurement at the proximal medial tibia, and the angle of joint space convergence (mTFA, mL DFA, MPTA, and JLCA, respectively). The results demonstrated that the automatic segmentation showed an average similarity of 89–97% with manual segmentation, was 3.44 times faster than manual measurements, and exhibited interobserver correlations between manual and automatic measurements ranging from 0.7126 to 0.9695, indicating both accuracy and time efficiency [14].

Erne et al. developed an artificial intelligence system based on UNet and Mask-RCNN for the analysis of full-length radiographs. Initially, the authors employed a Mask-RCNN architecture for anatomical structure segmentation. Subsequently, four anatomically targeted UNet-derived networks were utilized to automatically identify key anatomical reference points. The concluding phase leveraged these data to compute the lower-limb alignment angles mL DFA, MPTA, LD TA, ATFA, and HKA. The detection rates for these angles ranged from 92.4% to 98.9%, and the correlation analyses comparing inter-observer and observer-AI assessments in both preoperative and postoperative images yielded values between 0.80 and 1.0. The study demonstrates that the proposed algorithm operates with high accuracy and efficiency, closely matching measurements obtained by clinicians [15].

Bernard et al. created an automated system to perform angular measurements in lower extremity deformity assessment. During their research, a machine learning algorithm was trained using anatomical landmarks annotated by orthopedic surgeons on a total of 1,175 full-length lower limb radiographs. The algorithm detected 28

anatomical points in each image and calculated the HKA, MPTA, LDFA, and JLCA angles in approximately 48 seconds. When the algorithm's measurements were compared with those of the surgeons, the mean difference across all angles was found to be about 1.9 degrees. In addition, intraclass correlation coefficients for all angles were reported to exceed 95% [16].

Jo and colleagues investigated lower extremity alignment patterns in knee osteoarthritis patients through full-leg radiographic analysis. Their research employed a deep learning algorithm designed to identify anatomical landmarks and quantify angular measurements. The algorithm underwent training with 11,212 full-leg radiographs, incorporating 15 anatomical reference points, while calculating MPTA, mL DFA, HKAA and JLCA angles. The results demonstrated that the interobserver ICC values for all angle measurements were 0.98 or higher. Measurements performed by the model showed mean absolute errors below  $0.52^\circ$  for all angles except JLCA, with outlier rates below 4.26% [17].

Kunze et al. created an artificial intelligence-driven computational method for assessing lower extremity length and knee axis alignment in individuals diagnosed with knee osteoarthritis. The study included full-leg radiographs of 1,011 patients. A deep neural network architecture was designed to automatically measure the hip knee ankle (HKA) angle, femoral and tibial bone lengths, and limb length discrepancy. When compared with radiologists' measurements, the ICC values for HKA and LLD ranged from 0.87 to 1.00 [18].

Nguyen et al. developed an approach utilizing convolutional neural networks (CNNs) for automatically measuring key lower-limb angles, such as HKAA, FAMA, MPTA, and ML DFA. Their study utilizes a dual phase, decentralized convolutional neural network architecture. During the initial phase, the model approximately detects regions of interest (ROIs), while in the second phase, it localizes points within these ROIs. A dataset of 400 images was used for the first-stage model, and 17,000 images were used for the second-stage model. Evaluation on 80 test images showed that in 82.3% of cases, the deviation between automatic and manual measurements was less than 1.5 degrees [19].

In a study by Pei and colleagues, introduced an artificial intelligence-based approach to quantify the HKA angle from X-ray images. This technique relies on segmenting these three anatomical regions through a neural network, followed by identification of the central point within each segmented area. The HKA angles were subsequently computed based on the spatial coordinates of these identified points. The study, conducted on 796 images, demonstrated a high agreement between automatically measured HKA angles and manual measurements performed by three orthopedic specialists, with an ICC of 0.999 and a mean difference of  $0.49^\circ$  [20].

Tanner et al. proposed a neural network-based framework for automated computation of hip knee ankle angle measurements. Their study employed 1,221 radiographic images, allocated in a 70:20:10 ratio for training, validation, and testing purposes, respectively. Their model demonstrated a mean average precision (mAP) of 0.985 in the test cohort, while the intraclass correlation (ICC) comparing the algorithm's outputs with manual assessments was 0.987 for both operated and non-operated limbs. The standard errors of the automatic measurements were  $0.515^\circ$  for the operated limb and  $0.403^\circ$  for the non-operated limb. Compared to segmentation-based methods, the proposed algorithm offered a faster and more practical solution [21].

Tsai developed an artificial neural network approach for automated measurement of the hip knee ankle alignment in pediatric cases. In this investigation, six anatomical reference points; bilateral femoral head centers, intercondylar eminences, and distal tibial articular surfaces were labeled and incorporated into the model through 2D heatmap representations. Assessment performed on 528 radiographic images from 517 subjects revealed an average absolute deviation in predictions of approximately 0.94 degrees with a standard deviation of 0.84 degrees. The intraclass correlation coefficient comparing the algorithm's measurements with expert annotations was documented as 0.974 [22].

Wang and colleagues created an artificial intelligence-based approach for the automatic quantification of FTA and HKA angles. In their study, two separate models were trained using 6,149 FTA and 2,351 HKA posteroanterior knee radiographs obtained from the OAI dataset repository. The radiographic images were partitioned with a 70:15:15 ratio for model development, validation, and evaluation purposes. Model performance was assessed using the mean squared error loss, yielding absolute errors of  $0.8^\circ$  for FTA and  $1.7^\circ$  for HKA [23].

Lower extremity deformity analysis assists in evaluating appropriate angles for preoperative surgical planning through the quantification of angular relationships among the knee, hip, as well as foot. This analysis also enables the prediction of deformities that may develop due to osteoarthritis or post-traumatic conditions. Due to the time-consuming nature of this analysis, systems capable of providing rapid results have been proposed. This research presents a deep learning-driven pose/keypoint identification framework. A software infrastructure has been developed, the model has been implemented, and a system has been created whereby clinicians can obtain outputs simply by providing an X-ray image to the system.

### III. MATERIALS AND METHODS

#### A. Dataset

For this study, 927 images were randomly chosen from the publicly accessible Osteoarthritis Initiative (OAI)

database. The visibility and completeness of the edges of femoral head region, greater trochanter, knee joint, and ankle region bones were considered critical for accurate AI-based predictions. For example, in the assessment of the mechanical lateral distal femoral angle (mLDFA), clear visualization of the distal femoral condyles ensures proper vector placement and precise angle measurement. All images were therefore carefully reviewed, and 153 images were excluded because at least one bone line in each relevant region was not clearly visible, because one or more key bones (femoral head, greater trochanter, knee, or ankle) were not present in the full-leg images, or because the patient had prosthetic implants in the lower extremity. As examples of excluded images, Figure 5 shows the absence of the femoral head, while Figure 6 illustrates that the patient's lower extremity is not clearly visible.



Figure 5: Example of an excluded image from the dataset

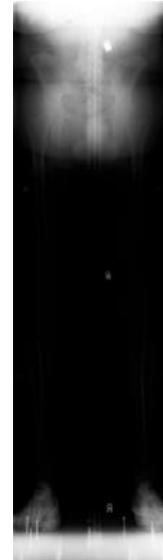


Figure 6: Example of an excluded image from the dataset

### B. Data Preprocessing

Data used in studies may contain various errors and issues, including noise, missing parts, image clarity problems, distortions, and other systematic errors. Data preprocessing refers to the procedures applied to make datasets usable before any analysis is conducted. During preprocessing, various techniques are employed to address these issues. These techniques include rotation, flipping, scaling, Gaussian filtering, color and format conversions, segmentation and masking to isolate regions of interest, and histogram equalization to enhance contrast [24].

During the data preprocessing step, a total of 927 radiographic images were analyzed. The obtained radiographic images were processed using the Contrast Limited Adaptive Histogram Equalization (CLAHE) technique, which is commonly employed in medical imaging. Global histogram equalization is a method utilized to amplify image contrast and, consequently, improve its clarity. Its main drawback is that the equalization is applied uniformly across the entire image. In the CLAHE method, however, the image is divided into tiles, and the contrast of each tile is recalculated individually [25].

After applying the histogram equalization technique, the images were reviewed, and those with insufficient quality were excluded. Figures 7 and 8 illustrate examples of X-ray images with and without the CLAHE enhancement applied to the dataset.

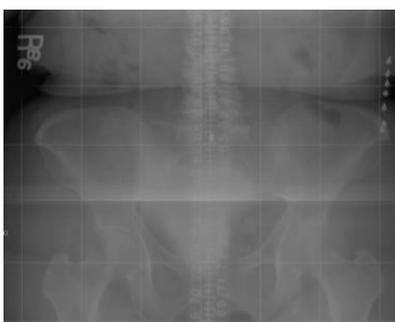


Figure 7: Before CLAHE

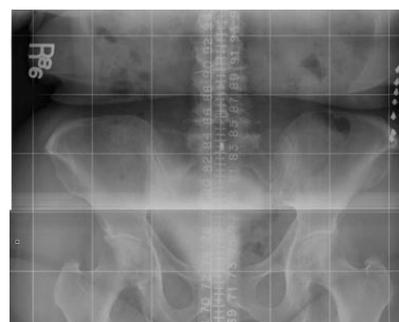


Figure 8: After CLAHE

### C. Deep Learning

Deep learning, functioning as a specialized branch of artificial intelligence, is a method characterized by an enhanced capacity for learning. Unlike traditional neural networks, deep learning networks contain hidden neurons and vary in the number of layers, allowing them to learn more complex relationships. In deep learning, the layers are sequentially connected, with each layer receiving the output from the preceding layer as its input [26]. Through this convolution process, each layer extracts edges and features from the images and transmits them to the subsequent layer while updating weights. Common types of layers include convolutional layers, max-pooling layers, fully connected layers, and dropout layers [27].

In this study, the You Only Look Once (YOLO) architecture, which is widely recognized for its application in computer vision tasks, was utilized.

#### *You Only Look Once (YOLO)*

YOLO is an object detection method that performs end-to-end optimization for a single neural network problem and provides direct detection [28]. What distinguishes YOLO from other algorithms is its ability to perform object detection in real time [29]. Beyond object detection, YOLO also handles other computer vision applications such as instance segmentation, oriented object detection, pose and keypoint prediction and classification. In this study, the YOLOv11x-Pose model, developed for pose/keypoint estimation, was used.

#### *Pose Estimation*

The primary goal of pose and keypoint estimation studies is to detect objects within an image and identify specific points on these instances. This technique is particularly utilized in fields such as robotics [30], sports [31] and healthcare [32]. There are two approaches that can be used for pose and keypoint estimation: the Top-Down and Bottom-Up methods.

##### The Top-Down Approach

One of the widely used approaches in pose and keypoint estimation is the top-down method. In technologies employing this method, the first step involves detecting humans or objects in an image, using models such as Faster R-CNN. Subsequently, keypoint estimation is performed for each detected bounding box. This process increases the overall computational load. Technologies such as CPN and Hourglass are examples of this approach. Generally, top-down methods require object detection and cropping operations, which result in significantly slower inference times [33, 34].

##### The Bottom-Up Approach

Bottom-up algorithms simultaneously identify all humans or objects in an image and then group the detected points to form a skeleton. This approach employs a heatmap-based technique that estimates the likelihood of each pixel containing a keypoint. The positions of the keypoints are determined using Non-Maximum Suppression (NMS), which acts as a filter to remove redundant predictions. Technologies such as OpenPose, HigherHRNet, and DEKR utilize this method. In general, bottom-up methods operate faster than top-down methods. However, their performance is somewhat lower due to challenges encountered during the grouping process [33, 34].

#### *Pose Estimation with YOLOv11x-Pose*

This model is a deep learning approach based on the bottom-up method that can perform object detection and keypoint estimation simultaneously. Unlike traditional bottom-up methods that rely on heatmaps, YOLOv11x-Pose eliminates the heatmap generation step. Instead, it detects the keypoints together with the object's bounding box in a single step, thereby reducing processing time [33].

### D. Data Annotation

During the data annotation phase, the keypoints required to measure the mLPFA, mL DFA, mMPTA, and mLDTA angles were determined with reference to the study organized by Çakmak et al. [10]. The annotated keypoints are illustrated in Figures 9, 10, 11, 12, 13, 14, 15, and 16. Specifically:

- For the femoral head center class used in the mLPFA angle, a total of 2 keypoints were annotated, locating the center on each side.
- For the greater trochanter class used in the mLPFA angle, a total of 2 keypoints were annotated, marking the trochanteric apex bilaterally.
- For the knee class used in the mLPFA, mL DFA, MPTA, and LD TA angles, a total of 12 keypoints were annotated: the distal femoral center and the two femoral condyles, along with the proximal tibial center and the corresponding tibial plateaus on both sides.
- For the ankle class, four keypoints were assigned to the distal tibia, with two points designated for each

leg to measure LDTA and MPTA.

To ensure that data remained secure during the annotation process and that the environment allowed for pose and keypoint labeling without external sharing, several tools were evaluated. Based on this assessment, the Computer Vision Annotation Tool (CVAT), which can be deployed on a local server, was identified as a suitable solution.

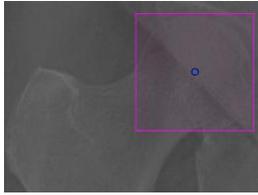


Figure 9: Center of the femoral head



Figure 10: Center of the femoral head

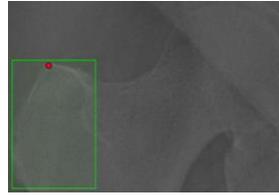


Figure 11: Greater trochanter.

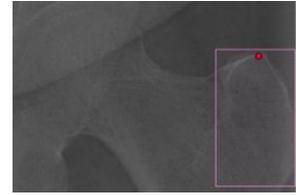


Figure 12: Greater trochanter

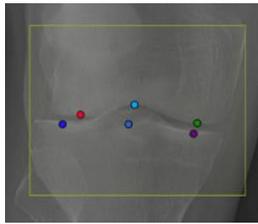


Figure 13: Knee

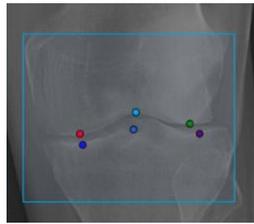


Figure 14: Knee

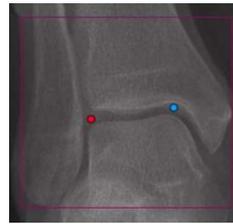


Figure 15: Ankle

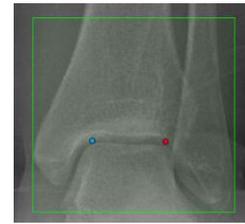


Figure 16: Ankle

In the images annotated with CVAT, the labeling format “CVAT for images 1.1” was used. Since CVAT’s XML format is not directly supported by YOLO, a conversion tool was developed to adapt the annotations to the YOLO keypoint labeling format.

*E. Evaluation of Model Performance*

Evaluating model performance is a step undertaken to determine the extent to which the model’s outputs correspond to the actual values. A confusion matrix is frequently used for performance evaluation. The confusion matrix summarizes the model’s correct and incorrect classifications using four concepts: True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN) [35-37].

Table 1: Confusion Matrix.

		Predicted	
		Positive	Negative
Actual	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

- True Positive (TP) indicates the count of positive samples correctly identified as positive.
- False Positive (FP) represents negative samples that were incorrectly labeled as positive.
- False Negative (FN) shows positive instances that were erroneously predicted as negative.
- True Negative (TN) refers to the number of negative cases correctly recognized as negative [37].

Using these four metrics, the performance of the model can be evaluated in detail. Performance indicators like precision and recall are frequently used metrics for such evaluation.

Precision represents the fraction of correct positive predictions relative to the total positive predictions made. In simple terms, this value is obtained by dividing the true positives by the combined count of true and false positives. This metric indicates the accuracy of the model’s positive predictions, with higher values reflecting a lower rate of incorrect detections [38].

$$Precision (P) = \frac{TP}{TP + FP} \tag{1}$$

Recall represents the ratio of correctly identified positive predictions to all genuine positive instances present in the dataset. This metric is computed by dividing the true positive count by the combined total of true positives and false negatives. A high recall value indicates that the model has a low rate of missing objects in the image [38].

$$Recall (R) = \frac{TP}{TP + FN} \quad (2)$$

In tasks such as object detection, where both class and location predictions are made, additional metrics such as IoU, mAP50, and mAP50-95 are also used.

The Intersection over Union (IoU) metric is computed by dividing the area of intersection by the area of union. It represents the accuracy of the model in detecting objects by measuring the overlap between the predicted bounding box and the ground truth box [38].

mAP50, computed at an intersection over union level of 0.5, represents the averaged precision measure that quantifies the degree of overlap between the predicted bounding regions and the ground truth. It evaluates the accuracy with which the model's estimated regions correspond to the actual object positions [38].

mAP50-95 provides a stricter and more realistic evaluation compared to mAP50. This metric evaluates IoU values from 0.5 to 0.95 at intervals of 0.05 and computes the mean of the obtained scores. A high mAP50-95 value indicates that the model can draw predicted bounding boxes very close to the true object locations, reflecting high positional accuracy and overall precision [38].

#### *F. Technologies Used In Software*

The primary objective of this step was to enable healthcare professionals to upload radiographic images to an interface and quickly obtain outputs that assist in the deformity analysis previously described, by drawing the necessary lines and calculating angles on the X-ray images.

C# was chosen along with the .NET Framework for the software due to its robust infrastructure. For image processing, the EmguCV library, which serves as a wrapper for OpenCV, was utilized. After applying the image processing operations on the relevant images, the YOLO command-line interface was used during the prediction phase.

##### *.NET Framework*

Developed by Microsoft, the .NET Framework is a foundational architecture that enables the development of applications across desktop, web, and mobile platforms. It simplifies application development and, for example, allows deployment to multiple platforms using the C# programming language [39].

The .NET Framework addresses performance issues associated with scripting languages. Before execution, it compiles code into machine language through an underlying system called the Common Language Runtime (CLR), enabling more efficient applications. Additionally, it provides a standard for developers in Windows and web-based applications, ensuring secure integration [40].

##### *EmguCV*

OpenCV functions as an open-source framework that includes well-known algorithms designed for computer vision tasks. It contains modules for image processing, object recognition, feature extraction, and deep neural networks, among others [41-45].

The EmguCV library is a .NET wrapper for OpenCV, which is frequently used in computer vision applications. It allows the functions and tools within the OpenCV library to be used in .NET-supported programming languages [46].

In this study, EmguCV was employed to apply CLAHE on radiographic scans submitted by healthcare professionals, mark relevant regions between bones, and display angle measurement results directly on the visualized areas of interest.

##### *YOLO Command-Line Interface (YOLO CLI)*

The YOLO command-line interface allows models developed by YOLO to be executed directly from the terminal without requiring any additional environment.

In the C# forms-based image processing software developed in this study, a CLI-based communication was implemented to perform object detection and present the results to the user using the YOLO algorithm.

Since YOLO11 does not provide a library specifically for the .NET Framework, an intermediary environment was established to facilitate communication while maintaining modularity and independence. This setup eliminates dependency on a specific YOLO version and allows switching to a different detection model with minimal modifications to the CLI commands, simplifying implementation.

### G. Implementation of the Software

During the system design phase, a needs analysis was first conducted, and based on the identified requirements, an object-oriented infrastructure was planned. The system was structured to perform YOLO-based keypoint detection, calculate angles from these keypoints, and visualize the results through on-image drawings. In the angle calculation process, three keypoints were used, and the angle between two vectors was computed using Formula 1.

In the angle calculation step, three anatomical keypoints (A, B, and C) were used. In this configuration, point B represents the vertex of the angle, whereas points A and C correspond to the endpoints of the two vectors forming the angle. After detecting the keypoint coordinates, the vectors  $\overrightarrow{BA}$  and  $\overrightarrow{BC}$  were constructed by translating points A and C relative to point B. The angle between these two vectors was then computed using the dot-product formulation of the cosine of the angle, and the resulting value was converted from radians to degrees. The mathematical expression employed in this process is provided in Formula 3.

$$\theta_{angle} = \cos^{-1} \left( \frac{(A - B) * (C - B)}{\|A - B\| * \|C - B\|} \right) * \frac{180}{\pi} \quad (3)$$

Furthermore, to ascertain the center of the distal tibia, the keypoints situated on the medial and lateral sides of the distal tibia were utilized. The central point derived from these two keypoints was calculated based on their coordinates, and this midpoint was defined as the distal tibia center. The method for calculating the distal tibia center is presented in Formula 4. Here,  $B$  denotes the distal tibia center,  $B_1$  represents the medial distal tibia keypoint, and  $B_2$  represents the lateral distal tibia keypoint.

$$B_{distal\ tibia\ center} = \left( \frac{B_1^x + B_2^x}{2}, \frac{B_1^y + B_2^y}{2} \right) \quad (4)$$

### Software Workflow

The software was built using C# and the EmguCV library, following an object-oriented and readable architecture. The system automatically performs mechanical axis drawings and angle calculations on the X-ray images uploaded by the user. The workflow proceeds as follows:

1. The uploaded image is backed up with a unique identifier, and a reference image is prepared for YOLO.
2. When the “Start” button is clicked, the prediction process begins using the trained model via the YOLO command-line interface. The model performs keypoint detection, and the results are saved to a text file. The system then reads this file to retrieve the predicted coordinates and confidence scores.
3. The predicted keypoints and class information are drawn on the image using the EmguCV library. Keypoints are represented by circles, and mechanical axes are shown as lines connecting these points. Subsequently, angle calculations are performed using the obtained keypoint coordinates.
4. The processed image is transferred to the interface as a bitmap, and the calculated angle values are displayed as text on the panel.

### Software UI

The user interface is responsible for facilitating interaction between the user and the software [47, 48]. In this study, the interface was designed to be simple and user-friendly. Personnel can upload X-ray images to the system and view the results directly on the screen. The interface of the developed software is shown in Figure 17.

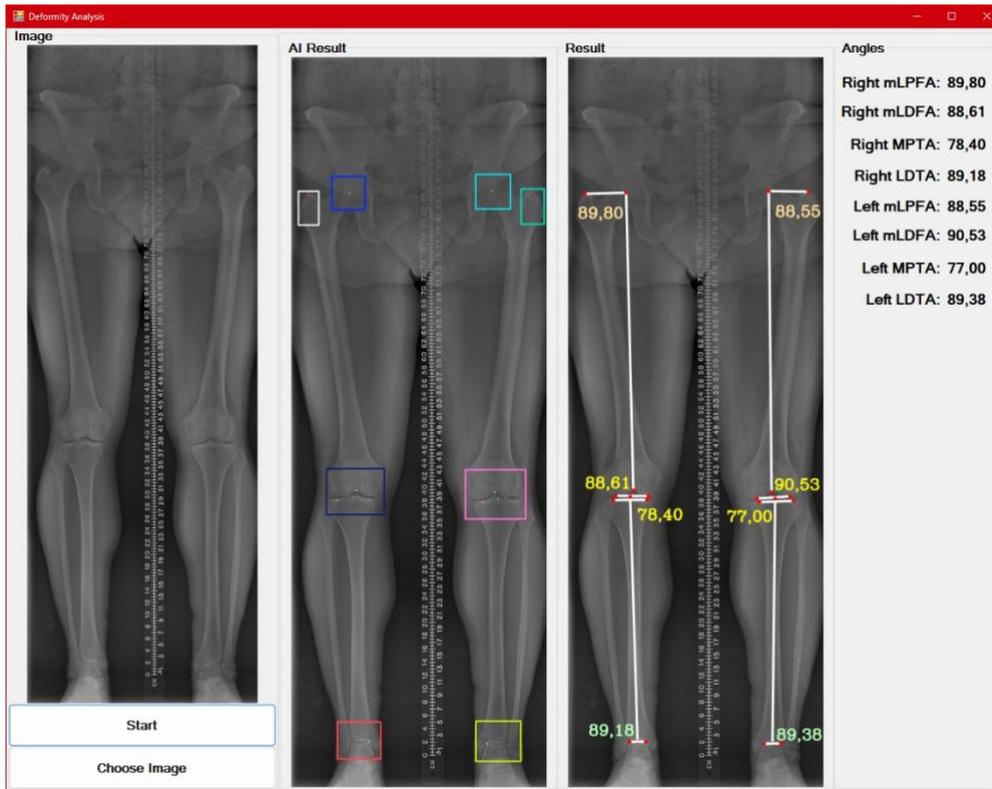


Figure 17: Software UI

#### IV. RESULTS

The YOLOv11x-POSE architecture was employed for model development. Throughout the training process, images were resized to 896×896 pixels and processed in batches of twenty samples across one hundred training iterations. The optimization process utilized the AdamW algorithm, configured with an initial learning rate of 0.000833 and a momentum parameter set to 0.9. To improve computational performance, Automatic Mixed Precision (AMP) was implemented alongside eight parallel processing threads. The dataset was preprocessed and prepared according to the YOLO annotation framework before being fed into the model. Training was conducted on Google Colab platform, leveraging an NVIDIA A100 graphics processing unit supported by 83.5 GB of system memory and 40 GB of dedicated video memory. Evaluation was performed on real images not included in the training process.

This study utilized a total of 774 images. The dataset distribution consisted of 620 images (80%) allocated for training purposes, while 77 images (10%) were designated for testing and an additional 77 images (10%) for validation. Table 2 presents the model's training performance metrics.

Table 2: Performance metrics from model training

Metric	Bounding Box (%)	Pose (%)
Precision(P)	99.61	99.61
Recall(R)	99.90	99.90
mAP50	99.45	99.45
mAP50-95	79.98	97.86

Examining the model's confusion matrix, each keypoint was predicted with high accuracy. Only a single misclassification occurred in the class corresponding to the left greater trochanter, while no errors were observed in the other classes. This indicates that the model can detect anatomical regions with high accuracy and reliability. This confusion matrix is visualized in Figure 18.

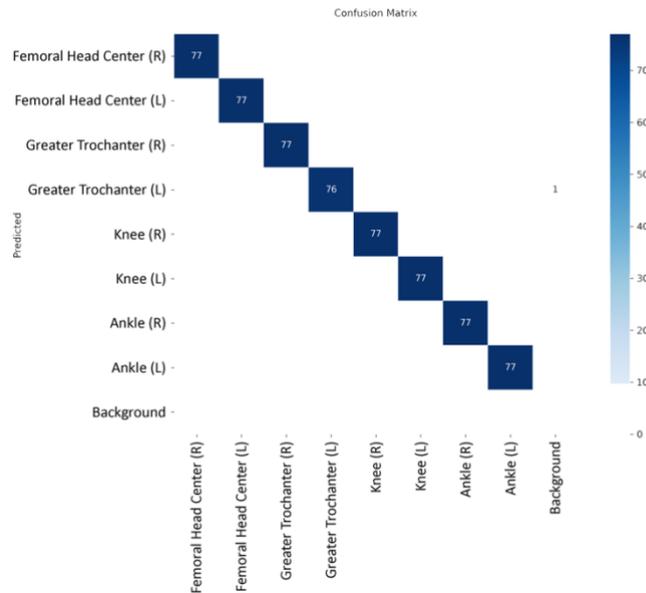


Figure 18: Confusion Matrix

Upon examining the comprehensive assessment of the proposed model, all loss metrics exhibited consistent reduction. The validation losses closely followed the training losses, indicating that overfitting did not occur and that the training progressed in a balanced manner. These results demonstrate that the model exhibits stable performance with high accuracy in both object and pose detection. Figure 19 presents the graphical representations of loss functions and performance indicators.

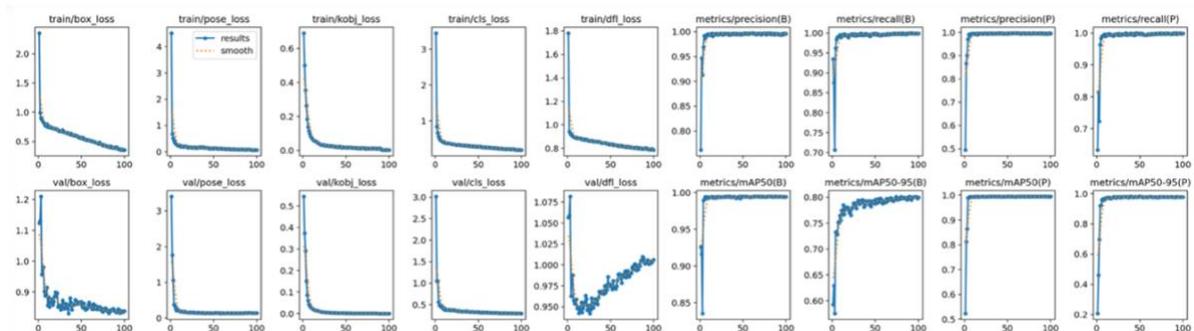


Figure 19: Training/validation loss curves alongside object and pose detection metrics.

### V. LIMITATIONS

Several limitations of this study warrant consideration. Firstly, images of patients with prostheses were excluded, which may limit the applicability of the model to such cases. Second, the dataset did not include labels for specific deformities such as genu valgum or genu varum; therefore, only images of patients with full-leg radiographs were included. Third, the analysis was restricted to mechanical axis angles, namely mLPFA, mLDFa, mMPTA, and mLDTA, limiting the scope of angular measurements. Finally, the computed keypoints were limited to those covering the center of the femoral head, greater trochanter, knee joint, and ankle joint. Future studies could expand the dataset to include prosthetic cases and pathological deformities, as well as extend the keypoint set and angular measurements to provide a more comprehensive evaluation of lower extremity alignment.

### VI. CONCLUSION

Deep learning technology provides significant benefits in medical image analysis, particularly for the early diagnosis of diseases and the acceleration of treatment-related processes. This research presents an automated system based on the YOLOv11x-Pose model designed to streamline and automate lower extremity deformity analysis. Unlike previous studies in the literature, this approach focused not only on object detection but also on pose and keypoint detection, enabling a more detailed analysis of the bones.

Training conducted on 774 images achieved high accuracy values. Looking at the mAP50 metrics, a value of 99.45% was obtained for bounding box detection, and 99.45% for keypoint detection. These results indicate that the developed system can accurately detect bones and reliably mark the keypoints necessary for deformity analysis. The high-performance rates and the developed software infrastructure suggest that this system could be effectively applied to the analysis of other orthopedic problems in the future. Furthermore, the system has the potential to be extended to other anatomical regions and types of deformities.

#### ACKNOWLEDGMENT

Data used in the preparation of this manuscript were obtained and analyzed from the controlled access datasets distributed from the Osteoarthritis Initiative (OAI), a data repository housed within the NIMH Data Archive (NDA). OAI is a collaborative informatics system created by the National Institute of Mental Health and the National Institute of Arthritis, Musculoskeletal and Skin Diseases (NIAMS) to provide a worldwide resource to quicken the pace of biomarker identification, scientific investigation and OA drug development. Dataset identifier: 10.15154/7xkm-mn77.

#### REFERENCES

- [1]. Aytaç, Z., İ. Iseri, and B. Dandıl, *Derin Öğrenme Kullanarak Tiroid Kanseri Teşhisi*. Avrupa Bilim ve Teknoloji Dergisi, 2021(29): p. 292-298.
- [2]. Hoşgör, H. and H. Güngördü, *Sağlıkta yapay Zekanın kullanım alanları üzerine nitel bir araştırma*. Avrupa Bilim ve Teknoloji Dergisi, 2022(35): p. 395-407.
- [3]. Güzel, Ş., H.A. Dömbekci, and F. Eren, *Yapay zekânın sağlık alanında kullanımı: Nitel bir araştırma*. Celal Bayar Üniversitesi Sağlık Bilimleri Enstitüsü Dergisi, 2022. **9**(4): p. 509-519.
- [4]. Robbins, C.A., *Deformity reconstruction surgery for Blount's disease*. Children, 2021. **8**(7): p. 566.
- [5]. Paley, D., *Principles of deformity correction*. 2014: Springer.
- [6]. Lott, A., et al., *Around-the-knee osteotomies part II: Surgical indications, techniques and outcomes—State of the art*. Journal of ISAKOS, 2024. **9**(4): p. 658-671.
- [7]. White, G.R. and G.A. Mencia, *Genu valgum in children: diagnostic and therapeutic alternatives*. JAAOS-Journal of the American Academy of Orthopaedic Surgeons, 1995. **3**(5): p. 275-283.
- [8]. Brooks, W.C. and R.H. Gross, *Genu varum in children: diagnosis and treatment*. JAAOS-Journal of the American Academy of Orthopaedic Surgeons, 1995. **3**(6): p. 326-335.
- [9]. Schröter, S., et al., *Lower limb deformity analysis and the planning of an osteotomy*. The journal of knee surgery, 2017. **30**(05): p. 393-408.
- [10]. Çakmak, M. and K. Özkan, *Alt ekstremite deformite analizi (I)*. Totbid dergisi, 2005. **4**: p. 50-62.
- [11]. Tack, A., B. Preim, and S. Zachow, *Fully automated assessment of knee alignment from full-leg X-rays employing a "YOLOv4 And Resnet Landmark regression Algorithm" (YARLA): data from the Osteoarthritis Initiative*. Computer Methods and Programs in Biomedicine, 2021. **205**: p. 106080.
- [12]. Schock, J., et al., *Automated analysis of alignment in long-leg radiographs by using a fully automated support system based on artificial intelligence*. Radiology: Artificial Intelligence, 2020. **3**(2): p. e200198.
- [13]. Wilhelm, N.J., et al., *Multicentric development and validation of a multi-scale and multi-task deep learning model for comprehensive lower extremity alignment analysis*. Artificial Intelligence in Medicine, 2024. **150**: p. 102843.
- [14]. Lee, H.S., et al., *Automated analysis of knee joint alignment using detailed angular values in long leg radiographs based on deep learning*. Scientific Reports, 2024. **14**(1): p. 7226.
- [15]. Erne, F., et al., *Automated artificial intelligence-based assessment of lower limb alignment validated on weight-bearing pre-and postoperative full-leg radiographs*. Diagnostics, 2022. **12**(11): p. 2679.
- [16]. Bernard de Villeneuve, F., et al., *An artificial intelligence based on a convolutional neural network allows a precise analysis of the alignment of the lower limb*. International Orthopaedics, 2023. **47**(2): p. 511-518.
- [17]. Jo, C., et al., *Deep learning-based landmark recognition and angle measurement of full-leg plain radiographs can be adopted to assess lower extremity alignment*. Knee Surgery, Sports Traumatology, Arthroscopy, 2023. **31**(4): p. 1388-1397.
- [18]. Kunze, K.N., et al., *Radiographic findings involved in knee osteoarthritis progression are associated with pain symptom frequency and baseline disease severity: a population-level analysis using deep learning*. Knee Surgery, Sports Traumatology, Arthroscopy, 2023. **31**(2): p. 586-595.
- [19]. Nguyen, T.P., et al., *Intelligent analysis of coronal alignment in lower limbs based on radiographic image with convolutional neural network*. Computers in biology and medicine, 2020. **120**: p. 103732.
- [20]. Pei, Y., et al., *Automated measurement of hip-knee-ankle angle on the unilateral lower limb X-rays using deep learning*. Physical and Engineering Sciences in Medicine, 2021. **44**(1): p. 53-62.
- [21]. Tanner, I.L., et al., *Developing a computer vision model to automate quantitative measurement of hip-knee-ankle angle in total hip and knee arthroplasty patients*. The Journal of Arthroplasty, 2024. **39**(9): p. 2225-2233.
- [22]. Tsai, A., *A deep learning approach to automatically quantify lower extremity alignment in children*. Skeletal Radiology, 2022. **51**(2): p. 381-390.
- [23]. Wang, J., et al., *Predicting hip-knee-ankle and femorotibial angles from knee radiographs with deep learning*. The Knee, 2023. **42**: p. 281-288.
- [24]. Maharana, K., S. Mondal, and B. Nemade, *A review: Data pre-processing and data augmentation techniques*. Global Transitions Proceedings, 2022. **3**(1): p. 91-99.
- [25]. Musa, P., F. Al Rafi, and M. Lamsani, *A Review: Contrast-Limited Adaptive Histogram Equalization (CLAHE) methods to help the application of face recognition*. in *2018 third international conference on informatics and computing (ICIC)*. 2018. IEEE.
- [26]. Deng, L. and D. Yu, *Deep learning: methods and applications*. Foundations and trends® in signal processing, 2014. **7**(3-4): p. 197-387.

- [27]. Doğan, F. and İ. Türkoğlu, *Derin öğrenme algoritmalarının yaprak sınıflandırma başarımlarının karşılaştırılması*. Sakarya University Journal of Computer and Information Sciences, 2018. **1**(1): p. 10-21.
- [28]. Redmon, J., et al. *You only look once: Unified, real-time object detection*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- [29]. Tan, F.G., et al., *Derin öğrenme teknikleri ile nesne tespiti ve takibi üzerine bir inceleme*. Avrupa Bilim ve Teknoloji Dergisi, 2021(25): p. 159-171.
- [30]. Jain, H.P., et al. *Real-time upper-body human pose estimation using a depth camera*. in *International Conference on Computer Vision/Computer Graphics Collaboration Techniques and Applications*. 2011. Springer.
- [31]. Chen, S. and R.R. Yang, *Pose trainer: correcting exercise posture using pose estimation*. arXiv preprint arXiv:2006.11718, 2020.
- [32]. Rohan, A., et al., *Human pose estimation-based real-time gait analysis using convolutional neural network*. IEEE access, 2020. **8**: p. 191542-191550.
- [33]. Jiang, Y., et al., *YOLO-Rlepose: improved YOLO based on Swin transformer and Rle-Oks loss for multi-person pose estimation*. Electronics, 2024. **13**(3): p. 563.
- [34]. Maji, D., et al. *Yolo-pose: Enhancing yolo for multi person pose estimation using object keypoint similarity loss*. in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2022.
- [35]. Doğanalp, T., *Tümör Tanılı Farklı Sekans Manyetik Rezonans Görüntülerinin Derin Öğrenme Modelleri ile Sınıflandırılması*. 2022.
- [36]. Aksoy, B., H.D. Halis, and O.K.M. Salman, *Elma bitkisindeki hastalıkların yapay zekâ yöntemleri ile tespiti ve yapay zekâ yöntemlerinin performanslarının karşılaştırılması*. International Journal of Engineering and Innovative Research, 2020. **2**(3): p. 194-210.
- [37]. Potur, E.A. and N. Erginel, *Kalp yetmezliği hastalarının sağ kalımlarının sınıflandırma algoritmaları ile tahmin edilmesi*. Avrupa Bilim ve Teknoloji Dergisi, 2021(24): p. 112-118.
- [38]. Mullins, C.C., et al., *Optimizing data collection requirements for machine learning models in wild blueberry automation through the application of DALL-E 2*. Smart Agricultural Technology, 2025. **10**: p. 100764.
- [39]. Wicaksana, B.A. and R.F. Sari. *Implementing text information display of detected color for partially color blinded person using NET platform and EmguCV library*. in *ICIMU 2011: Proceedings of the 5th international Conference on Information Technology & Multimedia*. 2011. IEEE.
- [40]. Gittleman, A. and A. Gittleman, *Computing with C# and the .NET Framework*. 2011: Jones & Bartlett Publishers.
- [41]. Golovnin, O. and D. Rybnikov. *Benchmarking of Feature Detectors and Matchers using OpenCV-Python Wrapper*. in *2021 International Conference on Information Technology and Nanotechnology (ITNT)*. 2021. IEEE.
- [42]. Hasan, R.T. and A.B. Sallow, *Face detection and recognition using opencv*. Journal of Soft Computing and Data Mining, 2021. **2**(2): p. 86-97.
- [43]. Giri, S., et al. *Emotion detection with facial feature recognition using CNN & OpenCV*. in *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*. 2022. IEEE.
- [44]. Shubho, F.H., et al. *Real-time traffic monitoring and traffic offense detection using YOLOv4 and OpenCV DNN*. in *TENCON 2021-2021 IEEE Region 10 Conference (TENCON)*. 2021. IEEE.
- [45]. Emami, S. and V.P. Suci, *Facial recognition using OpenCV*. Journal of Mobile, Embedded and Distributed Systems, 2012. **4**(1): p. 38-43.
- [46]. Magdin, M., et al., *Comparison of Multilayer Neural Network Models in Terms of Success of Classifications Based on EmguCV, ML, NET and Tensorflow. Net*. Applied Sciences, 2022. **12**(8): p. 3730.
- [47]. Dillon, A., *User interface design*. 2003, London: Macmillan.
- [48]. Morales Diaz, L.V. *What is a user interface, again? A survey of definitions of user interface: Our shared and implicit understanding of the concept of user interface*. in *Proceedings of the 9th Mexican International Conference on Human-Computer Interaction*. 2022.

## RESUME

### Melih GULEÇ

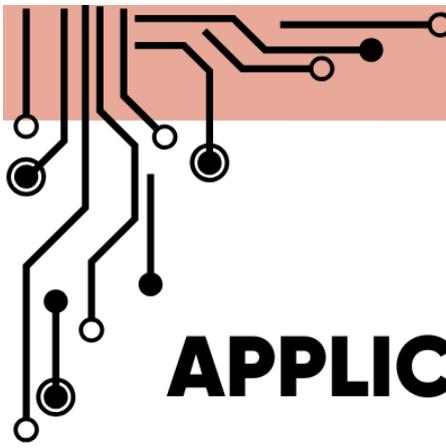
Melih GÜLEÇ received his bachelor's degree in computer engineering from the Faculty of Technology at Selçuk University in 2023. In the same year, he began his postgraduate studies in the Department of Computer Engineering at the Institute of Natural and Applied Sciences, Selçuk University. Upon completing his undergraduate studies, he started working as a Backend Developer at Macellan. His areas of work include backend technologies, big data, real-time data processing, artificial intelligence, and image processing.

### Nurettin DOĞAN

Nurettin DOĞAN is a Professor at the Department of Computer Engineering, Faculty of Technology, Selçuk University. He received his Master's and PhD degrees from Ankara University. He has published papers in various national and international journals on image encryption, approximate analytic serial solutions methods and artificial intelligence.

**Hasan Erdinc KOCER**

He completed his undergraduate studies in Electronics and Computer Education at Marmara University in 1998. He received his master's degree from the Computer Systems Education Department of the Institute of Science at Selçuk University. In 2007, he completed his PhD in the Electrical-Electronics Engineering Department of the Institute of Science at Selçuk University. He is currently working as a Full Professor in the Department of Electrical-Electronics Engineering at the Faculty of Technology, Selçuk University.



# INNOVATION AND APPLICATIONS IN ENGINEERING

Engineering is the cornerstone of modern technological advancement, bridging creativity and functionality. "Innovative Approaches in Applied Engineering Fields" brings together the latest research, cutting-edge methodologies, and transformative solutions from leading professionals and academics.

## FIELDS

- Artificial Intelligence
- Industry 4.0 Technologies
- Sustainable Engineering
- Real-world case studies in various engineering domains.

---

This book serves as a valuable resource for engineers, researchers, and students who aim to stay ahead in the rapidly evolving world of engineering. Uncover new horizons in innovation, efficiency, and sustainable solutions for a better future!



SELÇUK  
UNIVERSITY  
PRESS



978-975-448-252-2

# INNOVATION AND APPLICATIONS IN ENGINEERING

Engineering is the cornerstone of modern technological advancement, bridging creativity and functionality. "Innovative Approaches in Applied Engineering Fields" brings together the latest research, cutting-edge methodologies, and transformative solutions from leading professionals and academics.

## FIELDS

- Artificial Intelligence
- Industry 4.0 Technologies
- Sustainable Engineering
- Real-world case studies in various engineering domains.

This book serves as a valuable resource for engineers, researchers, and students who aim to stay ahead in the rapidly evolving world of engineering. Uncover new horizons in innovation, efficiency, and sustainable solutions for a better future!



INNOVATION AND APPLICATIONS IN ENGINEERING



SELÇUK  
UNIVERSITY  
PRESS

2025



# INNOVATION AND APPLICATIONS IN ENGINEERING

## EDITORS:

Prof. Dr. Nurettin DOĞAN  
Res. Asst. Yusuf ERYEŞİL

