BAJECE

e-mail: editor@bajece.com

Balkan Journal of Electrical & Computer Engineering

An International Peer Reviewed, Referred, Indexed and Open Access Journal

www.bajece.com

Vol : 13
No : 3
Year : 2025
ISSN : 2147 - 284X

It is abstracted and indexed in, Index Google Scholarship, the PSCR, Cross ref, DOAJ, Research Bible, Indian Open Access Journals (OAJ), Institutional Repositories (IR), J-Gate (Informatics India), Ulrich's, International Society of Universal Research in Sciences, DRJI, EyeSource, Cosmos Impact Factor, Cite Factor, SIS SIS Scientific Indexing Service, IIJIF, iiiFactor. ULAKBİM-TR Dizin.

Aim & Scope
The journal publishes original papers in the extensive field of Electrical-Electronics and Computer engineering. It accepts contributions which are fundamental for the development of electrical engineering, computer engineering and its applications, including overlaps to physics. Manuscripts on both theoretical and experimental work are welcome. Review articles and letters to the editors are also included.
Application areas include (but are not limited to): Electrical & Electronics Engineering, Computer Engineering, Software Engineering, Biomedical Engineering, Electrical Power Engineering, Control Engineering, Signal and Image Processing, Communications & Networking, Sensors, Actuators, Remote Sensing, Consumer Electronics, Fiber-Optics, Radar and Sonar Systems, Artificial Intelligence and its applications, Expert Systems, Medical Imaging, Biomedical Analysis and its applications, Computer Vision, Pattern Recognition, Robotics, Industrial Automation.

BAJECE
Balkan Journal of Electrical & Computer Engineering
An International Peer Reviewed, Indexed and Open Access Journal

ISSN: 2147- 284X
Vol: 13
No : 3
Year: September 2025

CONTENTS

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING
(An International Peer Reviewed, Indexed and Open Access Journal**)**
Contact: Batman University Department of Electrical-Electronics Eng.
Bati Raman Campus Batman-Turkey
Web: https://dergipark.org.tr/en/pub/bajece
https://www.bajece.com **e-mail:** bajece@hotmail.com

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 3, September 2025
Research Article

The Impact of Artificial Intelligence on Sentiment Analysis Detection in Music Reviews

Murat Simsek and Bugra Kagan Kayhan

***Abstract*—This study aims to perform sentiment and content analysis of Spotify user reviews using machine learning and deep learning methods. The goal is to better understand users' experiences and satisfaction. The study employs various machine learning and deep learning techniques to identify the emotional tendencies in user reviews and analyze the relationship between these tendencies and content features. The performance of these methods is evaluated using various metrics such as accuracy, precision, recall, and F1-score. By identifying the strengths and weaknesses of each method, the study determines which techniques are more effective in specific situations. The results provide valuable insights for improving the quality of music streaming services and enhancing user experience. This study aims to help service providers increase user satisfaction by gaining a better understanding of user feedback. Additionally, these analyses are expected to provide valuable data for future improvements in music streaming services. Thus, it will be possible to continuously improve user experiences and enhance service quality.**

Index Terms—**Sentiment Analysis, Machine Learning, Deep Learning, Large Language Model, Natural Language Processing, LSTM.**

I. Introduction

MUSIC IS an integral part of our society as a universal way to express and understand emotions. With the advancement of technology, music streaming services have become platforms that profoundly impact users' music experiences. As of March 2022, Spotify is one of the largest providers in this field, boasting 422 million monthly active users and 182 million paying subscribers. Users express their experiences with the Spotify app and their satisfaction through various ratings and reviews.

When users leave reviews, they not only provide likes or star ratings but also add comments that include their own expressions. This indicates that the number of stars or a single like button is not the only indicator of user satisfaction; the content of the written comment is also very significant. Hence, it is necessary to conduct sentiment analysis on all reviews to determine the emotional content expressed in these comments.

Artificial intelligence (AI)-based sentiment analysis involves the use of machine learning and deep learning techniques to automatically identify emotional tendencies in text data.
The integration of artificial intelligence (AI) into sentiment analysis for music reviews has seen noteworthy advancements, particularly with the adoption of machine learning and deep learning models The credibility of AI systems, as explored by Khan and Mishra, is crucial for facilitating positive consumer-AI interactions and enhancing trust in AI-generated insights[1]. Furthermore, the epistemic implications of AI, as discussed by Miragoli, underscore the importance of addressing structural biases to prevent epistemic injustices in AI applications[2]. These insights collectively underscore the transformative potential of AI in music sentiment analysis while highlighting areas that require further exploration and refinement. Recent studies highlight the efficacy of techniques such as convolutional neural networks (CNN) and recurrent neural networks (RNN), which have significantly enhanced the interpretation of emotive language in music critiques[3][4]. Moreover, transformer-based models like BERT and GPT have been pivotal in achieving more precise context analysis, thereby capturing nuanced sentiments that conventional models might overlook[5]. These advancements have not only improved accuracy but have also expanded the scope of sentiment analysis to allow for real-time processing and deeper insights into consumer preferences[3]. Nevertheless, challenges remain, particularly concerning data bias and the interpretability of models.

Academic studies in this field have primarily focused on analyzing large amounts of text data, such as social media posts, product reviews, and customer feedback[6-7]. Machine learning algorithms, such as Naive Bayes, Support Vector Machines (SVM), and Logistic Regression, are commonly used methods in sentiment analysis [8]. In recent years, deep learning models, especially Long Short-Term Memory (LSTM) and Bidirectional Encoder Representations from Transformers (BERT), have achieved notable successes in sentiment analysis, identifying emotional tendencies in text data with higher accuracy rates and a better understanding of the contextual meaning of language [9]. AI-based sentiment analysis has made significant contributions across a wide range of applications, from commercial uses to social research [10].

Literature studies have shown that LSTM (Long Short-Term

Murat Şimşek, is with Department of Artificial Intelligence Engineering of Ostim Technical University, Ankara, Turkey,(e-mail: murat.simsek@ostimteknik.edu.tr).
https://orcid.org/0000-0002-8648-3693

Buğra Kağan Kayhan, is with Department of Software Engineering of Ostim Technical University, Ankara, Turkey, (e-mail: 220904011@ostimteknik.edu.tr).
https://orcid.org/0009-0008-0501-3982

Manuscript received May 30, 2024; accepted Apr 27, 2025.
DOI: 10.17694/bajece.1492287

Memory) models are used for emotion analysis from music. In one study, the aim was to predict the emotional values of musical clips over time and determine the subsequent emotional value in a time series. Models were trained on music clips annotated with levels of valence and arousal using the Emotions in Music database. Mel Spectrograms extracted from music clips were used as input data, and the model's accuracy rates were evaluated using the mean squared error (MSE) metric. The results indicate that the LSTM model is effective in emotion analysis from music [11].

A study using data mining methods to analyze the emotional effects of music on people examined the impact of rhythm and timbre features extracted from music tracks on human emotions. The data consisted of 593 songs from the Mulan database and 72 musical features related to these songs. The songs were classified into six different emotion categories (sadness, joy, anger, fear, surprise, and calm). Data mining algorithms such as Random k-Labelsets (RAkEL) and Multi-Label k-Nearest Neighbors (MLkNN) were used to solve the multi-label classification problem. While the RAkEL algorithm achieved an accuracy of 78.8% and a Hamming loss of 21.2%, the MLkNN algorithm provided an accuracy of 80.4% and a Hamming loss of 19.6% [12].

A study utilizing natural language processing (NLP) and machine learning (ML) techniques to perform sentiment analysis on music reviews from Pitchfork.com examined the relationship between album reviews by Pitchfork critics and the numerical scores given to these reviews. The data was taken from a Kaggle dataset containing 18,376 observations. Preprocessing steps included converting texts to lowercase, removing stopwords, cleaning punctuation, and lemmatization. Feature engineering included extracting features such as review length, percentages of positive, negative, and neutral sentiments, rates of long words, counts of album/artist names, and similarity to top reviews. Models such as L_1 and L_2regularized linear regression, SVM regressor, and Stochastic Gradient Descent (SGD) regression were used for training. Models were evaluated using GridSearchCV for hyperparameter optimization. Performance evaluation used metrics such as mean squared error (MSE), mean absolute error (MAE), and R^2. Results indicated a 30% improvement in R^2 score with feature engineering, with the best performance provided by the linear regression model.

A study using natural language processing (NLP) and machine learning methods to perform sentiment analysis on user reviews of musical instruments on Amazon used a dataset of 10,262 reviews aiming to classify the reviews as positive or negative. Preprocessing steps included converting text to lowercase, removing punctuation, tokenization, removing stopwords, and lemmatization. Reviews were analyzed using the SpaCy library and classified using a Convolutional Neural Network (CNN) model. During training, ratings between 1-3 stars were labeled as negative, while 4-5 stars were labeled as positive. After training, the model's accuracy, precision, recall, and F1 score metrics were evaluated, achieving 78% precision and 79% F1 score. This shows that deep learning techniques can be effectively used for sentiment analysis of Amazon musical instrument reviews [13].

The studies reviewed have assessed that machine learning, deep learning, and recently, large language models play an effective role in sentiment analysis. This study aimed to perform sentiment and content analysis of Spotify user reviews using machine learning and deep learning methods. Sentiment analysis involves classifying text data as positive, negative, or neutral and plays a significant role in evaluating dynamic data sources containing large amounts of text data, such as social media and customer feedback. The goal was to automatically detect emotional tendencies in user reviews and analyze content features.

This study presents a comprehensive approach to sentiment analysis by integrating machine learning and deep learning techniques to analyze Spotify user reviews, thereby providing a multifaceted perspective on user sentiments. It rigorously evaluates various models, including LSTM and BERT, using performance metrics such as accuracy, precision, recall, and F1-score, allowing for a detailed comparison of their effectiveness in capturing emotional content. The study also addresses class imbalance through the application of the Synthetic Minority Over-sampling Technique (SMOTE), enhancing model robustness. Moreover, the findings offer industry-relevant insights, enabling music streaming platforms like Spotify to improve user satisfaction by understanding feedback trends. Finally, the research identifies challenges and suggests optimizations, providing a foundation for scalable sentiment analysis models adaptable to various music platforms.

II. MATERIALS AND METHODS

Spotify is a leading music streaming platform with over 422 million monthly active users worldwide, of which 182 million are paid subscribers. Users frequently express their experiences and opinions on the platform through reviews. These reviews contain valuable feedback about the quality of musical works and artist performances. Performing sentiment analysis on these comments is critical for measuring user satisfaction and understanding trends in the music industry as well as changes in user preferences. Particularly, accurate analyses can enhance music production and marketing strategies, enabling artists and record companies to respond more effectively to user demands.

A. Exploratory Data Analysis

In this study, a dataset containing user reviews from the Spotify application was used. This dataset, titled "Spotify App Reviews," was sourced from Kaggle and collected via scraping methods from the Google Store by M. Faarisul Ilmi [14]. The dataset comprises 61,594 rows and 5 columns.

- Time_submitted: Indicates the time frame in which the review was submitted.
- Review: Contains the text of the review.
- Rating: Shows the score given by the user (ranging from 1 to 5).
- Total_thumbsup: Indicates how many people found the review helpful.

- Reply: Contains the response to the review.

In the mentioned columns, all except for Time_submitted and Reply were used in the exploratory data analysis section. An overview of the dataset has been conducted, and the dataset information is in Table 1.

TABLE I
DATASET INFORMATION

Column	Non-Null Count	Dtype	Entries	Memory Usage
Time_submitted	61594 non-null	object	0 to 61593	2.3+ MB
Review	61594 non-null	object	0 to 61593	2.3+ MB
Rating	61594 non-null	int64	0 to 61593	2.3+ MB
Total_thumbsup	61594 non-null	int64	0 to 61593	2.3+ MB
Reply	216 non-null	object	0 to 61593	2.3+ MB

When Table 1 is examined, it is observed that there are only 216 data entries in the 'Reply' column, but considering the dataset has 61,594 rows, it is understood that most of the data for this feature is missing. When sentiment analysis from comments is targeted, the impact of this feature on sentiment analysis has been evaluated as weak, and it has been discarded along with the 'Time Submitted' column as part of Feature Reduction.

In the visualization related to our target column 'Rating', there are comments with ratings ranging from 1 to 5 stars. In Fig. 1, the number of comments varying from 1 to 5 stars in the Rating column has been shown.

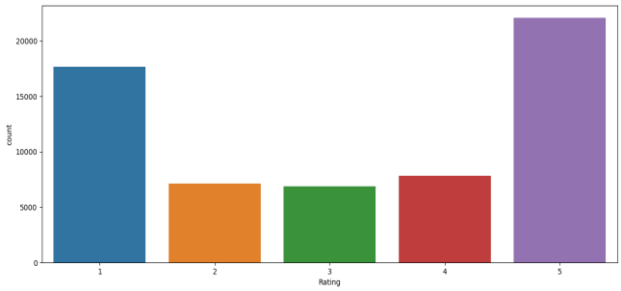


Fig.1. Histogram Chart According to Number of Stars in Rating Column

When Fig.1 is examined, comments range from 1 to 5 stars. In this context, 4 and 5 stars are considered positive, 3 stars neutral, and 1 and 2 stars negative. This transformation of star ratings is graphically represented in Fig.2.

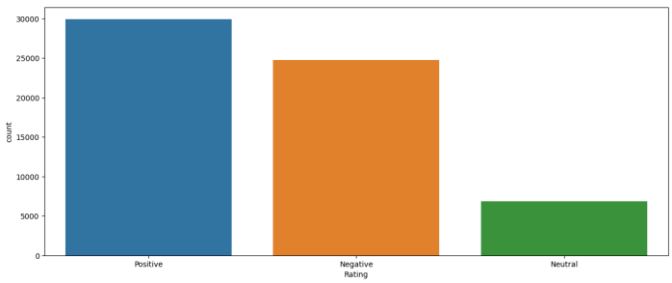


Fig.2. Star Transformation in Rating Feature

B. *Text Preprocessing*

Text preprocessing is a crucial step in natural language processing (NLP) projects to make data analyzable. This process includes methods like converting to lowercase, removing punctuation, tokenization, removing stopwords, and lemmatization. Each step aims to simplify the textual complexity to prepare the dataset for better analytical performance. Removing punctuation marks and converting text to lowercase are particularly emphasized, as they help eliminate noise and standardize the text for machine learning models, which is vital for consistent and effective processing [15].

In the realm of sentiment analysis and other NLP applications, techniques such as tokenization and lemmatization play significant roles. Tokenization segments text into manageable units, enhancing the machine learning models' ability to learn patterns within the text effectively. Lemmatization, on the other hand, delves deeper into the semantic aspect by reducing words to their base or dictionary forms, thus aiding in more accurate text classification and sentiment analysis by focusing on the semantic and contextual values of words[16]. After these preprocessing steps, tools like the Python Word Cloud library are employed to visualize frequently occurring words, providing insights into the predominant themes within the text data. The Word Cloud, which consists of frequently occurring words using the Python Word Cloud library, is shown in Fig 3.

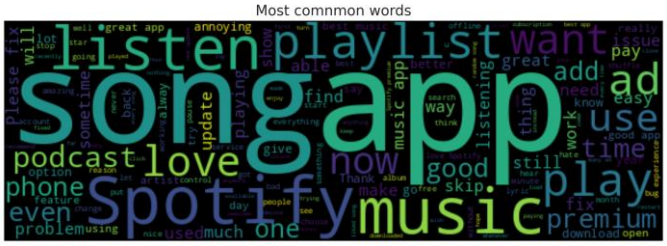


Fig.3. Histogram Chart According to Number of Stars in Rating Column

Different text encoding processes have been performed under data preprocessing and compared in terms of performance. Text encoding is a crucial process in modeling text data as numerical vectors, which allows for direct modeling of text data. This document discusses two primary text encoding methods utilized in the thesis: Count Vectorization and TF-IDF (Term Frequency-Inverse Document Frequency).

Count Vectorization is a technique that converts text documents into vectors based on word counts. Each document is represented as a vector where each word's frequency in the document determines the vector's values. This method is widely used in natural language processing (NLP) tasks such as text classification [17]. An example sentence, "The quick brown fox jumped over the lazy dog," would be represented as a vector [2, 1, 1, 1, 1, 1, 1, 1], corresponding to the word counts in the sentence [18].

TF-IDF, on the other hand, assigns weights to words based on their frequency in a document and their rarity across all documents. This method addresses the issues of over-

representation of common words in Count Vectorization, providing a more balanced measure of word importance. The TF-IDF calculation at the character level allows for more granular text analysis [19].

TF-IDF (Term Frequency-Inverse Document Frequency) is a numerical statistic that is intended to reflect how important a word is to a document in a collection or corpus. It is calculated as:

$$\mathrm{TF-IDF}(t, d, D) = \mathrm{TF}(t, d) \times \mathrm{IDF}(t, D) \quad (1)$$

where:

TF(t, d) = (Number of times term t appears in document d) / (Total number of terms in document d)

IDF(t, D) = log(Total number of documents / Number of documents containing term t)

This value increases proportionally to the number of times a word appears in the document but is offset by the frequency of the word in the corpus. This helps to adjust for the fact that some words appear more frequently in general.

To address class imbalance in the dataset, the Synthetic Minority Over-sampling Technique (SMOTE) was employed. SMOTE generates synthetic instances of minority classes by interpolating between existing instances.

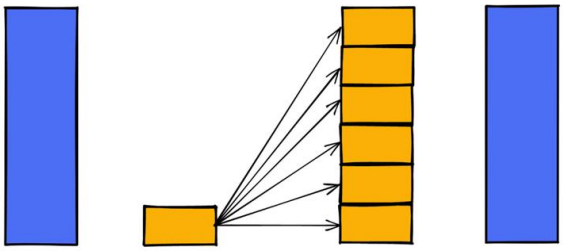


Fig.3. Upsampling of the SMOTE algorithm

In Fig.3., the yellow column in represents the minority class. Upsampling is done with the SMOTE algorithm to increase the minority class to the number of other classes. In this context, the distribution of classes before using the SMOTE algorithm and the distribution of classes after using this algorithm are shown in Fig. 4.

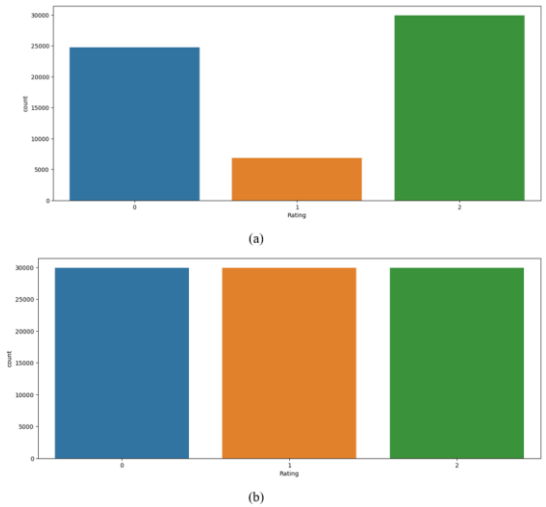


Fig.4. (a) Distribution of the Rating feature before the SMOTE process (b) Distribution of the Rating feature after the SMOTE process

After the SMOTE process, the number of elements in all classes was 29,937. Thus, the imbalance in the data set was eliminated. The machine learning methods used in this study, LSTM and the large language model BERT, are explained in the section III. Mathematical Background.

III. Mathematical Background

A. Sentiment Analysis using Machine Learning Algortihms

After the application of the Synthetic Minority Over-sampling Technique (SMOTE), the number of elements for each class in the dataset was balanced to 29,937, effectively eliminating the previously existing class imbalance. This balance allowed for a more equitable basis for training subsequent machine learning models. Following this preprocessing step, a predictive model was developed utilizing several well-established machine learning algorithms.

The algorithms employed included Logistic Regression, K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Decision Tree, Random Forest, and XGBoost. Each of these methods brings unique strengths to a machine learning task. Logistic Regression, often used for binary classification problems, models the probabilities for classification tasks by creating a linear decision boundary. KNN classifies data based on the most similar historical examples in the feature space, making it highly interpretable. SVM constructs a hyperplane in a high-dimensional space to separate different classes with a maximum margin, thus effective in non-linear separation problems. Decision Trees segment the data into branches to form a tree for prediction, which is simple to understand and interpret. Random Forest is an ensemble of Decision Trees, typically used to improve classification accuracy through voting from different trees. Lastly, XGBoost is a gradient boosting framework that uses a sequence of decision trees, where each tree corrects the errors of the previous ones, often achieving superior accuracy. The performance of these models was evaluated using classification metrics such as Accuracy, Precision, Recall, and F1-Score, providing a comprehensive measure of model efficacy across various aspects of predictive validation.

B. Sentiment Analysis using LSTM Deep Learning Algorithm

In this study, sentiment analysis was performed using machine learning methods as well as the deep learning technique known as Long Short-Term Memory (LSTM) algorithm. The deep learning method employed here was compared with other approaches in the 'Findings and Discussion' section. LSTM, a special structure developed for Recurrent Neural Networks (RNNs), is particularly effective for time series data and sequence classification problems. It was proposed by Hochreiter and Schmidhuber in 1997 to address the challenges RNNs face in learning long-term dependencies [20].

Compared to standard RNNs, LSTMs possess the capability to retain information for extended periods. This capability is facilitated by the inclusion of three distinct gate structures within LSTM cells: the forget gate, the input gate, and the output gate. These gates regulate the internal state of the cell

and the output, thereby controlling the flow of information. The fundamental equations of LSTM are articulated as follows:

$$Forget\ Gate \quad ft = \sigma(Wf \cdot [ht-1, xt] + bf) \tag{2}$$
$$Input\ Gate \quad it = \sigma(Wi \cdot [ht-1, xt] + bi) \tag{3}$$
$$Cell\ state\ update \quad C\sim t = tanh(WC \cdot [ht-1, xt] + bC) \tag{4}$$
$$Final\ cell\ state \quad Ct = ft * Ct - 1 + it * C\sim t \tag{5}$$
$$Output\ gate \quad ot = \sigma(Wo \cdot [ht-1, xt] + bo) \tag{6}$$
$$Cell\ output \quad ot = \sigma(Wo \cdot [ht-1, xt] + bo) \tag{7}$$

σ represents the sigmoid activation function, tanh the hyperbolic tangent activation function, W and b denote the weight matrices and bias vectors, respectively [21]. The LSTM architecture is presented in Fig.5.

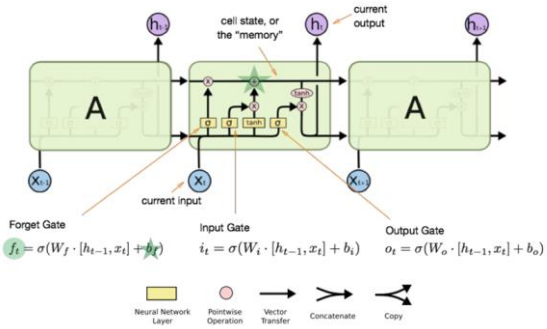


Fig.5. LSTM architecture

Fig. 5. illustrates the architecture of a Long Short-Term Memory (LSTM) network, highlighting the essential components involved in processing sequential data. An LSTM cell contains three main gates: the forget gate, input gate, and output gate, which collectively manage the flow of information through the cell. By iteratively updating the cell state, LSTMs maintain long-term dependencies, making them highly effective for tasks that involve sequential or time-dependent data, such as natural language processing and speech recognition.

C. Sentiment Analysis using BERT Large Language Model

Large Language Models (LLMs) are revolutionary artificial intelligence technologies in the field of Natural Language Processing (NLP). One such model, BERT (Bidirectional Encoder Representations from Transformers), was developed by Google in 2018 and has made significant advancements in language modeling [9]. BERT particularly utilizes a bidirectional Transformer architecture to achieve a deeper understanding of language. This model evaluates text in both left and right contexts simultaneously, offering much richer contextual representations compared to previous unidirectional models. The primary goal of BERT is to model the relationship of every word in a sentence with the words to its left and right at the same time. The model can be expressed as follows:

$$\mathrm{BERT(x)} = \mathrm{Transformer(x1{:}\,n)} \tag{8}$$

x1: n represents the sequence of words in the text. Fig.6 contains a diagram of the transformer architecture.

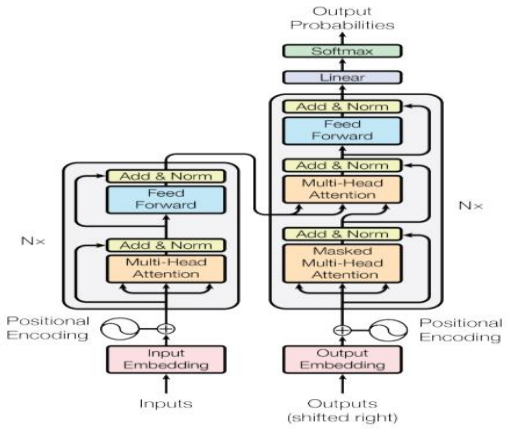


Fig.6. Transformer architecture [17]

Upon examining Fig.6, it is evident that the transformer architecture includes an Input and Positional Encoding section that converts the input sequence of words into fixed-size vectors (input embeddings). Positional encoding is used to indicate the order of words in the sequence. The Encoder part of the transformer architecture consists of six consecutive layers, each containing two sub-layers:

- *Multi-Head Attention:* Calculates the attention for each word of the input sequence towards the other words.
- *Feed-Forward Neural Network:* Applied independently to each word. The Decoder of the transformer architecture consists of six consecutive layers, each with three sub-layers:
- *Masked Multi-Head Attention:* Attention is applied to the previous words of the input sequence.
- *Encoder-Decoder Attention:* Attention is applied to the output of the Encoder.

In the output layer of the transformer architecture, the output taken from the last layer is passed through a linear and softmax layer, obtaining a probability distribution among possible words.

The BERT model, especially in NLP applications such as sentiment analysis, has demonstrated significant success. Sentiment analysis is the process of automatically detecting the emotional tone within texts, and BERT is highly effective in understanding the emotional context of texts. Once pretrained on a large dataset, the model can be fine-tuned to identify different emotional classes [23]

Introduced in 2017 by Vaswani and colleagues, the transformer architecture has sparked a revolution in deep learning models, providing a foundational approach for language models. Built upon the attention mechanism, this architecture calculates the relationships between all words in a sequence in parallel, offering a faster and more effective language modeling process. The core building block of the transformer is the attention mechanism, which calculates the relationship of each element in an input sequence with all other elements. This calculation begins with the generation of "query," "key," and "value" vectors for each component of the input sequence [22].

These vectors are computed as follows:

$$Q = XWQ \tag{9}$$
$$K = XWK \tag{10}$$

$$V = XWV \quad (11)$$

Examining Equation (9-11), Q, K, and V respectively represent the query, key, and value vectors; WQ, WK, and WV represent the learnable weight matrices. Attention scores are calculated by the dot product of query and key vectors and then normalized using the softmax function:

$$\text{Attention}(Q,K,V) = \text{softmax}(\frac{QK^T}{\sqrt{d_k}})V \quad (12)$$

In Equation (12), d_k is the dimension of the key vectors and is used as a scaling factor to balance the calculations. Transformer models significantly reduce training time due to their parallel computing capabilities, unlike RNN and LSTM-based models. They can also model long-range dependencies more effectively, improving performance in long sequences. The transformer architecture has achieved significant success in areas such as language modeling, machine translation, and text generation [9].

IV. CALCULATIONS RESULTS

In the section entitled "III. Mathematical Background," the outcomes derived from three distinct methodologies are elucidated through tabulated representations. This section provides a comprehensive comparative analysis, evaluating each method individually as well as in relation to one another, facilitating a deeper understanding of their respective merits and interrelations within the context of the study.

To assess the impact of Text Encoding techniques on performance metrics, an initial comparison of machine learning algorithms was conducted using the Count Vectorization method. The findings from this comparison are presented in Table-2.

TABLE II
RESULTS OF MACHINE LEARNING METHODS USING COUNT VECTORIZATION

Method	Accuracy	Precision	Recall	F1-Score
CV-LR	0.72	0.73	0.72	0.72
CV-KNN	0.47	0.61	0.47	0.44
CV-DT	0.59	0.60	0.60	0.60
CV-RF	0.70	0.71	0.70	0.70
CV-SVM	0.51	0.62	0.51	0.49
CV-XGBoost	**0.78**	**0.79**	**0.78**	**0.78**

When Table 2 is examined, it is understood that the CV-XGBoost method has the most successful performance parameters among the machine learning methods used by text coding using Count Vectorization.

The comparison table for machine learning methods with TF-IDF vectorization and performance metrics is included in Table-3.

TABLE III
RESULTS OF MACHINE LEARNING METHODS USING TF-IDF

Method	Accuracy	Precision	Recall	F1-Score
TF-IDF-LR	0.68	0.68	0.68	0.68
TF-IDF-KNN	0.42	0.58	0.57	0.57
TF-IDF-DT	0.53	0.54	0.54	0.54
TF-IDF-RF	0.68	0.72	0.69	0.69
TF-IDF-SVM	0.49	0.49	0.49	0.49
TF-IDF-XGBoost	**0.76**	**0.77**	**0.76**	**0.76**

When Table 3 is examined, it is understood that the method with the highest performance metrics is the TF-IDF-XGBoost method.Within the scope of the article, sentiment analysis was conducted using the LSTM algorithm as a deep learning method. This deep learning approach was implemented using Python programming, utilizing the TensorFlow library. For data preprocessing, the NLTK library was employed. Subsequently, the dataset was divided into two parts: training and testing. In this study, 80% of the dataset was allocated for training and 20% for testing. The text data were tokenized using a Tokenizer, limited to 50,000 words (num_words=50000). Additionally, an out-of-vocabulary token '<OOV>' was specified for words do not present in the text. After the word index was created by the Tokenizer, the texts were sequenced and converted into arrays of uniform length using the pad_sequences function. This process was performed for both the training and testing sets.

Labels were one-hot encoded using the LabelBinarizer class from the sklearn library. This encoding facilitates the numerical representation of classes, enabling the model to learn these classes more easily.

Subsequently, the construction of the LSTM model was addressed. At this stage, the LSTM model was built using the following layers:

- *Embedding Layer:* Creates an 8-dimensional embedding matrix based on the total number of words.
- *Bidirectional LSTM Layer:* Contains a bidirectional LSTM layer with 16 units.
- *Dropout Layers:* 50% dropout is applied to prevent overfitting.
- *Dense Layers:* Includes a 16-unit layer with relu activation followed by an output layer with softmax activation for three classes. The summary of the established LSTM Model is presented in Fig.6.

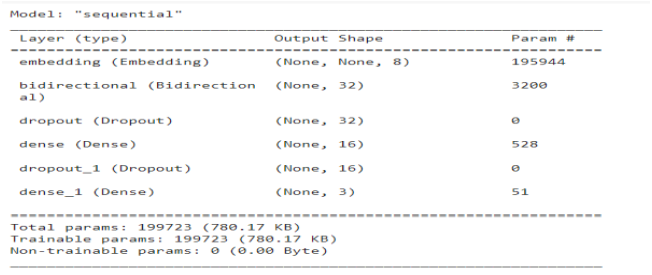


Fig.6. LSTM Model Summary

The model was compiled using the Adam optimization algorithm (learning_rate=0.0001). The loss function selected was categorical_crossentropy. The performance of the model was evaluated using accuracy, precision, recall, and AUC metrics.

The model was trained over 25 epochs on the training data (train_padded and train_labels). Validation of the model was conducted using the test data (test_padded and test_labels). The results obtained during the model training are presented in Table 4.

In the first epoch, the training data loss was calculated as 0.9893, accuracy as 0.4848, precision as 0.5813, recall as

0.1441, and AUC as 0.6938. For the validation data, the loss was observed as 0.8762, accuracy as 0.6192, precision as 0.7014, recall as 0.3521, and AUC as 0.7888. These results indicate that at the initial stage, the model is at the beginning of its learning process and its performance needs optimization. In the second and third epochs, the losses decreased to 0.8080 and 0.6815, respectively, while accuracy values increased to 0.6777 and 0.7563. Notably, in the second epoch, the rise in validation accuracy to 0.7592 demonstrates an improvement in the model's generalization ability. In the third epoch, the drop in validation loss to 0.6114 indicates better performance on both the training and validation data.

TABLE IV
PERFORMANCE VALUES OF THE LSTM MODEL IN THE TRAINING PHASE

Epoch	Loss	Accuracy	Precision	Recall	AUC	Val_ Loss	Val_ Accuracy	Val_ Precision	Val_ Recall	Val_ AUC
1	0.989	0.4848	0.5813	0.144	0.693	0.876	0.6192	0.7014	0.3521	0.7888
2	0.808	0.6777	0.7579	0.457	0.827	0.661	0.7592	0.7902	0.712	0.8854
3	0.681	0.7563	0.7961	0.664	0.876	0.6114	0.7777	0.7954	0.7489	0.8996
4	0.643	0.7784	0.809	0.709	0.889	0.595	0.7824	0.8012	0.758	0.9047
5	0.617	0.7879	0.8181	0.731	0.897	0.587	0.7856	0.8026	0.7591	0.9073
6	0.599	0.7948	0.822	0.744	0.903	0.580	0.7886	0.8077	0.7643	0.9094
7	0.588	0.7982	0.8274	0.753	0.907	0.576	0.7908	0.8092	0.7677	0.911
8	0.574	0.8022	0.8327	0.759	0.911	0.575	0.7898	0.8095	0.766	0.9116
9	0.563	0.8058	0.8352	0.764	0.915	0.576	0.7902	0.8066	0.7683	0.9115
10	0.554	0.8092	0.8375	0.769	0.918	0.575	0.7891	0.8064	0.7664	0.9123
11	0.547	0.8112	0.8407	0.772	0.929	0.579	0.7882	0.8041	0.7655	0.911
12	0.541	0.8139	0.843	0.776	0.923	0.577	0.7877	0.8055	0.7661	0.912
13	0.530	0.817	0.8455	0.781	0.925	0.580	0.7861	0.8057	0.7641	0.9116
14	0.524	0.8191	0.8476	0.783	0.927	0.581	0.7863	0.8058	0.7622	0.9116
15	0.515	0.8209	0.8503	0.787	0.93	0.583	0.7863	0.8054	0.7615	0.9113
16	0.509	0.8221	0.8523	0.789	0.931	0.587	0.7846	0.8044	0.7625	0.9111
17	0.503	0.8241	0.8543	0.791	0.934	0.609	0.7852	0.8022	0.7632	0.9096
18	0.497	0.8243	0.8552	0.792	0.935	0.599	0.7837	0.804	0.7605	0.91
19	0.491	0.8272	0.857	0.795	0.937	0.594	0.7807	0.8009	0.7567	0.9095
20	0.486	0.8274	0.8585	0.796	0.938	0.616	0.7824	0.8011	0.7602	0.9088
21	0.482	0.8292	0.8608	0.798	0.939	0.607	0.7795	0.7997	0.754	0.9083
22	0.478	0.8292	0.8616	0.798	0.941	0.624	0.779	0.7981	0.7549	0.9077
23	0.469	0.8313	0.8639	0.801	0.943	0.6355	0.7774	0.7986	0.7535	0.9062
24	0.467	0.8317	0.8655	0.801	0.943	0.6337	0.776	0.7986	0.7529	0.9056
25	0.461	0.8329	0.8681	0.803	0.945	0.6372	0.7771	0.7972	0.7536	0.9064

In the seventh and eighth epochs, the model achieved the lowest loss values (0.5882 and 0.5743, respectively) and high accuracy values (0.7982 and 0.8022). The decrease in validation losses to 0.5769 and 0.5754 suggests that the model has found a good balance during the training process. During this period, the precision and recall values for the validation data also remained high. In the later epochs (20-25), the gap between training and validation losses narrowed. For instance, in the 25th epoch, the training loss was recorded at 0.4618, while the validation loss was 0.6372. This indicates that the model has reached a stable learning process, reducing the risk of overfitting. The high levels of accuracy, precision, and recall also imply that the model's generalization ability is maintained.

Throughout the training process, a consistent decrease in loss values and an increase in accuracy values were observed. The diminishing performance gap between the training and validation data indicates an improvement in the model's generalization ability and a successful completion of the learning process. Additionally, the increase in precision, recall, and AUC values suggests that the model's classification performance has improved. These results demonstrate that the model can be effectively used for text classification tasks.

Sentiment analysis detection was performed using a BERT (Bidirectional Encoder Representations from Transformers) based model. The BERT model was downloaded from TensorFlow Hub and used with preprocessing layers. These layers facilitate the preprocessing of text data and its processing by the BERT encoder. Text data is passed through the BERT preprocessing layer to make it suitable for the BERT encoder. The output from the BERT encoder, obtained using 'pooled_output', serves as a summarized representation of the text data. This output is then passed through deep neural network (DNN) layers for classification. Firstly, a 20% dropout was applied to help prevent overfitting. Subsequently, a dense layer with a softmax activation function was added. This layer was configured to perform a multi-class classification with three classes.

During the compilation of the model, 'categorical_crossentropy' was selected as the loss function, and the Adam optimization algorithm was used as the optimizer (learning_rate=0.0005). The performance of the model was evaluated using accuracy, precision, recall, and AUC metrics. The model was trained on the training data (X_train and train_labels) for 25 epochs with mini-batch sizes of 32. The validation data (X_test and test_labels) was used for validation. This process enabled the model to optimize its learning and performance for the text classification task. The resulting model demonstrated good performance with high accuracy, precision, recall, and AUC values on both the training and validation data. The results obtained during the model training are presented in Table 5.

Upon examining Table 5, it is observed that in the first epoch, the training data loss was calculated as 0.7804, accuracy as 0.6793, precision as 0.7276, recall as 0.5869, and AUC as 0.8348. For the validation data, the loss was observed as 0.7028, accuracy as 0.7341, precision as 0.7668, recall as 0.6856, and

AUC as 0.8688. These results indicate that at the initial stage, the model is at the beginning of its learning process and its performance needs optimization. In the second and third epochs, the losses decreased to 0.6963 and 0.6799, respectively, while accuracy values increased to 0.7238 and 0.7302. The increase in validation accuracy to 0.7479 and 0.7514 demonstrates an improvement in the model's generalization ability. In the third epoch, the drop in validation loss to 0.6457 indicates better performance on both the training and validation data.

TABLE V
PERFORMANCE VALUES OF THE LSTM MODEL IN THE TRAINING PHASE

Epoch	Loss	Accuracy	Precision	Recall	AUC	Val_ Loss	Val_ Accuracy	Val_ Precision	Val_ Recall	Val_ AUC
1	0.780	0.6793	0.7276	0.586	0.834	0.702	0.7341	0.7668	0.685	0.868
2	0.696	0.7238	0.7675	0.660	0.87	0.661	0.7479	0.7971	0.684	0.886
3	0.679	0.7302	0.7718	0.670	0.875	0.645	0.7514	0.7923	0.705	0.890
4	0.670	0.7339	0.775	0.678	0.879	0.639	0.7517	0.8	0.685	0.892
5	0.669	0.7344	0.7748	0.680	0.879	0.632	0.7538	0.7987	0.700	0.894
6	0.661	0.7364	0.7759	0.683	0.881	0.632	0.7566	0.7916	0.719	0.893
7	0.664	0.7362	0.7752	0.682	0.881	0.625	0.7588	0.8051	0.706	0.897
8	0.661	0.7384	0.7785	0.685	0.8831	0.6272	0.758	0.7942	0.7176	0.896
9	0.666	0.7334	0.774	0.6807	0.881	0.6229	0.76	0.8097	0.7	0.8992
10	0.662	0.7386	0.7771	0.685	0.882	0.622	0.7607	0.8123	0.695	0.899
11	0.661	0.7379	0.7778	0.685	0.883	0.624	0.758	0.797	0.714	0.896
12	0.663	0.7374	0.7769	0.641	0.882	0.623	0.7592	0.8048	0.707	0.898
13	0.664	0.7367	0.7759	0.683	0.882	0.621	0.7594	0.8006	0.714	0.898
14	0.6628	0.7375	0.7764	0.685	0.882	0.625	0.7544	0.805	0.686	0.897
15	0.663	0.735	0.7759	0.683	0.882	0.621	0.7612	0.7982	0.718	0.898
16	0.6601	0.7387	0.7773	0.684	0.883	0.619	0.7599	0.8098	0.702	0.899
17	0.6623	0.7375	0.7769	0.686	0.882	0.619	0.7592	0.8067	0.705	0.899
18	0.69	0.7378	0.7774	0.685	0.883	0.620	0.7565	0.8012	0.705	0.898
19	0.660	0.7368	0.7769	0.685	0.883	0.618	0.7609	0.8121	0.703	0.900
20	0.669	0.7372	0.7772	0.683	0.883	0.622	0.7579	0.7994	0.712	0.897
21	0.661	0.7364	0.7756	0.684	0.882	0.620	0.7594	0.8046	0.704	0.898
22	0.662	0.737	0.7765	0.685	0.882	0.620	0.7591	0.8062	0.703	0.898
23	0.662	0.7376	0.7751	0.683	0.882	0.617	0.7605	0.8057	0.711	0.899
24	0.663	0.7349	0.7741	0.681	0.882	0.62	0.7593	0.8082	0.705	0.899
25	0.662	0.7376	0.7767	0.685	0.882	0.624	0.7605	0.7903	0.725	0.896

In the seventh and eighth epochs, the model achieved the lowest loss values (0.6646 and 0.6610, respectively) and high accuracy values (0.7362 and 0.7384). The decrease in validation losses to 0.6256 and 0.6272 suggests that the model has found a good balance during the training process. During this period, the precision and recall values for the validation data also remained high. In the later epochs (20-25), the gap between training and validation losses narrowed. For instance, in the 25th epoch, the training loss was recorded at 0.6629, while the validation loss was 0.6249. This indicates that the model has reached a stable learning process, reducing the risk of overfitting. The high levels of accuracy, precision, and recall also imply that the model's generalization ability is maintained.

Throughout the training process, a consistent decrease in loss values and an increase in accuracy values were observed. The diminishing performance gap between the training and validation data indicates an improvement in the model's generalization ability and a successful completion of the learning process. Additionally, the increase in precision, recall, and AUC values suggests that the model's classification performance has improved.

V. CONCLUSION

In the studies conducted within the scope of this research, the first step involved comparing machine learning methods using text encoding methods, namely Count Vectorization and TF-IDF, to perform sentiment analysis on Spotify comments. In the second step, sentiment analysis was conducted by optimizing the hyperparameters of the LSTM deep learning method. Finally, the recent focus on large language models was addressed, and sentiment analysis was performed using the BERT model. These methods were compared both individually and against each other to determine the most suitable approach for the specific problem.

To evaluate the performance of the methods used in this study, specific performance metrics were employed. The analyses and results obtained based on these metrics are summarized below. The performance results obtained using Count Vectorization (CV) and TF-IDF vectorization in the sentiment analysis conducted with machine learning methods are presented in Table 5 and Table 6. The CV-XGBoost method, with the highest performance values, stood out as the most successful method with 78% accuracy, 79% precision, 78% recall, and 78% F1-score. Similarly, the TF-IDF-XGBoost method also demonstrated high performance, achieving 76% accuracy, 77% precision, 76% recall, and 76% F1-score.

Sentiment analysis was performed using the LSTM algorithm as the deep learning method. Throughout the training process, a consistent decrease in loss values and a consistent increase in accuracy values were observed. The lowest loss values for the LSTM model were recorded as 0.5882 and 0.5743 in the seventh and eighth epochs, respectively, while the highest accuracy values reached 80.22%. These results indicate an

improvement in the model's generalization ability and classification performance.

In the sentiment analysis conducted using the BERT-based model, the model exhibited high performance in terms of accuracy, precision, recall, and AUC metrics. The model was effective in understanding the emotional context of the texts, and a consistent increase in accuracy values was observed. Throughout the 25 epochs of training, the accuracy and precision values remained high, demonstrating that the model maintained its generalization ability. The BERT language model achieved 73% accuracy.

When comparing the methods used in this study, the LSTM deep learning method emerged as the one with the highest performance. However, it is thought that fine-tuning large language models, which are a current focus of intense research, with localized information specific to the subject can achieve higher performance. In future studies, pre-trained language models can be customized and used for specific topics. Future work could focus on several key areas to enhance the scope and impact of this study. One promising direction is to AI-based model specifically for sentiment analysis within the music streaming domain, potentially improving its accuracy in detecting complex emotional nuances in user reviews. Additionally, incorporating data from multiple music streaming platforms could broaden the analysis, allowing for a more comprehensive understanding of user sentiment across diverse user bases and feedback types. Furthermore, exploring advanced model optimization techniques to balance computational efficiency with performance could facilitate the deployment of these models in real-time applications. By expanding the dataset and refining model parameters, future research could provide more robust insights into user satisfaction and emotional engagement in digital music platforms.

REFERENCES

[1] Khan, A. W., & Mishra, A. (2023). AI credibility and consumer-AI experiences: a conceptual framework. Journal of Service Theory and Practice.

[2] Miragoli, M. (2024). Conformism, Ignorance & Injustice: AI as a Tool of Epistemic Oppression.*Episteme*.

[3] Zheng, J. (2024). Music Sentiment Analysis and its Application in Music Therapy Based on AI Technology. *International Journal of Maritime Engineering*

[4] Chen, Y., & Sun, Y. (2024). The Usage of Artificial Intelligence Technology in Music Education System Under Deep Learning. *IEEE Access, 12*, 130546-130556

[5] Pires, I. M., Zafar, S., Iqbal, K., Sharif, M., Shah, Y. A., Khalil, A., Irfan, M. A., & Rosak-Szyrocka, J. (2024). Attention-aware with stacked embedding for sentiment analysis of student feedback through deep learning techniques. *PeerJ Computer Science, 10*.

[6] Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. Foundations and Trends® in Information Retrieval, 2(1–2), 1–135. DOI: 10.1561/1500000011

[7] Liu, B. (2012). Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies, 5(1), 1–167. DOI: 10.2200/S00416ED1V01Y201204HLT016.

[8] Jain, N., Kumar, S., & Fernandes, S. L. (2019). Machine Learning Techniques for Sentiment Analysis: A Review. In: Das, H., & Pattnaik, P. (Eds.), Computational Intelligence in Data Mining. Advances in Intelligent Systems and Computing, vol 711. Springer, Singapore. DOI: 10.1007/978-981-13-1810-8_19

[9] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint arXiv:1810.04805. DOI: 10.48550/arXiv.1810.04805

[10] Zhang, L., Wang, S., & Liu, B. (2020). Deep Learning for Sentiment Analysis: A Survey. IEEE Transactions on Affective Computing, 12(2), 314-332. DOI: 10.1109/TAFFC.2020.2984262

[11] Jhanji, R. (2024). Emotion Analysis from Music Using LSTM Models and Mel Spectrograms. Journal of Music Data Science, 8(1), 123-134. DOI: 10.1016/j.jmds.2024.01.001

[12] Martin-Gomez, A., Garcia, J., & Lopez, V. (2018). Comparative Study of Multi-Label Classification Algorithms for Emotion Recognition. Journal of Machine Learning Research, 19(45), 1-25.

[13] Tilloo, R., Patel, S., & Shah, A. (2021). Sentiment Analysis of Amazon Musical Instrument Reviews Using CNN and NLP Techniques. International Journal of Data Science and Analytics, 12(3), 234-245. DOI: 10.1007/s41060-021-00259-9

[14] Website. [Online]. Available: https://www.kaggle.com/datasets/mfaaris/spotify-appreviews-2022

[15] Bird, S., Klein, E., & Loper, E. (2009). Natural Language Processing with Python. O'Reilly Media.

[16] Hovy, E., & Lavid, J. (2010). Towards a 'science' of corpus annotation: A new methodological challenge for corpus linguistics. International Journal of Translation, 22(1), 13-36

[17] Manning, C. D., Raghavan, P., & Schütze, H. (2008). Introduction to Information Retrieval. Cambridge University Press

[18] Rajaraman, A., & Ullman, J. D. (2011). Mining of Massive Datasets. Cambridge University Press

[19] E. Cambria, B. Schuller, Y. Xia, and C. Havasi, "New Avenues in Opinion Mining and Sentiment Analysis," Intelligent Systems, IEEE , vol.28, no.2, pp. 15-21, 2013

[20] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735-1780

[21] Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. Neural Computation, 12(10), 2451-2471.

[22] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems

[23] Sun, C., Qiu, X., Xu, Y., & Huang, X. (2019). How to fine-tune BERT for text classification?. China National Conference on Chinese Computational Linguistics.

BIOGRAPHIES

Murat ŞİMŞEK completed his undergraduate and graduate education at Kırıkkale University, Department of Electrical and Electronics Engineering. He completed his master's degree on the effect of super-resolution methods on increasing the resolution of hyperspectral images. He completed his doctorate at Ankara University, Department of Electrical and Electronics Engineering. He completed his doctorate on real-time target detection in hyperspectral images. Throughout his career, he has worked in many different positions as an electrical and electronics engineer, electronic technologies specialist, artificial intelligence specialist and project manager. He has academic studies on image processing and artificial intelligence. He is head of Artificial Intelligence Engineering Department of Ostim Technical University.

Buğra Kağan Kayhan completed his undergraduate at Ufuk University, Department of Management Information Systems. He completed his master's degree on Emotion and content analysis of music data using machine learning methods at Ostim Technical University Software Engineering Department. Throughout his career, he has worked in many different positions as an software technologies specialist, artificial intelligence specialist.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 3, September 2025
Research Article

Calculation of Power Losses for SiC MOSFET Based 3-Phase 3-Level T-Type Inverter

Erkan Deniz and Berkan Turan

***Abstract*— In recent years, interest in highly efficient and compact power converters in power electronics applications has been increasing day by day. In this study, a SiC MOSFET-based 3-level (3L) T-Type inverter (TNPC) is proposed to obtain a high-efficiency and compact converter. Considering the MSCSM120HRM163AG-SiC-MOSFET intelligent power modules (IPM) developed by Microchip for 3L T-Type inverter, the power losses of a 3-phase 3-level T-Type inverter are calculated in MATLAB environment. To demonstrate the efficiency of SiC MOSFET based T-Type Inverter, a 3-phase 3L T-Type inverter feeding a vector-controlled 3-phase PMSM is simulated using MATLAB/Simulink and Simscape blocks. In the simulation, the PMSM is operated at different speed references under almost full load (42.09 Nm). While the PMSM operates at 3000 rpm under 40Nm load, the voltage-current waveforms of the SiC MOSFETs in the T-Type inverter are obtained. Using these waveforms and the data from the data sheets of the IPMs, the power losses of the 3L TNPC inverter are calculated for different switching frequencies. The Space Vector PWM method used to generate 50 kHz PWM signals for the 3L T-Type inverter also ensures that the dc-link capacitor voltages remain balanced. In addition, the variations of line-to-line voltage and dc-link capacitor voltages of the inverter and the variations of speed, torque, rotor position, d-q currents, and stator currents of the PMSM are given.**

Index Terms—**Three-level T-Type inverter, SiC MOSFET, power loss calculation, PMSM, space vector PWM (SVPWM).**

I. Introduction

FOR NEARLY a quarter century, Si-based multilevel PWM converters have been used in many applications, including motor drives, energy storage systems, static VAr compensation, HVDC, renewable energy systems, battery chargers in electric vehicles, microgrids, etc. Because, they have advantageous features such as being suitable for the use of lower-voltage rating switches, low dv/dt ratio that results from the reduced voltage stress, low common mode voltage, output voltage-current waveforms with lower THD, and smaller output filter size [1,2]. The low dv/dt minimizes Electromagnetic Interference (EMI) issues. However, increasing the number of levels in MLIs leads to an increase in the number of switches, which increases the design and control complexity and power losses of the converter [3]. Therefore, there is a growing interest in 3-level (3L) converters in low and medium voltage power electronics applications because 3L topologies require less devices and gate drivers that do not cause control complexity and large size. Multilevel converters come in three conventional topologies: cascade, capacitor-clamped, and diode-clamped. With this, as it eliminates the need for multiple isolated dc-link sources and large clamping capacitors, the 3L diode-clamped topology is widely preferred in industrial applications. There are three widely recognized types of diode clamping converters: conventional neutral point clamped (NPC) [4], active NPC (ANPC), and T-Type. While ANPC is proposed in [5] to solve the problem of unbalanced power losses between semiconductors in the NPC circuit, T-Type converter is proposed as a result of studies carried out to reduce the number of semiconductors in the NPC circuit in [6]. 3L T-type topology provides maximum efficiency and lower cost compared to other topologies such as 3L NPC, 3L ANPC, and 3L flying capacitor due to its simple operating principle, less power devices and isolated gate driver requirements, lower power losses, and smaller converter size [7-9]. Nevertheless, the semiconductors (S_1, S_4) in the legs connected to the +dc and -dc busbars of the 3L T-type converter must have higher-rated voltages than those in the other 3L topologies [10]. The efficiency of 3L NPC and T-type inverters is evaluated for switched reluctance motor drives in hybrid electric vehicles and it was shown that the T-type inverter provides superior energy savings [11]. In [12], the power loss comparison of Si IGBT-based 3L NPC and T-type inverters is investigated for different modulation techniques. Since there are not two series-connected switches in the inverter leg, the power losses of the T-type are lower than those of the NPC. Low power losses extend the life of the converter and reduce the size of the heat sinks. Moreover, smaller and less expensive filters can be employed to filter EMI noise since the T-type inverter permits reduced EMI emissions [13]. However, 3L-TNPC suffers from high power loop leakage inductance due to the complex power loop structure like other three-level topologies. To reduce the leakage inductances, a SiC-MOSFET based 3L-TNPC power modules are usually preferred. In [14], a SiC-MOSFET-based 3L-TNPC power module with hybrid packaging method is proposed.

The more compact and highly efficient power electronic converters, which are the basic building blocks of electrical energy conversion applications, are designed, the more certain it is that the size of the entire system will be significantly reduced and its efficiency increased [15]. Using wide bandgap

Erkan Deniz, is with Department of Electrical Electronics Engineering Firat University, Elazığ, Turkey,(e-mail: ededeniz@firat.edu.tr).
https://orcid.org//0000-0002-9048-6547

Berkan Turan, is with TCDD-Republic of Turkey State Railways, Elazığ, Turkey, (e-mail: turanberkan23@gmail.com).
https://orcid.org/0000-0003-4304-0099

Manuscript received Jan 14, 2025; accepted Feb 9, 2025.
DOI: 10.17694/bajece.1619631

(WBG) SiC MOSFETs for the 3L T-Type inverter can enable the creation of a compact and highly efficient power converter. Because, SiC MOSFETs offer higher breakdown voltages, higher switching frequencies, lower switching losses, higher operating temperatures, low turn-on resistances, low stray inductance, and smaller chip size compared to Si-based semiconductors, which allows for increased power density and reduced size of passive components in converters [16-18]. Using gate drivers with features such as “noise robustness, fast response protection, and accurate slew-rate control” will enable full utilization of SiC MOSFETs. Nevertheless, it raises concerns regarding electromagnetic compatibility (EMC) since high switching frequencies generate more overshoot, causing EMI noises, and making interference optimization difficult in printed circuit board (PCB) layout and EMI filter design [19,20]. Therefore, all power converters using high switching frequencies require the use of EMI filters. However, EMI filter design is a very complex and specialized subject of study and expertise depending on the converter type and PCB layout.

EMI, which emerges in power converters in two forms, such as *Conducted* through power traces and *Radiated* through electromagnetic fields, can hamper the proper operation of both the converter itself and nearby electronic devices. The process of designing filters for the suppression of EMI noises starts with the identification of EMI sources caused by high-frequency switching, fast current (di/dt) and voltage (dv/dt) changes in the converter power circuit. EMI sources vary for different PCB layouts and different PWM methods, even for the same converter. For this purpose, all parasitic capacities between SiC-MOSFETs and +dc, -dc, neutral busbars, and heat sinks are calculated using the designed PCB layout [21]. In the next step, based on an applicable regulatory standard such as IEC, CISPR, FCC, maximum admissible EMI levels at various frequencies should be determined, and frequency ranges around the switching harmonics where noise is most prominent should be defined. For example, according to standards such as CISPR 16-1-2 and IEC 61000-6-3, the frequency range in which the electromagnetic noise of a device radiated over electrical lines must be limited is 150 kHz to 30 MHz [22].In the last stage, common mode and differential mode filters are usually preferred and designed. Inductors in these filters are selected according to their impedance characteristics at noise frequencies, while capacitors are selected for their high-frequency filtering properties. Simulation tools are used to verify the performance of the filter, followed by prototype development and EMI testing to verify its effectiveness [23].

In this work, a SiC-MOSFET based 3L T-Type topology is proposed to obtain a highly efficient and compact dc/ac converter for low-voltage (<1kV) power electronics applications. The switching frequency is chosen as 50kHz to reduce the size of the passive components and switching losses in the converter circuit. In order to demonstrate the high efficiency operation of the inverter circuit, a 3L T-type inverter driving a 3-phase PMSM is simulated in MATLAB. In order to reduce EMI issues while utilizing the high switching frequency, Microchip MSCSM120HRM163AG-SiC-MOSFET intelligent power modules (IPMs) are considered for the phase legs of the 3L T-Type inverter. The conduction and switching losses of the 3L TNPC are calculated by taking into account the current-voltage waveforms of the SiC-MOSFETs and the information on the data sheets of the SiC-MOSFETs in the most demanding operating condition of the inverter. For a more accurate calculation, the average and effective values of the currents flowing through the SiC-MOSFETs and body diodes are also measured for each half of the switching period. Similar measurements are made for the voltages at the terminals of the SiC-MOSFETs and gap diodes.

II. Operation of 3-Level T-Type Inverter Topology

In high-power applications and especially at high switching frequencies, the large number of power devices on PCBs increases EMI noises. This causes overshoots and ringing in the converter output voltage. For this reason, as a result of studies carried out to reduce the number of semiconductors for 3-level converters, the first T-Type inverter was proposed in [6]. A 3L T-Type inverter, also known as a Neutral Point Piloted (NPP) inverter, consists of a split dc-link and twelve MOSFETs with body diodes, as seen in Figure 1. There are two series switches on each phase leg as in a two-level inverter, and the middle point of one leg is connected to the middle of the dc-link via a bidirectional power device. In Fig.1, S_{1X} and S_{4X} have twice the rated voltage of switches S_{2X} and S_{3X} connected to the neutral point [24].

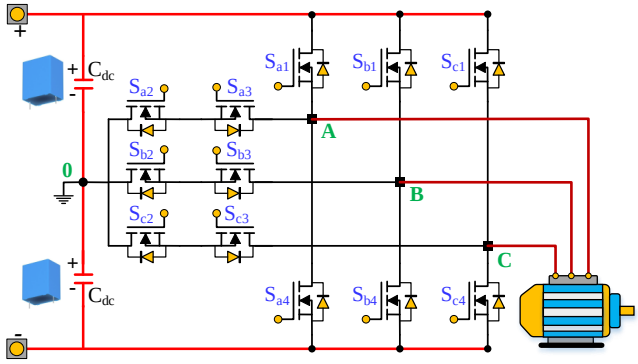


Fig.1. Circuit of 3-phase 3L T-Type inverter

TABLE I
SWITCHING STATES OF A 3L T-TYPE INVERTER

		Power Devices				V_{out}
		S_{1X}	S_{2X}	S_{3X}	S_{4X}	V_{A0}
Switching States	P	1	1	0	0	$+V_{dc}/2$
	O	0	1	1	0	0
	N	0	0	1	1	$-V_{dc}/2$

The switching states of the 3L T-Type inverter and the corresponding output phase-neutral voltage values are given in Table I. The output of a phase leg is connected via S_{1X} and S_{4X} to the dc-link's positive (P) and negative (N) voltage levels, and via S_{2X}-S_{3X} to the zero (O) level [25]. As seen in Table I, inverter phase-neutral voltage is obtained in 3-levels as $-V_{dc}/2$, 0 and $+V_{dc}/2$ for the three possible switching states shown as N, O, and P. The output phase-phase voltage of the inverter has five levels at $-V_{dc}$, $-V_{dc}/2$, 0, $+V_{dc}/2$, and $+V_{dc}$ values. The positive and negative current pathways during the change from the P to the O switching state are shown in Fig.2 [24].

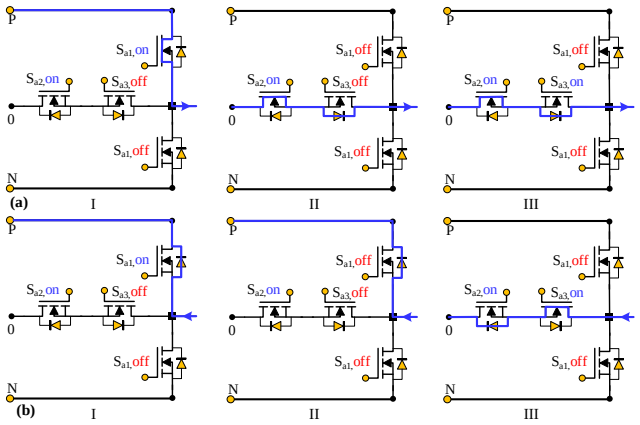


Fig.2. Current commutation during transition from switching state P to 0 (a) positive I_{out} , (b) negative I_{out}

III. SVPWM TECHNIQUE FOR 3L T-TYPE INVERTER

The Space Vector PWM (SVPWM) technique for a voltage source inverter is based on representing the ac desired output voltages from the inverter with a reference voltage space vector in α-β space. The instantaneous values of desired ac output voltages can be used to determine the $\vec{V}_{ref}$ vector's amplitude and phase angle. If the voltages are sinusoidal and balanced, the vector will rotate at a constant angular velocity and will have a constant amplitude [26, 27]. Considering the 3-phase legs of the inverter, the 3L T-Type inverter has 27 switching states, and each of these switching states can be represented as a vector using Eq.(1). As a result, a total of 27 voltage vectors are formed, as seen in Table II. Figure 3 shows the space-vector diagram of the 3L T-Type inverter, which is created when all voltage vectors are positioned in the α-β plane [28]. As seen in Figure 3, while the voltage vectors are constant, the reference voltage vector rotates with an angular speed ω. The space-vector diagram has six sectors, each of which is separated into four regions, as shown in Fig.3.

TABLE II
VOLTAGE VECTORS CORRESPONDING TO SWITCHING STATES

Type	Vectors						Length
Zero	PPP	OOO	NNN				**0**
Small	POO	PPO	OPO	OPP	OOP	POP	$V_{dc}/3$
	ONN	OON	NON	NOO	NNO	ONO	
Medium	PON	OPN	NPO	NOP	ONP	PNO	$V_{dc}/\sqrt{3}$
Large	PNN	PPN	NPN	NPP	NNP	PNP	$2V_{dc}/3$

$$\vec{V} = \frac{2}{3}\left(V_a\, e^{j0} + V_b\, e^{j2\pi/3} + V_c\, e^{-j2\pi/3}\right) \qquad (1)$$

The calculation of voltage vectors' dwell times for a 3L T-Type inverter is based on the principle of "*Volt-Second Balancing*" principle [29]. Therefore, the reference vector is formed by the three closest vectors. For example, when the reference vector $\vec{V}_{ref}$ is located in region-2 of Sector-1, the three closest vectors are $\vec{V}_1$, $\vec{V}_2$, and $\vec{V}_7$. In this case, voltage-time equalization is as shown in Eq.(2).

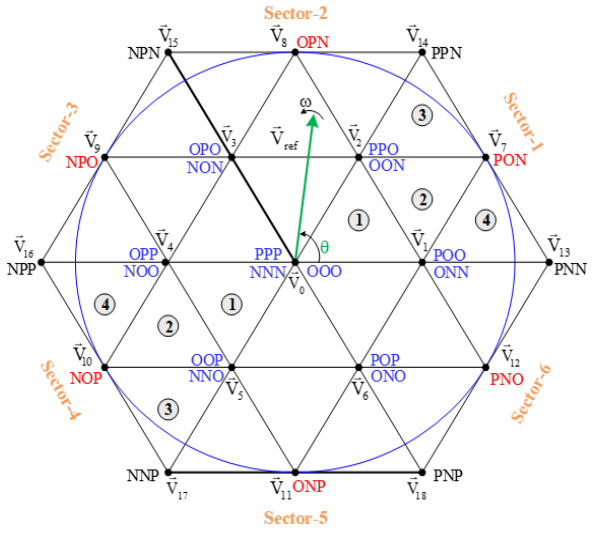


Fig.3. Space-vector diagram for 3L T-Type inverter

$$\begin{aligned} \vec{V}_1.T_a + \vec{V}_7.T_b + \vec{V}_2.T_c &= \vec{V}_{ref}.T_s \\ T_a + T_b + T_c &= T_s \end{aligned} \qquad (2)$$

Here, T_a, T_b, and T_c are the dwell times of $\vec{V}_2$, $\vec{V}_7$ and $\vec{V}_{14}$, respectively. By writing voltage-second balancing equations for other regions in Sector-1 and substituting vector values in these equations, dwell times for all regions can be found as given in [28]. After calculating the dwell times for each sector and region, PWM signals should be generated by determining the application order of space vectors according to the sector and region where $\vec{V}_{ref}$ is located. In this study, the order of application of vectors for minimum THD is determined [29]. For region-3 and region-4 in Sector-2, the application order of space vectors and accordingly the switching schemas of phases a, b, and c of the T-Type inverter are given in Fig.4.

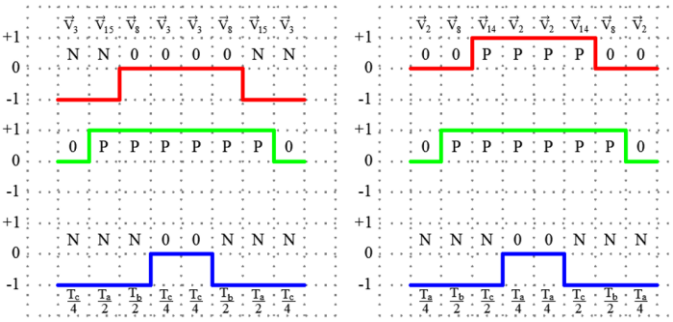


Fig.4. Switching states of inverter phase legs (a) region-3, (b) region-4

IV. FIELD ORIENTED CONTROL FOR 3-PHASE PMSM

When forming the dynamic mathematical model of a PMSM, the components of the machine can be expressed in three different reference frameworks such as stator reference frame (3-phase model), fixed 2-phase reference frame (2-phase α-β model), and rotor reference frame (2-phase d-q model) [30, 31]. Since stator winding inductances vary depending on the rotor

position, the rotor-referenced d-q model is preferred in this study. In this model, the stator currents are converted into virtual two-phase d-q axis currents that rotate with the electrical speed of the rotor. Thus, inductance values that do not change depending on the position of the rotor are obtained for the counterpart of the stator windings on the d-q axis set. Since the rotor speed is equal to the synchronous speed, this model is also called the d-q model with synchronous reference [32, 33]. The Kirchhoff voltage law can be used to derive the PMSM equations, which are provided in Eq. (3) and (4).

$$u_d = R_s\, i_d + \frac{d}{dt}\psi_d - \omega_r\, \psi_q \tag{3}$$

$$u_q = R_s\, i_q + \frac{d}{dt}\psi_q - \omega_r\, \psi_d \tag{4}$$

Where ω_r is the electrical rotor frequency and R_s is the stator winding resistance. The linkage fluxes of the d- and q-axes, $\vec{\psi}_d$ and $\vec{\psi}_q$, are given in Eqs. (5) and (6).

$$\psi_d = \psi_m + L_d i_d \tag{5}$$

$$\psi_q = L_q i_q \tag{6}$$

Where ψ_m is the flux linkage due to the rotor magnets and its derivative is zero since it is a constant value. The electromagnetic torque for a PMSM is expressed in terms of the d and q-axis components as follows:

$$T_e = \frac{3}{2} P(\psi_d i_q - \psi_q i_d) \tag{7}$$

Where P is the pole pair number. To achieve a basic control similar like a dc motor, I_q and I_d are controlled by a vector control algorithm.

Field Oriented Control (FOC) is the most widely known vector control technique. The basic principle of FOC is the separation of 3-phase stator currents into two separate d-q components that produce torque and flux. Thus, magnetic flux and torque control are controlled independently of each other. Torque and flux are controlled by the stator current components in the rotor reference plane, I_d and I_q, respectively. However, the effectiveness of the I_d component in control is eliminated because permanent magnets supply the flux in PMSMs. Therefore, controlling just I_q will be sufficient in moment control. To obtain the components of the stator currents in the rotor reference plane (I_d-I_q) in FOC, position information about the rotor is required. The block diagram of the FOC for a PMSM is given in Figure 5 [23, 31]. As seen in Fig.5, the speed of the motor is measured with an encoder and compared with the reference speed. The error is applied to a PI speed controller and the reference I_{q_ref} is obtained from the controller output. Since the flux for the motor is provided by permanent magnets, the reference value I_{d_ref} is taken as zero. Besides, after the 3-phase stator currents (I_{abc}) of the motor are measured with sensors, d-q current components (I_d-I_q) are obtained using the Park Transformation. These current components are compared with the reference current components (I_{d_ref}, I_{q_ref}), and the current errors are applied to PI controllers. At the output of the PI controllers, the components of the reference voltages (V_{d_ref}, V_{q_ref}) that the T-Type inverter is required to produce are obtained. These reference voltages are then converted into 3-phase reference voltages for the three-level SVPWM algorithm using the Inverse Park Transform. As a result, the 3L T-Type inverter is switched with SVPWM signals, and the voltages required by the PMSM for different speed references and load cases are generated.

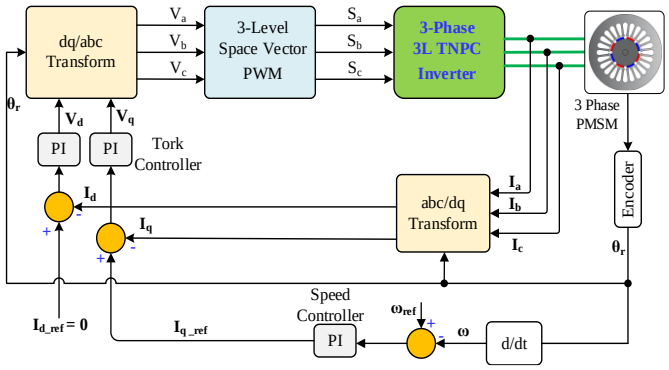


Fig.5. The block diagram of the FOC for a PMSM

V. MATLAB SIMULATION OF 3L T-TYPE INVERTER SUPPLYING 3-PHASE PMSM

In this study, the field-oriented vector control of a 3-phase PMSM fed by a 3L T-Type inverter is simulated by using MATLAB/*Simulink* and *Simscape Electrical* blocks. The general view of simulation is given in Figure 6. The eighteen SiC MOSFETs are used to create a 3-phase 3L T-Type inverter, which is switched by 50kHz 3-level SVPWM signals. In simulation, a PMSM with 8 poles, 3000 rpm, and 42.09 Nm nominal values is preferred as a motor. Other parameters of the PMSM are given in *Appendix-A*.

In the simulation, the variations of the motor speed, torque, rotor position, stator currents, inverter line-to-line voltage, and inverter dc-link capacitor voltages are analyzed for different reference speeds. The motor is continuously loaded with a constant load of 40 Nm. As seen in Figure 7, the motor reaches the reference speed given as 1500 rpm at t=0 in approximately 0.1s. The PMSM smoothly follows this reference speed which is kept constant until t=1.0s. Then the reference speed is changed to 3000 rpm at t=1.0s and -3000 rpm at t=2.0s. The vector control enabled the engine to reach these reference speeds in a very short time and to follow the reference speed stably. The change of moment produced by PMSM is given in Figure 8.As seen in Figure 8, it is seen that the PMSM produces high moments in transients, while it produces a constant moment of approximately 40 Nm in steady states. Also, at the reference speed change of -3000 rpm at t=2.0s, the PMSM produces a negative moment.

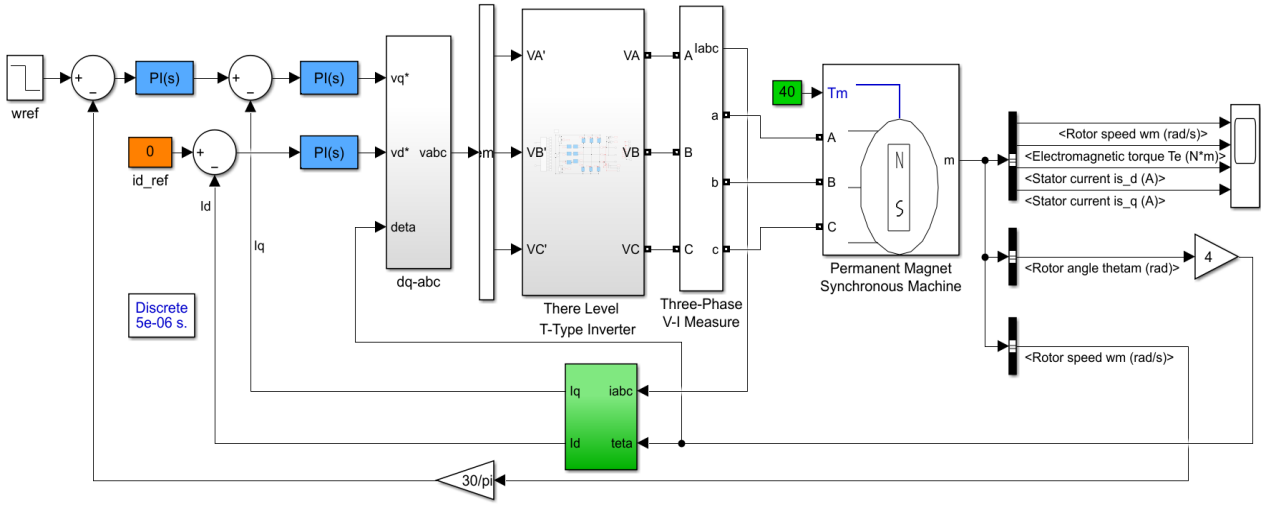


Fig.6. MATLAB simulation of PMSM fed by 3L T-Type inverter

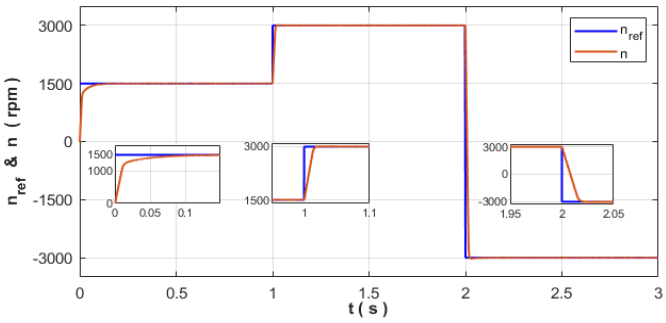


Fig.7. The speed of PMSM for Te= 40N.m and n_{ref} = 1500rpm, ±3000rpm

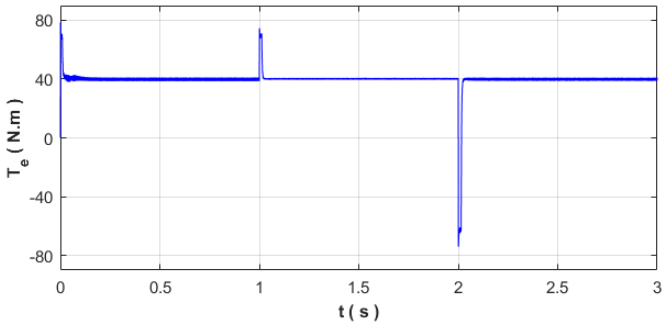


Fig.8. Tork of PMSM for T_e= 40N.m and n_{ref} = 1500rpm, ±3000rpm

The rotor's electrical position angle (Theta) is given in Figure 9. As seen from the figure, the theta passes through the zero point every 2π radians. In the reference speed changes at t=0, 1.0s and 2.0s, the FOC changed the rotor's position angle appropriately and quickly, as seen in Figure 9. The rotor's electrical angle has a frequency of 100Hz for 1500 rpm and 200 Hz for ±3000 rpm.

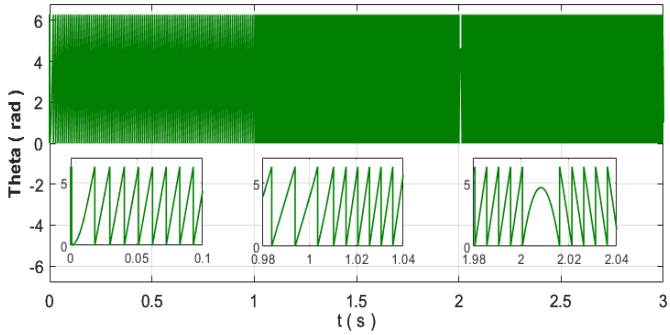


Fig.9. The rotor's position angle for T_e= 40N.m and n_{ref} = 1500rpm, ±3000rpm

For reference speeds of 1500 rpm, and ±3000 rpm under T_e=40 Nm load, the waveforms of the flux component I_d and moment component I_q of the current are given in Figure 10 and the 3-phase stator currents with THD analysis are given in Figure 11. As can be seen from the figures, both stator currents and I_q current reached high values during the transient times. However, in steady states, both remain constant at their nominal values. Besides, as expected, the value of the I_d component is not affected by the speed changes and remains at zero all the time. The total harmonic distortion of stator currents is obtained as 1.67% for 1500 rpm speed reference at t=0.5s and 1.36% for +3000 rpm reference speed at t=1.5s.

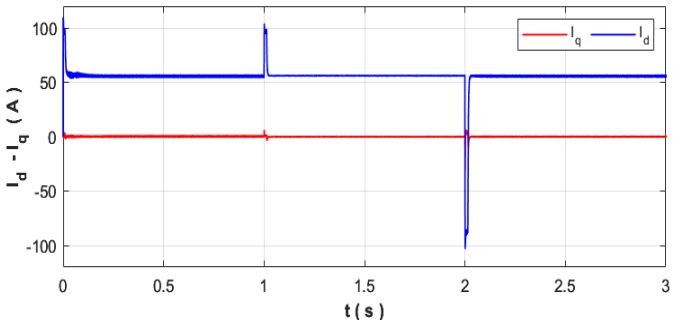


Fig.10. The I_d-I_q components for T_e= 40N.m and n_{ref} = 1500rpm, ±3000rpm

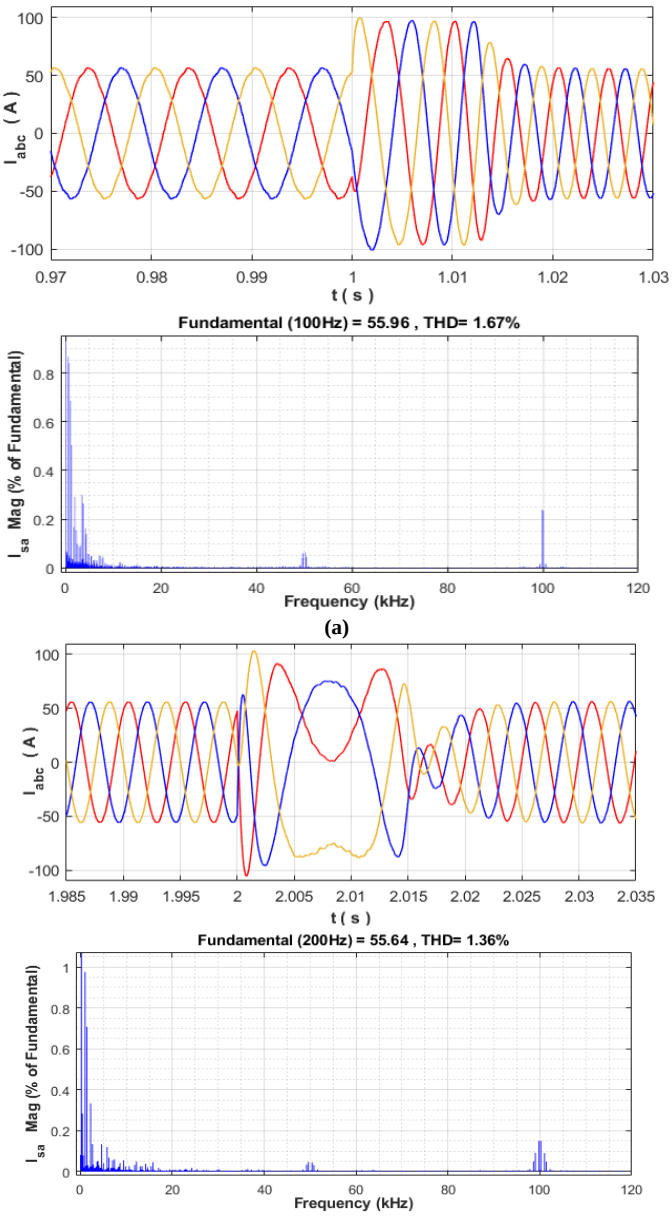


Fig.11. The stator currents for T_e= 40Nm, **a)** transition from +1500rpm to +3000rpm and THD analysis for t=0.5s, **b)**+3000rpm to -3000rpm and THD analysis for t=1.5s

The waveforms of the line-to-line voltages produced by the 3L T-Type inverter feeding the PMSM are given in Figure 12 (a) and (b). As seen from the figures, the inverter produces a three-level line-to-line voltage of 100 Hz for a speed of 1500rpm up to t =1s, while it produces a five-level voltage of 200Hz after reaching the reference speed of 3000 rpm. While it produces a three-level voltage of 100Hz during its transition to -3000 rpm reference speed at t=2s, it produces a five-level voltage of 200 Hz again when it reaches the reference speed. The harmonic analysis of the line-to-line voltage generated by the inverter while the PMSM is running at 3000rpm is given in Figure 13. As seen in the figure, the 3L T-Type inverter produces the fundamental harmonic has 269V amplitude and 200Hz frequency with a THD of 43.69%.

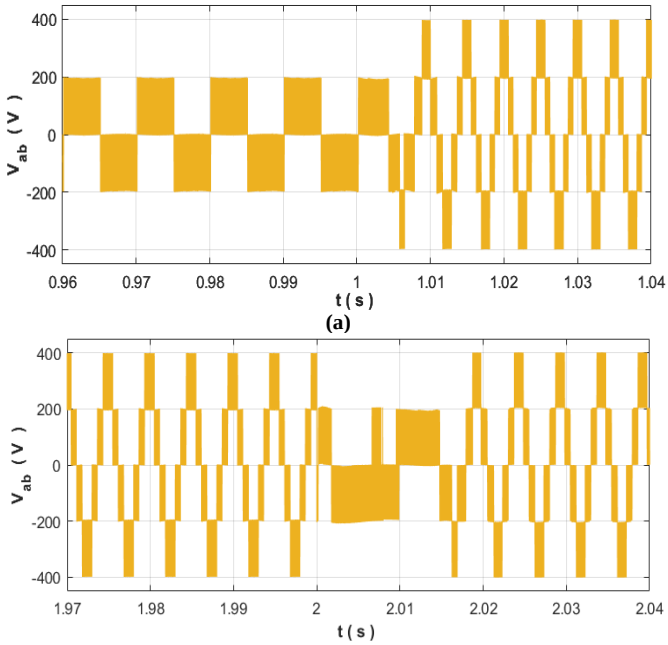


Fig.12. The line-to-line voltages of 3L T-Type inverter for transitions **a)** from 1500 to 3000 rpm, **b)** from 3000 to -3000 rpm

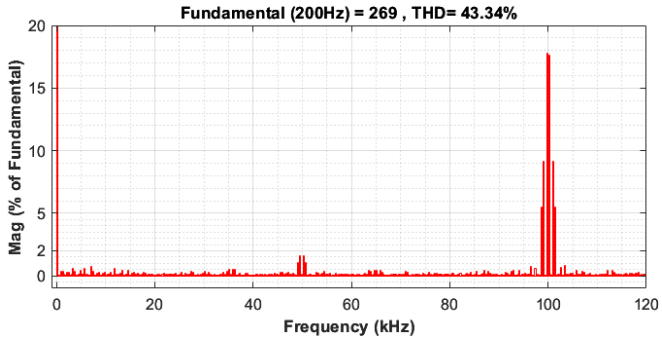


Fig.13. The THD analysis of line-to-line voltage (t=1.5 s)

The voltage change of the dc-link capacitors of the 3L T-Type inverter during the operation of the PMSM at different reference speeds is given in Figure 14. As seen in the figure, the voltage on each capacitor remained approximately constant at 200V, which is half of the dc voltage, in all operating states. While the fluctuations in capacitor voltages are a maximum of 10V during reference speed changes, it is seen that there is less fluctuation in steady states.

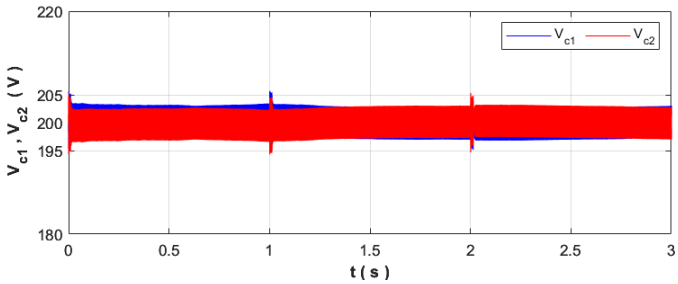


Fig.14. The dc-link capacitor voltages of T-type inverter

VI. Calculation of Power Losses for SiC MOSFET Based 3L T-type Inverter

The calculation of power losses for a power electronic converter is certainly done before starting the practical design. One of the most used methods is the calculation of semiconductor switches using datasheet information and any simulation program. The main purpose of these calculations is to obtain information about the efficiency of the entire system and to determine the size of the cooling system and fans to be used for the semiconductors in the application.

To calculate the efficiency of the 3L T-Type based 3-phase PMSM driver, it is necessary to calculate the inverter power losses. Nevertheless, gate drive losses are very small and often neglected. To calculate power losses, the current and voltage waveforms of all semiconductors and datasheet parameters such as turn-on resistance, turn-on and turn-off times, etc. are needed. In this study, the MSCSM120HRM163AG-SiC-MOSFET power module manufactured by Microchip for 3L T-Type applications is chosen for one phase leg of the inverter. The PCB layout's stray inductances will surely be reduced by using IPM modules rather than discrete MOSFETs. This module has four SiC MOSFETs, two (S_1-S_4) rated at 1200V, 138A (80°C), and two (S_2-S_3) rated at 700V, 98A (80°C). Body diodes (S_1, S_2, S_3, S_4) have similar values. While making this selection, the current value that the PMSM can draw in transient situations such as speed changes under full load is also taken into account.

In the simulation, the input power of the 3L T-Type inverter for 400V dc voltage is measured as P_{inp}=13.04kW, and the 3-phase ac output power of the inverter is measured as P_{out}=12.91kW as shown in Figure 15. In this case, the 3-phase T-type inverter has a power loss of about 1% (130W). However, to verify this power measurement, the power losses in the inverter circuit must also be calculated.

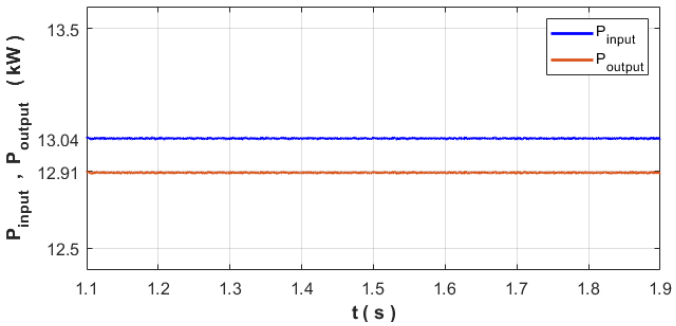


Fig.15. Input dc power and output 3-phase ac power of 3L T-type inverter

The IPM's datasheet parameters are used for the turn-on-resistance and voltage drop values of the MOSFETs and body diodes in the simulation. When the PMSM is operating at a 40Nm load and +3000rmp speed, the PWM signal, current and voltage waveforms for all semiconductors are measured for one period of voltage. As an example, the waveforms for S_{a1} and D_{a1} are given in Figure 16 (a) and for S_{a2} and D_{a2} in Figure 16 (b). To calculate the conduction and switching losses of each SiC MOSFET and conduction losses of body diodes, the average and rms values of MOSFET and body diodes current for both halves of the switching period are calculated separately

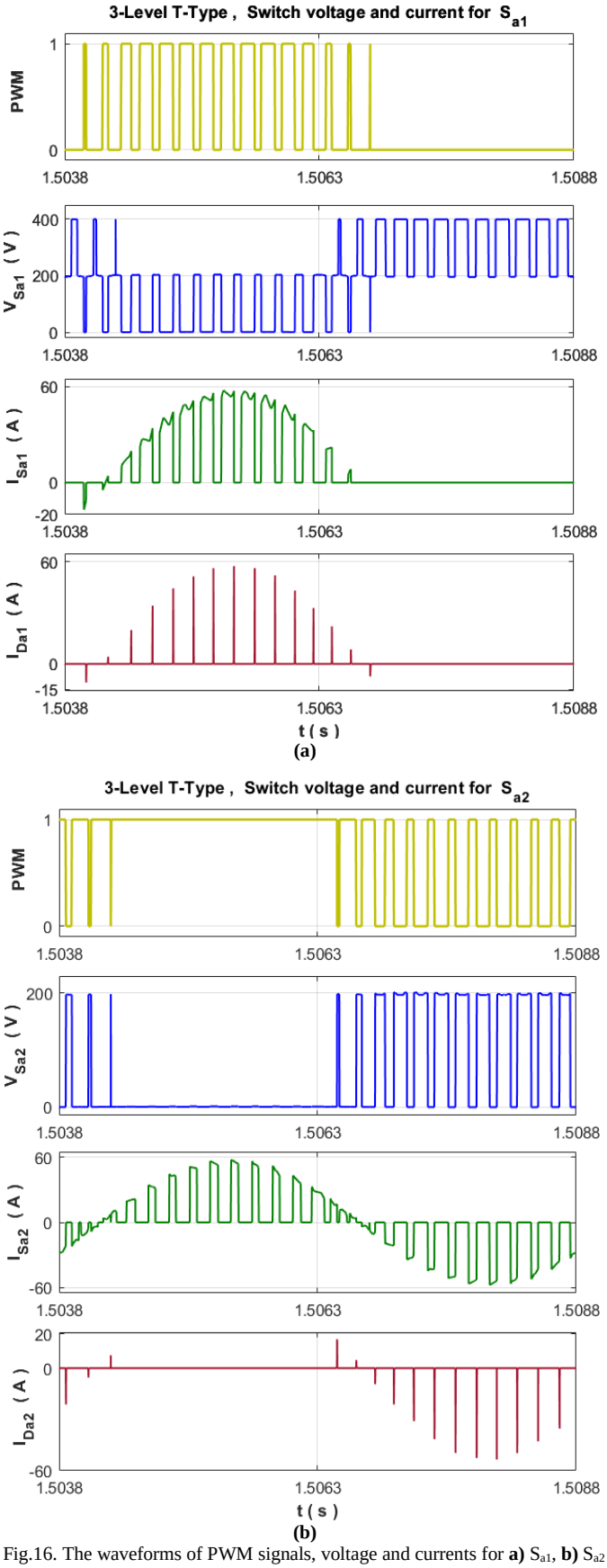


Fig.16. The waveforms of PWM signals, voltage and currents for **a)** S_{a1}, **b)** S_{a2}

TABLE III
I_{AVG} AND I_{RMS} VALUES FOR SiC MOSFET AND BODY DİODES

t	1.5038s – 1.5063s		1.5063s – 1.5088s	
	avg	rms	avg	rms
I_{Sa1}	21.62	31.50	0.027	1.41
I_{Sa2}	12.74	23.81	12.76	23.81
I_{Sa3}	12.72	23.76	12.76	23.81
I_{Sa4}	0.052	1.5	21.50	31.47
I_{Da1}	0.09	2.1	0.00023	0.16
I_{Da2}	0.0037	0.27	0.08	1.86
I_{Da3}	0.08	1.86	0.0037	0.27
I_{Da4}	0.002	0.22	0.092	2.10

in simulation and given in Table 3. These current values and datasheet parameters such as the MOSFETs' drain-source on resistance, turn-on delay time, rise time, turn-off delay time, and fall time values for 150°C junction temperature are substituted into Eq. (8)-(10) and the conduction and switching losses are calculated. This is because these values for SiC MOSFETs are the largest for a junction temperature of 150°C. This is also true for body diodes. Thus, for the most difficult operating conditions, such as large current and maximum junction temperature allowed in the datasheet, almost the largest possible power losses in the power circuit of the converter are calculated. In order to reduce the size of the passive components in the power circuit of the 3L T-Type inverter while not causing excessive EMI noise, the switching frequency is selected as 50kHz. Because overwhelming EMI noise makes the inverter hard to control and necessitates additional filters and an expensive PCB design. Conduction losses for the body diodes are also computed by substituting into Eq. (11) with the average current, rms current, voltage drop, and turn-on resistance values for a junction temperature of 150°C. The switching losses of the body diodes are neglected as they are very small. The total power loss in one period for 1-phase and 3-phase legs of T-Type is summarized in Table 4. The table shows that total losses calculated for 3-phase T-Type are about 1.0% for 50kHz. This result is compatible with the above measurement result. The bar representation of power losses in semiconductors is given in Figure 17. It is clearly seen that conduction losses are larger than switching losses in SiC MOSFETs. The body diode losses are quite small.

$$P_{Conduction\_S} = I_{rms}^2 \times R_{DS\,on}\,(T_j = 150\,^{\circ}C) \tag{8}$$

$$P_{on\_S} = \frac{1}{2} \times I_{DS} \times V_{DS} \times \frac{I_{DS}}{I_{DS\,Test}} \times \frac{V_{DS}}{V_{DS\,Test}} \times t_{on} \times f_s \tag{9}$$

$$P_{off\_S} = \frac{1}{2} \times I_{DS} \times V_{DS} \times \frac{I_{DS}}{I_{DS\,Test}} \times \frac{V_{DS}}{V_{DS\,Test}} \times t_{off} \times f_s \tag{10}$$

$$P_{Conduction\_Diode} = (I_{rms}^2 \times r_D) + (I_{avg} \times V_D) \tag{11}$$

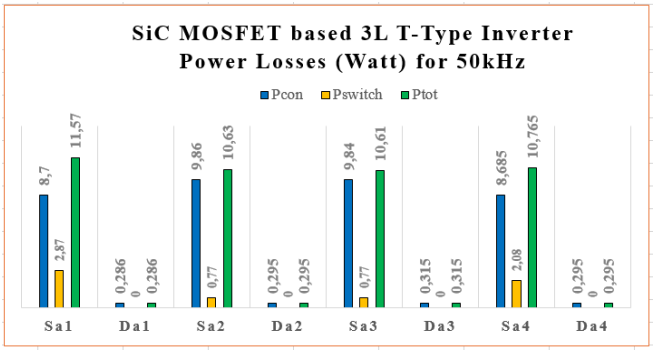


Fig.17. SiC MOSFET based 3L T-Type inverter power losses for 50kHz

TABLE IV
THE CONDUCTION AND SWITCHING LOSSES FOR 3L T-TYPE INVERTER WITH 50kHz

	1st Half Period			2nd Half Period			f_s=50kHz
	P_{Con}	P_{on}	P_{off}	P_{Con}	P_{on}	P_{off}	$P_{average}$
S_{a1}	17,36	1,84	2,3	0,034	0	0	21,53/2=10,767
D_{a1}	0,57	0	0	0,002	0	0	0,572/2=0,286
S_{a2}	9,86	0,41	0,36	9,86	0,41	0,36	21,26/2=10,63
D_{a2}	0,02	0	0	0,57	0	0	0,59/2=0,295
S_{a3}	9,82	0,41	0,36	9,86	0,41	0,36	21,22/2=10,61
D_{a3}	0,57	0	0	0,06	0	0	0,63/2=0,315
S_{a4}	0,04	0	0	17,33	1,85	2,31	21,53/2=10,765
D_{a4}	0,01	0	0	0,58	0	0	0,59/2=0,295
						Total losses for 1-leg of 3L T-Type	**42,772W**
						Total losses for 3-leg of 3L T-type	**128,32W**

$P_{out} = 13.040W - 128{,}32W = 12.911{,}68\,W$

$\eta\% = (12.911{,}68\,/13.040) \times 100 = 99\%$

VII. CONCLUSION

In this study, a compact and high-efficiency SiC MOSFET-based 3-level T-Type driver is proposed for PMSMs, which are widely used in many applications today. In this work, a SiC MOSFET based 3-level T-Type topology is proposed to build a compact and highly efficient dc/ac inverter that can be used in many low voltage (<1kV) applications. 3L T-Type already has a compact structure compared to other 3-level topologies. Also, due to a higher switching frequency operation, SiC-based T-Type inverter can become significantly more compact as they require much smaller passive elements and cooling arrangements. Since the Drain-Source on resistance of the preferred SiC MOSFETs is very small (17.5mΩ, 150C), the conduction losses of the 3L T-Type inverter are also small. The total power dissipation of the 3-phase 3L T-type inverter is calculated in detail for one period of the voltage for the worst case operation. The mathematically calculated power losses are approximately 1% and are in agreement with the measured power losses in the simulation. However, the power losses of all SiC MOSFETs are approximately equal, indicating that the power loss distribution is balanced.

APPENDIX A

Motor parameter values

Stator resistance R_a=0.0485Ω, d and q-axis inductance L_d= L_q=0.395mH, Moment of inertia J=0.0027$kg.m^2$, Friction coefficient B=0.0004924 $N.m.s$, Moment coefficient k_t=0.7164 N.m/A, Magnetic flux ϕ=0.1194$weber$, Electromagnetic moment T_e= 42.09Nm, Rated speed n_r=3000rpm, Number of pole pairs P=4, dc-link voltage V_{dc}=560V, Rotor Type = Round, Back emf= sinusoidal

ACKNOWLEDGMENT

The studies presented in this article were carried out within the scope of Berkan Turan's master thesis titled "Thesis Number:683667, Vector control of PMSM fed by three-phase three-level T-Type inverter, 2021". The advisor of the master's thesis is Assoc. Prof. Dr. Erkan Deniz

REFERENCES

[1] W. da S. Lima et al. "A bidirectional isolated integrated ac–dc converter based on an interleaved 3-level T-type power converters." IEEE Access, vol. 9, 2021, pp 142754-142767.

[2] R. Phukan et al. "Characterization and mitigation of conducted emissions in a SiC based three-level T-Type motor drive for aircraft propulsion." IEEE Transactions on Industry Applications, vol. 59.3, 2023, pp 3400-3412.

[3] E. Deniz, H. Altun. "Bes Seviyeli Izole DA Kaynakli Kaskat Inverterin SDGM Teknigi Ile Kontrolu." SAU Journal of Science, vol. 11. 1, 2007, pp 1–9.

[4] A. Nabae, I. Takahashi, H. Akagi. "A new neutral-point-clamped PWM inverter." IEEE Transaction Industrial Application, vol. IA. 17(5), 1981, pp 518–523.

[5] T. Bruckner, S. Bernet, H. Guldner. "The active NPC converter and its loss-balancing control." IEEE Transaction Ind. Electronics, vol. 52. 3, 2005, pp 855–868.

[6] M. Schweizer, I. Lizama, T. Friedli and J. W. Kolar, "Comparison of the chip area usage of 2-level and 3-level voltage source converter topologies." IECON 2010 - 36th Annual conference on IEEE Industrial Electronics Society. Glendale, AZ, 2010.

[7] Z. Wang, "Design and validation of a high-power, high density all silicon carbide three-level inverter." Ph.D Thesis, Electrical Engineering, University of Arkansas, 2021.

[8] Z. Wang, Y Wu, et al. "Busbar design and optimization for voltage overshoot mitigation of a silicon carbide high-power three-phase TNPC inverter." IEEE Transactions on Power Electronics, vol. 36.1, 2021, pp 204-214.

[9] A. I. Emon, Z. Yuan, A. B. Mirza, A. Deshpande, M. U. Hassan and F. Luo, "1200 V/650 V/160 A SiC+Si IGBT 3L Hybrid T-Type NPC Power Module With Enhanced EMI Shielding," in IEEE Transactions on Power Electronics, vol. 36, no. 12, 2021, pp 13660-13673.

[10] H. Peng, et al. "Improved space vector modulation for neutral-point balancing control in hybrid-switch-based TNPC neutral point-clamped inverters with loss and common-mode voltage reduction." CPSS Transactions on Power Electronics and Applications, vol. 4.4, 2019, pp 328-338.

[11] P. Azer, J. Bauman, "An asymmetric three-level T-type converter for switched reluctance motor drives in hybrid electric vehicles." 2019 IEEE Transportation Electrification Conference and Expo (ITEC), Detroit, MI, 2019.

[12] D. Xiao, J. Bauman, "Power loss comparison between three-level T-type and NPC converters with SVPWM and MPCC modulation schemes in electric vehicles." 2020 IEEE Transportation Electrification Conference and Expo (ITEC), Chicago, IL, 2020.

[13] R. A. Ghaderloo, C. Singhabahu, R. Resalayyan and A. Khaligh, "Selection of three-phase inverter topology and optimization of the pcb layout of a T-type bridge-leg for power dense motor drive application." 2024 4th International Conference on Smart Grid and Renewable Energy (SGRE), Doha, Qatar, 2024.

[14] Z. Yuan et al., "A 1.2 kV 400A SiC-Mosfet based 3L-TNPC power module with improved hybrid packaging method for high-density applications," 2021 IEEE Applied Power Electronics Conference and Exposition (APEC), Phoenix, AZ, USA, 2021.

[15] M. H. Adeli, E. Deniz, N. Altin, S. Ozdemir and A. Nasiri, "Compact GaN-Based 25kW, 480V Three-Level Active Front End Rectifier," 2024 IEEE Sixth International Conference on DC Microgrids (ICDCM), Columbia, SC, USA, 2024. doi: 10.1109/ICDCM60322.2024.10665131.

[16] A. Nasiri, M. H. Adeli, E. Deniz. "Active Neutral Point Clamped Electric Drive Converters Using WBG Devices." *Wide Bandgap Power Electronics, Emerging Converter Technologies and Applications*, edited by Isik C. Kizilyalli, Z. John Shen, Thomas M. Jahns, and Daniel W. Cunningham, Springer Cham, pp. 417-451, August, 2025. https://doi.org/10.1007/978-3-031-78631-0_16

[17] Z. Yuan et al., "A 1.2 kV 400A SiC-MOSFET Based 3L-TNPC Power Module with Improved Hybrid Packaging Method for High-Density Applications," 2021 IEEE Applied Power Electronics Conference and Exposition (APEC), Phoenix, AZ, USA, 2021.

[18] A. I. Emon et al, "1200 V/650 V/160 A SiC+Si IGBT 3L Hybrid T-Type NPC Power Module With Enhanced EMI Shielding," IEEE Transactions on Power Electronics, vol. 36.12, 2021, pp. 13660-13673.

[19] M. H. Adeli, N. Altin, E. Deniz, and A. Nasiri, "EMI Modeling of PCB-Based Three-Level Active Neutral-Point-Clamped GaN Converter." IEEE Applied Power Electronics Conference and Exposition (APEC2025), Atlanta, GA, USA, March 16-20, 2025. (Accepted).

[20] R. A. Ghaderloo, C. Singhabahu, R. Resalayyan and A. Khaligh, "Selection of Three-phase Inverter Topology and Optimization of the PCB Layout of a T-type Bridge-Leg for Power Dense Motor Drive Application," 2024 4th International Conference on Smart Grid and Renewable Energy (SGRE), Doha, Qatar, 2024.

[21] "PCB design techniques to reduces EMI", www.altium.com, 2024.

[22] "Conducted and Radiated Emissions Testing", Tektronix Application Note, pp 1-34

[23] R. Oberhuber. "Low-EMI designs for isolated ADC signal-chain solutions", Analog Design Journal, Texas Instrument, 2024.

[24] B. Turan, "Üç fazlı üç seviyeli T-Tipi inverterden beslenen sürekli mıknatıslı senkron motorun vektör kontrolü." Yüksek Lisan Tezi, Fırat Üniversitesi, Fen bilimleri Enstitüsü, Elektrik Elektronik Mühendisliği Teknolojileri Programı, 2021.

[25] C. Mohan, "A generalized predictive controlled TNPC power inverter with a deterministic dc-link capacitor voltage balancing approach." Master Thesis, Universidad De Sevilla Escuela Tecnica Superior De Ingenieria, Sevilla, 2015.

[26] E. Deniz, "ANN-based MPPT algorithm for standalone photovoltaic PMSM drive system without dc-dc boost converter." IENSC2018

International Engineering and Natural Sciences Conference, Diyarbakir, Turkiye, 14-17 November 2018.
[27] E. Deniz, "Medium and large vector-based SVPWM technique for five-phase two-level inverter." European Journal of Technique (EJT), vol.11.2, pp 209-216, 2021. https://doi.org/10.36222/ejt.1029249
[28] E. Deniz, O Aydogmus, "Comparison of asynchronous motor drives fed by three-level H-bridge inverter using SPWM and SVPWM. 6th International Advanced Technologies Symposium, Elazig/Turkey, pp. 448–454, May 16-18, 2011.
[29] B. Wu, *High-Power Converters and AC Drives*, IEEE Press, Hoboken, New Jersey, USA, 2006.
[30] A. Gündoğdu, "Matris konvertörden beslenen sabit mıknatıslı senkron motorun vektör kontrolünün sayısal benzetimi." Yüksek Lisans Tezi, Fırat Üniversitesi Fen Bilimleri Enstitüsü, 2005.
[31] Ö. Aydoğmuş, "Matris Çevirici ile beslenen sürekli mıknatıslı senkron motor sürücü tasarımı ve algılayıcısız hız denetimi." Doktora Tezi, Fırat Üniversitesi, Fen bilimleri Enstitüsü, 2011.
[32] A. Gundogdu, H. Altun, "Simulation of a RL load fed by a matrix converter with Matlab/Simulink." Journal of The Faculty of Engineering and Architecture of Gazi University, vol. 24.2), 2009, pp 199-207.
[33] A. Gundogdu, H. Kizmaz, R. Celikel, M. Yilmaz, "Speed sensorless adaptive power control for photovoltaic-fed water pump using extended Kalman–Bucy filter." Energy Reports, 10, 2023, pp. 1785–1795.

BIOGRAPHIES

Erkan Deniz was born in 1977. He received the B.S., M.S., and Ph.D. degrees from Firat University, Elazig, respectively in 2001, 2005, and 2011. From 2001 to 2012, he was a Research Assistant with the electrical teaching department. Since 2018, he has been an Associate Professor with the Electrical Electronics Engineering Department, Firat University. His research interests include power electronics, SiC MOSFETs, GaN FETs, pulse width modulation techniques, power quality, D-STATCOM, grid connected converters, energy storage systems, motor control, machine learning, CNN.

Berkan Turan was born in 1992. He received the B.S. from Gazi University, Ankara in 2014 and the M.S. from Fırat University, Elazığin 2021. He has been working as an electrical engineer at the Turkish State Railways since 2018. His research interests include grid connected converters, motor drivers.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 3, September 2025

Research Article

Detection of Epileptic Seizures with Different Machine Learning Algorithms Using EEG Signals in Daily Life

Temel Sonmezocak and B. Koray Tuncalp

***Abstract*—Today, Electroencephalography (EEG) is commonly used as a diagnostic tool for epilepsy. In this study, an effective method for diagnosing epileptic seizures in non-clinical settings is proposed. To evaluate the performance of this method, EEG data from 7 pediatric patients at Boston Children's Hospital were analyzed using Decision Tree (DT), Linear Discriminant (LD), Naive Bayes (NB), Support Vector Machine (SVM), and K-Nearest Neighbor (KNN). The time and frequency characteristics of the EEG signals were compared. Experimental results show that epileptic seizures can be determined effectively with 100% accuracy by using only 3 channels (FP1-F7, FP2-F4 and T8-P8) with mean amplitude, mean frequency, median frequency and variance features with SVM, KNN or DT.**

Index Terms—**Electroencephalography, epileptic seizure, machine learning, median frequency, time domain analysis.**

I. Introduction

EPILEPTIC SEIZURE is known as the unusual electrical activities of the brain that occur in the central nervous system. During these activities, patients may experience loss of consciousness, confusion, tongue biting, falling down and related injuries [1]. It is not easy to detect when seizures will start. Seizures can be occurred unexpectedly in every moment of daily life, so this situation is pretty troubling especially for children and people with mental disabilities. It can be classified using machine learning algorithms for electrical activities that lead to the emergence of symptoms that occur during seizures [1], [2] and [3]. However, it is still difficult to diagnose. In this context, researchers have worked on the determination of deterministic features with a series of non-linear time series analyzes based on EEG measurements of seizures [4], [5] and [6].

Temel SÖNMEZOCAK, is with Department of Electrical and Electronic Engineering, Faculty of Engineering, Yeni Yuzyil University, Istanbul, Turkey, (e-mail: temel.sonmezocak@yeniyuzyil.edu.tr).

https://orcid.org/ 0000-0003-4831-9005

B. Koray TUNCALP, is with Department of Electrical Electronics Engineering University of Halic, Istanbul, Turkey,
(e-mail: koraytuncalp@halic.edu.tr).

https://orcid.org/ 0000-0003-1632-8900

Manuscript received Nov 03, 2024; accepted Jan 09, 2025.
DOI: 10.17694/bajece.1577914

EEG signals are uninterrupted waveforms composed of fluctuations in potential differences across time. EEG plots depict a voltage signal that varies over time. Each data point represents the instantaneous voltage value at evenly spaced time intervals [2] and [7]. In this system, 21 electrodes are placed around the skull [8].

In the literature, there are investigations focused on identifying epileptic seizures through EEG measurements employing these electrode structures. Notably, various convolutional neural network (CNN) architectures rooted in deep learning have been utilized for this purpose [9]. From these CNN structures, 95% accuracy performances can be obtained with AlexNet and 94.17% accuracy performances with GoogleNet [2] and [10]. However, the seizure moment can be detected with 100% accuracy in EEG signals by obtaining many frequency and time characteristics from many different channels of EEG signals with different frequency regions. Especially K-nearest neighbour (KNN) and SVM algorithms are successful in this regard [11], [12] and [13]. One of the most important advantages of this study in comparison with the similar studies in the literature is the ability for determining the instantaneous seizures that patients may encounter in daily life with high accuracy by using the time and frequency characteristics obtained from the most effective frequency range of the EEG signals with the least effective number of channels and the most effective frequency range of these channels. Therefore, in cases of seizures that may occur in the daily life of the patients, a 1 second delay is also important in terms of performing the transactions quickly with even less storage space. In this way, using EEG data obtained by sampling rate of 256 Hz from 16 different channels of 7 pediatric epilepsy patients by Children's hospital of Boston-MIT, each signal was divided into one second segment. The time and frequency characteristics of the signals were obtained for each separated second segment. These were obtained as mean absolute value (MAV), mean frequency (MNF), median frequency (MDF) and variance (VAR) [11], respectively. Pearson correlation relationships were examined for each channel by applying comparatively statistical analysis of data. Thus, the most effective features for each channel were determined and their performances for determining seizure occurrence moments were comparatively determined by employing 5 different machine learning algorithms (DT, SVM, LD, NB and KNN) [14]. By using the achieved

statistics of most effective 3 channels FP1-F7, FP2-F4 and T8-P8, seizure occurrence moments can be determined with 100% efficiency with MNF, MDF, MAV and VAR features, SVM, KNN, DT and LD.

The structure of this article is outlined as follows: Section 2 introduces the materials and methods employed in identifying epileptic seizure states. Experimental results, depicted through graphs and tables, are presented in Section 3, while Section 4 delves into the implications of these findings. Section 5 serves as the conclusion of the article.

II. Materials and Methods

A. EEG Dataset and Preprocessing

In this study, the EEG data of 7 patients in Boston-MIT Children's hospital with ages between 1.5 and 22 were investigated. These records consist of EEG data from 16 different channels with the sample rate of 256 Hz. The delta, theta, alpha and beta components of an example signal taken from the data set are given in Figure 1.

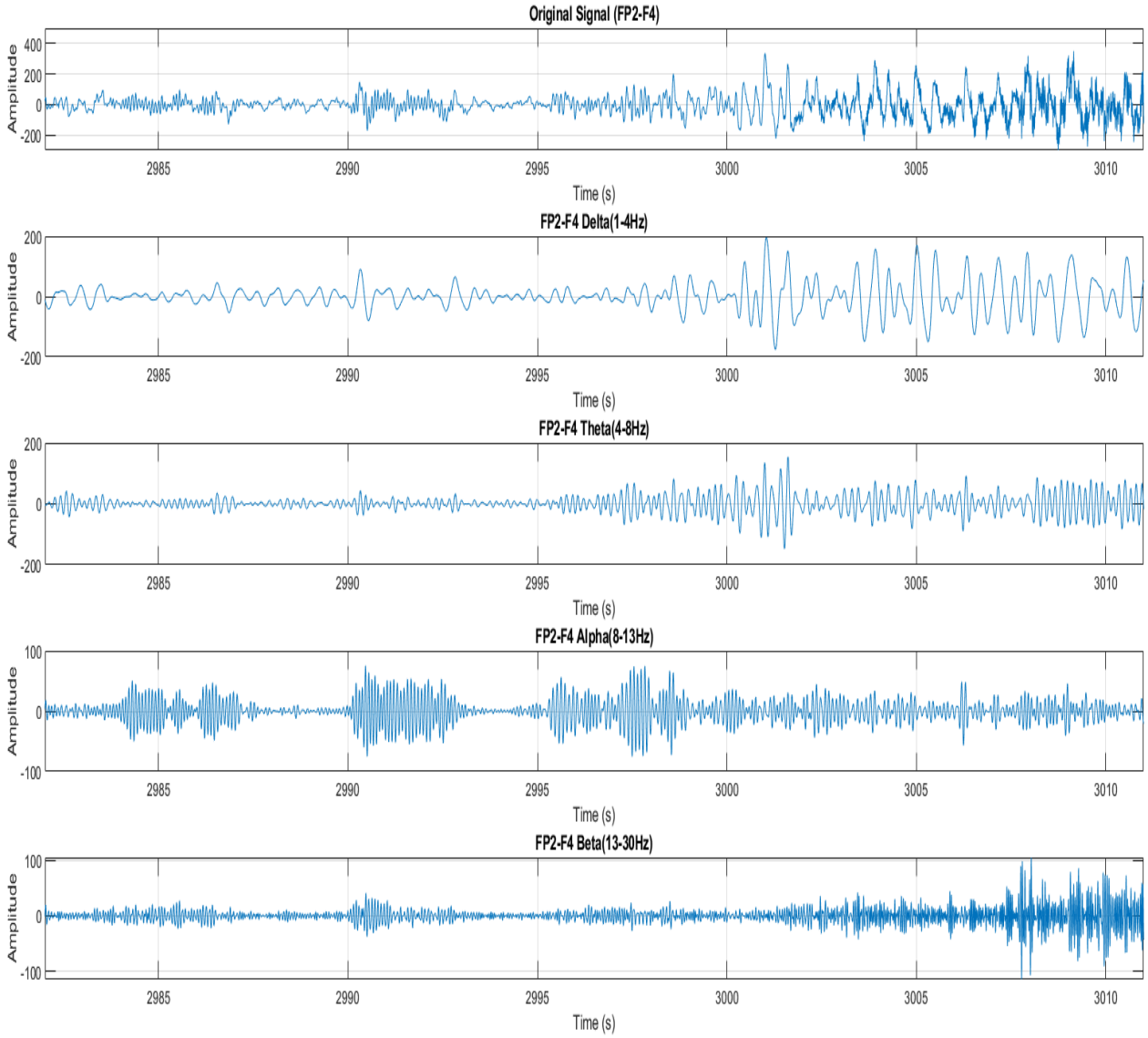


Fig.1. Time-frequency plots illustrating the EEG signal obtained from the FP2-FP4 channel during an epileptic seizure, depicting different frequency regions [2]

Beta waves (12-30Hz) typically manifest during periods of wakefulness and intense cognitive activity, characterized by their high frequency. Alpha waves (8-12Hz) are commonly observed when the brain is in a state of relaxation, displaying a lower frequency [2] and [15]. Theta waves (4-8Hz) are generally present during light sleep, while Delta waves (0.1-4Hz) exhibit the lowest frequency and are typically associated with deep sleep. The human EEG predominantly encompasses signals within the 1-30Hz frequency range. Although there is some indication that higher frequencies may convey significant neurophysiological information, the majority of EEG studies focus on signals within the 1-30Hz frequency range [2] and [15].

The alpha, theta, delta, and beta frequency bands were analyzed for signals with a sample rate of 256 Hz for each patient. To remove high-frequency components from the EEG data in these signals, a 2nd-order low-pass Butterworth digital

filter was applied. However, for EEG recordings, the frequency regions of occurring seizures were determined using Fast Fourier Transform (FFT) and significant frequency regions were determined using MATLAB R2021b software. The block diagram of the proposed algorithm is given in Figure 2.

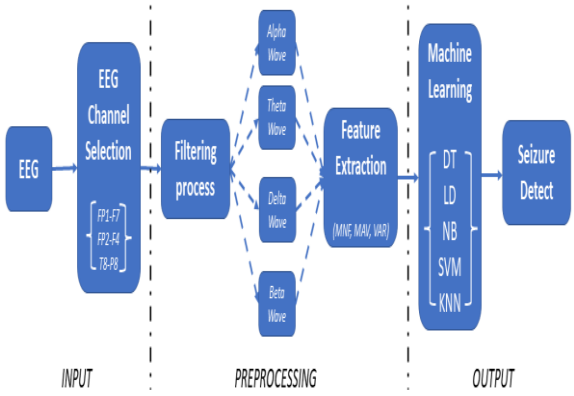


Fig.2. Block diagram of the system

B. E Feature Extraction of EEG Signals

Four distinct features were derived from the filtered EEG signals. MNF and MDF primarily encompass information associated with the amplitude and frequency of a signal [14], [16], [17] and [18].

$$MNF = \sum_{j=1}^{M} f_j P_j \Big/ \sum_{j=1}^{M} P_j \,, \tag{1}$$

In Eq. (1), MNF is an average frequency value calculated by dividing the sum of the multiplications of the power spectrum intensity values of the signal up to J=1, 2,…,M' and the values of each instantaneous frequency by the intensity values of the same spectrum. fj represents its frequency in the jth frequency band, Pj is the power spectrum of the signal in the jth frequency band, and M is the length of the frequency band. MDF is a frequency where the spectrum is divided into two regions with equal amplitudes [2] and [19].

$$\sum_{j=1}^{MDF} P_j = \sum_{j=MDF}^{M} P_j = \frac{1}{2}\sum_{j=1}^{M} P_j \,, \tag{2}$$

$$MAV = \frac{1}{N}\sum_{i=1}^{N} \left| X_i \right| , \tag{3}$$

Another widely employed parameter is MAV, which encapsulates crucial information regarding the average amplitude of the signal [14]. The MAV features of EEG signals are computed using Eq. (3). VAR is another significant feature used to assess the power of signals, defined as the mean square of the deviation values of signals [11], [14], [19] and [20].

$$VAR = \frac{1}{N-1}\sum_{i=1}^{N} X_i^{\,2} \,, \tag{4}$$

In Eq. (3) and (4), X_i represents the signal in segment *i*, *N* denotes the length of the EEG signal [14]. These obtained features MNF, MDF, VAR and MAV are classified in two classes as non-seizure (0) and seizure (1), in terms of entry features.

C. Classification of EEG Signals

In this study, five different machine learning algorithms (KNN, DT, NB, SVM, and LD) [14] are employed for the classification of EEG signals [21] and [22]. MNF, MDF, VAR, and MAV inputs are classified into two classes: non-seizure (0) and seizure (1). Each classifier undergoes training through cross-validation with a k-fold value of 5. To assess the performance of the introduced method, classification accuracy was calculated using Eq. (5) [14] and [23].

$$Accuracy(\%) = \frac{1}{N}\sum_{k=1}^{N}\left[\frac{TP+TN}{TP+TN+FP+FN}\right]\times 100 \,, \tag{5}$$

In this context, N represents the number of classes. True positive (TP) and true negative (TN) denote the number of correctly classified samples, while false positive (FP) and false negative (FN) represent the number of misclassified samples [11] and [23].

III. Results and Discussing

The received signals were plotted for 15 seconds before the seizure and 15 seconds at the time of the seizure. The EEG signals of all channels for the sample patient are given in Figure 3. The obtained results show that the amplitude of the EEG signals increases at the time of the seizure. Recently studies show that the moment of seizure can be detected with 100% accuracy by employing EEG signals, specially by using the K-nearest neighbor (KNN) and SVM algorithms [10] and [11]. Also, there are studies based on deep learning architecture (CNN) [9] and [10].

However, all these studies are clinical studies for the use of channels collectively. In this study, all channels were already examined separately and together, and there were significant increases in the amplitude values of all channels at the time of seizure. However, the primary objective of this study is to determine the most effective channels (FP1-F7, FP2-F4 and T8-P8) taken from the frontal and temporal regions of the human brain, which will not disturb the person in daily life, to determine the seizure quickly and without requiring storage space. The frequency analysis of the original signal of the sample patient taken from the FP2-F4 channel and the frequency analysis of the delta, theta, alpha and beta components of this signal are given in Figure 4.

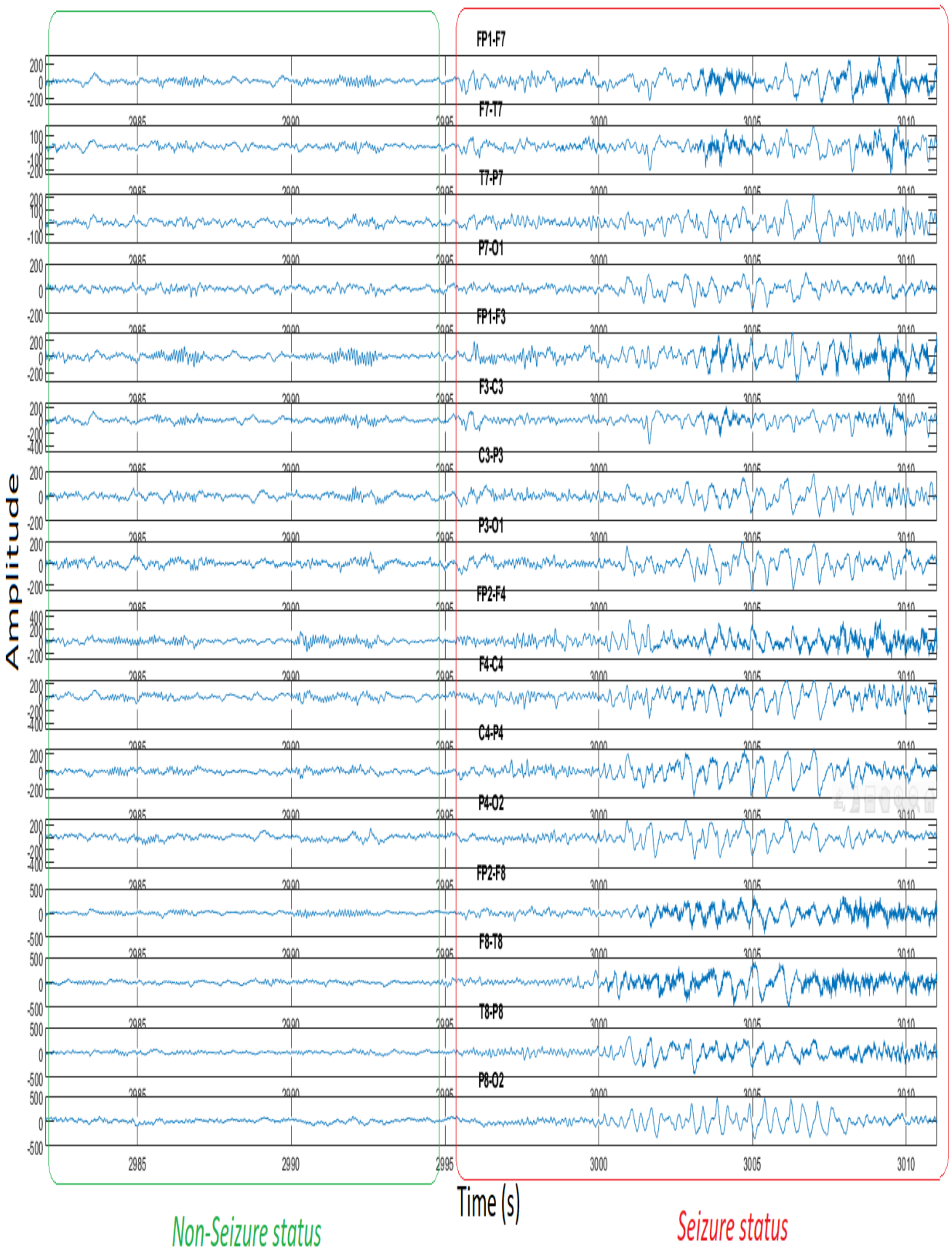


Fig.3. Amplitude-time graphs of EEG signals of all 16 channels of an 11-year-old girl with epilepsy

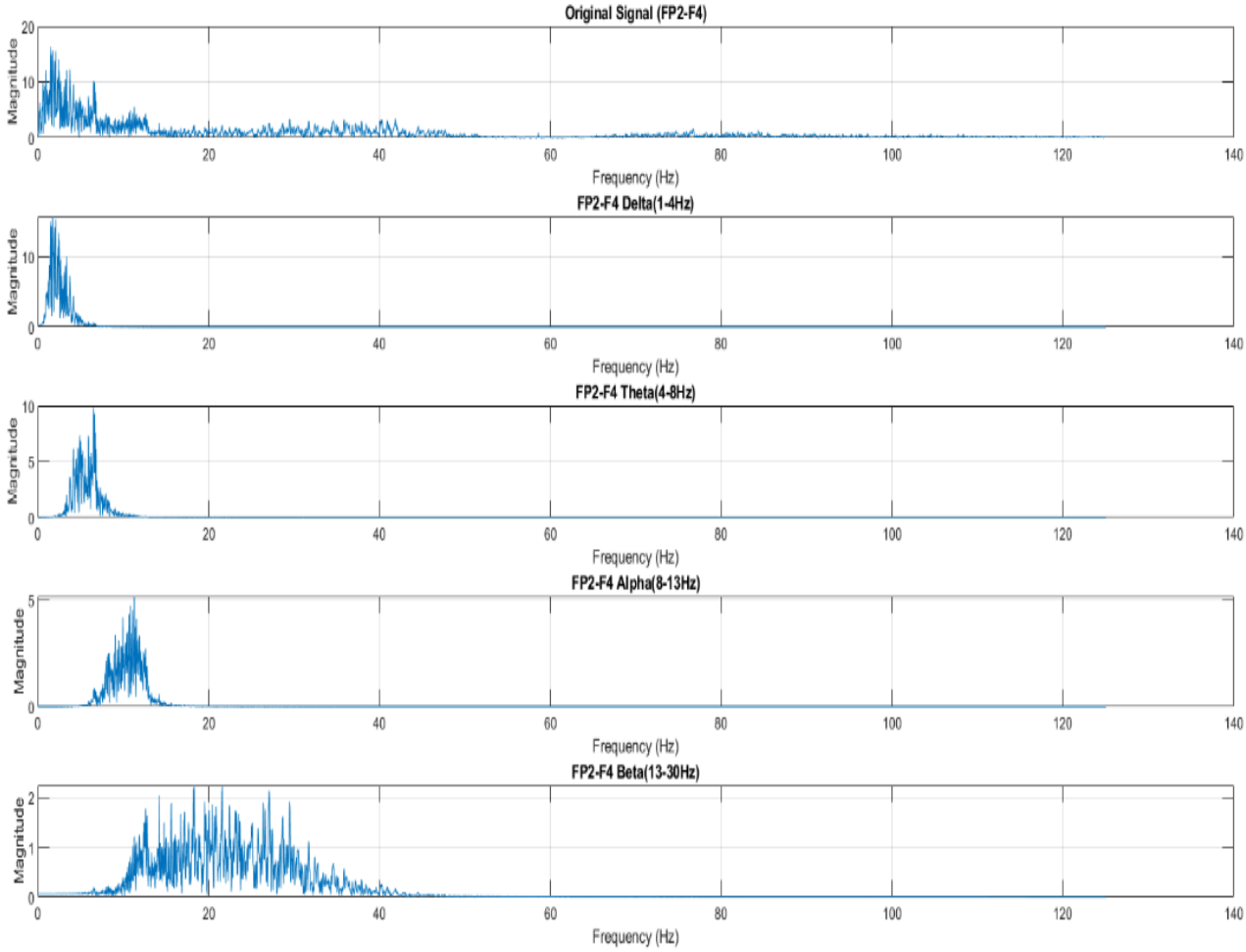


Fig.4. EEG frequency spectrum depicting the different frequency regions of the FP2-F4 channel during an epileptic seizure in a patient

As seen in Figure 4, this study investigates distinctive features of frequency spectra to achieve effective classification by analyzing all channels for each patient across various frequency ranges. Upon scrutiny of this figure, it becomes apparent that the notable frequency range extends up to the corner frequency (Fc) of 30 Hz.

The results of Pearson correlation for the selection of the most effective channels in terms of the impact of input features on the output are presented in Table I.

The obtained results indicate that the FP2-F4 channel and the T8-P8 channel are the most prominent channels showing rhythmic activations, while the FP1-F7 channel is sensitive to high-amplitude deviations [26]. When examining the frequency spectra in this study, these three channels are prominently observed. Statistically, in Pearson correlation analysis, it is observed that the features MAV, MNF, MDF, and VAR, which are solely utilized with these three channels, exhibit high correlation in determining seizure and non-seizure states.

Frequency spectra of EEG recordings containing all channels for each patient have been obtained. Figure 5 displays the frequency spectra encompassing a total of 30 seconds, including the first 15 seconds before the seizure, the seizure moment, and the last 15 seconds after the seizure, from EEG recordings of a sample patient comprising FP1-F7, FP2-F4, and T8-P8 channels.

TABLE I
THE PEARSON CORRELATION RELATIONSHIPS OF EEG CHANNELS

EEG Channel Selection	FP1-F7 F7-T7 T7-P7 P7-O1 FP1-F3 F3-C3 C3-P3 P3-O1 FP2-F4 F4-C4 C4-P4 P4-O2 FP2-F8 F8-T8 T8-P8 P8-O2	FP1-F7 T7-P7 FP1-F3 FP2-F4 C4-P4 FP2-F8 F8-T8 T8-P8	FP1-F7 FP1-F3 FP2-F4 T8-P8	FP1-F7 FP2-F4 T8-P8
Signal Features	***Pearson Coef.***	***Pearson Coef.***	***Pearson Coef.***	***Pearson Coef.***
MAV	0,789	0,881	0,964	0,999
MNF	0,147	0,394	0,599	0,949
MDF	0,394	0,708	0,977	0,999
VAR	0,494	0,704	0,676	0,854

The obtained results indicate that the FP2-F4 channel and the T8-P8 channel are the most prominent channels showing rhythmic activations, while the FP1-F7 channel is sensitive to high-amplitude deviations [26]. When examining the frequency spectra in this study, these three channels are prominently observed. Statistically, in Pearson correlation analysis, it is observed that the features MAV, MNF, MDF, and VAR, which are solely utilized with these three channels, exhibit high correlation in determining seizure and non-seizure states. Frequency spectra of EEG recordings containing all channels for each patient have been obtained. Figure 5 displays the frequency spectra encompassing a total of 30 seconds, including the first 15 seconds before the seizure, the seizure moment, and the last 15 seconds after the seizure, from EEG recordings of a sample patient comprising FP1-F7, FP2-F4, and T8-P8 channels. While the seizure states are pronounced in these channels, the frequency response results, particularly in the low-frequency regions (up to 10 Hz), are provided in Figure 6.

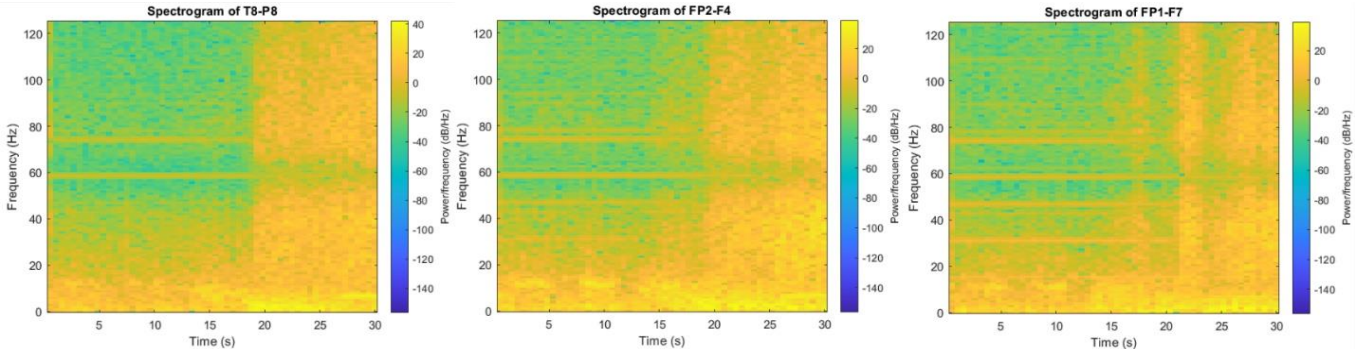


Fig.5. Frequency spectrums of 3 channels showing the seizure moment and pre-seizure situation

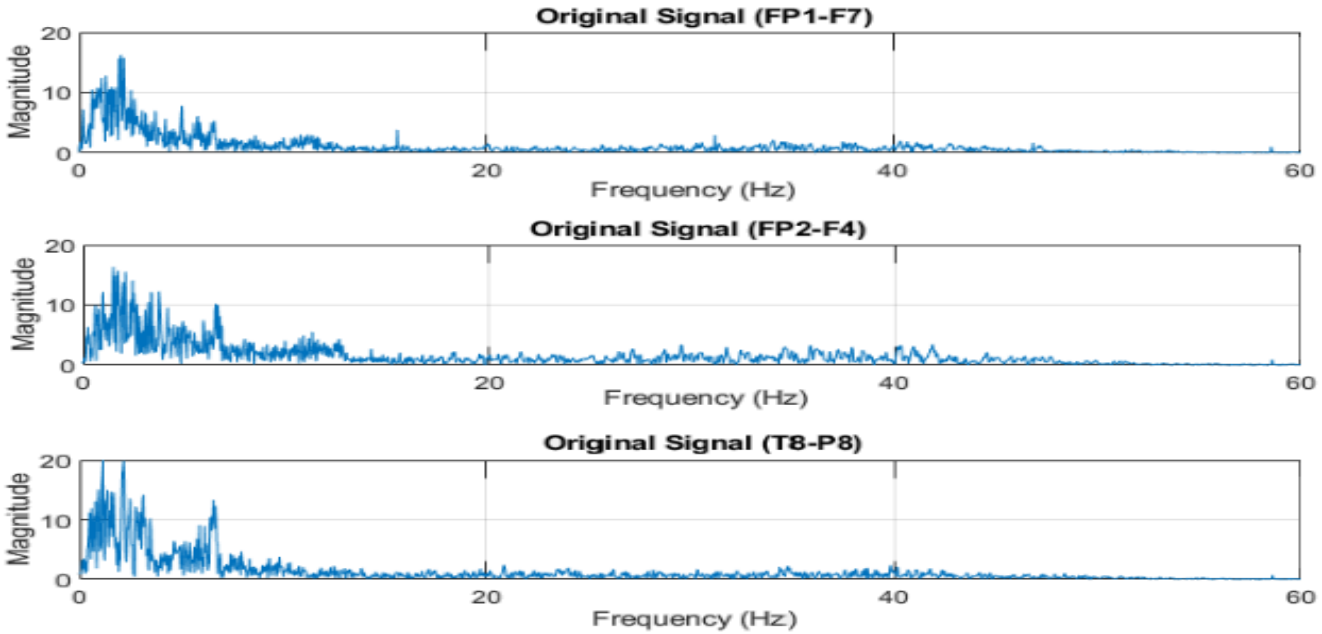


Fig.6. Frequency domain FFT graphs of 3 channels showing the seizure moment and pre-seizure situation

As observed in Figure 6, High amplitudes are present in low frequencies during seizure moments, accompanied by a significant variation in amplitude values. Therefore, MAV, MNF, MDF, and VAR features have been selected. Thus, with minimal features and considering the minimum number of channels, effective seizure detection has been achieved. The recordings obtained for pre-seizure and seizure states have been classified using Machine Learning algorithms in 1-second segments. This is because the primary aim of this study is to detect seizure states in real-time data with 1-second intervals, considering daily life scenarios. For this purpose, different classifiers have been employed and compared using the FP1-F7, FP2-F4, and T8-P8 channels. During the training of these classifiers, cross-validation (k-fold: 5) has been chosen. Subsequently, performance metrics have been determined using 1000 different data points representing seizure and non-seizure states that were reserved for testing. The test results are provided in Table 2.

The accuracy performance of classifications increases as the

number of features increases. By utilizing all four of these features together, KNN, SVM, and DT classifiers can achieve 100% performance. The performance results, including ROC curves and Confusion matrices obtained by using these four features together, are further depicted in Figure 7.

While the performance values in Figure 7(b) and (c) are lower when using LD and NB, as observed in Figure 7(a), when SVM, KNN, or DT are used, seizure states have been accurately detected with 100% accuracy using the MAV, MNF, MDF, and VAR features. Upon examining the ROC curves, it is evident that the area under the curve (AUC) values are 1.0, indicating the most effective classification performance.

This is particularly significant for wearable headgear systems in everyday life. Notably, even though this study utilizes a ready-made EEG dataset comprising 16 channels obtained from 7 pediatric patients by Children's Hospital Boston, and employs MATLAB R2021b, the data has been transformed into 1-second segments with time and frequency features (MAV, MNF, MDF, and VAR) to simulate real-life scenarios. The resulting model, trained with K-fold: 5 cross-validation, is well-suited for testing with structures that mimic real-world conditions.

TABLE II

ACCURACY PERFORMANCES USİNG DİFFERENT CLASSİFİERS WİTH FP1-F7, FP2-F4, T8-P8

Used Classifier	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)	Accuracy (%)
Fine Tree	*100%**	*90,5%***	*97,6%****	*99,8%*****	*100%******
Medium Tree	*100%**	*90,5%***	*97,6%****	*99,8%*****	*100%******
Coarse Tree	*100%**	*90,5%***	*97,6%****	*100%*****	*99,9%******
Linear Disc.	*71,4%*	*71,4%***	*95,2%****	*92,9%*****	*100%******
Gaussion NB	*85,7%**	*81%***	*100%****	*100%*****	*97,6%******
Kernel NB	*85,7%**	*78,6%***	*100%****	*97,6%*****	*90,5%******
SVM Linear	*100%**	*88,1%***	*97,6%****	*97,6%*****	*100%******
SVM Qaudratic	*100%**	*83,3%***	*100%****	*100%*****	*100%******
SVM Fine Gausion	*100%**	*90,5%***	*99,8%*****	*100%*****	*99,8%******
SVM Medium Gausion	*100%**	*88,1%***	*97,6%****	*97,6%*****	*100%******
SVM Coarce Gausion	*100%**	*69%***	*92,9%****	*88,.1%*****	*100%******
SVM Kernel	*100%**	*83,3%***	*97,6%****	*100%*****	*99,9%******
KNN fine	*100%**	*88,1%***	*97,6%****	*97,6%*****	*97,6%******
KNN medium	*100%**	*85,7%***	*97,6%****	*100%*****	*100%******
KNN weighted	*100%**	*90,5%***	*97,6%****	*100%*****	*99,8%******

**Used with MAV, MNF, MDF and VAR*
***Used only MDF*
**** Used only VAR*
***** Used with MNF, VAR*
****** Used with MAV, MNF*

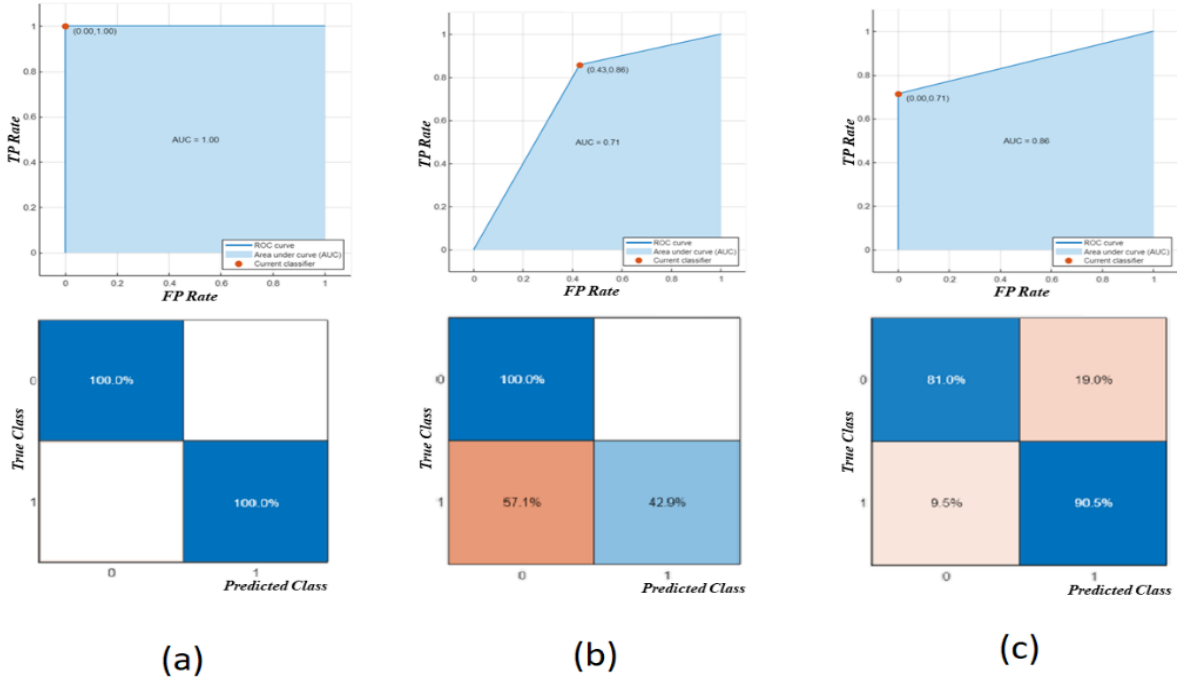


Fig.7. Accuracy performance results (with MAV, MNF, MDF and VAR features); (a) Results for SVM, KNN or DT, (b) Results for LD, (c) Results for NB

IV. CONCLUSION

In this study, EEG signals sampled at 256Hz were utilized from 7 child patients with ages ranging between 7 to 12 years, obtained by Children's Hospital Boston. Time and frequency features of these signals were comparatively examined. The experimental results demonstrate that epileptic seizures can be effectively identified with 100% accuracy using only 3 channels (FP1-F7, FP2-F4, and T8-P8) along with features like mean amplitude, mean frequency, median frequency, and variance, when employed with SVM, KNN, or DT. As a result, this allows for the diagnosis of daily occurring seizures without disturbing individuals, making it suitable for wearable systems geared towards instantaneous seizure detection with fewer channels, thereby requiring less storage space and enabling rapid real-time identification of seizures.

REFERENCES

[1] N. Shamriz, M. Yaganoglu. "Classification of Epileptic Seizure Dataset Using Different Machine Learning Algorithms and PCA Feature Reduction Technique", Journal of Investigations on Engineering & Technology, Vol. 4, iss. 2, pp. 47-60, 2021.

[2] T. Sonmezocak, G. Guler, M. Yildiz. "Classification of Resampled Pediatric Epilepsy EEG Data Using Artificial Neural Networks with Discrete Fourier Transforms", ELEKTRONIKA IR ELEKTROTECHNIKA, Vol. 29, pp. 19-25, 2023.

[3] K. Rasheed, A. Qayyum, J. Qadir, et al. "Machine Learning for Predicting Epileptic Seizures Using EEG Signals: A Review", IEEE Reviews in Biomed. Engineering, vol. 14, pp. 139-155, 2020.

[4] S. A. R. B. Rombouts, R. W. M. Keunen and C. J. Stam. "Investigation of nonlinear structure in multichannel EEG", Phys Lett A, vol. 202, pp. 352-358, 1995.

[5] F. H. Lopes Da Silva, J. P. Pijn, D. Velis, P. C. G. Nijssen. "Alpha rhythms: noise, dynamics and models", Int J Psychophysiology, vol. 26, pp. 237-249, 1997.

[6] W. S. Pritchard, D. W. Duke, K. K. Krieble. "Dimensional analysis of resting human EEG II: surrogate-Data testing indicates nonlinearity but not low-dimensional chaos", Int J Psychophysiology, 1995.

[7] S. Mahmud, M. S. Hossain, M. E. H. Chowdhury, M. B. I. Reaz. "MLMRS-Net: Electroencephalography (EEG) motion artifacts removal using a multi-layer multi-resolution spatially pooled 1D signal reconstruction network", Neural Comput. & Applications, pp. 8371–8388, 2023.

[8] S. A. Taywade, R. D. Raut. "A Review: EEG Signal Analysis with Different Methodologies", IJCA Proceedings on National Conference on Innovative Paradigms in Eng. and Tech., vol. 6, pp. 29 – 31, 2014.

[9] M. Eberlein, et al. "Convolutional Neural Networks for Epileptic Seizure Prediction", IEEE International Conference on Bioinformatics and Biomedicine. 2018, pp. 2577-2582, 2018.

[10] M. S. Munoz, C. E. S. Torres, R. Salazar-Carera, D. M. Lopez, R. Vargas-Canas. "Digital Transformation in Epilepsy Diagnosis Using Raw Images and Transfer Learning in Electroencephalograms", Sustainability, 14(18), 2022.

[11] T. Sonmezocak, S. Kurt. "Detection of lower jaw activities from micro vibration signals of masseter muscles using MEMS accelerometer", Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, Vol. 11, pp. 476-484, 2023.

[12] P. K. Sethy, M. Panigrahi, K. Vijayakumar, S. K. Behera. "Machine learning based classification of EEG signal for detection of child epileptic seizure without snipping", Int J Speech Technology, 2021.

[13] M. SAVADKOOHI, T. Oladunni, L. Thompson. "A machine learning approach to epileptic seizure prediction using Electroencephologram (EEG) Signal", Biocybernetics and Biomedical Engineering, pp. 1328-1341, 2020.

[14] T. Sonmezocak, S. Kurt. "Machine learning and regression analysis for diagnosis of bruxism by using EMG signals of jaw muscles", *Biomedical Signal Processing and Control*, Vol. 69, 2021.

[15] K. Zeng, J. Yan, Y. Wang, A. Sık, G. Ouyang and X. Li. "Automatic detection of absence seizures with compressive sensing EEG", Neurocomputing. 2016, vol. 171, pp. 497-502, 2016.

[16] S. Karlsson, B. Gerdlle. "Mean frequency and signal amplitude of the surface EMG of the quadriceps muscles increase with increasing torque-a study using the continuous wavelet transform", *J. Electromyogr. Kinesiology.* 2001, pp. 131–140, 2001.

[17] A. Georgakis, A., L. K. "Stergioulas and G. Giakas. Fatigue analysis of the surface EMG signal in isometric Constant force contractions using the average instantaneous frequency", IEEE Trans. *Biomed. Engineering.* 2003, pp. 262– 265.

[18] M. A. Oskoei, H. Hu. "Support vector machine-based classification scheme for myoelectric control applied to upper limb". *IEEE Transactions on Biomedical Engineering.* 2008, 1956-1965.

[19] A. Phinyomark, P. Phukpattaranont, C. Limsakul. "Feature reduction and selection for EMG signal classification", *Expert Systems with Applications*, vol. 39, pp. 7420–7431, 2012.

[20] S. P. Arjunan, D. K. Kumar. "Decoding subtle forearm flexions using fractal features of surface electromyogram from single and multiple sensors", *Journal of NeuroEngineering and Rehabilitation*. 2010, 7(1):53.

[21] A. N. Kamarudin, T. Cox and R. Kolamunnage-Dona. "Time-dependent ROC curve analysis in medical research: current methods and applications", *BMC Med. Res. Methodology*, 2017.

[22] M. Maeda, T. Yamaguchi, S. Mikami, W. Yachida, T. Saito, T. Sakuma, H. Nakamura et al., "Validity of single-channel masseteric electromyography by using an ultraminiature wearable electromyographic device for diagnosis of sleep bruxism", *Journal of Prosthodontic Research*. 2020, vol. 64, pp. 90–97.

[23] M. V. Artega, J. C. Castiblanco, I. F. Mondragon, J. D. Colorado, C. Alvarado-Rojas. "EMG-driven hand model based on the classification of individual finger movements", *Biomed. Signal Process. And Control*. 2020.

[24] R. Asif, A. Merceron and M. K. Pathan. "Predicting Student Academic Performance at Degree Level: A Case Study", *I. J. Intell. Syst. And App*., 2015, pp. 49–61.

[25] J. Too, A. R. Abdullah, N. M. Saad and W. Tee. "EMG feature Selection and classification using a pbest-guide binary particle swarm optimization", *Computation*, vol. 7, iss. 1, 2019.

[26] A. Shoeb, J. Guttag. "Application of Machine Learning To Epileptic Seizure Detection", *Appearing in Proceedings of the 27th International Conference on Machine Learning*, Haifa, Israel. 2010. pp. 975-982.

BIOGRAPHIES

Temel SONMEZOCAK was born in Edirne in 1978. He graduated his B.S. degree from Electrical Education Department of Technical Education Faculty of University of Marmara, Istanbul, in 2001. In 2015, he completed his M.Sc. degree at Istanbul Aydin University, Department of Electrical and Electronics Engineering. In 2021, He completed his PhD education in Yıldız Technical University Electronics Engineering program. He is currently working as Assistant Prof. Dr. in the Department of Electrical & Electronics Engineering at Doğuş University.

His research interests include circuits, systems, biomedical, sensor networks, lighting systems, mechatronics and instrumentation.

B. Koray TUNCALP was born in Istanbul in 1962. He received his B.S. degree from Electrical Education Department of Technical Education Faculty of University of Marmara, Istanbul, in 1984 and his M.Sc. and Ph.D. degree from Electrical Education Department of Institute of Pure and Applied Sciences, University of Marmara, Istanbul, in 1988 and 1999 respectively.

From 1984 to 1993, he worked as a research assistant, from 1993 to 1999 as a lecturer, from 1999 to 2001 as an Assistant Prof. Dr., from 2001 to 2003 as an Associate Prof. Dr., from 2003 to 2014 as an Prof. Dr. in two departments for University of Marmara. Since 2015 he has been working as an Prof. Dr. for Electrical and Electronics Engineering Department of School of Engineering, also is Director of Vocational School of Halic University.

His research interests include Electrical Energy Measurement, Mechatronics, Instrumentation, E-learning, Vocational and Technical Education.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 3, September 2025
Research Article

Performance Analysis of Fault-Tolerant Power Control Strategy for Bidirectional Power Transfer in DC Microgrid

Adnan Tan

***Abstract*—This study proposes a fault-tolerant converter control strategy for a DC microgrid designed to maintain continuous power supply to sensitive loads under converter failure conditions. The proposed control method is formed from constant current (CC) and constant voltage (CV) controllers with a fault detection method. In normal operating conditions, the proposed controller operates in CC control mode to control the power flow between the battery and the dc grid. The fault detection method observes grid voltage in real-time to detect grid voltage disturbance on the grid. If a voltage disturbance is detected, the fault detection method switches the controller to CV mode to maintain a stable bus voltage on the dc microgrid. Performance evaluations conducted via Matlab/Simulink demonstrate the effectiveness of the proposed method in stabilizing load voltage, managing battery charging/discharging efficiently, and enhancing system reliability without additional control complexities. This approach provides a significant advancement toward stable, efficient, and resilient DC microgrid operations.**

Index Terms—**Battery charging, DC microgrid, Bidirectional DC–DC converter, Converter-failure resiliency, Power management.**

I. Introduction

MICROGRIDS that incorporate distributed energy resources offer numerous advantages that conventional centralized power systems are often unable to provide. These systems are characterized by enhanced reliability, operational efficiency, and flexibility. Moreover, by utilizing clean energy sources such as solar and wind, microgrids contribute to a significant reduction in greenhouse gas emissions [1]. The operational costs are typically lower due to the integration of various energy technologies. Given that electrical energy is a fundamental necessity for modern society, the environmental impact of traditional fossil fuel-based power generation. Especially, its contribution to global warming and pollution has become a critical concern. In response, the deployment of microgrids powered by various energy sources has gained significant attention. As microgrids generate and supply electricity independently through distributed generation units, considerable research efforts have focused on their management, aiming to enhance system stability and optimize operational performance [2].

Microgrids are generally categorized as either alternating current (AC) or direct current (DC) systems [3]. Compared to their AC counterparts, DC microgrids offer a simplified power conversion process, which typically results in reduced power losses and facilitates easier control, as they do not require reactive power compensation or frequency regulation [4]. In a typical DC microgrid, multiple power electronic converters are interconnected through a common bus, with each converter independently designed to operate reliably. However, when multiple converters function concurrently within the same system, unforeseen dynamic interactions can arise. These interactions may adversely impact the performance of other converters connected to the common bus, potentially compromising the overall system stability. Consequently, there is a critical need for effective strategies to mitigate inter-converter interference and enhance the stability and reliability of the DC microgrid [5].

DC microgrids present a promising solution for the efficient integration of renewable energy sources, energy storage systems, and electrical loads, particularly at the distribution or sub-transmission levels. They have emerged as a practical alternative to traditional DC distribution systems, which typically involve multiple stages of power electronic conversion. A DC microgrid is a locally controllable system capable of operating in both grid-connected and stand-alone (islanded) modes, enabling it to function independently of the main transmission grid when necessary [6]. Despite these advantages, DC microgrids also encounter several challenges that must be addressed to ensure robust performance. These include: (1) maintaining power supply reliability; (2) mitigating instability issues arising from the constant power load characteristics of point-of-load converters and the dynamic interactions among multiple power converters; (3) achieving seamless integration with conventional AC distribution networks; (4) maintaining power balance between distributed energy sources and varying load demands; and (5) ensuring high-efficiency, high-performance power conversion [6, 7]. Nevertheless, DC microgrids offer several advantages over AC systems. They eliminate the need for complex reactive power control, reduce the number of power conversion stages.

Adnan Tan, is with Department of Electrical & Electronics Engineering Çukurova University, Adana, Türkiye,(e-mail: atan@cu.edu.tr).
https://orcid.org/0000-0002-5227-2556

Manuscript received Jun 18, 2025; accepted Jul 30, 2025.
DOI: 10.17694/bajece.1722413

Thereby, it provides the improving overall efficiency and avoiding issues related to frequency stability and harmonic distortion [8].

In DC microgrids, bidirectional DC–DC converters are often essential to facilitate power flow to, from, or between various energy storage components [9]. As a result, the integration of such converters between energy storage systems and DC buses operating at different voltage levels has emerged as a significant area of research. In many applications, these converters are required to achieve high voltage conversion ratios between input and output terminals. Bidirectional DC–DC converters play a critical role in regulating the internal DC bus voltage and maintaining power balance within the microgrid [9]. Consequently, there is a growing demand for converter topologies that can present efficient and reliable bidirectional power flow management among diverse storage systems [10].

Numerous control strategies have been developed to ensure the reliable and efficient operation of microgrids under both grid-connected and stand-alone modes. In the literature, some studies have addressed and examined similar issues. For instance, in the study [11], the researchers propose a fault-tolerant converter topology in conjunction with a fault diagnosis algorithm for a bidirectional DC–DC converter utilized in a battery energy storage system. The proposed approach enables the detection of open-circuit faults and facilitates dynamic reconfiguration of the converter topology to ensure safe and uninterrupted system operation while maintaining connectivity with the DC microgrid [11]. In another study, a power management control strategy designed to maintain power balance within a stand-alone DC microgrid is introduced, wherein the renewable energy source and the battery energy storage system are integrated through dedicated power electronic converter interfaces [9]. Another study aims to present a reliability assessment methodology for the aggregate power conversion unit within an off-grid PV–battery-based DC microgrid, specifically targeting dynamic and transient operating conditions in local and rural energy communities [12]. In addition, the study proposes a complex power-based approach encompassing both real and imaginary components to develop effective protection schemes for rapid short-circuit fault detection and isolation within a DC microgrid [13].

This study proposes a fault-tolerant power control strategy for a DC microgrid architecture where a sensitive load is normally fed by the main DC bus connected to the AC grid through an AC/DC converter. The proposed control methodology incorporates a hybrid regulation framework composed of Constant Current (CC) and Constant Voltage (CV) controllers, coordinated through an integrated fault detection algorithm. During nominal operating conditions, the control system functions in CC mode to effectively manage bidirectional power flow between the battery energy storage system (BESS) and the DC microgrid. Simultaneously, the fault detection mechanism performs real-time monitoring of the DC bus voltage to promptly identify grid disturbances. Upon detecting any voltage anomaly indicative of a fault condition, the controller transitions to CV mode to ensure voltage stability and preserve the operational integrity of the DC bus throughout the disturbance period.

In the proposed system study, a dynamic and automatic control strategy between energy storage charging and discharging modes based on DC bus conditions is proposed. In addition, the use of constant current charging and constant voltage discharging modes optimizes both battery life and load protection. Compared to existing approaches, this control method aims to provide both continuity and system efficiency without requiring additional control circuits. In summary, this work contributes to the DC microgrid field as follows:

- Designing a fault-tolerant control strategy for the bidirectional converter (buck for charging, boost for discharging),
- Implementing converter-failure detection by using DC-bus voltage to change DC-DC converter operation mode,
- Maintaining load voltage stability at 48 V during converter-failure conditions.

The current study is organized as follows: In Section 2, the system design is explained and detailed in step-by-step. In Section 3, the control scheme for charging/discharging modes is clarified. Then, the performance findings are presented in Section 4. And finally, a Conclusions section is located to explain concluding remarks in Section 5.

II. The System Design

In this section, the system structure of DC microgrid consisting an energy storage system (ESS), and a bidirectional DC-DC converter and a load connected to common bus, as presented in Figure 1.

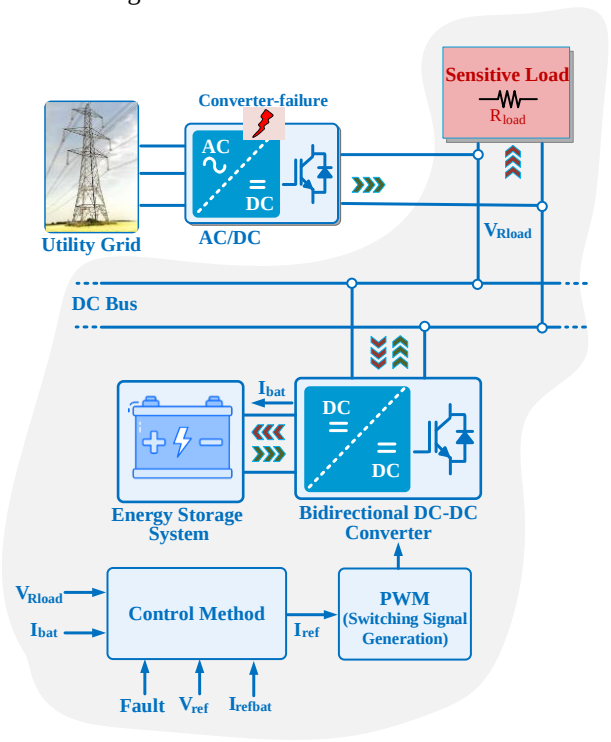


Fig. 1. The designed system consisting of a bidirectional DC-DC converter, battery and sensitive load

When a converter-failure occurs and disconnects from the electrical grid, the ESS supplies an electrical power to maintain stable voltage at the load-side. DC voltage is continuously measured and bidirectional converter is controlled by a method that monitors voltage and battery current. The control system generates a reference current and uses PWM signals to adjust the DC-DC converter's operation accordingly.

A. Bidirectional DC-DC Converter Design

A bidirectional buck-boost converter facilitates the two way power transfer of electrical energy between two distinct DC voltage levels, functioning either as a step-down (buck) or step-up (boost) converter. The topology incorporates both power switches and diodes, with the switches operating in an alternating manner. One switch is only active at any given time while the diodes guide the current flow [14]. The converter operates in two principal modes: boost mode, where the output voltage exceeds the input voltage, and buck mode, where the output voltage is lower than the input.

In practical operation, the converter functions in two distinct modes based on converter-failure conditions within the DC microgrid. The direction of energy transfer and the selection between buck and boost operation are governed by the states of the power switches S1 and S2 [14]. The various operational behavior of the converter under different conditions is summarized in Table 1.

Under normal operating conditions, the bidirectional converter functions in buck mode to charge the battery. In this operational state, switch S1 is activated while switch S2 remains inactive, as illustrated in Figure 2(b). As the switching cycle progresses, the current through the inductor decreases, and the energy previously stored in inductor L is transferred to the battery. Additionally, the bidirectional converter facilitates the transfer of electrical power from the DC bus to the battery for charging process.

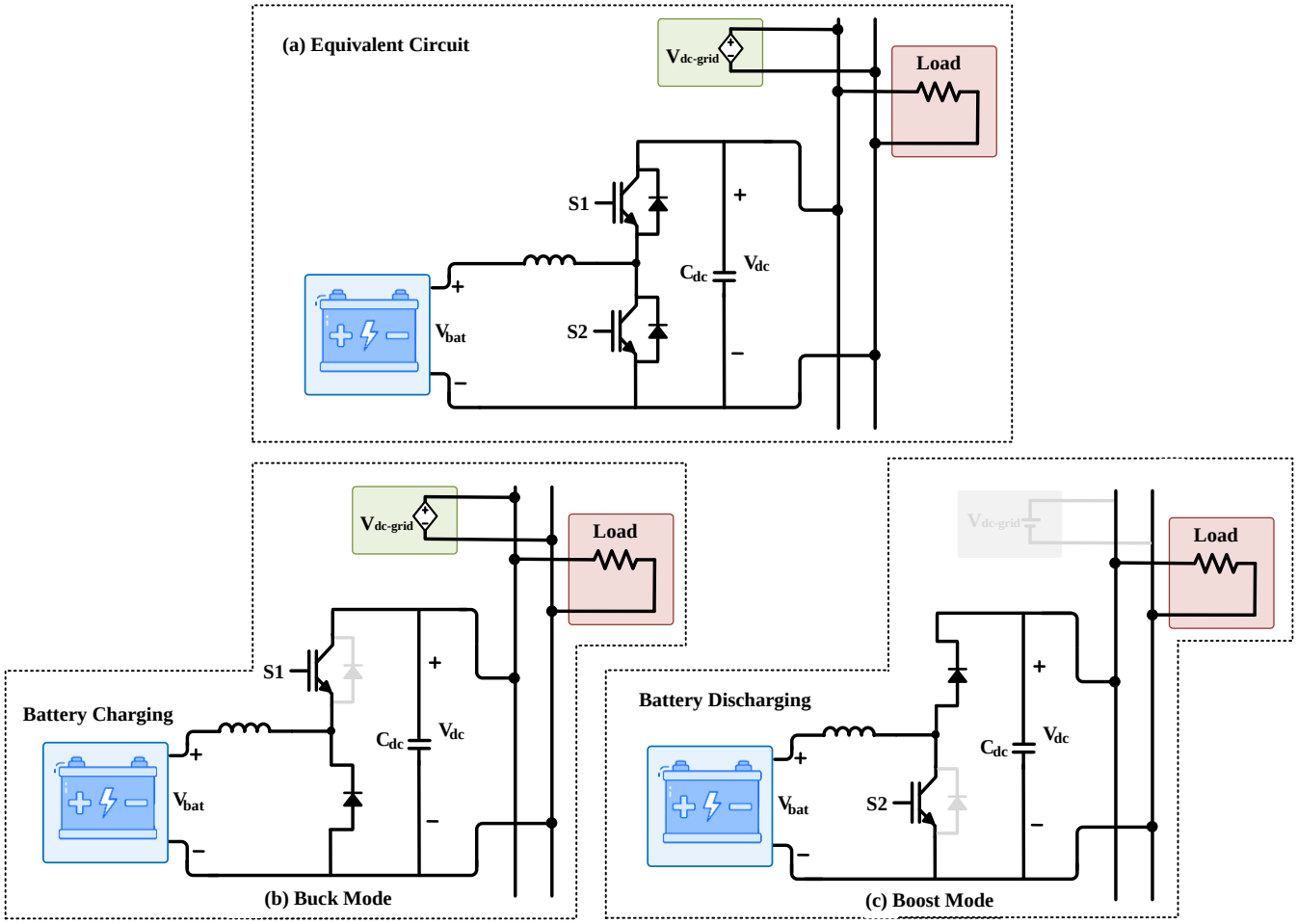


Fig. 2. (a) equivalent circuit, (b) buck mode, and (c) boost mode of the designed DC-microgrid system

Table 1. The operating specifications of converter circuit in the designed system

Mod	Operating mode	Power Transfer	Operating mode	Duty cycle
Mod 1	Charging	DC Grid to Battery	Buck	D_{buck}
Mod 2	Discharging	Battery to DC load	Boost	D_{boost}

During a converter-failure, the converter transitions to boost mode to discharge the battery. In this scenario, switch S2 is activated, while switch S1 assumes a passive state, as shown in Figure 2(c). The converter then supplies power from the battery to the DC load. Furthermore, in instances where renewable energy production is insufficient, the bidirectional converter ensures the regulation of power flow and stabilization of the DC bus voltage during both charging and discharging operations.

In Equation 1, the D values for buck and boost modes are calculated as follows:

$$D = \begin{bmatrix} D_{buck} \\ D_{boost} \end{bmatrix} = \begin{bmatrix} V_{bat}/V_{dc} \\ 1 - (V_{bat}/V_{dc}) \end{bmatrix} \quad (1)$$

C_{dc} capacitance value is calculated by using Equation 2.

$$C_{dc} = (I_{dc-load} D)/(f_s \Delta V_{dc}) \quad (2)$$

In the designed charging system, the switching frequency is 10 kHz. In the system, C_{dc} is selected as 1.2 mF and L value is selected as 0.1 mH by using equations.

B. Battery Unit

In this study, the battery module is modeled using a generic dynamic framework capable of representing the behavior of widely used rechargeable battery types. The equivalent circuit diagram of the battery model is illustrated in Figure 3. The proposed system incorporates a lithium-ion battery pack. Under buck mode operation, the battery pack functions in charging mode, wherein the DC link serves as the input to the bidirectional converter and the battery acts as the load on the output side. Conversely, in boost mode, the battery discharges to supply electrical energy to the system.

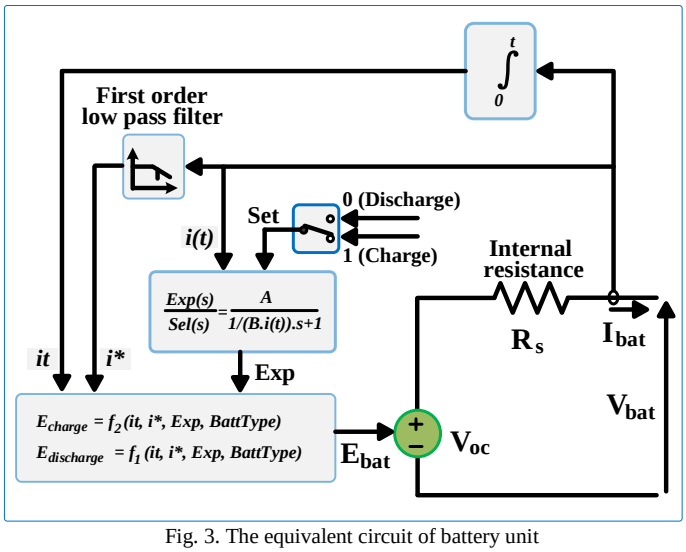


Fig. 3. The equivalent circuit of battery unit

In the proposed system, a lithium-ion battery with a rated capacity of 720 Ah is designed and implemented. The specific parameter values of the battery are presented in Table 2. A comprehensive description of the circuit configuration can be found in Ref. [15]. For lithium-ion batteries, the design criteria for the battery voltage (V_{bat}) under both charging and discharging conditions are determined based on the following set of equations.

The Equation 3 explains the discharging state (i*>0).

$$E_{discharge} = f_1(it, i^*, i) = E_0 - K.\frac{Q}{Q-it}.i^* - K.\frac{Q}{Q-it}.it + A.exp(-B.it) \quad (3)$$

The Equation 4 explains the charging state (i*<0).

$$E_{charge} = f_2(it, i^*, i) = E_0 - K.\frac{Q}{it+0.1Q}.i^* - K.\frac{Q}{Q-it}.it + A.\exp(-B.it) \quad (4)$$

In the aforementioned equations, E_0 represents the constant voltage, expressed in volts (V). The parameter K denotes the polarization constant, measured in volts per ampere-hour (V/Ah). Q corresponds to the maximum capacity of the battery, given in ampere-hours (Ah). The term A refers to the exponential voltage component (V), while B is the exponential capacity parameter, expressed in inverse ampere-hours Ah^{-1} [15, 16].

Based on the lithium-ion battery parameters outlined in Table 3, the battery pack is configured as an energy storage unit and implemented within a host software environment. The model is characterized by a nominal voltage of 24 V, a charge capacity of 30 Ah, and a dynamic response time of 1 second. Additionally, the battery includes an internal resistance of 0.008 Ω and is initialized at a state of charge (SoC) of 60%.

Table 2. The parameters of the lithium-ion battery in the designed energy supply unit

Parameter	Value	Unit
Nominal voltage	24	[V]
Rated capacity	30	[Ah]
Initial state of charge	60	[%]
Cut off voltage	18	[V]
Fully charged voltage	27.9357	[V]
Battery response time	1	[s]
Internal resistance	0.008	[Ω]

C. DC-Grid and Sensitive Load

The system considered in this study is designed to be fed from the DC grid and the sensitive load is fed using a DC-bus providing 48 V DC voltage. The sensitive load consists of a 10 ohm resistance and it means that a current of 4.8 A passes through the load with approximately 230 W consumed power rating. In case of any converter-failure situation in the system, the ESS is activated to ensure that the load is fed continuously and stably. The energy storage unit works integrated with the grid via a bidirectional DC-DC converter and monitors the changes in the load voltage with the help of a control algorithm

and supports the load if necessary.

III. Charging/Discharging Control Methodology

The detailed fault-tolerant control architecture used in the designed system is illustrated in Figure 4. The primary objectives of the overall control scheme are as follows: (i) to maintain continuous power delivery to the load by regulating the DC-link voltage under both normal and converter-failure conditions, (ii) to enable battery charging with a constant current during normal operating conditions, and (iii) to facilitate bidirectional power transfer via the bidirectional DC–DC converter [17]. The proposed charging and discharging strategy employs a multi-loop control framework, which integrates power flow management, constant current charging control, and DC-link voltage regulation. This control structure is designed to operate adaptively, allowing real-time transitions between control loops to respond dynamically to changing system conditions.

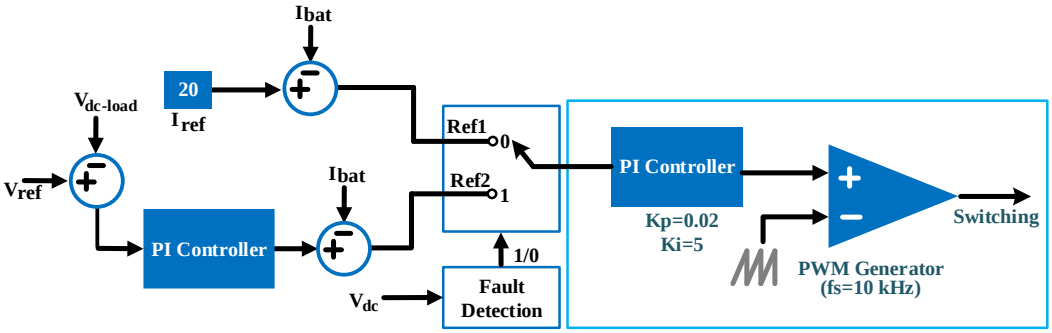


Fig. 4. The control method for battery charging and load voltage regulation

Continuous power delivery to the load is ensured through the DC-link control loop, which remains active at all times. In contrast, the activation of the additional control loops depends on the instantaneous power generation capability of the battery. Power transfer from the DC-side to the battery is governed by a constant current charging algorithm. A shared characteristic of all the aforementioned control loops is their reliance on pulse-width modulation techniques. In this context, the duty cycle of the power electronic switches is dynamically adjusted to regulate power flow by controlling the voltage levels of both the output DC-link and the battery-connected DC busbar.

A PI voltage controller is employed to regulate the output voltage of power converters by continuously comparing the measured output with a predefined reference voltage. It utilizes both proportional and integral control actions to compensate for discrepancies arising from fluctuations in the input voltage or variations in load conditions, thereby ensuring a stable and consistent output voltage [18]. In contrast, a PI current controller is designed to regulate the current within the circuit by monitoring the actual current and adjusting it in accordance with the error between the measured value and the desired reference. This approach allows for precise current control, contributing to the overall stability and performance of the power conversion system. To charge battery at reference value (I_{ref}), the controller adjusts battery current (I_{bat}) accordingly via D_{buck}.

$$Ref_1(t) = Error_1 = I_{ref} - I_{bat}(t) \quad (5)$$

$$D_{buck}(t) = K_p Ref_1(t) + K_i \int Ref_1(t)dt \quad (6)$$

In charging process, I_{ref} value is set to 20 A. By comparing I_{ref} value, battery current is permanently measured. PI control loop is used to adjust the D value to provide I_{bat} current at desired value.

The DC-load control loop generates a reference signal by comparing the actual load voltage with a predefined reference voltage (V_{ref} =48 V) and feeding the resulting error into a PI controller. This reference signal is then further compared with the battery current, and the resulting error is processed through an additional PI controller to produce a reference signal for switching process. During the discharging phase, the converter operates according to a discharging algorithm, while the charging unit maintains a constant output voltage, denoted as V_{ref}, to ensure stable charging conditions.

The error value between V_{ref} and $V_{dc-load}$ is given as:

$$Error_2(t) = V_{ref} - V_{dc-load}(t) \quad (7)$$

The reference value and D_{boost} value of discharging mode is:

$$Ref_2(t) = \left[\left(K_p Error_2(t) + K_i \int Error_2(t)dt\right) - I_{bat}(t)\right] \quad (8)$$

$$D_{boost} = K_p Ref_2(t) + K_i \int Ref_2(t)dt \quad (9)$$

In this operating mode, the converter is controlled to maintain the output voltage at the reference level of 48 V. During converter-failure conditions, the duty cycle is modulated to ensure that $V_{dc-load}$ remains at its desired value.

IV. Performance Results

In this study, the performance results of the charging and discharging states of the bidirectional dc-dc converter in the event of a converter-failure in the dc microgrid are analyzed. In this context, the performance analysis is performed using the Matlab/Simulink program. The simulation time is set to 5 seconds and the resolution time is selected as 10 us.

The characteristics of the lithium-ion battery used in the system is introduced in Figure 5. Nominal current discharge characteristics is presented for discharge curve (blue), nominal area (gray) and exponential area (yellow) at13.0435 A. Nominal area is defined as the region where the battery operates mostly at constant voltage and provides most of its capacity. As the current increases, the discharge time shortens and the battery discharges faster. For example, at 50 A, the battery is completely discharged in ~0.5 hours. At lower currents, the battery lasts longer (~3.5 hours) and the discharge time is longer.

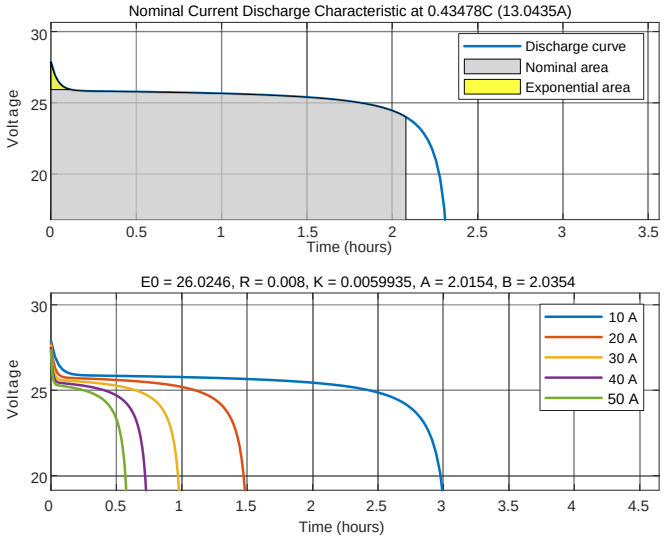


Fig. 5. Discharging characteristics of lithium-ion battery in the system

The results of charging and discharging the battery in a certain time interval in the system are shown in Figure 6. The upper graph shows the battery current (I_{bat}), the lower graph shows the battery voltage (V_{bat}). In the first case, the battery current is around -20 A (negative current) while connected to the grid, which indicates that the battery is being charged. The actual charging current is measured in the range of 19.4–20.6 A and is emphasized in the results. At t=2 s, a converter-failure occurs in the DC grid-side and the battery transitions from charge to discharge. The system operates in discharge mode between T=2 s and t=5 s, and the battery current is measured around 8.6-9.8 A, and the battery current is also compared with the reference current value of 9.2 A.

The power waveforms on the grid, battery and load sides are presented in Figure 7. In the normal grid case, it is seen that 780 W power is provided. In this case, the instantaneous power values on the load and battery sides are seen to be 230 W and 525 W, respectively. After t=2 s, the grid has no supplied-power (its value is 0 W) due to converter-failure. In the converter-failure case, the battery goes into discharge mode and delivers power to the load side. The load draws approximately the same power, and it is also confirmed that the system works effectively.

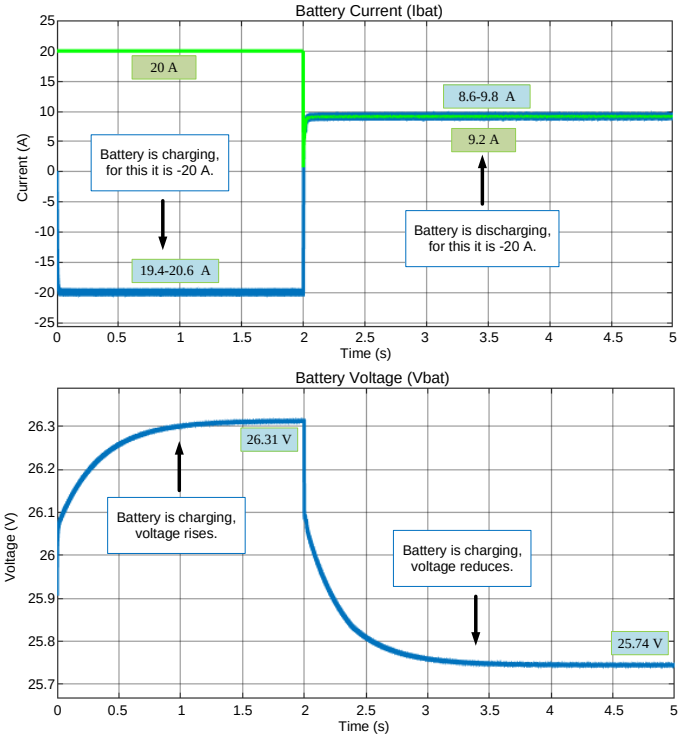


Fig. 6. Battery current and voltage waveforms during charging and discharging

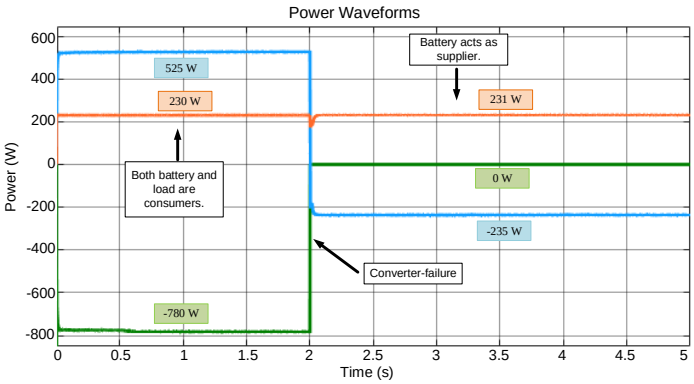


Fig. 7. Power waveforms of grid, load and battery-side during charging and discharging

Figure 8 shows the SoC variations of the battery connected to DC-DC converter during both charging and discharging modes. During the normal operation conditions, the DC-DC converter operates in charging mode and SoC increases from 60% up to 60.04%. After converter failure occurs at t=2s, the DC-DC converter seamlessly transfer to discharging operation to stabilize the voltage on DC grid and supply the load. After the failure, the SoC decreases almost to 60.01 at the end of the operation. This SoC results confirm that the designed system charges/discharges the battery properly and ensures continuous and stable energy delivery under converter fault conditions.

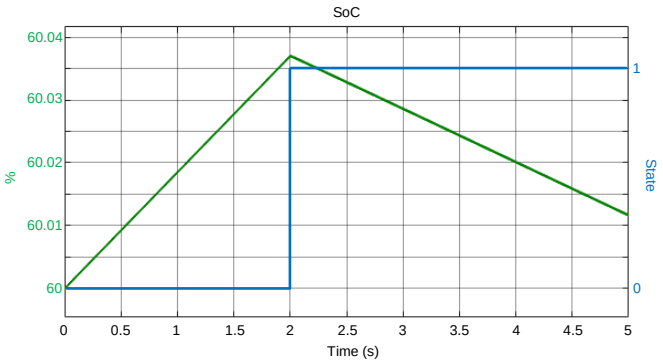


Fig. 8. SoC values during charging and discharging

V. CONCLUSIONS

In this study, a fault-resilient control strategy combining Constant Current (CC) and Constant Voltage (CV) modes has been developed for efficient power management and enhanced stability in DC microgrids. The proposed method enables seamless transition between CC and CV modes through a real-time fault detection algorithm that monitors grid voltage disturbances. The proposed method has proven its effectiveness in maintaining stable voltage levels at sensitive loads, preserving power continuity under varying operational conditions. The implementation of constant current charging and constant voltage discharging optimizes battery performance, extending battery lifespan while enhancing overall system efficiency. Simulation results validated the proposed strategy's reliability, highlighting its suitability for practical DC microgrid applications. Future work may explore adaptive algorithms for even finer control under dynamic load variations and integration with diverse renewable energy sources.

REFERENCES

[1] K. Paul et al., "Optimizing sustainable energy management in grid connected microgrids using quantum particle swarm optimization for cost and emission reduction," Scientific Reports, vol. 15, no. 1, p. 5843, 2025.

[2] G. S. Thirunavukkarasu, M. Seyedmahmoudian, E. Jamei, B. Horan, S. Mekhilef, and A. Stojcevski, "Role of optimization techniques in microgrid energy management systems—A review," Energy Strategy Reviews, vol. 43, p. 100899, 2022.

[3] S. Ferahtia, A. Houari, T. Cioara, M. Bouznit, H. Rezk, and A. Djerioui, "Recent advances on energy management and control of direct current microgrid for smart cities and industry: A Survey," Applied Energy, vol. 368, p. 123501, 2024.

[4] A. W. Adegboyega, S. Sepasi, H. O. R. Howlader, B. Griswold, M. Matsuura, and L. R. Roose, "DC Microgrid Deployments and Challenges: A Comprehensive Review of Academic and Corporate Implementations," Energies, vol. 18, no. 5, p. 1064, 2025.

[5] S. S. Rangarajan et al., "DC microgrids: a propitious smart grid paradigm for smart cities," Smart Cities, vol. 6, no. 4, pp. 1690-1718, 2023.

[6] F. S. Al-Ismail, "DC microgrid planning, operation, and control: A comprehensive review," IEEE Access, vol. 9, pp. 36154-36172, 2021.

[7] B. Aluisio, M. Dicorato, I. Ferrini, G. Forte, R. Sbrizzai, and M. Trovato, "Planning and reliability of DC microgrid configurations for Electric Vehicle Supply Infrastructure," International Journal of Electrical Power & Energy Systems, vol. 131, p. 107104, 2021.

[8] S. Sarangi, B. K. Sahu, and P. K. Rout, "A comprehensive review of distribution generation integrated DC microgrid protection: issues, strategies, and future direction," International journal of energy research, vol. 45, no. 4, pp. 5006-5031, 2021.

[9] S. J. Ríos, D. J. Pagano, and K. E. Lucas, "Bidirectional power sharing for DC microgrid enabled by dual active bridge DC-DC converter," Energies, vol. 14, no. 2, p. 404, 2021.

[10] S. Punna, R. Mailugundla, and S. R. Salkuti, "Design, analysis and implementation of bidirectional DC–DC converters for Hess in DC microgrid applications," Smart Cities, vol. 5, no. 2, pp. 433-454, 2022.

[11] M. S. Mahdavi, M. S. Karimzadeh, T. Rahimi, and G. B. Gharehpetian, "A fault-tolerant bidirectional converter for battery energy storage systems in DC microgrids," Electronics, vol. 12, no. 3, p. 679, 2023.

[12] T.-C. Kuo, T. T. Pham, D. M. Bui, P. D. Le, T. L. Van, and P.-T. Huang, "Reliability evaluation of an aggregate power conversion unit in the off-grid PV-battery-based DC microgrid from local energy communities under dynamic and transient operation," Energy Reports, vol. 8, pp. 5688-5726, 2022.

[13] S. K. Prince, S. Affijulla, and G. Panda, "Protection of DC microgrids based on complex power during faults in on/off-grid scenarios," IEEE Transactions on Industry Applications, vol. 59, no. 1, pp. 244-254, 2022.

[14] M. Berger, I. Kocar, E. Farantatos, and A. Haddadi, "Modeling of Li-ion battery energy storage systems (BESSs) for grid fault analysis," Electric Power Systems Research, vol. 196, p. 107160, 2021.

[15] M. M. Savrun, M. İnci, and M. Büyük, "Design and analysis of a high energy efficient multi-port dc-dc converter interface for fuel cell/battery electric vehicle-to-home (V2H) system," Journal of Energy Storage, vol. 45, p. 103755, 2022.

[16] J. Bilansky, M. Lacko, M. Pastor, A. Marcinek, and F. Durovsky, "Improved digital twin of li-ion battery based on generic matlab model," Energies, vol. 16, no. 3, p. 1194, 2023.

[17] W. Christopher et al., "A bidirectional DC/DC converter for renewable energy source-fed EV charging stations with enhanced DC link voltage and ripple frequency management," Results in Engineering, vol. 24, p. 103469, 2024.

[18] T. Chen et al., "An optimized bidirectional buck–boost converter for DC bus voltage stabilization in new generation poloidal field power supply," Energy Reports, vol. 8, pp. 188-200, 2022.

BIOGRAPHY

Adnan Tan received the M.Sc. and Ph.D. degrees from the Department of Electrical and Electronics Engineering, Çukurova University, Adana, Turkey, in 2011 and 2015, respectively. He has been working as an Assistant Professor with the Electrical and Electronics Engineering Department, Çukurova University, since 2015. His research interests include power electronic converters, utility applications of power electronics, electrical power quality, energy storage systems, and microgrid.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 3, September 2025
Research Article

Design and Optimization of Oxide Material-Based Single-Mode Rib Waveguides for Photonic Integrated Circuits

Huriye Gencal, Mustafa Demirtas and Umut Aydemir

***Abstract*— The primary objective of this study is to design single-mode rib waveguide structures using oxide-based materials. To achieve this goal, the waveguide dimensions including width and height were optimized to ensure reliable single-mode operation. Optimization was performed using the particle swarm optimization (PSO) algorithm. The effective refractive index (n_{eff}), a critical parameter for waveguide design, was calculated and evaluated to confirm suitability for single-mode propagation. Additionally, the influence of various oxide materials commonly used in passive photonic components, including integrated optical waveguides, was comparatively analyzed.**

Simulation results provided a comparison of effective refractive indices for rib waveguides fabricated from different oxide materials. Optimal geometric parameters ensuring minimal propagation losses and maximum mode purity were identified through these simulations. Overall, this work offers a comprehensive optimization strategy for oxide-based rib waveguides, contributing valuable insights toward enhancing the performance of integrated photonic devices.

Index Terms— **Rib waveguide, oxide-based materials, single mode, propagation losses, effective refractive index (n_{eff}), integrated photonics.**

I. Introduction

PHOTONIC INTEGRATED CIRCUIT (PIC) technologies play a crucial role in fields such as communication systems, sensing platforms, and quantum information processing, thanks to their advantages including high bandwidth, low power consumption, and integration capability and maintenance of modal purity. In this context, single-mode waveguides are particularly advantageous because they exclusively support the fundamental mode, thereby eliminating unwanted effects such as dispersion and inter-modal interference, resulting in high-quality signal transmission [1].

While silicon-based waveguides have traditionally been widely utilized, their limitations, including the absence of second-order optical nonlinearity and relatively high optical losses, have increased interest in alternative material platforms [2], [3]. Among these, oxide materials with varying refractive indices, notably amorphous Al_2O_3, HfO_2, and TiO_2 have attracted considerable attention for next-generation dielectric waveguide cores owing to their wide bandwidth, thermal stability, and compatibility with CMOS fabrication processes. Typically combined with SiO_2, substrates, these materials enable the formation of high-index contrast structures, thus ensuring strong mode confinement. In this work, the selection of Al_2O_3, HfO_2, and TiO_2 enables a direct comparison across low, medium, and high refractive index regimes. This approach allows a systematic investigation of how index contrast affects the confinement factor, while also taking advantage of each material's proven CMOS compatibility, strong thermal stability, and mature fabrication methods, qualities that make them strong candidates for practical implementation in integrated photonic systems.

Al_2O_3 facilitates low-loss optical transmission in the 1.48–1.61 μm wavelength range and is particularly notable for its amorphous Özden et al. demonstrated that Al_2O_3 rib waveguides, with dimensions around 0.5 μm thickness and approximately 3.5 μm width, could achieve polarization-insensitive single-mode transmission for both TE and TM polarizations [4]. Further research from the same group showed that a slab thickness of approximately 125 nm limits the number of supported modes and minimizes losses, achieving propagation losses as low as ~0.04 dB/cm using ALD [5]. Additionally, erbium-doped Al_2O_3 waveguides have been developed into integrated optical amplifiers exhibiting a net gain of 13.7 dB/cm [6]. Hafnium dioxide (HfO_2) offers a higher refractive index (~2.0), enabling more compact waveguide designs. In particular, Jaramillo et al. introduced an HfO_2 – Al_2O_3 composite platform deposited via ALD that demonstrated single-mode waveguide losses around 0.25 dB/cm at 729 nm and 2.6 dB/cm at 405 nm, along with intrinsic quality factors up to 2.6×10^6 in visible microresonator devices [7]. Titanium dioxide (TiO_2), with a refractive index reaching approximately 2.3, provides strong modal

Huriye Gencal, is with Department of Electrical Electronics Engineering University of Bursa Uludag University, Bursa, Turkey, (e-mail: huriyegencal@gmail.com).
https://orcid.org/0000-0002-6460-1177

Mustafa Demirtas, is with Department of Electrical Electronics Engineering University of Bursa Uludag University, Bursa, Turkey, (e-mail: mustafademirtas@uludag.edu.tr).
https://orcid.org/0000-0001-6832-4341

Umut Aydemir, is with Department of Electrical Electronics Engineering University of Bursa Uludag University, Bursa, Turkey, (e-mail: umutaydemir@uludag.edu.tr).
https://orcid.org/0000-0001-5396-4610

Manuscript received Jul 02, 2025; accepted Aug 18, 2025.
DOI: 10.17694/bajece.1733035

confinement, especially beneficial for visible wavelength applications. However, the high index introduces strong dispersion and surface scattering effects, necessitating precise control of single-mode conditions. Evans et al. achieved single-mode TiO_2 waveguides by employing core films 260 nm thick and widths between 0.8–1.2 µm, reporting losses of ~1.2 dB/cm at 1550 nm. However, in cases where higher surface quality was not achieved, edge roughness could elevate losses significantly (4–28 dB/cm) [8].

The above studies demonstrate that high refractive index oxide materials can effectively produce single-mode rib waveguides. Nevertheless, careful optimization of geometric parameters such as rib width, slab thickness, and total waveguide height is essential for precise control of mode number [1], [9]. Recently, heuristic methods such as particle swarm optimization (PSO), integrated with optical simulation tools, have emerged as powerful approaches for designing high-performance single-mode waveguides [10], [11]. In this study, the optimum geometric parameters ensuring single-mode operation in rib waveguides made from Al_2O_3, HfO_2, and TiO_2 were determined based on effective refractive index (n_{eff}) analyses, and their performance was evaluated through simulations. The ultimate goal is to contribute systematically to the design of low-loss, high modal purity waveguides suitable for integrated photonic systems.

II. DESIGN AND OPTIMIZATION METHODOLOGY

In the analysis of optical waveguides, certain fundamental parameters are essential for characterizing and comparing device performance. [12] The effective refractive index (n_{eff}) is defined as the ratio of the mode's propagation constant (β) to the free-space wave number (k_0) expressed as:

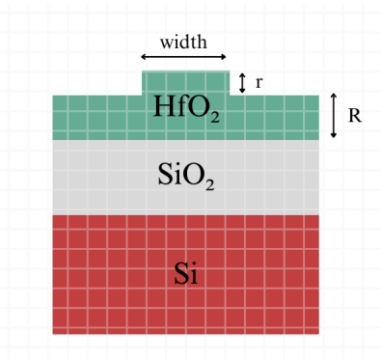
$$n_{eff} = \beta / k_0 \quad (1)$$

Another important metric is the confinement factor (Γ), which represents the proportion of the optical field energy confined within the core region of the waveguide. It can be calculated using:

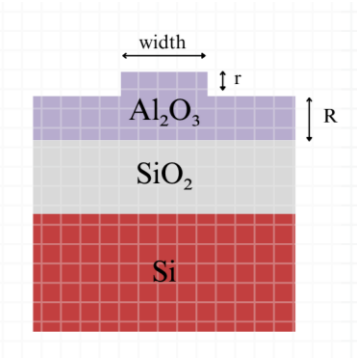
$$\Gamma = \frac{\iint_{core} |E|^2 \, dxdy}{\iint_{total} |E|^2 \, dxdy} \quad (2)$$

where represents the mode's electric field intensity. These parameters form the basis for interpreting the simulation results presented in this study. The designs focused on rib-structured optical waveguides were numerically designed using high refractive index oxides, specifically TiO_2 [8], HfO_2 [7], and Al_2O_3 [5]. The waveguide structures were constructed on a Si substrate, incorporating a SiO_2, buffer layer beneath the core region. Each core was formed by the selected dielectric oxide, with the rib region defined on top of a supporting slab layer. Fig. 1 illustrates schematic cross-sectional views of the rib waveguide geometries developed for each material. The physical characteristics of these waveguides are described by key geometrical parameters, including rib slab height (R), rib height (r), and waveguide width.

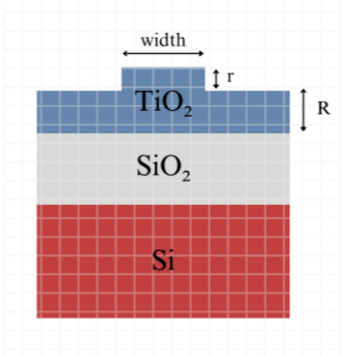
The optical performance of the waveguides was investigated using the Lumerical MODE simulation software, employing the Eigenmode Solver module, which is based on the finite element method (FEM). The simulations were conducted at a wavelength of 1550 nm, a value selected for its critical importance in optical communication applications. For each material, wavelength-dependent refractive index values reported in the literature were utilized. To prevent unphysical reflections at the simulation boundaries, perfectly matched layer (PML) boundary conditions were implemented. All essential simulation parameters used in this study are summarized in Table I. Critical optical properties, such as the effective refractive index (n_{eff}) and propagation losses, were analyzed in detail based on the simulation results.



(a)



(b)



(c)

Fig.1. Schematic cross-sectional views of rib waveguides based on (a) HfO_2, (b) Al_2O_3, and (c) TiO_2.

TABLE I Key parameters used in waveguide simulations

Parameter	Value
Wavelength (λ)	1550 nm
Core Materials	TiO_2, HfO_2, Al_2O_3
Cladding (Lower Layer)	SiO_2
Bottom Substrate	Si
Rib Height (r)	0.3–0.7 μm (optimized)
Slab Height (R)	0.05–0.3 μm (optimized)
Waveguide Width	0.2–1.5 μm (optimized)
Simulation Method	Eigenmode Solver
Boundary Conditions	PML (Perfectly Matched Layer)

The analyses were conducted to ensure exclusive support for the fundamental mode while effectively suppressing higher-order modes. In this context, the optimization of structural parameters, specifically the rib width and height, was carried out using the PSO algorithm implemented within the MATLAB environment. PSO is a heuristic optimization technique in which multiple candidate solutions, referred to as particles, explore the solution space by adapting their positions based on both their individual optimal experiences (*pBest*) and the globally best-known solution within the swarm (*gBest*) [11], [12].

Each particle in the PSO approach represents a unique combination of waveguide dimensions, such as width and height. Throughout successive iterations, the particles iteratively adjust their positions, generating new candidate solutions informed by their historical best solutions as well as the optimal solutions identified by the collective swarm.

An automated optimization process was implemented by establishing a bidirectional interface between Lumerical MODE and MATLAB software to determine the optimal geometric parameters of rib-structured oxide-based waveguides. For each material (TiO_2, HfO_2, and Al_2O_3), rib width and height parameters were optimized using the PSO algorithm, with the process extending up to 150 iterations to ensure optimal single-mode operation conditions [13].

As illustrated in detail in Fig. 2, the PSO algorithm initially generated random parameter sets (rib width and height) for each particle in the MATLAB environment, defining these as vectors [11]. These parameters were automatically transferred to the Lumerical MODE software through an Application Programming Interface (API). Subsequently, Lumerical MODE generated corresponding waveguide geometries and performed modal analyses using its Eigenmode Solver module. This analysis computed the supported modes and their associated effective refractive indices (n_{eff}), as well as spatial distributions of the optical field and mode profiles. Following this step, the optical parameters computed within the Lumerical environment were transmitted back to MATLAB via the API.

Each particle's solution was then evaluated using a fitness function. This fitness function was designed with a criterion that required only the fundamental mode to have n_{eff}> 1.44, while higher-order modes were constrained to $n_{eff}\leq$ 1.44. If second or higher-order modes exceeded this threshold, a penalty term was activated, significantly increasing the total cost function and consequently excluding such structures from consideration. Thus, the algorithm favored configurations that strictly supported single-mode propagation.

Throughout each iteration, both the particles individual best-known solutions (*pBest*) and the overall population's best solution (*gBest*) were updated, subsequently generating new geometric parameters informed by these optimal results. This iterative optimization loop, sustained for 150 iterations, successfully identified optimal rib width and height parameters for each material that ensured single-mode propagation and optimal effective refractive indices. Consequently, systematic design strategies for oxide-based waveguides with high index contrast, low loss, and high modal purity were effectively established.

III. RESULT&DISCUSSION

Due to its high refractive index (~2.4 at 1550 nm), TiO_2 exhibits strong mode confinement, and PSO identified optimal waveguide dimensions as a width of 1.36 μm, a slab height (R) of 0.5489 μm, and a total height of 0.7 μm. The obtained objective function value of -2.2687 indicates exclusive fundamental mode propagation with significant mode confinement, highlighting the suitability of this geometry for compact, high-performance photonic integration. In contrast, HfO_2, having a moderately high refractive index (~1.87 at 1550 nm), offers smoother mode transitions compared to TiO_2. The optimized geometry for HfO_2 was determined to be a width of 1.5 μm, slab height of 0.5668 μm, and total height of 0.7 μm. These parameters suggest that HfO_2 performs optimally at slightly larger dimensions, effectively satisfying single-mode conditions. Its gentle modal transition characteristic could be particularly advantageous for applications requiring low back-reflection. Al_2O_3, which has a relatively lower refractive index (~1.65 at 1550 nm), is commonly employed in passive circuits aimed at minimizing optical losses. The PSO optimization for Al_2O_3 resulted in an optimal waveguide width of 1.5 μm, a slab height of 0.4996 μm, and a total height of 0.7 μm. Although Al_2O_3 exhibits limited mode confinement due to its lower index contrast, it offers low dispersion and broad-bandwidth capabilities, and its high thermal stability makes it suitable for integration into sensor and laser systems.

The parameters obtained from pso algorithm-based optimization for each of the three materials are summarized in Table II. Comparing these waveguide designs, TiO_2 provided the strongest mode confinement and yielded the highest objective function value. Meanwhile, HfO_2 exhibited balanced mode propagation with a larger rib height at similar widths, whereas Al_2O_3 maintained single-mode operation despite lower index contrast at the width limit. These findings

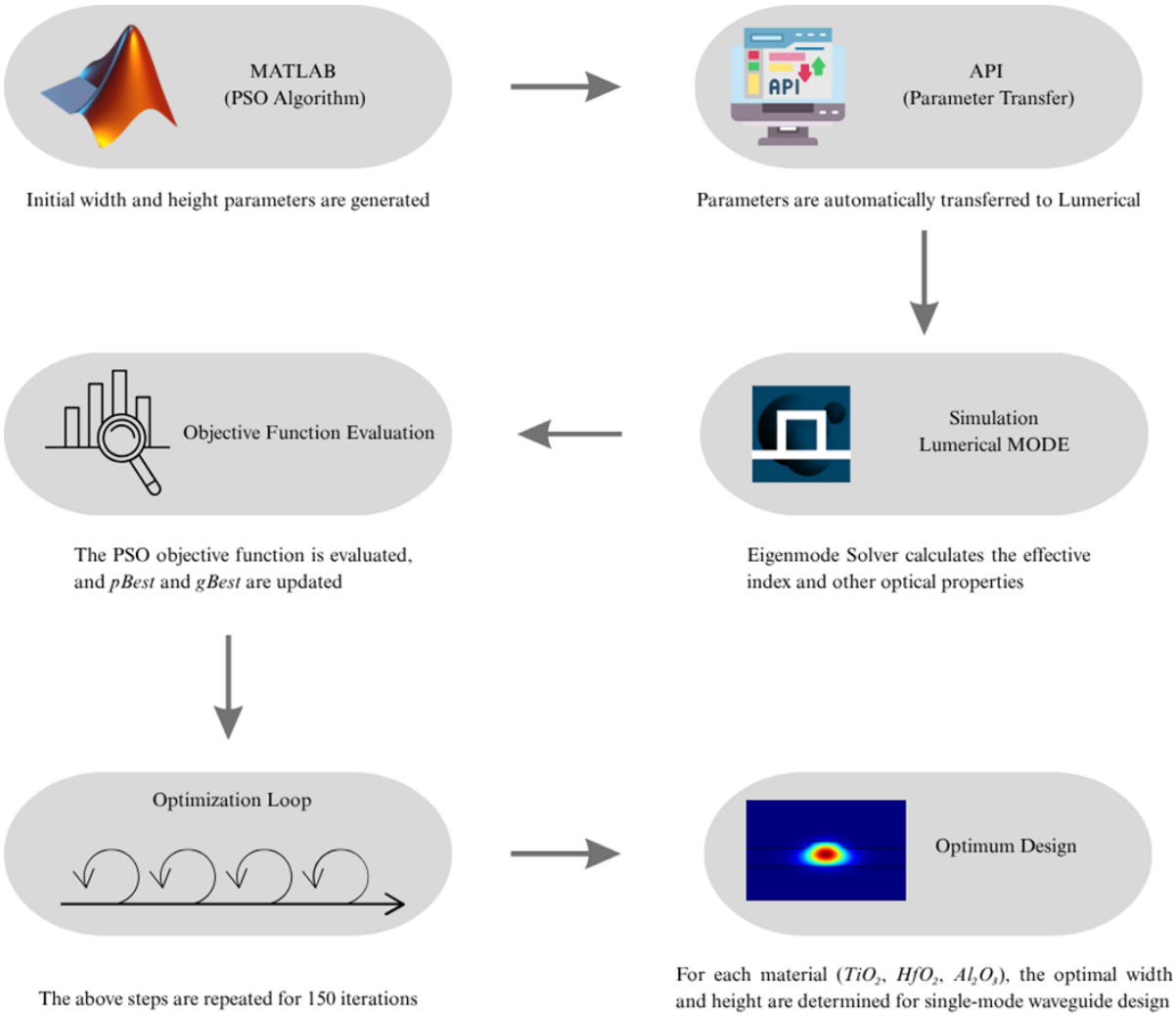


Fig.2. PSO-based waveguide design loop via MATLAB–Lumerical interface

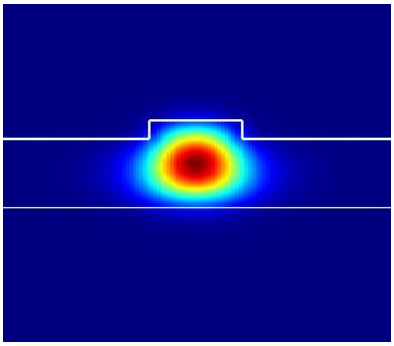
demonstrate that each material offers distinct advantages TiO_2 for compact mode confinement, HfO_2 for balanced guidance, and Al_2O_3 for low-loss broadband transmission allowing optimized designs tailored to specific application requirements.

TABLE II Optimal Design Parameters and Confinement Efficiencies for Single-Mode Rib Structures

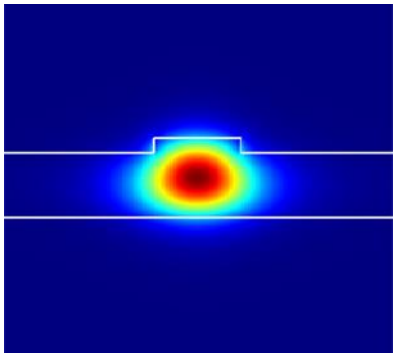
Material	Optimal Width	Optimal R	Total Height	Objective Function	Confinement Factor
TiO_2	1.36	0.5489	0.7	-2.2687	94.54%
HfO_2	1.5	0.5668	0.7	-1.7302	84.8%
Al_2O_3	1.5	0.4996	0.7	-1.6081	76.86%

In optical waveguides, the confinement factor is a critical parameter that indicates the fraction of light confined within the waveguide core, directly reflecting the optical guiding efficiency of the structure. A high confinement factor contributes to reduced optical losses and enhanced transmission efficiency. Through our optimization studies, confinement factors of 94.54% for TiO_2, 84.8% for HfO_2, and 76.86% for Al_2O_3 were achieved. These results highlight that TiO_2 provides the strongest mode confinement due to its high refractive index, while HfO_2 demonstrates balanced confinement performance. Although the confinement factor for Al_2O_3 is slightly lower compared to the other materials, its structure remains beneficial for applications requiring low optical loss and broad bandwidth capabilities.

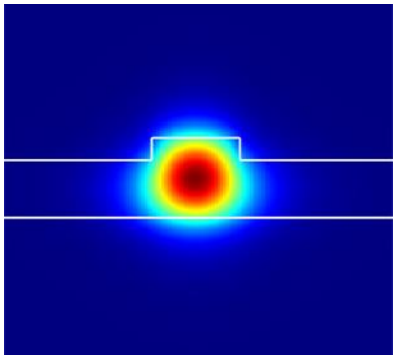
In this study, the modal analysis of the waveguides focused primarily on TE polarization. While both TE and TM modes were evaluated during the analysis, the optimization process resulted in final designs that exclusively support the fundamental TE_0 mode. For the optimized geometries, TM and higher-order modes remain in cutoff. This outcome confirms that the waveguides operate with high modal purity in TE polarization.



(a)



(b)



(c)

Fig.3. Electric field distributions of the fundamental mode in rib waveguide structures optimized with the following materials: a) TiO_2, b) HfO_2, and c) Al_2O_3.

The electric field profiles shown in Fig. 3 illustrate how the fundamental mode (TE_0) is confined within the core region of waveguides optimized for each material. In the TiO_2 -based structure, the maximum electric field intensity is highly concentrated at the center of the waveguide core, with minimal leakage to the surrounding regions. This result is consistent with TiO_2's high refractive index, providing strong mode confinement and aligning with the observed high confinement factor (94.54%). For the waveguides made of HfO_2 and Al_2O_3, the fundamental mode remains effectively confined within the core; however, the electric field distribution exhibits a smoother transition, extending slightly into the surrounding regions. Particularly for Al_2O_3, the comparatively lower refractive index results in a modestly greater mode leakage beyond the core.

Overall, the electric field profiles obtained for the rib waveguides optimized with these three different materials demonstrate effective confinement of the mode within the core region, thus ensuring single-mode operation. Variations in material properties notably influence mode confinement and field distribution, significantly impacting the determination of optimal design parameters for specific application scenarios.

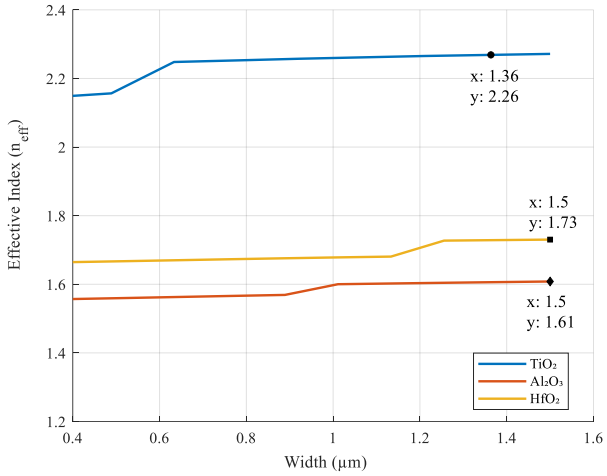


Fig.4. Effective Refractive Index Variation with Width for TiO_2, HfO_2, and Al_2O_3

To investigate the effect of waveguide width on the effective refractive index (n_{eff}), simulations were conducted in which the width parameter was systematically varied over a defined range for three different materials: TiO_2, HfO_2, and Al_2O_3. As shown in Fig. 4, the resulting plots reveal a general trend where n_{eff} increases with increasing width. At narrower widths, the mode confinement within the waveguide core is insufficient, resulting in lower n_{eff} values and greater leakage of the optical field into the surrounding regions. As the width increases, the optical mode becomes more tightly confined within the core, causing a sharp rise in n_{eff}. This trend continues up to a threshold value corresponding to the upper limit of single-mode operation, beyond which the waveguide begins to support higher-order modes, thus violating the single-mode condition. Through PSO-based optimization, the optimal width values were determined for each material, ensuring that only the fundamental mode is supported and that mode confinement is maximized. These optimal points are also marked in Fig. 4 to highlight the transition region. In Figure 4, the blue curve representing the TiO_2 waveguide shows that although effective index values above 1.326 are observed for larger widths, this corresponds to the onset of higher-order mode formation. The objective of this study is to strictly maintain single-mode operation; therefore, the maximum n_{eff} value ensuring the single-mode condition for TiO_2 was set at 1.326. The results indicate that TiO_2 achieves optimal single-mode operation at comparatively smaller widths, while $Al_2O_3$and HfO_2 require broader geometries to reach their respective optimal confinement conditions. These findings underscore the critical influence of material refractive index and geometry on the

confinement characteristics and allowable single-mode range of the waveguide.

In conclusion, detailed analysis of n_{eff} as a function of width clearly visualized in Fig. 4 along with identification of the single-mode boundary, is essential for achieving both optical efficiency and modal purity in waveguide design. The optimized widths derived from this study ensure maximum mode confinement and robust single-mode performance for each material-geometry combination.

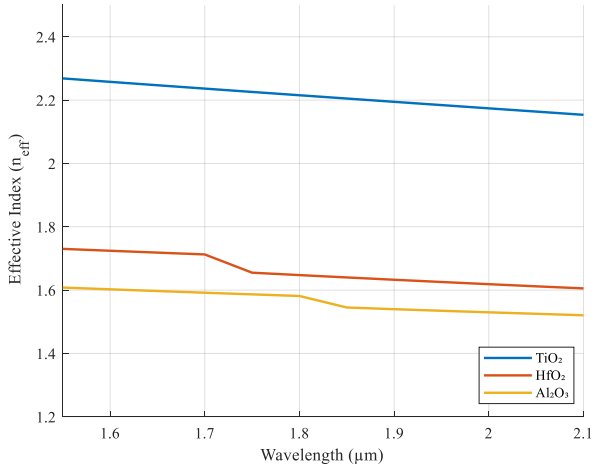


Fig.5. Effective Refractive Index Variation with Wavelength for TiO_2, HfO_2, and Al_2O_3

In the conducted simulations, the wavelength was varied from 1.55 μm to 2.1 μm to analyze the effective refractive index (n_{eff}) behavior of waveguide structures composed of three different materials: TiO_2, HfO_2, and Al_2O_3. For the wavelength sweep simulations, the refractive indices of the materials were not assumed constant; instead, wavelength-dependent refractive index data reported in the literature for each oxide material were used. As illustrated in Fig. 5, the results indicate that n_{eff} gradually decreases with increasing wavelength for all materials. Notably, at a wavelength of 1.55 μm, each material exhibits its highest n_{eff} value, suggesting that the waveguide most effectively supports the fundamental mode and achieves optimal mode confinement at this wavelength. As the wavelength approaches 2.1 μm, a noticeable decline in n_{eff} is observed. These trends are directly linked to the variation in refractive index among the materials; TiO_2, with its high index contrast, maintains effective mode confinement more robustly even at longer wavelengths.

Generally, the reduction in effective refractive index with increasing wavelength corresponds to weaker mode confinement within the waveguide core and increased optical leakage into the cladding or surrounding media. Therefore, achieving the highest possible TiO_2 at the target application wavelength is crucial for optimal waveguide performance. In conclusion, the TiO_2 -based design demonstrates the strongest mode confinement and highest effective refractive index at 1.55 μm, while HfO_2 and Al_2O_3 exhibit more significant performance degradation as the wavelength increases. These findings, as presented in Fig. 5, underscore the importance of material selection and geometric optimization tailored to the operating wavelength in the design of advanced photonic circuits.

IV. Conclusion

This study presents a comparative analysis of oxide-based dielectric materials specifically TiO_2, HfO_2, and Al_2O_3 focused on the design of single-mode rib waveguides. Utilizing a custom-developed API interface between Lumerical MODE and MATLAB, the PSO algorithm was employed to efficiently determine the optimal width and height parameters for each material. The primary objective was to ensure that the waveguide supports only the fundamental mode, thereby enhancing modal purity.

Simulation results reveal that TiO_2, owing to its high refractive index, delivers superior mode confinement and a compact mode profile. HfO_2 exhibits a more balanced modal distribution, whereas Al_2O_3, despite a broader mode profile, offers advantages for broadband and low-loss applications. Across all materials, the effective refractive index (n_{eff}) was observed to decrease gradually with increasing wavelength, reaching peak values at 1.55 μm.

These findings underscore the critical role of material selection and structural optimization in determining the optical performance of waveguides. The comprehensive comparison provided in this work offers a unique and practical perspective for researchers and designers aiming to achieve single-mode operation in integrated photonic circuits. Moreover, the methodology and results presented here are readily adaptable to other material systems or device geometries, making the approach broadly applicable to advanced photonic design efforts.

References

[1] S. P. Pogossian, L. Vescan, and A. Vonsovici, "The single-mode condition for semiconductor rib waveguides with large cross section," *J. Light. Technol.*, vol. 16, no. 10, pp. 1851–1853, 1998, doi: 10.1109/50.721072.

[2] X. Colin Tong, *Advanced Materials for Integrated Optical Waveguides*, vol. 46. [Online]. Available: http://www.springer.com/series/4076

[3] R. Soref, "The past, present, and future of silicon photonics," *IEEE J. Sel. Top. Quantum Electron.*, vol. 12, no. 6, pp. 1678–1687, Nov. 2006, doi: 10.1109/JSTQE.2006.883151.

[4] A. Özden, M. Demirtaş, and F. Ay, "Polarization insensitive single mode Al2O3 rib waveguide design for applications in active and passive optical waveguides," *J. Eur. Opt. Soc.*, vol. 10, Jan. 2015, doi: 10.2971/jeos.2015.15005.

[5] M. Demirtas, C. Odaci, N. K. Perkgoz, C. Sevik, and F. Ay, "Low Loss Atomic Layer Deposited Al2O3 Waveguides for Applications in On-Chip Optical Amplifiers," *IEEE J. Sel. Top. Quantum Electron.*, vol. 24, no. 4, Jul. 2018, doi: 10.1109/JSTQE.2018.2825880.

[6] M. Demirtas and F. Ay, "High-Gain Er3+:Al2O3On-Chip Waveguide Amplifiers," *IEEE J. Sel. Top. Quantum Electron.*, vol. 26, no. 5, Sep. 2020, doi: 10.1109/JSTQE.2020.3002656.

[7] O. Jaramillo, V. Natarajan, H. M. Rivy, J. Tensuan, L. Massai, and K. K. Mehta, "HfO 2-based platform for high-index-contrast visible

and UV integrated photonics," 2025, doi: 10.1364/OL.553552.

[8] C. C. Evans, C. Liu, and J. Suntivich, "Low-loss titanium dioxide waveguides and resonators using a dielectric lift-off fabrication process," *Opt. Express*, vol. 23, no. 9, p. 11160, May 2015, doi: 10.1364/oe.23.011160.

[9] Y. Zhou, L. Feng, and J. Sun, "Large single-mode rib waveguide in lithium niobate on insulator," in *6th International Symposium on Advanced Optical Manufacturing and Testing Technologies: Optoelectronic Materials and Devices for Sensing, Imaging, and Solar Energy*, Y. Jiang, J. Yu, and Z. Wang, Eds., SPIE, 2012, p. 84190Y. doi: 10.1117/12.975644.

[10] D. Wang, D. Tan, and L. Liu, "Particle swarm optimization algorithm: an overview," *Soft Comput.*, vol. 22, no. 2, pp. 387–408, 2018, doi: 10.1007/s00500-016-2474-6.

[11] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, 1995, pp. 1942–1948 vol.4. doi: 10.1109/ICNN.1995.488968.

[12] A. W. Snyder and J. D. Love, *Optical waveguide theory*, vol. 175. Chapman and hall London, 1983.

[13] J. Kennedy and R. Eberhart, "Particle Swarm Optimization," *Ind. Electron. Handb. - Five Vol. Set*, pp. 1942–1948, 1995, doi: 10.1007/978-3-319-46173-1_2.

BIOGRAPHIES

Huriye Gencal received her B.Sc. degree in Electrical and Electronics Engineering from Manisa Celal Bayar University, Manisa, Turkey, in 2019. She completed her M.Sc. degree in Electrical and Electronics Engineering at Bursa Technical University, Bursa, Turkey, in 2021. She is currently pursuing her Ph.D. degree at the Department of Electrical and Electronics Engineering, Bursa Uludag University, Turkey. Her research interests include photonic integrated circuits and optical waveguide design.

Mustafa Demirtaş received the B.Sc. degree in electrical and electronics engineering and industrial engineering from Dumlupınar University, Kütahya, Turkey in 2009 and 2010, respectively, and the M.Sc. and Ph.D. degrees in electrical and electronics engineering from the Anadolu University, Eskişehir, Turkey, in 2014 and 2019, respectively. From 2012 to 2019, he was a Research Assistant with the Department of EEE at Anadolu University and Eskişehir Technical University. He is an Assistant Professor with the Department of EEE at the Bursa Uludag University. His research interests include atomic layer deposition and related thin film process, novel 2D materials, integrated optics/photonic devices, and on-chip solid-state amplifiers and lasers.

Umut Aydemir received the M.S. and Ph.D. degrees from the Department of Physics, Faculty of Arts and Sciences, Gazi University, Ankara, Turkey, in 2010 and 2015, respectively. He is working as an Assoc. Professor with the Department of Electrical and Electronics Engineering, Uludag University, Turkey. His research areas involve semiconductor material based microelectronic and photonic devices.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 3, September 2025
Research Article

Lumped-Parameter Transmission Line Modeling Using MATLAB/Simulink

Ilker Ari, Ramazan Menak

***Abstract*—Accurate modeling of transmission lines is fundamental for predicting their dynamic behavior under various operating and fault conditions. This study presents a comprehensive investigation into the transient performance of a single-phase lumped-parameter transmission line using MATLAB/Simulink. The research investigates the impact of section number (N) in π-section representations on waveform fidelity and propagation characteristics. A systematic simulation framework was developed to analyze voltage and current waveforms at both sending and receiving ends, capturing key phenomena such as energization-induced oscillations, propagation delays, and the progression toward steady-state conditions. Unlike conventional tools such as ATP-EMTP, the proposed modeling approach enables flexible parameter adjustment without manual reconstruction of circuit topology, offering significant advantages in scalability and computational efficiency. The results provide deeper insight into the relationship between model segmentation and transient accuracy, offering practical guidance for both academic studies and engineering applications in power system design, stability assessment, and protection scheme development.**

Index Terms— **Analysis, Modeling, Simulink, Transients, Transmission line.**

I. INTRODUCTION

ELECTRICAL energy is recognized as a strategic resource that forms the backbone of modern societies' technological, economic, and social infrastructures. In particular, the sustainability of industrial production, transportation systems, information and communication technologies, and daily residential activities is directly dependent on the continuous and high-quality supply of electrical power. Therefore, ensuring the delivery of electrical energy in accordance with the principles of reliability, efficiency, and continuity is of critical importance for both individual consumers and national economies. Contemporary energy systems are complex and expansive structures composed of numerous components that span wide geographic areas from generation points to end-users. These systems require the integrated management of not only

electrical quantities but also economic, environmental, and operational parameters. The increasing demand for energy, the integration of renewable resources, and the advancement of smart grid technologies further necessitate the operation of energy infrastructures in a more flexible, observable, and stable manner. In this context, the stages of generation, transmission, and distribution of electrical energy must be approached as interconnected and complementary processes. Enhancing the overall security and performance of the power system can only be achieved through the holistic modeling, monitoring, and optimization of these stages. Analytical, modeling, and optimization studies conducted within the various subfields of electrical engineering play a crucial role in improving system efficiency and ensuring the long-term sustainability of electrical networks.

Transmission lines are essential infrastructure components that enable the reliable and efficient transfer of electrical energy over long distances from generation sources to consumption centers at high voltage levels. The planning and operation of these lines are inherently dependent on numerous engineering parameters that directly influence the overall performance of power systems. To ensure effective energy transmission, factors such as system voltage level, total line length, conductor type and size, reactive power compensation strategies, and system stability must be carefully considered. A holistic evaluation of these variables is crucial for maintaining sustainable and optimal operating conditions. Steady-state analyses are commonly employed in power system studies to assess whether the system can operate under stable and balanced conditions for a given loading scenario. These analyses typically involve the evaluation of voltage profiles, active and reactive power flows, voltage drops, and transmission losses, offering critical insights into the system's operational performance. While steady-state analysis plays a significant role in long-term planning, operational strategy, and power quality assurance, relying solely on these conditions is insufficient for comprehensive system evaluation. In practice, power systems are frequently exposed to transient phenomena induced by external or internal disturbances, such as lightning strikes, short circuits, sudden load variations, and switching operations. These events result in abrupt changes in system behavior, leading to high-magnitude voltage and current surges over very short time intervals. Such conditions can impose significant electrical and

Ilker Ari, is with Department of Motor Vehicles and Transportation Technologies, Siirt University, Türkiye, (e-mail: ilker.ari@siirt.edu.tr).
https://orcid.org/0000-0002-1171-5533
Ramazan Menak, is with Department of Electrical Electronics Engineering, Siirt University, Türkiye, (e-mail: ramazanmenak@siirt.edu.tr).
https://orcid.org/0000-0003-3223-4808
Manuscript received Aug 09, 2025; accepted Sep 09, 2025.
DOI: 10.17694/bajece.1761735

mechanical stresses on system components, potentially threatening system reliability. Therefore, in addition to steady-state performance assessment, the transient behavior of transmission lines must be thoroughly analyzed to ensure the overall integrity and resilience of the electrical grid. In transient regime analyses of transmission lines, modeling the transmission lines is a critical step. The models used to accurately represent transmission lines vary depending on the frequency ranges they focus on and the accuracy of the analysis.

The analysis of transient regimes occurring in transmission lines is carried out in both the time and frequency domains in order to accurately reveal the dynamic responses of the system. Time domain analyses allow for the observation of the direct time-dependent effects of transient events, while frequency domain approaches are important for examining the system's responses to its frequency components. Frequency domain methods are particularly effective in determining frequency-dependent parameters, and calculations performed in this domain are often converted back to the time domain using inverse Fourier or inverse Laplace transforms. This enables the analysis of time-sensitive effects such as sudden voltage changes, reflections, and refractions in transmission lines with high accuracy through frequency-based modeling. The complementary use of these two fields enables a more robust interpretation of transient regimes, both physically and mathematically. It is also important to model the transmission line correctly in order to perform transient regime analyses. In this context, transmission lines are modeled as lumped or distributed parameters.

In [1], both lumped parameter line and distributed parameter line models were used to model the transmission line. The state-space method was used to model the lumped parameter line, while the inverse Laplace transform was used to model the distributed parameter line. In [2], a lumped parameter transmission line was modeled using the state-space method. Inductor currents and capacitor voltages were selected as state variables. The modeling was performed by adding corona loss to the line used in the study. In [3], a modified lumped parameter line model was used for transient regime analysis. In the study, a damping resistance was added to the traditional lumped parameter model to reduce numerical oscillations. The damping resistance was connected in series to the shunt capacitor in the π-segment. In [4], a study was conducted to include the frequency dependence of longitudinal parameters in the lumped parameter line model. The state-space method was used to calculate the voltage and current values of the line consisting of cascaded π circuits. EMTP was used to verify the model and method used in the study, and the results were found to be consistent with each other. In [5], transient regime analysis of non-uniform single-phase bulk parameter lines was performed using the state variable method. The state equations were obtained as a linear set of first-order differential equations and solved using MATLAB to obtain the transient response in the time domain. The results of the proposed method were validated using s-domain simulations of the distributed parameter transmission line and results obtained using the inverse Laplace transform. In [6], a single-phase transmission line was modeled in two different ways, namely as a lumped parameter model and a distributed parameter model, and the two models were compared. The results of the study indicate that the performance of the lumped parameter line model depends on the line length, number of segments, and frequency band in the simulations. Additionally, it was observed that spurious oscillations were present in the lumped parameter model, whereas no such oscillations were detected in the distributed parameter model. In [7], a method in the frequency domain was proposed for transient regime analysis in power distribution networks. The working methodology is based on modeling electrical components in the frequency domain and adding the corresponding admittance to the system matrix. Depending on the size of the system matrix, simulations can take several hours, so network reduction techniques were applied. This has simplified the system and accelerated simulation times. The Kron reduction method was used to reduce the admittance matrix. After the matrix reduction process, the numerical inverse Laplace transform was applied to obtain the system's response in the time domain. When the results of the proposed method were compared with those obtained from ATP-EMTP, it was found that the results were consistent with each other. In [8], a modified implementation of the Folded Line Equivalent (FLE) model for three-phase transmission lines in ATP-EMTP is introduced. By employing an orthogonal transformation matrix and mode decomposition via Clarke's matrix, the proposed Modified FLE (MFLE) allows bidirectional transformations using ideal transformers. Unlike traditional models derived from the method of characteristics, MFLE enables accurate simulations even with time steps larger than the line's propagation delay. Comparative results with the JMarti and Universal Line Model demonstrate that MFLE maintains accuracy while significantly reducing computational time, making it well-suited for transient analysis of short lines in large-scale networks. In [9], the Descriptor State-Space (DSE) approach was proposed for formulating circuit equations in EMT simulations. This method has enabled the automatic derivation of equations from netlist data without requiring intermediate matrix operations, thereby making it suitable for large-scale systems. Compared to the traditional Companion Circuit (CC) method, the DSE approach has allowed direct analysis of the system's eigenvalues; however, it has also been noted to require longer simulation times. In addition, a no-delay interface between DSE-based modules and CC-based solvers has been developed, providing support for modular and parallel simulation frameworks. In [10], A nonlinear system of equations has been proposed to estimate the parameters of a generic three-phase transmission line by utilizing modal-phase relations and terminal phasor measurements. The approach employs the Newton–Raphson method to solve the system and determine the transmission line parameters (TLPs) with high accuracy. The relative error analysis has confirmed the effectiveness of the method and highlighted the impact of lumped element modeling. Simulation results have shown that the technique is both robust

and efficient across lines of varying lengths. In [11], modeling techniques for transmission lines loaded with time-varying series capacitors have been explored using both the FDTD method and a block-diagram-based approach implemented in MATLAB Simulink. In the FDTD method, the incremental transmission line model has been modified to include lumped capacitors with time-varying characteristics. In parallel, the RF Blockset library in Simulink has been used to design RF circuits with variable lumped elements. The comparison of both methods has revealed consistent and promising results. Furthermore, the time-varying capacitor-loaded line has been utilized to demonstrate mixing behavior, and the influence of physical parameters on modulation performance has been analyzed. In [12], a transmission line with distributed parameters was modeled using the state-space method, and a transient regime analysis was performed. In this study, capacitor voltages and inductor currents were selected as state variables. A disc-type transformer was connected to the end of the transmission line, and the voltage distribution across the transformer windings was examined under different conditions. The proposed method was compared with ATP-EMTP, and it was found that the results were in agreement with each other.

When reviewing the literature, it is observed that transmission line modeling and transient regime analysis have been important research topics from the past to the present. Based on this point, in this study, a lumped parameter transmission line was modeled and analyzed using MATLAB/Simulink. The study consists of the following main sections: introduction, transmission lines, modeling of transmission line, results, and conclusions.

II. TRANSMISSION LINES

A. *Transmission Line Parameters*

The conductors that make up power transmission lines have specific resistance (r), inductance (l), capacitance (c), and conductance (g) values per unit length. These line parameters are not concentrated at specific points but are distributed continuously and evenly along the transmission line.

The effective resistance value of a transmission line is calculated as given in equation (1).

$$R_{AC} = \frac{P_{loss}}{I^2}\,\Omega \tag{1}$$

The effective resistance value is obtained by dividing the power loss in the conductor by the square of the current. The resistance value of the conductor against direct current is calculated as shown in equation (2).

$$R_{DC} = \rho\frac{d}{A}\,\Omega \tag{2}$$

In equation (2), A represents the cross-sectional area of the conductor, d represents the length of the conductor, and ρ represents the resistivity of the conductor.

The inductance of a transmission line depends on the type of material used, the structure of the conductors, and the relative positions of the phase conductors. The inductance of one phase of the transmission line is calculated as follows.

$$l = 2x10^{-7}\left(\ln\frac{D}{r'}\right)[H/m] \tag{3}$$

In Equation (3), D and r' represent the distance between conductors and the effective radius of the conductor, respectively. If three phase conductors are placed in an equilateral triangle on a pole, no transposition is required, and the geometric mean distance (GMD) is equal to the distance between the phases. However, if the phase conductors are not arranged in an equilateral triangle, transposition is applied for lines longer than 50 km. After calculating the GMD using Equation (4), the inductance of each phase of the bundle conductor line is calculated using Equation (5). The GMR given in the equations represents the geometric mean radius of the conductors.

$$GMD = \sqrt[3]{D_{ab}D_{bc}D_{ac}} \tag{4}$$

$$l = 2x10^{-7}\left(\ln\frac{GMD}{GMR}\right)[H/m] \tag{5}$$

The unit capacitance per phase in three-phase lines with transposed bundle conductors is calculated as follows.

$$c = \frac{2\pi\varepsilon_0}{\ln\frac{GMD}{GMR}}[F/m] \tag{6}$$

Transmission lines have a certain amount of leakage conductivity. This leakage conductivity represents the power loss between conductors and between the conductor and the ground. Leakage current and corona in insulators are some of the causes of these power losses. The value of leakage conductivity varies depending on the insulating material used, frequency, and atmospheric conditions [13].

B. *Lumped Parameter Transmission Lines*

When modeling lumped-parameter transmission lines, they are divided into numerous sections. Each section consists of its own total series resistance (R), series inductance (L), shunt capacitance (C), and shunt conductance (G) parameters [1]. The sections are identical and are modeled as cascades. Transmission lines of this type can be modeled in four different ways: T, Π, Γ, and Ꞁ, as shown in Fig. 1.

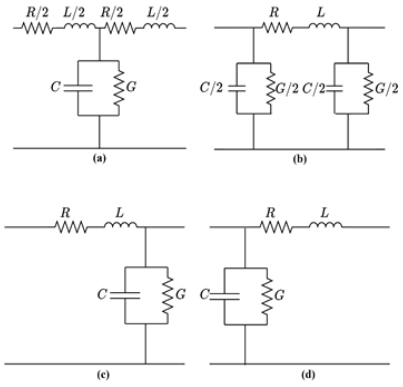


Fig. 1. Lump parameter line models (a) T-circuit (b) Π-circuit (c) Ꞁ-circuit (d) Γ-circui

Let us assume that a transmission line with lumped parameters consists of *N* sections, each of which consists of a Π-type circuit. In this case, the resulting circuit model is shown in Fig. 2.

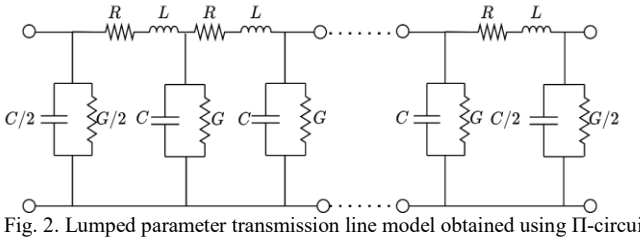


Fig. 2. Lumped parameter transmission line model obtained using Π-circuit

When modeling transmission lines with lumped parameters, the *R, L, C*, and *G* values for each section are obtained by dividing the total resistance, inductance, capacitance, and conductivity values of the line by the number of sections, or in other words, by dividing them by the number *N*. Therefore, the line parameter value in each section will change depending on the change in the number *N*.

Transmission lines are divided into three categories based on their length: short, medium, and long lines. Transmission lines with a length between 0 and 80 km are referred to as short transmission lines. When modeling this type of short transmission line, the shunt capacitance is neglected, and the line is modeled using only the series impedance value consisting of resistance and inductance. This model is generally used to represent electrical distribution networks. The two-port equivalent circuit of the short transmission line is shown in Fig. 3.

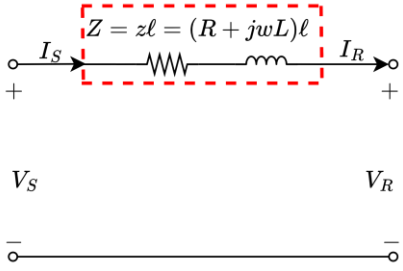


Fig. 3. Equivalent circuit of a short transmission line model

V_s, V_r, I_s, and I_r in the Fig. 3. represent the sending end voltage, receiving end voltage, sending end current, and receiving end current of the line, respectively. The equation for the given line model is as follows.

$$\begin{bmatrix} V_S \\ I_S \end{bmatrix} = \begin{bmatrix} 1 & Z \\ 0 & 1 \end{bmatrix} \begin{bmatrix} V_R \\ I_R \end{bmatrix} \tag{7}$$

The ABCD parameters of the short transmission line are as follows.

$$\begin{aligned} A &= D = 1 \\ B &= Z \\ C &= 0 \end{aligned} \tag{8}$$

Transmission lines with a length of more than 80 km and less than 250 km are referred to as medium-length transmission lines. As the length of the transmission line increases, the line charging current becomes significant, and shunt capacitance is taken into account. Therefore, a medium-length transmission line is modeled using the total shunt admittance and total impedance connected in series to the circuit. Medium-length transmission lines with lumped parameters are modeled in two ways: the Nominal Π-circuit model and the Nominal T-circuit model. In the nominal Π-circuit representation, the total series impedance is placed in the middle of the line, while the shunt admittance is divided into two equal parts and placed at both ends of the line. The equivalent circuit of the Π-circuit model is shown in Fig. 4.

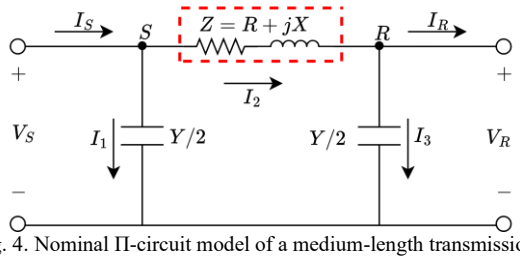


Fig. 4. Nominal Π-circuit model of a medium-length transmission line

The equations obtained by applying Kirchhoff's current and voltage laws to the equivalent circuit shown in Fig. 4. are given below.

$$V_S = ZI_2 + V_R = Z\left(V_R\frac{Y}{2} + I_R\right) + V_R = \left(\frac{YZ}{2} + 1\right)V_R + ZI_R \tag{9}$$

$$I_S = I_1 + I_2 = I_1 + I_3 + I_R = \frac{Y}{2}V_S + \frac{Y}{2}V_R + I_R \tag{10}$$

Substituting equation (9) into equation (10), equation (10) is rearranged to become equation (11).

$$I_S = \frac{Y}{2}\left[\left(\frac{YZ}{2} + 1\right)V_R + ZI_R\right] + \frac{Y}{2}V_R + I_R = Y\left(\frac{YZ}{4} + 1\right)V_R + \left(\frac{YZ}{2} + 1\right)I_R \tag{11}$$

The ABCD parameters of the nominal Π-circuit are written as shown below.

$$\begin{aligned} A = D &= \left(\frac{YZ}{2} + 1\right) \\ B &= Z \\ C &= \left(\frac{YZ}{4} + 1\right) \end{aligned} \tag{12}$$

Transmission lines that are 250 km or longer are called long transmission lines. Long transmission lines are generally modeled as distributed parameters, but they can also be modeled using Π equivalent circuits. For the equivalent Π circuit of the long transmission line model, Z' and Y' are used instead of Z and Y in the nominal Π circuit representation. A visual representation of this model is provided in Fig. 5.

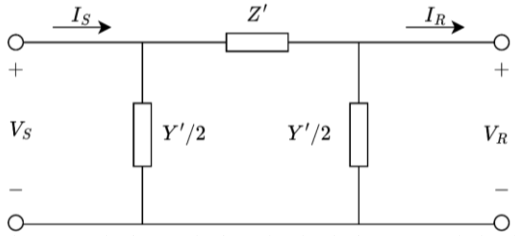


Fig. 5. Nominal Π-equivalent circuit of a long transmission line

The ABCD parameters of the Π-circuit model of the long transmission line are as follows:

$$A = D = 1 + \frac{Y'Z'}{2}$$
$$B = Z' \tag{13}$$
$$C = Y'\left(1 + \frac{Y'Z'}{4}\right)$$

Here;

$$Z' = Z_C \sinh(\gamma d) = Z\frac{\sinh(\gamma d)}{\gamma d}$$
$$\frac{Y'}{2} = \frac{1}{Z_C}\tanh(\gamma d/2) = \frac{Y}{2}\frac{\tanh(\gamma d/2)}{(\gamma d/2)} \tag{14}$$

In the equations given above, z and y are unit length parameters, while $Z=zd$ and $Y=yd$.

C. Reflection and Refraction in Traveling Waves

When a traveling wave on a transmission line encounters an impedance discontinuity at a specific point, a portion of the wave is reflected while the remainder continues to propagate into the new medium. Part of the wave remains in the original medium, whereas the other part advances through the region with different electrical characteristics. The phenomenon of wave refraction arises due to variations in material properties as the wave transitions from one segment of the line to another. If multiple impedance regions are present along the line, the wave may undergo changes in direction and exhibit variations in amplitude. The equations governing wave reflection and refraction are presented below.

$$\Gamma = \frac{Z_L - Z_C}{Z_L + Z_C} \tag{15}$$

$$T = \frac{2Z_L}{Z_L + Z_C} \tag{16}$$

Here;
Γ: Reflection coefficient,
T: Transmission coefficient,
Z_L: Load impedance at the end of the transmission line.
Z_C: Characteristic impedance of the transmission line.

The impedance observed at the input of an infinitely long transmission line is referred to as the characteristic impedance. This impedance is inherently determined by the physical structure of the line, and these physical properties can be expressed in terms of electrical quantities. By taking the line parameters into account, the characteristic impedance can be accurately calculated. According to equations (15) and (16), if $Z_L=Z_C$, the entire wave is transmitted and no reflection occurs. However, if $Z_L \neq Z_C$, a reflection occurs and part of the wave is returned. If there is an open circuit or short circuit at the end of a transmission line, total reflection occurs. In the case of an open-circuit termination ($Z_L \rightarrow \infty$), the reflected voltage wave has a positive polarity ($\Gamma=+1$), whereas in the case of a short-circuit termination ($Z_L=0$), the reflected voltage wave exhibits a negative polarity ($\Gamma= -1$). To analyze wave reflection and transmission phenomena, different methods can be employed in both the time and frequency domains. In the time domain, the instantaneous variations of voltage and current waves can be observed directly. In contrast, the frequency domain analysis, typically conducted using the Fourier transform, allows for examination of how the wave is decomposed into its constituent frequency components. Wave propagation and reflections can also be examined graphically. Using the Bewley Lattice Diagram method, the behavior of wave packets at specific locations along the line can be tracked by employing a time-position coordinate system. This graphical technique provides a clear visualization of how incident and reflected waves interact at impedance discontinuities over time.

III. MODELING OF TRANSMISSON LINE

Various methods and simulation programs are used for transmission line modeling and transient regime analysis. Programs such as PSS®E (Power System Simulator for Engineering), RTDS (Real-Time Digital Simulator), PSCAD (Power Systems Computer-Aided Design), and ATP-EMTP (Alternative Transients Program–Electromagnetic Transients Program) are some of the programs used for simulation and analysis. In addition, modeling can also be performed using the state-space method. In this study, the MATLAB/Simulink program is used for modeling transmission lines with lumped parameters.

In this study, a single-phase lumped parameter transmission line was used, and an R-L load was connected to the end of the line. The equivalent circuit of the transmission line used for modeling is shown in Fig. 6.

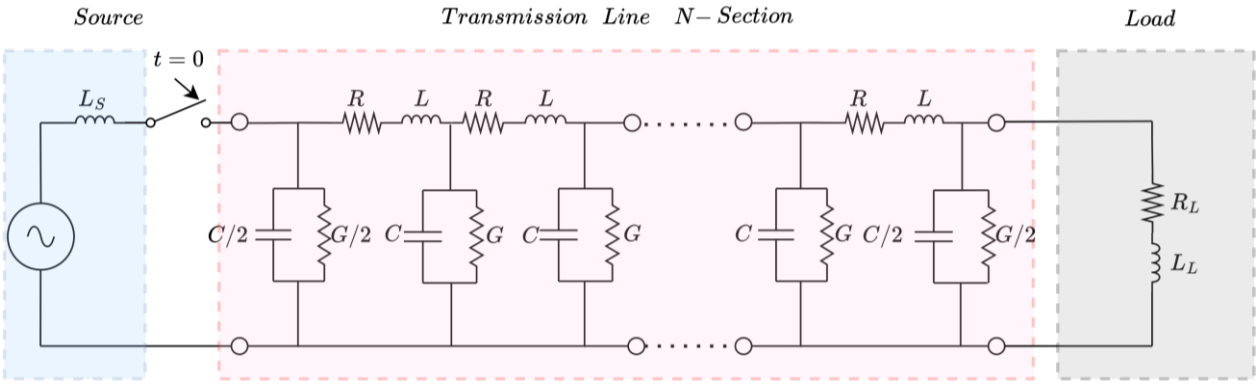


Fig. 6. Equivalent circuit of transmission lines with lumped parameters

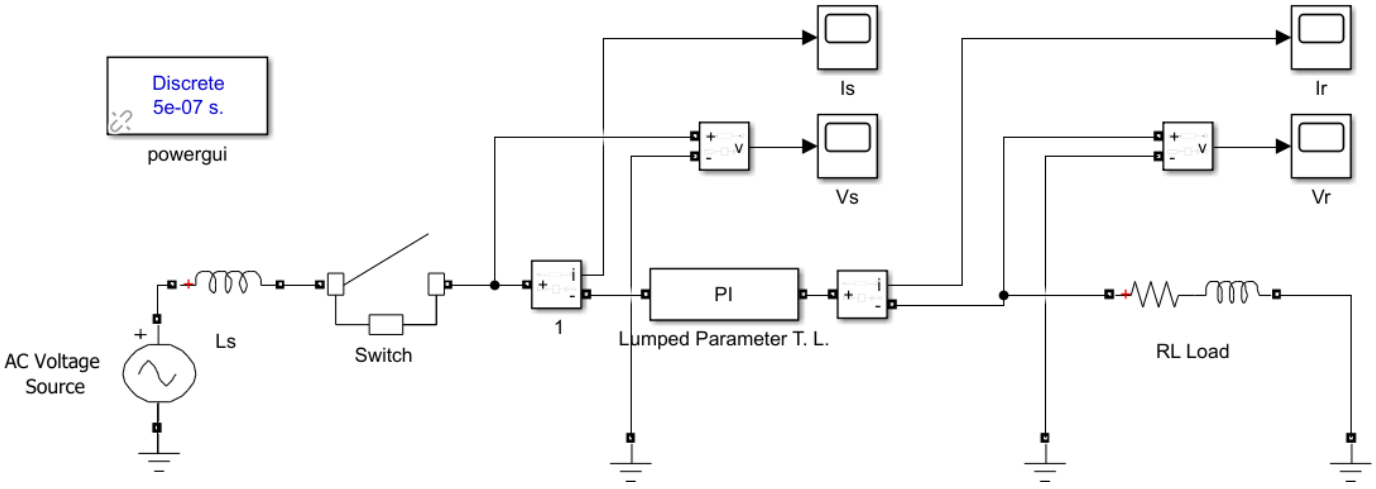


Fig. 7. MATLAB/Simulink model of a lumped parameter line

As can be seen in Fig. 6, the circuit used in the model consists of a cascade connection of multiple π-circuits. Each π-circuit represents one section. The number of sections is expressed by *N*. The L_S, R_L, and L_L in the circuit represent the source inductance, load resistance, and load inductance, respectively. An AC source was used as the voltage source. The equivalent circuit used in the study, modeled in MATLAB/Simulink, is shown in Fig. 7. According to the circuit, at *t=0 s*, the switch closes and the transmission line is energized. The parameters used in the model are given in Table I.

TABLE I
MODEL PARAMETERS

d	*180 km*
r	*0.045 Ω/km*
l	*0.76 mH/km*
c	*11 nF/km*
R_L	*1.1 k Ω*
L_L	*0.45 H*
L_S	*50 mH*
Source	*154 kV AC*
f	*50 Hz*
N	*1, 5, 20, 50*

In Table I, the parameters *d, r, l*, and *c* represent the length of the transmission line, series resistance, series inductance and shunt capacitance, respectively, while *f* represents the frequency and *N* represents the number of sections in the line. The step length (Δx) is 50 μs and The Tustin/Backward Euler (TBE) solver type has been used for the simulation.

When modeling in MATLAB/Simulink, there is no need to draw separate sections for each pi circuit. Modeling can be performed by entering the desired number of sections in the block parameter section. This provides an advantage when compared to a generally accepted transient regime simulation program such as ATP-EMTP. This is because in the ATP-EMTP program, when modeling transmission lines with lumped parameters, sections must be added manually for each π-circuit. The number of sections entered into the block parameters and the number of sections represented in the ATP-EMTP program are shown in Fig.8., Fig. 9. and Fig. 10.

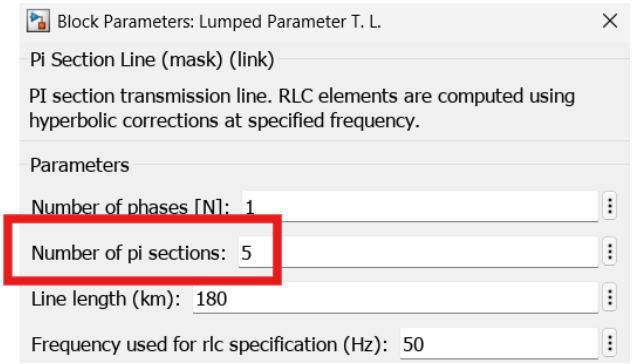


Fig. 8. PI block parameters in MATLAB/Simulink

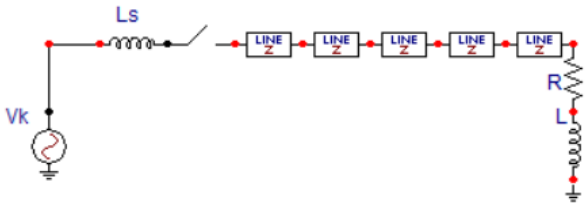


Fig. 9. ATP-EMTP model of a transmission line consisting of 5 sections

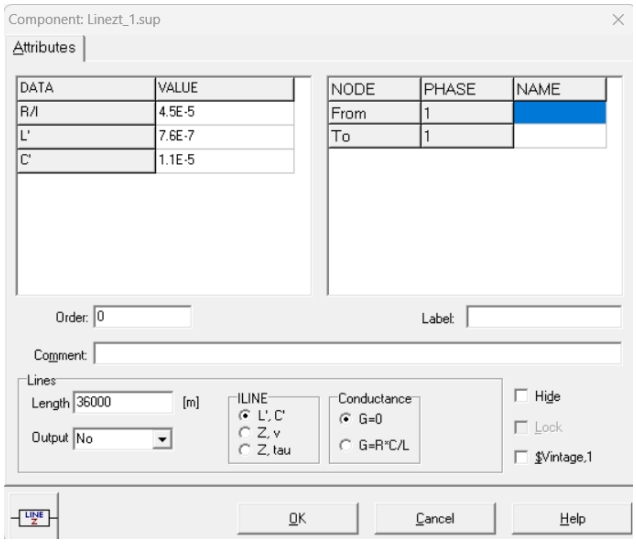


Fig. 10. Block parameters in ATP-EMTP

When examining Fig. 9. and Fig. 10., it can be seen that five π-circuits must be manually added to the circuit to model a line consisting of five sections, and that the line length for each π-circuit must be entered according to the *d/N* formula. Here, *d* is the total line length and *N* is the number of π-circuits. For a circuit consisting of 180 km and 5 sections, the line length should be entered as 180/5, which is 36 km. In this study, simulations were performed for different *N* values. Therefore, the use of MATLAB/Simulink is more advantageous for large *N* values.

IV. RESULTS

Transient simulations were performed using the lumped parameter transmission line model shown in Fig. 7. At *t=0*, the switch closes and the transmission line is energized. The sending end voltage (V_S) profile of the lumped parameter transmission line is shown in Fig. 11.

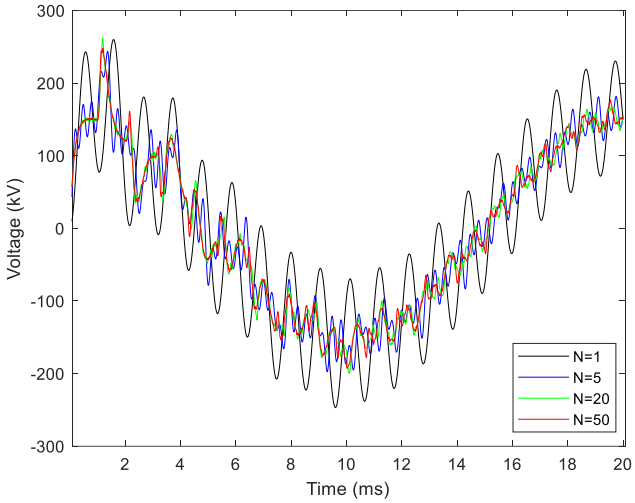


Fig. 11. Sending-end voltage waveforms for different N values

When examining Fig. 11., it is observed that different waveforms are produced for different N values. For N=1, the oscillation frequency in the wave is lower and the peak voltage value is higher than other N values. Here, it can be seen that as the N value increases, the voltage amplitude in the waveforms decreases. Despite using a voltage source of 154 kV, it was found that the sending-end voltage reached approximately 260 kV within the first few milliseconds after energization.

The voltage profile at the receiving-end (V_R) of the line is shown in Fig. 12. When the voltage-time profile at the receiving end of the line was examined, it was found that, unlike the voltage profile at the sending-end, the voltage wave reached the end of the line with a certain delay. This delay is related to the line length, wave propagation speed, and number of sections. The wave propagation speed depends on the line parameters. When examining waveforms, the time it takes for the wave to reach the end of the line varies as the N value changes. The voltage wave reached the end of the line after 0.11 ms for N=1, 0.31 ms for N=5, 0.44 ms for N=20, and 0.49 ms for N=50. The receiving-end voltage has taken values between 300 kV and 264 kV up to 2 ms for different N values. In the following period, these peak values have shown a decreasing trend.

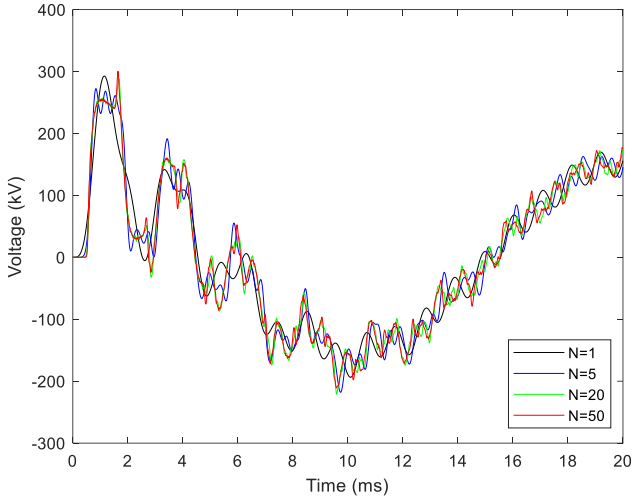


Fig. 12. Receiving-end voltage waveforms for different N values

The current (I_s) profile at the sending-end of the line is shown in Fig. 13.

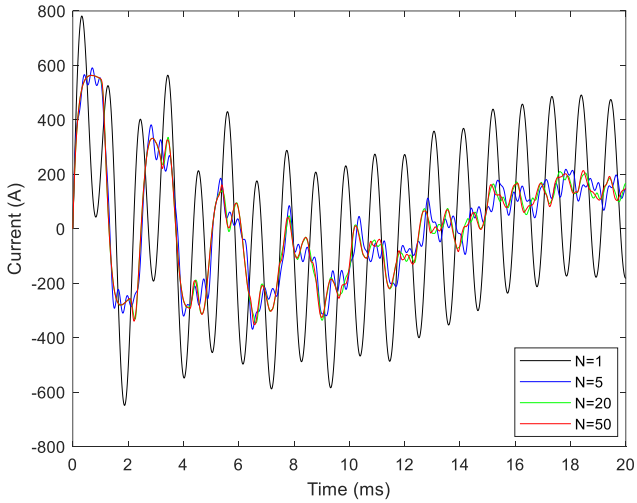


Fig. 13. Sending-end current waveforms for different N values

When the sending-end current waveform was examined, it was determined that peak current values were reached between approximately 500 A and 780 A for different N values. Over time, these peak values decreased significantly, especially for N=5, 20, and 50.

The receiving-end current (I_R) profile of the transmission line is shown in Fig. 14. When examining the receiving-end current waveforms, it was determined that the wave reached the receiving end of the line with a certain delay. After energization, the current waveform suddenly reached peak values of 204 A to 214 A after the delay for all N values. Here, the current waveforms for N=1 and N=50 are more similar to each other compared to the sending-end current profile. Current waves remain unstable for approximately the first 10 ms, but after 10 ms, the waves attempt to stabilize. Stabilization will occur over time.

Fig. 15. shows how the sending-end voltages attempt to steady state.

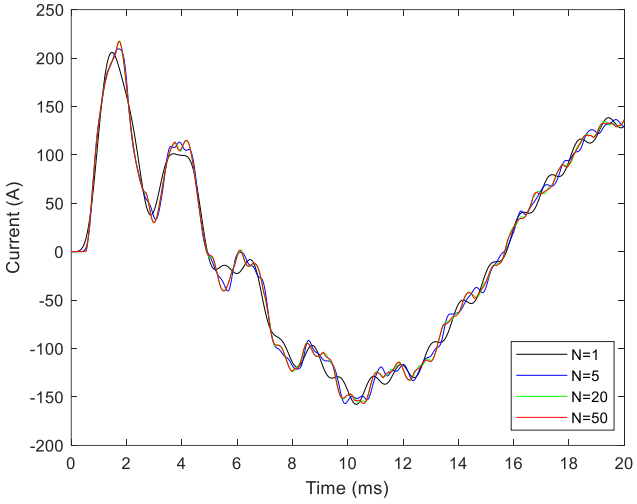


Fig. 14. Receiving-end current waveforms for different N values

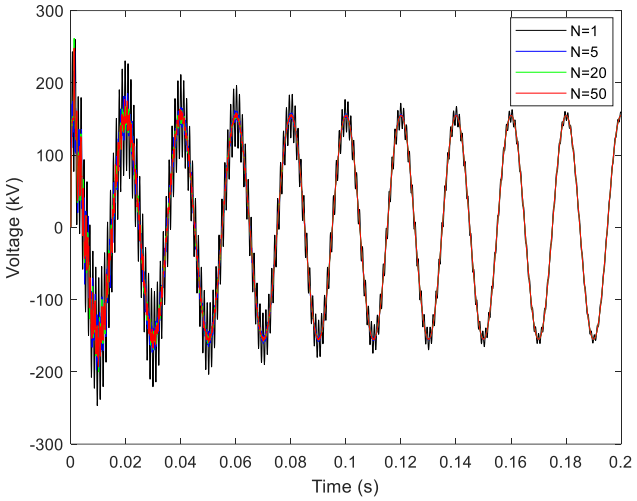


Fig. 15. Sending-end voltage waveforms for steady state

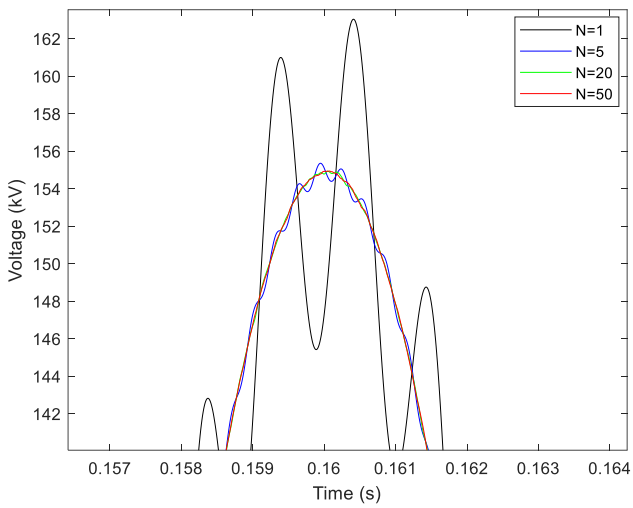


Fig. 16. Closer view of sending-end voltage waveforms for steady state

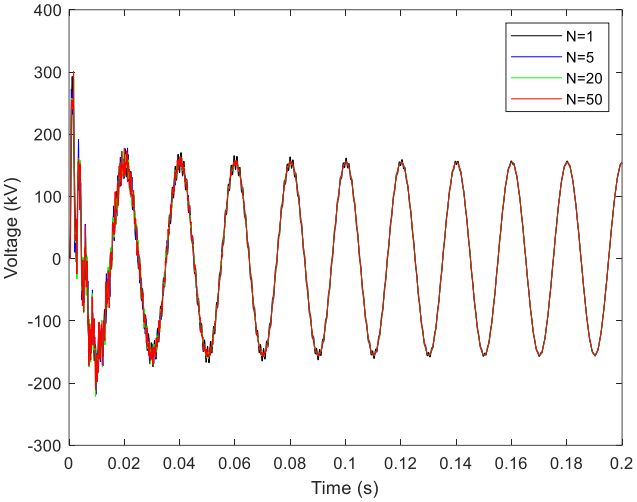


Fig. 17. Receiving-end voltage waveforms for steady state

When Fig. 15. is examined, it can be seen that all N values transitioned to a steady state at different times. The simulation time was set to 0.2 s in order to observe the steady state. Simulations were performed at the ms level for the analysis of transient states.

A closer view of the steady-state voltages at the sending-end is shown in Fig. 16. When the figure is examined, it is observed that the voltage waveform at the sending-end for N=50 reaches a steady state approximately 0.16 s later. The voltage level at this point is approximately 154 kV. At 0.16 s, the voltage waveform for N=1 has not yet reached a steady state and continues to oscillate at high amplitude. When examining the voltage waveforms for N=5 and N=20, it was observed that they had not yet transitioned to a steady state but were very close to it. The oscillations here, however, have lower amplitudes.

Fig. 17. shows how the receiving-end voltages attempt to steady state. When examining the steady-state form of the receiving-end voltage waveforms, oscillations were observed in the first milliseconds, with very high peaks detected. After approximately 0.16 seconds, it was determined that N=50 had reached a steady state, while other N values were close to a steady state.

When both the sending-end and receiving-end waveforms are examined, it is observed that oscillations occur along the line. The reason for these oscillations is that the line exhibits an LC filter effect. The delay time of the waves in the graphs, on the other hand, is related to the wave propagation speed and the length of the line.

V. CONCLUSION

In this study, a single-phase 154 kV lumped-parameter transmission line with an RL load was modeled and energized at t = 0 s. The advantages of the modeling approach used in this work over ATP-EMTP were highlighted. Voltage and current waveforms at both the sending and receiving ends of the line were analyzed. It was observed that oscillations occurred in the initial time interval due to energization, and that the waveforms approached a steady-state condition approximately 0.16 s after energization. Owing to the inherent characteristics of the line, the voltage and current waves reached the receiving end with a certain propagation delay. The lumped-parameter model of the transmission line inherently consists of multiple π-sections, each representing one segment of the line. To investigate the

effect of the number of sections on transient analysis, four different values of N were employed. The results indicate that increasing the number of sections yields more accurate results, producing waveforms that more closely match the expected physical behavior. MATLAB/Simulink provides significant convenience in modeling transmission lines and in the design of power systems. The program allows for the analysis of various load conditions and delayed switching operations. Future research directions may include extending the lumped-parameter transmission line model to a three-phase configuration, simulating the impact of lightning impulses on the line, integrating a transformer at the receiving end, investigating short-circuit faults at specific points along the line, and performing comparative studies with distributed-parameter line models.

REFERENCES

[1] M. Koksal, M. S. Mamis, "Remark on the Lumped Parameter Modeling of Transmission Lines." Electr. Mach. Power Syst., vol. 28, no. 6, pp. 565–575, Jun. 2000, doi: 10.1080/073135600268180.

[2] M. S. Mamis, "State-Space Transient Analysis of Single-Phase Transmission Lines with Corona," Proc. Int. Conf. Power Syst. Transients 2003 (IPST 2003), 2003.

[3] J. S. L. Colqui, A. R. J. De Araújo, and S. Kurokawa, "Improving the performance of a lumped transmission line model used in electromagnetic transient analysis," IET Gener. Transm. Distrib., vol. 13, no. 21, pp. 4942–4951, Nov. 2019, doi: 10.1049/iet-gtd.2018.6301.

[4] S. Kurokawa, F. N. R. Yamanaka, A. J. Prado, and J. Pissolato, "Inclusion of the frequency effect in the lumped parameters transmission line model: State space formulation," Electr. Power Syst. Res., vol. 79, no. 7, pp. 1155–1163, Jul. 2009, doi: 10.1016/j.epsr.2009.02.007.

[5] M. S. Mamis, "Lumped-parameter-based electromagnetic transients simulation of non-uniform singlephase lines using state variable method," IET Gener. Transm. Distrib., vol. 14, no. 23, pp. 5626–5633, Dec. 2020, doi: 10.1049/iet-gtd.2020.0454.

[6] A. R. J. Araújo, R. C. Silva, and S. Kurokawa, "Comparing Lumped and Distributed Parameters Models in," in 2014 Ieee Pes T&d Conference And Exposition, 2014, pp. 1–5.

[7] G. Bilal, P. Gomez, R. Salcedo, and J. M. Villanueva-Ramirez, "Electromagnetic transient studies of large distribution systems using frequency domain modeling methods and network reduction techniques," Int. J. Electr. Power Energy Syst., vol. 110, no. January, pp. 11–20, Sep. 2019, doi: 10.1016/j.ijepes.2019.02.043.

[8] J. S. L. Colqui, L. C. Timaná, P. T. Caballero, S. Kurokawa, and J. P. Filho, "A modified implementation of the Folded Line Equivalent transmission line model in the Alternative Transient Program," Electr. Power Syst. Res., vol. 211, no. November 2021, p. 108185, Oct. 2022, doi: 10.1016/j.epsr.2022.108185.

[9] A. Sinkar, H. Zhao, B. Qu, and A. M. Gole, "A Comparative Study of Electromagnetic Transient Simulations using Companion Circuits, and Descriptor State-space Equations," Electr. Power Syst. Res., vol. 198, no. April, p. 107360, 2021, doi: 10.1016/j.epsr.2021.107360.

[10] J. T. R. Pineda and S. Kurokawa, "Application of modal-phase domain relations to estimate transmission line parameters," IET Sci. Meas. Technol., vol. 13, no. 5, pp. 684–691, Jul. 2019, doi: 10.1049/iet-smt.2018.5428.

[11] A. Kumar, J. C. Dash, and D. Sarkar, "Computational Techniques for Design and Analysis of Time-Varying Capacitor Loaded Transmission Lines Using FDTD and Simulink," IEEE J. Multiscale Multiphysics Comput. Tech., vol. 7, pp. 228–235, 2022, doi: 10.1109/JMMCT.2022.3202990.

[12] İ. Arı and M. S. Mamiş, "Voltage Distribution on Transformer Windings Subjected to Lightning Strike Using State-Space Method," Appl. Sci., vol. 15, no. 3, p. 1569, Feb. 2025, doi: 10.3390/app15031569.

[13] D. Akmaz, "İletim Hatlarinda Ariza Yeri Ve Türünün Geçici Rejim Sinyalleri Ve Makine Öğrenme Algoritmalari İle Belirlenmesi," Ph.D. Dissertation, İnönü Universtiy, 2017.

BIOGRAPHIES

Dr. Ilker ARI was born in Siirt, Türkiye, in 1989. He received the B.Sc. degree in Electrical and Electronics Engineering from Çukurova University, Adana, Türkiye, in 2014, the M.Sc. degree in Electrical and Electronics Engineering from Siirt University, Siirt, Türkiye, in 2018 and Ph.D. degree in Electrical and Electronics Engineering from İnönü University, Malatya, Türkiye, in 2025. Since 2020, he has been a Lecturer with the Department of Motor Vehicles and Transportation Technologies, Siirt University. His research interests include power systems, line modeling and analysis.

Dr. Ramazan MENAK was born in Siirt, Türkiye, in 1993. He received the B.Sc. degree in Electrical and Electronics Engineering from İnönü University, Malatya, Türkiye, in 2015, the M.Sc. degree in Electrical and Electronics Engineering from Siirt University, Siirt, in 2018 and Ph.D. degree in Electrical and Electronics Engineering from İnönü University, Malatya, Türkiye, in 2025. Since 2017, he has been a Research Assistant with the Department of Electrical and Electronics Engineering, Siirt University. His research interests include control system design, robust controller design, and H-∞ norm based controller design.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 3, September 2025
Research Article

Detection of Bearing Faults from Vibration Signals

Eyyup Akcan

***Abstract*— Bearings are critical mechanical components in rotating machinery, playing a vital role in system safety and operational continuity. In this study, the Case Western Reserve University (CWRU) bearing dataset is used to perform fault classification using four machine learning algorithms: Random Forest, XGBoost, Support Vector Machine (SVM), and Naive Bayes. Based on statistical features extracted in the time domain, the performance of each model is evaluated using accuracy, precision, recall, and F1-score metrics. The results reveal that Random Forest and XGBoost algorithms achieved superior performance with 95.73% accuracy and 96% in precision, recall, and F1-score. The SVM model, with 93.73% accuracy, stands out as a robust alternative, while the Naive Bayes algorithm shows relatively lower performance with 92.40% accuracy. Additionally, an individual feature-based classification analysis indicates that standard deviation (sd) and root mean square (RMS) features contribute most significantly to model performance. This study provides a comprehensive performance analysis of traditional machine learning algorithms, offering a valuable reference for early and accurate detection of bearing faults.**

Index Terms— **Bearing fault diagnosis, Machine learning, Random Forest, XGBoost, Support Vector Machine, Naive Bayes, CWRU dataset.**

I. Introduction

BEARINGS are among the most fundamental and mission-critical components in rotating machinery, playing a pivotal role in ensuring mechanical performance, operational reliability, and safety.

Structurally, a typical bearing comprises an inner race, an outer race, rolling elements (usually balls or rollers) positioned between these races, and a cage that maintains the spacing and alignment of the rolling elements. This configuration enables the bearing to accommodate both radial and axial loads while minimizing friction, thereby contributing significantly to the energy efficiency of mechanical systems. Bearings are ubiquitous in various industrial domains, including aerospace, automotive, wind energy, railway transportation, industrial drives, and automated production systems. In such applications, they are often subjected to harsh operational environments characterized by variable loads, elevated temperatures, and persistent vibrations. These demanding conditions, over extended periods, give rise to a variety of structural degradations such as surface fatigue, micro-cracking, plastic deformation, pitting, and spalling. If left undetected, such defects can evolve rapidly and pose serious risks to the mechanical integrity and safety of the overall system. Empirical studies report that nearly 40% of failures in rotating machinery can be attributed to bearing faults. Despite their relatively low cost compared to other mechanical subsystems, bearings possess a strategically critical function. A single bearing malfunction has the potential to cause collateral damage to adjoining components such as shafts, gear trains, and motor assemblies, often resulting in costly production downtimes and substantial financial losses. Given these risks, real-time monitoring and early fault detection in bearings are indispensable for preventing unplanned outages and optimizing maintenance scheduling. Timely diagnosis of incipient faults not only mitigates the likelihood of catastrophic equipment failure but also reduces maintenance expenditures, enhances operational efficiency, and improves the overall reliability of the system [1-3].

Over time, bearings undergo deterioration as a result of prolonged exposure to adverse operating conditions, including excessive mechanical loads, irregular torque fluctuations, inadequate lubrication, and environmental contaminants such as dust and moisture. Additionally, sudden mechanical shocks can accelerate this degradation process. Among the most prevalent types of bearing failures are defects in the inner race, cracks in the outer race, surface wear or pitting on rolling elements, and deformation of the cage. As these faults evolve, they lead to increased vibrational and acoustic emissions, reduced mechanical efficiency, and compromised system stability. Ultimately, undetected or untreated bearing damage can culminate in critical equipment failures, unplanned production interruptions, and considerable economic consequences [4].

Rotating machinery is extensively utilized across a range of industrial domains, including aerospace, automotive, energy production, and transportation systems. Within these machines, rolling bearings serve as crucial mechanical elements due to their role in ensuring smooth and reliable operation. Studies have revealed that nearly 40% of mechanical malfunctions stem from bearing-related faults, a figure that may escalate to 90% in smaller-scale equipment. Such malfunctions can pose serious safety hazards, disrupt operations, and lead to substantial financial repercussions. Consequently, there is an increasing demand for reliable and automated bearing fault diagnosis systems. In this context, machine learning approaches have emerged as powerful

Eyyüp Akcan, is with Department of Motor Vehicles and Transportation Technologies, Siirt, Turkey,(e-mail: renaz82@gmail.com).
https://orcid.org/0000-0002-4133-4344

Manuscript received Aug 02, 2025; accepted Aug 22, 2025.
DOI: 10.17694/bajece.1757057

tools for fault detection, isolation, and classification, offering the advantage of minimizing reliance on domain-specific expertise or manual supervision. Traditional techniques such as artificial neural networks (ANN), support vector machines (SVM), k-nearest neighbors (k-NN), and principal component analysis (PCA) have been widely adopted in this domain. However, these methods often involve distinct steps for feature extraction and classification, where informative features must be manually derived from raw sensor signals. Moreover, their performance tends to decline when faced with large-scale or complex datasets, indicating limited scalability and generalization capacity. In light of these challenges, deep learning methodologies have garnered attention for their capability to autonomously learn discriminative representations from high-dimensional data while simultaneously performing classification tasks [5].

Rotating machinery is extensively utilized in modern industrial applications such as compressors, pumps, conveyor systems, and electric motors. Among the essential components in these machines, rolling bearings are pivotal in maintaining the system's reliability and operational integrity. Nevertheless, prolonged operational durations and exposure to severe environmental conditions render these components susceptible to degradation and eventual failure. As such, the development of advanced diagnostic methodologies for the early detection of bearing faults is vital to ensure equipment reliability, mitigate unplanned downtimes, and minimize economic losses. A variety of sensing techniques have been implemented in recent years for condition monitoring and fault diagnosis of rolling bearings. These include, but are not limited to, vibration analysis, acoustic emission detection, sound signature analysis, thermal imaging, and motor current signal analysis. Of these, vibration analysis remains the most prevalent and effective diagnostic approach due to its sensitivity to mechanical anomalies. It is particularly effective in identifying four principal fault categories: defects in the outer race, inner race, rolling elements, and cage structures. The presence of a fault typically induces repetitive transient pulses within the vibration signal, providing diagnostic cues for fault detection and classification [6].

Contemporary industrial machinery frequently operates in dynamic and often harsh environments, characterized by structural complexity and variable load conditions. These factors elevate the risk of minor defects escalating into systemic failures, thereby compromising equipment reliability and posing potential safety hazards. To mitigate such risks, accurate estimation of the remaining useful life (RUL) and reliable fault diagnosis of critical components—such as bearings—are essential. The recent progress in machine learning has facilitated the development of advanced diagnostic frameworks capable of achieving high accuracy in fault identification. In contrast to conventional methods, intelligent fault diagnosis systems offer automated feature extraction and integrated classification capabilities, reducing dependency on manual preprocessing. The effectiveness of these systems is influenced not only by the robustness of the algorithmic models but also by the quality and structure of the datasets employed. Despite widespread reliance on balanced datasets in experimental studies, this assumption rarely holds true in industrial environments, where data imbalance is a common issue—particularly due to the scarcity of fault-related data compared to normal operating conditions [7].

Conventional approaches to bearing fault detection predominantly rely on time-frequency domain analyses and handcrafted feature extraction. Methods such as Fast Fourier Transform (FFT), Wavelet Transform (WT), and Empirical Mode Decomposition (EMD) have been widely utilized for signal interpretation in diagnostic applications. While these techniques have demonstrated effectiveness in controlled scenarios, they often exhibit sensitivity to noise and require significant expert knowledge for parameter tuning and feature design. Such dependencies restrict their applicability in complex, real-world environments and hinder automation. As a result, there has been a paradigm shift toward data-driven techniques, particularly those based on machine learning (ML) and deep learning (DL), which offer more robust and scalable solutions by enabling end-to-end learning directly from raw or minimally processed sensor data [4].

Timely identification of bearing faults is essential to enhance system reliability and reduce operational maintenance expenses. Conventional diagnostic techniques primarily utilize signal processing approaches in the time, frequency, and time-frequency domains. Commonly applied methods include wavelet transform, short-time Fourier transform (STFT), empirical mode decomposition (EMD), variational mode decomposition (VMD), and continuous wavelet transform (CWT), which aim to reveal hidden fault characteristics within vibration signals. Despite their widespread use, these methods are often constrained by their dependency on manually selected parameters and prior domain expertise, which can limit adaptability and reduce performance under varying operating conditions [8,9].

In recent years, to address the limitations of conventional approaches, the adoption of machine learning and deep learning techniques in bearing fault diagnosis has gained significant momentum. While classical machine learning algorithms such as support vector machines (SVM), decision trees (DT), k-nearest neighbors (k-NN), and random forests (RF) continue to be utilized, more sophisticated deep learning architectures have demonstrated remarkable performance. These include convolutional neural networks (CNN), long short-term memory networks (LSTM), spiking neural networks (SNN), autoencoders, generative adversarial networks (GAN), and transformer-based models. Their ability to autonomously extract hierarchical features from raw data and adapt to complex fault patterns makes them powerful tools in modern fault diagnosis systems [10,11].

In conclusion, bearing fault diagnosis is of critical importance in modern industry. Studies conducted using the CWRU dataset have shown that machine learning and deep learning methods can be effectively applied in this domain. In this context, the present study utilizes the CWRU dataset

to extract time-domain features and comparatively analyze the classification performance of different machine learning algorithms (RF, SVM, XGBoost, NB).

II. LITERATURE REVIEW

In this section, the literature on bearing fault diagnosis is summarized.

Li et al. (2024) developed a novel deep learning framework to improve the accuracy of fault identification in essential mechanical systems. Addressing the deficiencies of conventional methods in extracting fault-related information from rotating machinery, the study introduced an attention-augmented Convolutional Neural Network (CNN) architecture, termed Attention Improved CNN (AT-ICNN). This architecture combines an enhanced convolutional module (IMConv) with a hybrid attention mechanism, allowing for the effective extraction of both local and global signal characteristics and emphasizing critical fault features. Consequently, the proposed model demonstrates superior fault classification capabilities. The AT-ICNN was validated on the widely-used Case Western Reserve University (CWRU) bearing dataset and an in-house laboratory dataset. It achieved classification accuracies of 98.12% and 98.72%, respectively, surpassing traditional and state-of-the-art methods by nearly 9%. These outcomes highlight the model's potential as a reliable and high-performance solution for fault diagnosis in industrial mechanical components [12].

Borghesani et al. (2023) sought to advance the transparent and theoretically grounded development of neural networks (NNs) specifically within the domain of Machine Condition Monitoring (MCM). They highlighted the structural parallels between NN layers and classical signal processing operations—such as filtering, decimation, and envelope extraction—and conducted a rigorous mathematical analysis of these similarities. Given that MCM frequently deals with signals characterized by periodic and cyclostationary behavior, the authors employed a Fourier-based framework to elucidate the influence of NN layers on such signal types. This analysis provided a foundation for establishing design heuristics and parameter optimization strategies tailored to MCM tasks, drawing upon long-standing insights from signal processing literature. The research aligns with contemporary movements in Explainable Artificial Intelligence (XAI) and Physics-Informed Neural Networks (PINNs), aiming to rationalize parameter choices without deviating from standard NN structures. The proposed methodology was substantiated via numerical simulations and experimental validation on the Case Western Reserve University (CWRU) bearing dataset. Overall, the study delivers a structured and interpretable framework for constructing neural networks optimized for MCM contexts [13].

Xu et al. (2021) introduced a hybrid deep learning framework designed to improve the accuracy of bearing fault diagnosis, which is vital for minimizing economic losses in industrial environments. Recognizing that many conventional deep learning models, such as CNN and gcForest, tend to underemphasize the feature extraction stage, the authors integrated these two architectures into a complementary hybrid model. In their approach, vibration signals from bearings were first converted into time-frequency representations via Continuous Wavelet Transform (CWT). These transformed images served as inputs to a Convolutional Neural Network (CNN), which extracted discriminative features related to fault characteristics. The extracted features were then input into a gcForest classifier for final classification. The model was validated using bearing fault datasets from both Case Western Reserve University (CWRU) and Xi'an Jiaotong University (XJTU-SY). Comparative experiments demonstrated that the CNN + gcForest combination significantly outperformed the individual models, offering superior fault detection accuracy. These findings underscore the model's potential applicability in real-world diagnostic systems by enhancing robustness and classification precision [14].

Similarly, Li (2024) introduced a novel deep learning architecture named DPW ATTCNN, specifically designed to facilitate fast and precise bearing fault diagnosis in a cost-effective and deployment-friendly manner. To minimize computational complexity, the model utilizes Depthwise Separable Convolution (DPW), which significantly reduces the number of trainable parameters without compromising feature extraction capabilities. Additionally, the incorporation of an Efficient Channel Attention (ECA) mechanism enables the model to capture and emphasize critical inter-channel dependencies, thereby improving its representational power. To enhance the model's performance across varying operational environments, Adaptive Batch Normalization (AdaBN) was employed, allowing effective domain adaptation. The model's performance was rigorously evaluated on the Case Western Reserve University (CWRU) bearing dataset, achieving an impressive classification accuracy of 99.58%. Even under high levels of additive noise, the model preserved over 95% accuracy, and domain transfer experiments yielded an average recognition rate of 97.35%. These results affirm DPW ATTCNN's high robustness and its capacity for generalization in real-world fault diagnosis scenarios [15].

Wu et al. (2024) presented a comprehensive approach to overcome two critical challenges in bearing fault diagnosis under varying operational environments: class imbalance and domain distribution divergence. These factors are known to degrade model generalization and often result in poor recognition of minority fault categories, thereby limiting diagnostic reliability. To address this, the authors developed a novel framework called Iterative Resampling Deep Decoupling Domain Adaptation (IRDDDA). The IRDDDA architecture integrates four key modules: a feature extractor, a domain discriminator, a label predictor, and a feature resampler, operating within a decoupled two-phase training paradigm. During the initial phase, domain-invariant representations are learned from imbalanced datasets across source and target domains. In the subsequent phase, a feature-wise resampling strategy is employed to alleviate

classification bias caused by class imbalance, while the upper layers of the feature extractor are iteratively refined to enhance model performance. The efficacy of the model was assessed using both the publicly available Case Western Reserve University (CWRU) dataset and experimental data acquired from the authors' custom-built bearing test platform. Empirical results demonstrated that IRDDDA consistently achieved higher and more balanced classification accuracy, particularly for underrepresented fault classes, and delivered robust generalization across domains. The study underscores the advantage of jointly addressing data imbalance and domain discrepancy in the development of reliable fault diagnosis systems [16].

Gu et al. (2022) introduced a novel hybrid methodology aimed at enhancing the precision of fault diagnosis in rolling bearings, which serve as essential components in mechanical manufacturing and transport systems. The primary objective was to effectively identify fault-related features from weak and noise-contaminated signals, particularly in cases involving small-sized datasets. The proposed method integrates multiple advanced techniques, including Variational Mode Decomposition (VMD), Continuous Wavelet Transform (CWT), Convolutional Neural Networks (CNN), and Support Vector Machines (SVM). Initially, vibration signals are preprocessed and decomposed into several Intrinsic Mode Functions (IMFs) via VMD. These IMFs are then converted into two-dimensional time-frequency representations using CWT. A CNN model, designed with carefully optimized hyperparameters, is employed to learn discriminative features from these images. Unlike conventional approaches that rely on Softmax layers for classification, this framework utilizes an SVM classifier in the final stage to improve robustness and accuracy. The proposed strategy was validated using both the Case Western Reserve University (CWRU) bearing dataset and a spindle fault dataset acquired from a dedicated test bench. Experimental evaluations revealed that the method achieved an average accuracy of 99.9% on the CWRU dataset and 90.15% on the spindle fault dataset. These results demonstrate the superior performance of the hybrid model over traditional CNN-based and feature engineering-based methods, confirming its potential in terms of diagnostic precision and generalization capacity across different fault scenarios [17].

Huang and Zhao (2024) presented an innovative deep learning architecture designed to enhance both the accuracy and efficiency of bearing fault diagnosis, with particular emphasis on the early identification of weak fault signals—an essential requirement for ensuring the dependable operation of rotating machinery. Their proposed method integrates three fundamental components. Firstly, an Improved Multi-Scale Feature Fusion Residual Network (IMSFFRN) is developed, which utilizes convolutional layers with diverse dilation rates to extract features at multiple scales. This configuration facilitates the effective fusion of intermediate representations, thereby enabling the model to capture more detailed and hierarchical characteristics from vibration signals. Secondly, acknowledging that different diagnostic features contribute unevenly to fault classification, the authors incorporated a Multiple-Winning Consciousness Self-Organizing Map (MCSOM) competitive mechanism. This layer enhances the discriminative capability of individual neurons by learning their sensitivity to particular fault types. Thirdly, to strengthen the model's generalization ability, a Support Vector Machine (SVM) is employed during the classification phase. The performance of the proposed framework was rigorously evaluated on three benchmark datasets: CWRU, PU, and SEU, where it achieved exceptional classification accuracies of 100%, 99.56%, and 100%, respectively—surpassing several state-of-the-art approaches. Furthermore, the model maintained high classification reliability under noisy conditions, highlighting its robustness. These findings confirm that the multi-scale competitive architecture is particularly suitable for early-stage fault detection in real-world industrial environments where signal degradation and noise are prevalent [18].

Zhang et al. (2024) proposed a robust and noise-resilient approach to tackle the significant challenges of condition monitoring and fault diagnosis of rolling bearings—key components in mechanical systems—under high-noise environments commonly encountered in industrial settings. Recognizing that traditional diagnostic models often struggle with the degradation of signal quality due to noise, the authors introduced a novel framework named Stochastic Resonance-assisted Deep Neural Network (SRDN). This method embeds the Stochastic Resonance (SR) mechanism within a Spiking Neural Network (SNN) architecture to amplify weak and noisy fault signals. By leveraging the SR effect, the model enhances the signal-to-noise ratio (SNR) of input signals before classification, enabling more accurate recognition of subtle fault characteristics. The SRDN model was tested on the widely used Case Western Reserve University (CWRU) bearing dataset, achieving an outstanding classification accuracy of 99.9%. More notably, even under severe noise conditions such as −8 dB SNR, the model sustained an accuracy exceeding 92%, indicating exceptional robustness and generalization. These results suggest that integrating SR mechanisms into neural network architectures significantly improves diagnostic performance in noisy environments and demonstrates high potential for deployment in real-world fault diagnosis systems where signal clarity cannot be guaranteed [19].

Han and Jeong (2020) highlighted the broader implications of bearing fault diagnosis, noting its influence not only on mechanical reliability but also on production efficiency and operational strategy. Their study sought to improve the practical applicability and robustness of deep learning models by simulating more realistic industrial conditions. While deep learning methods reported in the literature often achieve high diagnostic accuracy, these outcomes are typically derived from idealized, noise-free datasets collected in controlled simulator environments—conditions that rarely mirror actual factory settings. To bridge this gap, the authors augmented the widely used Case Western Reserve University (CWRU) bearing dataset with Gaussian noise to emulate the signal distortions commonly

found in industrial environments. They then proposed a Weighted Arithmetic Mean CNN Ensemble Model, which demonstrated enhanced classification performance compared to both standard CNN and unweighted ensemble models. Evaluation metrics, including accuracy and F1-score, confirmed the ensemble model's superiority in handling noisy and limited data. These results underscore the importance of robust ensemble strategies for industrial deployment, and suggest that the proposed method offers a practical and reliable solution for real-world fault diagnosis applications under non-ideal conditions [20].

Chen et al. (2023) addressed a critical challenge in intelligent fault diagnosis: the mismatch of data distributions resulting from varying operating conditions, which significantly hinders the performance of diagnostic models in real-world industrial settings. While conventional transfer learning techniques offer partial solutions to this issue, they often neglect the seamless integration of low-dimensional pre-learned features with high-dimensional deep representations, limiting their effectiveness. To resolve this limitation, the authors introduced a residual convolutional transfer learning framework guided by slow-varying features. Initially, Slow Feature Analysis (SFA) was utilized to extract latent representations that encapsulate the intrinsic and stable characteristics of the mechanical system. These features were then passed through a Residual Convolutional Network to learn deeper abstract representations. A bypass mechanism was employed to fuse low- and high-level features, ensuring comprehensive feature integration. Moreover, to mitigate domain discrepancies between the source and target data, the Maximum Mean Discrepancy (MMD) metric was applied in the Reproducing Kernel Hilbert Space (RKHS) to align distributions effectively. The framework was validated using the Xi'an Jiaotong University (XJTU) and Case Western Reserve University (CWRU) bearing datasets. Experimental outcomes demonstrated that the proposed method achieved over 99% classification accuracy across both datasets, confirming its strong transfer learning capability and resilience under dynamic operational environments. These findings highlight the framework's potential for deployment in industrial fault diagnosis applications involving significant domain shifts [21].

Gupta et al. (2024) introduced an innovative diagnostic approach termed the Discriminant Analysis-based Unimodality Test (DAT), aimed at enhancing the classification of time series data into unimodal or multimodal categories. Beyond this primary objective, DAT is also capable of detecting anomalies, estimating key statistical parameters, and identifying data skewness. The method is particularly effective in binary classification scenarios and demonstrates robust performance across both unimodal and multimodal datasets. In comparative analyses with established unimodality tests—specifically the dip test and the folding test—DAT consistently delivered superior accuracy and stability. To further validate its applicability, the authors extended DAT's functionality to fault detection tasks by evaluating its performance on the Case Western Reserve University (CWRU) bearing dataset. Five different machine learning classifiers were employed in this validation, and the method demonstrated exceptional diagnostic precision, successfully detecting bearing faults as small as 0.007 inches with a remarkable 99.999% accuracy. These results underscore DAT's potential as a powerful tool for industrial anomaly detection, offering significant improvements over traditional techniques, particularly in scenarios where high sensitivity to subtle fault signatures is required [22].

Hou et al. (2023) introduced a novel deep learning framework, named Diagnosisformer, which leverages a multi-feature parallel fusion strategy embedded within a Transformer-based architecture to improve the performance of bearing fault diagnosis. Aiming to overcome the limitations of conventional deep learning models—particularly their low diagnostic accuracy and limited robustness—the proposed model employs a structured approach for enhanced feature representation. Initially, Fast Fourier Transform (FFT) is used to extract frequency-domain features from raw vibration signals, followed by normalization and embedding steps. These processed features are then input into a parallel fusion encoder capable of simultaneously capturing both local and global signal characteristics. The resulting feature representations are passed through a cross-flipped decoder before reaching the classification module. The Diagnosisformer model was validated using two datasets: one collected from a laboratory-scale rotating machinery test rig and the publicly available Case Western Reserve University (CWRU) bearing dataset. The model achieved average classification accuracies of 99.84% and 99.85% on the respective datasets, surpassing benchmark methods. These findings affirm the model's superior diagnostic accuracy, generalization ability, and robustness under noisy conditions. The study ultimately highlights the potential of attention-based Transformer architectures as a robust and efficient solution for intelligent fault detection in industrial settings [23].

III. CWRU DATASET

The Case Western Reserve University (CWRU) bearing dataset has become a benchmark resource in the field of machine condition monitoring and predictive maintenance research. Its structured composition, high-fidelity vibration measurements captured under well-controlled experimental conditions, and detailed categorization of fault types and severities make it a widely adopted dataset for evaluating diagnostic algorithms. The dataset was generated using a custom-designed bearing fault test rig located in the Electrical Engineering Laboratory at Case Western Reserve University. As shown in Figure 3.1, the experimental system consists of a 2-hp induction motor, a torque transducer, a dynamometer, and two accelerometers strategically placed at the motor's drive end and fan end. Vibration data were typically recorded at high sampling rates of 12 kHz and 48 kHz, allowing for fine-resolution signal analysis suitable for fault detection tasks [24].

 ISSN: 2147-284X http://dergipark.gov.tr/bajece

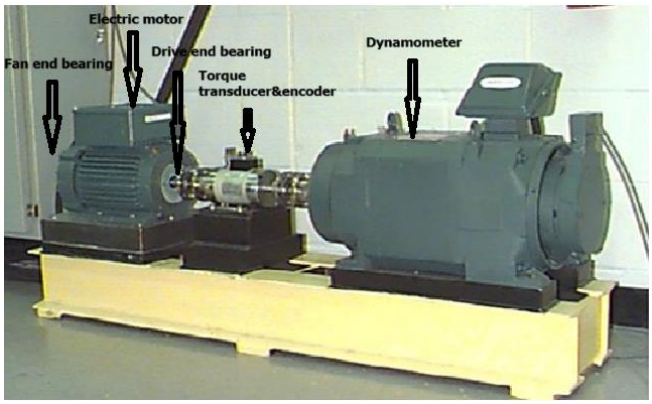


Figure 3. 1 CWRU Bearing Test Rig Experimental Platform

The bearing defects in the CWRU dataset were intentionally induced using the Electro-Discharge Machining (EDM) technique to simulate real-world fault scenarios in a controlled environment. These defects are systematically categorized into four primary classes: Normal (No Fault), Inner Race Faults, Outer Race Faults, and Ball (Rolling Element) Faults. To examine the effects of fault severity on diagnostic accuracy, each fault category was fabricated in three distinct sizes, corresponding to diameters of 0.007 inches, 0.014 inches, and 0.021 inches. This gradation of defect sizes enables comprehensive analysis of model sensitivity to varying levels of damage [25].

As shown in Figure 3.2, there are significant differences in vibration amplitude and waveform between the healthy condition and various fault types in the dataset (e.g., inner race, outer race, and rolling element faults). For instance, the ball fault in file *B021_1_227* is characterized by more irregular and abrupt amplitude changes over time; meanwhile, in *IR021_1_214*, the inner race fault presents with prominent vibration peaks. On the other hand, the *OR021_6_1_239* outer race fault scenario exhibits more periodic fault-induced impacts in the signal. In contrast, the healthy condition signal in *Time_Normal_1_098* displays a low-amplitude and more stable pattern.

The CWRU bearing dataset encompasses a total of 161 individual datasets, systematically grouped according to fault location and signal sampling frequency. These groups include normal (healthy) condition, 12kHz drive-end fault, 48kHz drive-end fault, and 12kHz fan-end fault scenarios. Each category contains three primary fault types: inner race faults (IF), ball (rolling element) faults (BF), and outer race faults (OF). Vibration signals were acquired under varying operational conditions, with motor speeds ranging from 1720 to 1797 revolutions per minute (RPM) and at load levels of 0, 1, 2, and 3 horsepower (HP). Furthermore, measurements were taken from multiple sensor placements on the experimental platform, including Centered (6:00 orientation), Orthogonal (3:00 orientation), and Opposite (12:00 orientation) positions, providing diverse perspectives of fault-induced signal characteristics [26].

The CWRU dataset is widely used for benchmarking the accuracy of machine learning and deep learning models.

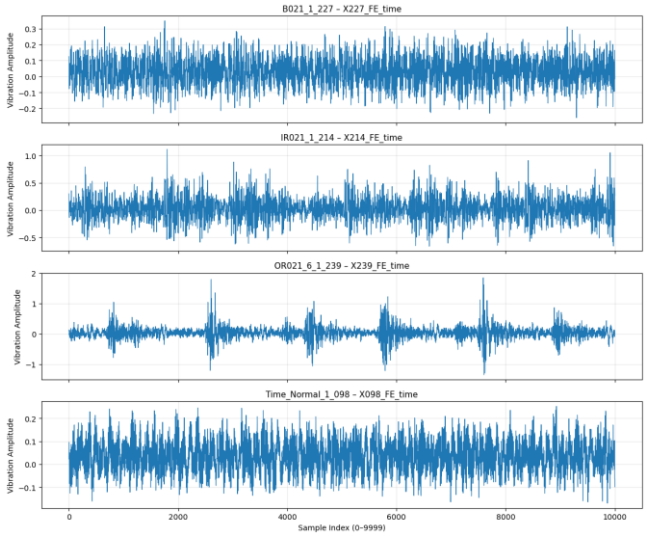


Figure 3. 2 Some example time-domain vibration signals from the CWRU bearing dataset.

Due to its high sampling rate and diversity in motor loads and fault sizes, it enables the development of classification, transfer learning, data augmentation, signal processing, and predictive maintenance algorithms. As one of the most cited open datasets in the field of predictive maintenance, the CWRU bearing dataset holds significant importance in both academic and practical research. It serves as a fundamental resource for deep learning-based bearing fault diagnosis studies, both at the initial stage and for advanced model comparisons.

IV. METHOD

A. SUPPORT VECTOR MACHINES (SVM)

Support Vector Machines (SVM) are a supervised learning-based, powerful classification method with high generalization capability. The main objective of SVM is to define an optimal hyperplane that separates examples belonging to different classes with the maximum margin. This method operates on a training dataset $D = \{(x_i, y_i)\}_{i=1}^{N}$ $x_i \in R^n$ as input vectors and $y_i = \in \{-1, +1\}$ as class labels. In linearly separable datasets, this hyperplane is represented by the following decision function:

$$f(x) = \langle w, x \rangle + b \qquad (1)$$

Here, www is the normal vector to the hyperplane, and bbb is the bias term. To maximize the margin width, SVM solves the following optimization problem:

$$\min_{w,b} \frac{1}{2}\|w\|^2 \quad \text{subject to} \quad y_i\left(w^T x_i + b\right) \geq 1, \quad \forall i \qquad (2)$$

Since real-world data are often not linearly separable, the model is converted into a soft-margin form. In this case, slack variables ξi that allow for classification errors and a regularization parameter C are introduced, leading to the following optimization formulation:

$$\min_{w,b,\xi} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i \quad \text{subject to}$$
$$y_i(w^T x_i + b) \geq 1-\xi_i,\ \ \xi_i \geq 0 \qquad (3)$$

The hyperparameter C provides a balance between model complexity and classification errors. Smaller values of C allow for a wider margin, while larger values aim to reduce classification errors.

A significant advantage of SVM is its capacity to solve non-linear classification problems. For this purpose, input data are transformed into a higher-dimensional feature space using kernel functions, making them linearly separable in that space. The kernel function enables the measurement of similarity without explicitly computing the transformation ϕ. One of the most widely used kernel functions is the Radial Basis Function (RBF) kernel, defined as:

$$K(x_i, y_i) = \exp\left(-\gamma\left\|x_i - x_j\right\|^2\right),\quad \gamma = \frac{1}{2\sigma^2} \qquad (4)$$

This kernel function is particularly well-suited for modeling non-linear patterns. During training, only the data points closest to the margin—called support vectors—determine the decision boundary. As a result, SVM offers a sparse model structure that is resistant to overfitting, even in high-dimensional datasets [27,28].

B. RANDOM FOREST (RF) METHOD

Random Forest (RF) is one of the ensemble learning algorithms developed by Leo Breiman, consisting of a combination of multiple decision trees. This method, which can be applied to both classification and regression problems, aims to reduce the bias of individual decision trees and enhance the model's generalization ability by training multiple trees independently and aggregating their results. The fundamental working principle of the RF algorithm is to construct each tree using bootstrap sampling of the training dataset and to perform node splitting based on a randomly selected subset of features. These strategies reduce the correlation between trees and significantly lower the risk of overfitting by increasing the diversity within the ensemble.

For classification problems, the RF algorithm performs majority voting based on the predictions obtained from each decision tree. Mathematically, this process is carried out over T decision tree classifiers such as $h_1(x), h_2(x), ...h_T(x)$,

$$\hat{y} = \mathrm{mode}\{h_1(x), h_2(x), ...h_T(x)\} \qquad (5)$$

It is expressed as such, where xxx represents the input vector, and the output of the RF model is represented by the result of the voting process. There are two main hyperparameters that significantly influence the performance of the RF algorithm:

• **n_estimators (T):** The number of trees in the forest. Increasing the number of trees generally results in more stable outcomes, but also increases computational cost.

• **max_features (m):** The number of features considered when splitting a node. This parameter helps reduce correlation between trees and improves the generalization capability of the model.

During the construction of each decision tree, impurity measures such as Gini impurity or entropy are commonly used for node splitting. The Gini impurity is calculated as follows:

$$Gini(D) = 1 - \sum_{i=1}^{C} p_i^{\,2} \qquad (6)$$

Here, D denotes the dataset at a given node, C represents the number of classes, and Pi is the proportion of samples belonging to the i^{th} class. The entropy criterion is defined as:

$$Entropy(D) = -\sum_{i=1}^{C} p_i \log_2(p_i) \qquad (7)$$

The RF algorithm stands out with its robustness against overfitting, effective performance on high-dimensional datasets, and ability to handle missing data. Moreover, it provides interpretability through feature importance rankings. Thanks to these advantages, RF is widely used in both classification and regression tasks. In recent years, RF-based approaches have demonstrated successful results in various application domains such as financial modeling, environmental systems, and subsurface resource classification, as highlighted in the literature [29,30].

C. NAIVE BAYES (NB) ALGORITHM

The Naive Bayes (NB) algorithm is a supervised classification method based on probability theory and rooted in Bayes' Theorem. The core assumption of this method is that all features used in classification are conditionally independent of each other. This assumption increases the simplicity of the model and reduces computational time; however, in some cases, it may also affect classification accuracy.

The Naive Bayes algorithm, for each class, relies on Bayes' Theorem to determine to which class the observed features belong when a feature vector $X = (x_1, x_2, ...x_n)$ is given:

$$P(C_k \setminus X) = \frac{P(X \setminus C_k).P(C_k)}{P(X)} \qquad (8)$$

Here:

$P(C_k \setminus X)$: It is the posterior probability of class Ck given the observation set X.

$P(X \setminus C_k)$: It is the likelihood of observing X given class Ck.

$P(C_k)$: It is the prior probability of class Ck.

$P(X)$: It is the marginal probability of the observed X.

Naive Bayes is called "naive" because it assumes that the features are conditionally independent of each other. Thanks

to this assumption, the conditional probability can be decomposed as follows:

$$P(X \setminus C_k) = \prod_{i=1}^{n} P(X_i \setminus C_k) \quad (9)$$

In this case, the classification decision is made by selecting the class with the highest probability, as shown below:

$$\hat{C} = \arg\max_{C_k} P(C_k).\prod_{i=1}^{n} P(X_i \setminus C_k) \quad (10)$$

The Naive Bayes (NB) algorithm offers several advantages, including short training time, high accuracy on small-sized datasets, and the ability to perform well even in high-dimensional feature spaces. Among its strengths are computational simplicity, fast model training, and low memory usage. However, its most significant drawback is the assumption of feature independence, which is not always realistic. As a result, the classification accuracy of the model may decline in some cases. Nonetheless, in certain fields—such as text mining and medical data analysis—this assumption does not pose a major problem [31-33].

D. XGBOOST ALGORITHM (EXTREME GRADIENT BOOSTING)

XGBoost (eXtreme Gradient Boosting) is a powerful ensemble learning algorithm based on decision trees, developed by Chen and Guestrin. It is a robust gradient-boosted decision tree (GBDT) method that constructs models in a sequential manner, where each new tree is optimized to correct the errors made by the previous trees, aiming to minimize the loss function. Compared to traditional GBDT approaches, XGBoost offers higher generalization performance, faster processing speed, and greater model flexibility.

XGBoost works by building decision trees sequentially, with each new tree focusing on minimizing the residual errors from the previous ones. The overall objective function optimized by the model can be expressed as follows:

$$L(\phi) = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{\,t}) + \sum_{k=1}^{t} \Omega(f_k) \quad (11)$$

Here, $l(y_i, \hat{y}_i)$ represents the loss function that measures the prediction error (e.g., log loss or squared error), while $\Omega(f_k)$ is the regularization term that controls the complexity of the model. This regularization term is formulated as follows:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \left\| w \right\|^2 \quad (12)$$

Here, **T** represents the number of leaves in the tree, **w** denotes the leaf scores, and **γ** and **λ** are the regularization coefficients applied to model complexity and weight magnitude, respectively. This structure helps prevent overfitting and enhances the generalization capability of the model.

The performance of the XGBoost model heavily depends on the proper tuning of its hyperparameters. In this study, the following key hyperparameters were optimized for model configuration:

- **n_estimators**: The total number of decision trees to be constructed. More trees generally lead to better generalization but increase computational cost.
- **max_depth**: The maximum depth of each tree. Greater depth can create complex decision boundaries but may also lead to overfitting.
- **learning_rate**: The learning rate; it limits the contribution of each individual tree to the final model.
- **subsample** and **colsample_bytree**: The subsampling ratios for data and features, respectively; used to reduce overfitting and increase diversity among trees.

These hyperparameters were optimized using grid search and cross-validation methods [34-35].

E. PERFORMANCE METRICS

In classification systems developed using machine learning and artificial intelligence techniques, performance metrics play a critical role in quantitatively evaluating a model's predictive capabilities. These metrics offer a comprehensive understanding of the model's general performance as well as its ability to correctly classify individual categories. Among the most commonly employed and informative evaluation measures are accuracy, precision, recall, and F1-score. Each of these metrics captures different aspects of classification quality and is elaborated upon in the following sections.

Accuracy

Accuracy refers to the proportion of correctly classified instances among the total number of predictions made by the model. While it offers a straightforward and easily interpretable measure of overall performance, it may not always reflect true effectiveness—particularly in scenarios involving imbalanced class distributions. For example, a model that consistently misclassifies minority class samples might still report high accuracy, thereby providing an overly optimistic evaluation of its classification performance.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (13)$$

Precision

Precision quantifies the accuracy of positive predictions by calculating the ratio of true positives to the total number of instances classified as positive by the model. This metric is particularly critical in domains where the consequences of false positives are significant—such as in industrial fault detection or medical diagnostics. A high precision value signifies that the model tends to be cautious in assigning positive labels, thereby ensuring that most identified positives genuinely belong to the target class.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (14)$$

Recall

Recall represents the percentage of true positive instances that the model successfully detects among all actual positive cases. This metric is especially vital in scenarios where missing a positive case—i.e., generating a false negative—can have serious consequences, such as in safety-critical systems or medical diagnoses. A high recall score reflects the model's strong ability to capture relevant positive examples, ensuring minimal oversight in identifying the target condition or class.

$$\text{Recall} = \frac{TP}{TP + FN} \qquad (15)$$

F1 Score

The F1 score serves as the harmonic mean of precision and recall, offering a balanced measure that considers both metrics simultaneously. It is especially advantageous in scenarios with imbalanced datasets or when it is equally important to minimize both false positives and false negatives. A high F1 score signifies that the model performs well in accurately identifying positive instances while also maintaining a low rate of misclassification.

$$\text{F1 Score} = 2x\frac{\text{Pr}\,ecision\,x\,\text{Re}\,call}{\text{Pr}\,ecision + \text{Re}\,call} \qquad (16)$$

Taken together, these metrics provide a more detailed and reliable evaluation of a classification model's overall performance and its ability to make accurate predictions across different classes [36-37].

V. FINDINGS

In this section, the performance of different machine learning algorithms used for bearing fault classification is evaluated. Based on signals obtained from the Case Western Reserve University (CWRU) bearing dataset, the Random Forest, XGBoost, SVM, and Naive Bayes algorithms were applied. Each model was analyzed in detail in terms of accuracy, precision, recall, and F1-score metrics. Additionally, performance comparisons were visualized and evaluated based on signal type and feature type. Assessing model performance not only at the algorithm level but also in terms of feature types (mean, max, min, sd, rms, skewness, kurtosis, crest, and form) is of great importance in identifying which statistical features are more decisive in fault diagnosis. The classification success rates obtained using the four machine learning methods are presented in Table 5.1.

TABLE 5. 1 PERFORMANCE COMPARISON OF CLASSIFICATION MODELS

Model	Accuracy	precision	recall	f1-score
Random Forest	0.9573	0.96	0.96	0.96
SVM	0.9373	0.94	0.94	0.94
Naive Bayes	0.9240	0.93	0.92	0.92
XGBoost	0.9573	0.96	0.96	0.96

As shown in Table 5.1, based on the performance results, the most successful classification models are Random Forest and XGBoost. Both models stand out with an accuracy of 95.73% and precision, recall, and F1-score values of 96%. These results indicate that the models demonstrate high performance in terms of making correct classifications overall.

The Support Vector Machine (SVM) algorithm, with an accuracy rate of 93.73%, emerges as a stable and reliable alternative. With each of its precision, recall, and F1-score values at 94%, SVM shows a strong capability to distinguish between classes and possesses a high generalization ability. In contrast, the Naive Bayes model performs lower compared to the other algorithms. Although it achieved 92.40% accuracy, along with 93% precision and 92% recall and F1-score, the model still demonstrates acceptable performance in basic classification tasks. However, due to its conditional independence assumption, this model may have limitations, especially when dealing with complex and high-dimensional datasets.

In conclusion, the evaluation results suggest that the Random Forest and XGBoost models provided the highest classification performance on the dataset used in this study and can be considered the most suitable methods for accurately diagnosing bearing faults.

The bar chart in Figure 5.1 below presents a comparison of the four classification models (Random Forest, SVM, Naive Bayes, and XGBoost) in terms of the four key performance metrics: accuracy, precision, recall, and F1-score. The graph provides a visual comparison of each model's performance level across each metric.

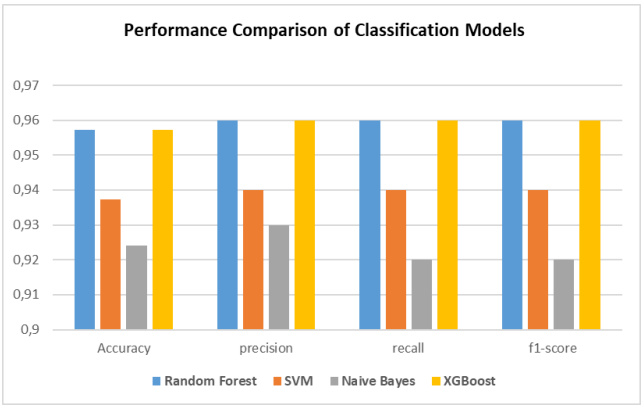


Figure 5. 1 Performance Comparison of Random Forest, SVM, Naive Bayes, and XGBoost Models

In Table 5.1, the overall performance of different classification algorithms was compared, and it was observed that the Random Forest and XGBoost models achieved the highest accuracy. This evaluation demonstrated the overall accuracy and classification capability of the models across the entire dataset. In Table 5.2, this general assessment is further detailed by presenting the performance metrics of the Random Forest model for each specific fault type (e.g., IR_007_1, Ball_021_1).

As shown in Table 5.2, the Random Forest model achieved 100% accuracy, precision, recall, and F1-score for the signals IR_007_1, IR_014_1, IR_021_1, and Normal_1. These results indicate that the model was able to classify these signal types almost flawlessly. In particular, the 100% recall rate for the "normal" signal suggests that the model is highly

successful in accurately identifying the healthy (non-faulty) condition.

TABLE 5. 2 SIGNAL-BASED CLASSIFICATION PERFORMANCE OF THE RANDOM FOREST MODEL

Signal	Accuracy	precision	recall	f1-score
Ball_007_1	0.96	0.97	0.96	0.97
Ball_014_1	0.88	0.92	0.88	0.90
Ball_021_1	0.8667	0.90	0.87	0.88
IR_007_1	1.00	1.00	1.00	1.00
IR_014_1	1.00	1.00	1.00	1.00
IR_021_1	1.00	1.00	1.00	1.00
Normal_1	1.00	1.00	1.00	1.00
OR_007_6_1	1.00	0.99	1.00	0.99
OR_014_6_1	0.8933	0.84	0.89	0.86
OR_021_6_1	0,9867	0.97	0.99	0.98

However, performance drops were observed for signals such as Ball_014_1, Ball_021_1, and OR_014_6_1. For example, in the Ball_021_1 class, the accuracy drops to 86.67%, and there are discrepancies between precision and recall values. This implies that the model misclassified some samples in these classes, confusing them with other fault types. Overall, while the model performs with very high accuracy on "inner race" (IR) and "outer race" (OR) faults, it shows variable performance on "ball" type faults depending on the signal characteristics.

Figure 5.2 below illustrates the classification accuracy rates obtained by the model for each bearing fault class. The model reaches 100% accuracy for IR_007_1, IR_014_1, IR_021_1, and Normal_1, demonstrating highly effective classification for these classes.

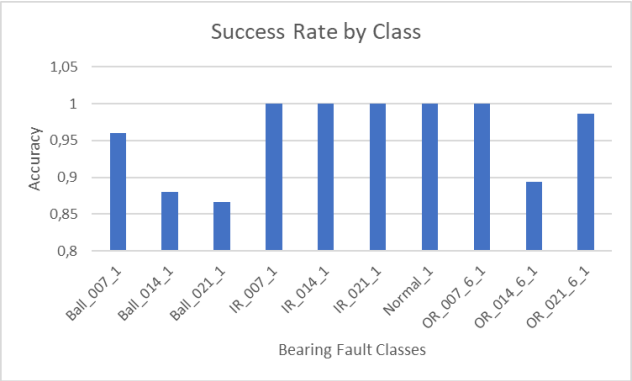


Figure 5. 2 Classification Accuracy Rates for Each Fault Signal Class

Confusion matrices are provided to illustrate the class-wise performance of different models. These matrices clearly reveal in which classes the models make more errors and in which they perform strongly. Figure 5.3 presents the confusion matrices for the models used.

As shown in Figure 5.3, the Random Forest and XGBoost models distinguish IR and OR signals with high accuracy, while the Naive Bayes model demonstrates comparatively lower performance for the Ball class than the other models. As shown in Table 5.3, the results indicate that the highest classification performance was achieved using the standard deviation (sd) feature. This feature demonstrated superior performance in terms of both accuracy and F1-score compared to the other features.

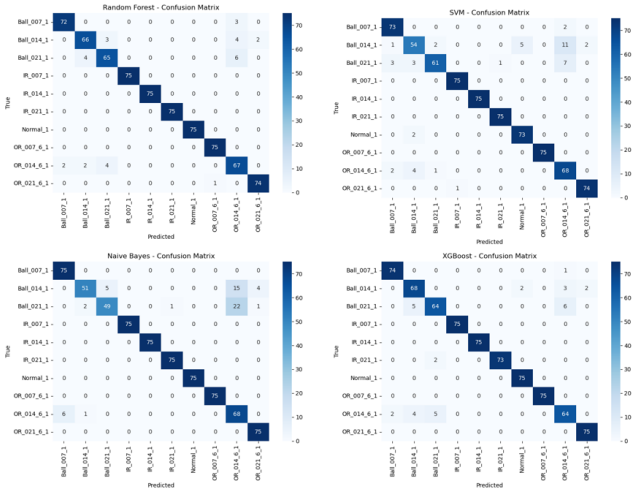


Figure 5. 3 Comparison of Confusion Matrices for Four Different Classification Models

An important contribution of this study is the performance analysis conducted at the feature level. Table 3 below presents the classification performance of fundamental statistical features used in bearing fault diagnosis, evaluated using the Random Forest algorithm. Each feature was individually tested in the model, and its performance was measured using accuracy, precision, recall, and F1-score metrics. The results are provided in Table 5.3.

TABLE 5. 3 FEATURE-BASED PERFORMANCE EVALUATION OF RANDOM FOREST CLASSIFICATION

Feature	Accuracy	precision	recall	f1-score
max	0.578261	0.574589	0.578261	0.575498
min	0.565217	0.551416	0.565217	0.556578
mean	0.285507	0.284134	0.285507	0.283265
sd	0.628986	0.634502	0.628986	0.629776
rms	0.615942	0.618286	0.615942	0.616016
skewness	0.24058	0.248969	0.24058	0.24235
kurtosis	0.386957	0.382938	0.386957	0.380831
Crest	0.275362	0.270234	0.275362	0.271135
form	0.457971	0.461526	0.457971	0.457787

Similarly, the Root Mean Square (rms) feature also yielded high performance, highlighting that the overall energy level of the signal plays a significant role in fault classification. On the other hand, features such as skewness and mean provided the lowest classification performance. This suggests that these features may be insufficient for distinguishing bearing faults and do not offer meaningful discrimination when used alone. Additionally, features like kurtosis and form factor were observed to exhibit moderate performance levels.

Figure 5.4 below presents the feature importance rankings in the classification processes of the RF and XGBoost models. This comparative analysis is particularly significant from a

computational cost perspective, as it enables the identification of the most effective features that enhance model performance.

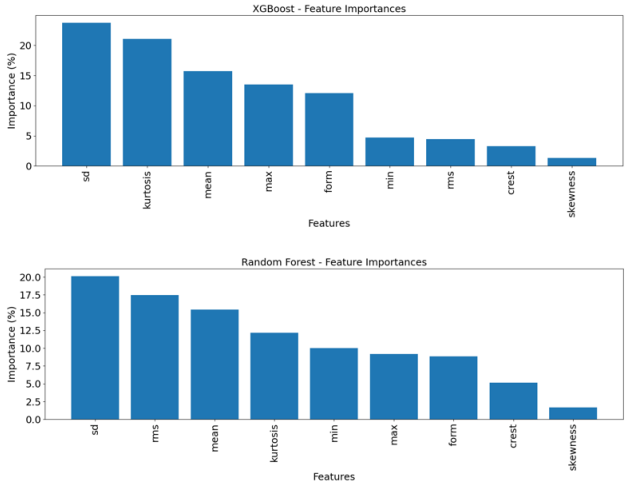


Figure 5. 4 Comparison of Feature Importance Rankings According to XGBoost and Random Forest Models

A common finding in both the RF and XGBoost models is that the standard deviation (sd) feature holds the highest importance score. This indicates that the "sd" feature, which represents the dispersion and variance in the data, is a highly decisive factor in distinguishing bearing faults. In the XGBoost model, features such as kurtosis, mean, max, and form exhibit high importance scores, whereas features like skewness, crest, rms, and min contribute relatively less. Notably, skewness is considered the least contributing feature by XGBoost. In the Random Forest model, following sd in importance are rms, mean, kurtosis, and min. Similarly, skewness and crest are identified as the least important features in this model as well.

VI. DISCUSSION AND CONCLUSİON

In this study, a comparative analysis of different machine learning algorithms for the diagnosis of bearing faults was conducted. As emphasized in the introduction, bearings are critical components for the continuity and safety of mechanical systems. Failure to detect bearing faults in a timely manner can lead to reduced operational efficiency and significant economic losses. To overcome the limitations of traditional signal processing-based methods, machine learning and deep learning techniques have gained increasing importance in recent years. The study utilized the Case Western Reserve University (CWRU) dataset and applied four classification algorithms: Random Forest (RF), Support Vector Machine (SVM), Naive Bayes (NB), and XGBoost. The findings revealed that Random Forest and XGBoost achieved the best performance, each with an accuracy of 95.73% and precision, recall, and F1-score values of 96%. These results indicate that both algorithms possess high generalization capability and strong discriminative power for fault classification. The SVM algorithm demonstrated stable performance with an accuracy of 93.73%, while the Naive Bayes algorithm performed comparatively worse, primarily due to its assumption of conditional independence between features. The results also showed notable differences in model performance depending on the fault type. Inner race (IR) and outer race (OR) faults could be classified with up to 100% accuracy, whereas ball-type faults exhibited decreased success rates depending on signal strength. This suggests that the signal characteristics of certain fault types are more easily distinguishable by classification models.

Moreover, the feature-level analysis highlighted that standard deviation (sd) and root mean square (RMS) are the most decisive features in bearing fault diagnosis. In contrast, features such as mean, skewness, and crest contributed less to classification performance. The feature importance rankings produced by both RF and XGBoost models supported these findings, confirming that features based on signal variance play a leading role in fault classification. In conclusion, this study not only demonstrated the effectiveness of various machine learning algorithms in bearing fault diagnosis but also systematically analyzed which signal features are more influential in the classification process. Ensemble learning-based models like Random Forest and XGBoost emerged as the most suitable methods due to their robust generalization capabilities and feature selection strengths. Future work may focus on hybridizing these models with deep learning architectures, testing on different datasets, and integrating into real-time predictive maintenance systems to further enhance the efficiency of bearing health monitoring.

REFERENCES

[1] Kaya, Y., Kuncan, M., Kaplan, K., Minaz, M. R., & Ertunç, H. M. (2021). A new feature extraction approach based on one dimensional gray level co-occurrence matrices for bearing fault classification. *Journal of Experimental & Theoretical Artificial Intelligence*, *33*(1), 161-178.

[2] Kaya, Y., Kuncan, F., & ERTUNÇ, H. M. (2022). A new automatic bearing fault size diagnosis using time-frequency images of CWT and deep transfer learning methods. *Turkish Journal of Electrical Engineering and Computer Sciences*, *30*(5), 1851-1867.

[3] Zhang, X., Zhang, M., Wan, S., He, Y., & Wang, X. (2021). A bearing fault diagnosis method based on multiscale dispersion entropy and GG clustering. *Measurement*, *185*, 110023.

[4] Du, Y., Geng, X., Zhou, Q., & Cheng, S. (2024). A fault diagnosis method for offshore wind turbine bearing based on adaptive deep echo state network and bidirectional long short term memory network in noisy environment. *Ocean Engineering*, *312*, 119101.

[5] Wang, P., Xiong, H., & He, H. (2023). Bearing fault diagnosis under various conditions using an incremental learning-based multi-task shared classifier. *Knowledge-Based Systems*, *266*, 110395.

[6] Wang, Z., Shi, D., Xu, Y., Zhen, D., Gu, F., & Ball, A. D. (2023). Early rolling bearing fault diagnosis in induction motors based on on-rotor sensing vibrations. *Measurement*, *222*, 113614.

[7] Gu, X., Yu, Y., Guo, L., Gao, H., & Luo, M. (2023). CSWGAN-GP: A new method for bearing fault diagnosis under imbalanced condition. *Measurement*, *217*, 113014

[8] Li, F., Wang, L., Wang, D., Wu, J., & Zhao, H. (2023). An adaptive multiscale fully convolutional network for bearing fault diagnosis under noisy environments. *Measurement*, *216*, 112993.

[9] Wen, L., Xie, X., Li, X., & Gao, L. (2022). A new ensemble convolutional neural network with diversity regularization for fault diagnosis. *Journal of Manufacturing Systems*, *62*, 964-971.

[10] Su, Z., Zhang, J., Tang, J., Wang, Y., Xu, H., Zou, J., & Fan, S. (2023). A novel deep transfer learning method with inter-domain decision discrepancy minimization for intelligent fault diagnosis. *Knowledge-Based Systems*, *259*, 110065.

[11] Meng, Z., He, H., Cao, W., Li, J., Cao, L., Fan, J., ... & Fan, F. (2023). A novel generation network using feature fusion and guided

adversarial learning for fault diagnosis of rotating machinery. *Expert Systems with Applications*, *234*, 121058.
[12] Li, X., Xiao, S., Zhang, F., Huang, J., Xie, Z., & Kong, X. (2024). A fault diagnosis method with AT-ICNN based on a hybrid attention mechanism and improved convolutional layers. *Applied Acoustics*, *225*, 110191.
[13] Borghesani, P., Herwig, N., Antoni, J., & Wang, W. (2023). A Fourier-based explanation of 1D-CNNs for machine condition monitoring applications. *Mechanical Systems and Signal Processing*, *205*, 110865.
[14] Xu, Y., Li, Z., Wang, S., Li, W., Sarkodie-Gyan, T., & Feng, S. (2021). A hybrid deep-learning model for fault diagnosis of rolling bearings. *Measurement*, *169*, 108502.
[15] Li, Y. (2024). An accurate lightweight algorithm for bearings fault diagnosis based on DPW ATTCNN model. *Physical Communication*, *66*, 102383.
[16] Wu, Z., Guo, J., Liu, Y., Li, L., & Ji, Y. (2024). An Iterative Resampling Deep Decoupling Domain Adaptation method for class-imbalance bearing fault diagnosis under variant working conditions. *Expert Systems with Applications*, *252*, 124240.
[17] Gu, J., Peng, Y., Lu, H., Chang, X., & Chen, G. (2022). A novel fault diagnosis method of rotating machinery via VMD, CWT and improved CNN. *Measurement*, *200*, 111635.
[18] Huang, Z., & Zhao, X. (2024). A novel multi-scale competitive network for fault diagnosis in rotating machinery. *Engineering Applications of Artificial Intelligence*, *128*, 107441.
[19] Zhang, X., Ma, Y., Pan, Z., & Wang, G. (2024). A novel stochastic resonance based deep residual network for fault diagnosis of rolling bearing system. *ISA transactions*, *148*, 279-284.
[20] Han, S., & Jeong, J. (2020). An weighted CNN ensemble model with small amount of data for bearing fault diagnosis. *Procedia Computer Science*, *175*, 88-95.
[21] Chen, S., Zheng, W., Xiao, H., Han, P., & Luo, K. (2023). A residual convolution transfer framework based on slow feature for cross-domain machinery fault diagnosis. *Neurocomputing*, *546*, 126322
[22] Gupta, A., Onumanyi, A. J., Ahlawat, S., Prasad, Y., Singh, V., & Abu-Mahfouz, A. M. (2024). DAT: A robust Discriminant Analysis-based Test of unimodality for unknown input distributions. *Pattern Recognition Letters*, *182*, 125-132.
[23] Hou, Y., Wang, J., Chen, Z., Ma, J., & Li, T. (2023). Diagnosisformer: An efficient rolling bearing fault diagnosis method based on improved Transformer. *Engineering Applications of Artificial Intelligence*, *124*, 106507.
[24] Liu, X., Wang, J., Meng, S., Qiu, X., & Zhao, G. (2023). Multi-view rotating machinery fault diagnosis with adaptive co-attention fusion network. *Engineering Applications of Artificial Intelligence*, *122*, 106138.
[25] Zhong, J., Zheng, Y., Ruan, C., Chen, L., Bao, X., & Lyu, L. (2025). M-IPISincNet: an explainable multi-source physics-informed neural network based on improved SincNet for rolling bearings fault diagnosis. *Information Fusion*, *115*, 102761.
[26] Liu, S., Li, J., Zhou, N., Chen, G., Lu, K., & Wu, Y. (2025). Intelligent fault diagnosis of rotating machine via Expansive dual-attention fusion Transformer enhanced by semi-supervised learning. *Expert Systems with Applications*, *260*, 125398.
[27] Zhang, W., Yu, B., Li, G., Zhuang, P., Liang, Z., & Zhao, W. (2024). Unified multi-color-model-learning-based deep support vector machine for underwater image classification. *Engineering Applications of Artificial Intelligence*, *138*, 109437.
[28] Akinola, I. T., Sun, Y., Adebayo, I. G., & Wang, Z. (2024). Daily peak demand forecasting using pelican algorithm optimised support vector machine (POA-SVM). *Energy Reports*, *12*, 4438-4448.
[29] Du, Q., & Zhai, J. (2024). Application of artificial intelligence Sensors based on random forest algorithm in financial recognition models. *Measurement: Sensors*, *33*, 101245.
[30] Chen, Y., Li, T., Fu, B., Xia, Q., Liu, Q., Li, T., ... & Huang, Y. (2024). Deposit type discrimination of Jiaodong gold deposits using random forest algorithm: Constraints from trace elements of pyrite. *Ore Geology Reviews*, *175*, 106343.
[31] Xi, W. (2024). Research on E-learning interactive English vocabulary recommendation education system based on naive Bayes algorithm. *Entertainment Computing*, *51*, 100732.
[32] Shang, Y. (2024). Prevention and detection of DDOS attack in virtual cloud computing environment using Naive Bayes algorithm of machine learning. *Measurement: Sensors*, *31*, 100991.
[33] Raj, S., Vishnoi, A., & Srivastava, A. (2024). Classify Alzheimer genes association using Naïve Bayes algorithm. *Human Gene*, *41*, 201309.
[34] Kan, X., Fan, Y., Zheng, J., Chi, C. H., Song, W., & Kudreyko, A. (2023). Data adjusting strategy and optimized XGBoost algorithm for novel insider threat detection model. *Journal of the Franklin Institute*, *360*(16), 11414-11443
[35] Wang, T., Bian, Y., Zhang, Y., & Hou, X. (2023). Classification of earthquakes, explosions and mining-induced earthquakes based on XGBoost algorithm. *Computers & Geosciences*, *170*, 105242.
[36] Kaya, Y., Kuncan, M., Akcan, E., & Kaplan, K. (2024). An efficient approach based on a novel 1D-LBP for the detection of bearing failures with a hybrid deep learning method. Applied Soft Computing, 155, 111438.
[37] Akcan, E., Kuncan, M., Kaplan, K., & Kaya, Y. (2024). Diagnosing bearing fault location, size, and rotational speed with entropy variables using extreme learning machine. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, *46*(1), 4.

BIOGRAPHIES

Eyyüp Akcan Dr. Lecturer. Assist.
Eyyüp Akcan is an academic specialised in the field of electrical and electronics engineering. He completed his undergraduate education at Siirt University, Department of Electrical and Electronics Engineering in 2019, completed his master's degree at Siirt University, Institute of Science and Technology in 2021, and completed his doctorate education at the Institute of Graduate Studies of the Batman university in 2025. He is currently working as a Dr. Lecturer in the of Motor Vehicles and Transportation Technologies, Siirt the University.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 3, September 2025
Research Article

DC-DC Converter Architectures

Abdurrahim Erat and Ahmet Mete Vural

***Abstract*—Direct Current (DC)-DC converter architectures constitute a fundamental component of modern power electronic systems, enabling efficient voltage regulation and energy transfer across various stages of electrical networks. The growing integration of renewable energy sources, the electrification of transportation, and the proliferation of portable electronic devices have significantly increased the demand for compact, reliable, and high-efficiency converter designs. Recent advancements in semiconductor technologies, particularly those based on wide-bandgap materials, have further accelerated the development and deployment of diverse converter topologies. This study provides a comprehensive review of both conventional and emerging DC-DC converter configurations, examining their operating principles, performance characteristics, suitability for specific applications, and design complexity. Furthermore, a bibliometric analysis has been conducted to identify prevailing research trends, influential contributors, and collaborative networks within the field. By critically evaluating the advantages and limitations of each topology, the study aims to support future research directions and technological innovations in power electronics.**

Index Terms— **Converter topologies, DC-DC converter, Isolated topologies, Non-isolated converters, Power electronics.**

I. Introduction

IN RECENT YEARS, the evolution of DC-DC converter topologies has been inspired by the rising requirement for superior efficiency, modular, compact, and reliable power conversion solutions across various applications such as green power mechanisms, electric vehicle technologies, high-voltage DC power networks, multi-terminal DC power grids, and smart power networks. Recent studies have focused on enhancing the performance of DC-DC converter topologies through innovative topological designs, advanced control techniques, and the integration of wide-bandgap semiconductor technologies. These developments have significantly improved the efficiency, compactness, and reliability of DC-DC converters, facilitating their widespread application in various industrial sectors.

Abdurrahim Erat, is with Department of Electrical and Electronics Engineering University of Gaziantep University, Gaziantep, Turkey, (e-mail: a.rahim_erat@sirnak.edu.tr).

https://orcid.org/0000-0003-0606-9132

Ahmet Mete Vural, is with Department of Electrical and Electronics Engineering University of Gaziantep University, Gaziantep, Turkey, (e-mail: mvural@gantep.edu.tr).

https://orcid.org/0000-0003-2543-4019

Manuscript Received Jun 04, 2025; Accepted Jul 15, 2025.
DOI: 10.17694/bajece.1713384

By utilizing an available DC power source, these converters adjust a fixed DC voltage to a regulated and variable level, as illustrated in Fig. 1. The widespread adoption of DC-DC converter architectures is primarily driven by advances in semiconductor device technologies, which serve as the foundational components of such systems.

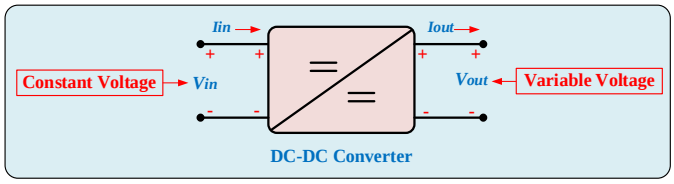


Fig.1. The schematic block configuration of a DC-DC converter

DC-DC converter technologies have recently become an exciting field of study for many researchers due to their numerous advantages and critical roles within various industrial environments. This study [1] reviews several DC-DC converter architectures and management strategies for electrifying transport. Particularly included are the difficulties and potential approaches for future research and development in creating innovative DC-DC power converter topologies. Due to several benefits, including increased longevity, ecological sustainability, lower maintenance levels, greater accessibility, and the capacity to generate higher amounts of electricity to meet load demands, photovoltaic-based power generation facilities have emerged as the most desirable option. This study [2] examines robust DC-DC converter configurations that facilitate the generation of energy systems powered by photovoltaics. The primary objective of this study is to provide an exhaustive guide to selecting the optimal converter topology for integrating solar energy. Various DC-DC converter structures are systematically evaluated based on key performance indicators, including supported power ranges, hardware intricacy, implementation costs, tracking capabilities, and overall energy conversion efficiency. To bridge the research gap in the domain of light charging of electrified vehicles, this paper [3] presents an extensive survey of the latest advancements in DC–DC converter architectures tailored for such applications. The distinctive operational features and limitations of each topology are critically analyzed. Moreover, the study identifies prevailing technical challenges and outlines prospective research pathways. The objective is to support future development efforts by offering insights that can inform the development and improvement of efficient charging systems for light electrified cars. For multi-terminal DC power networks, this research [4] provides a concise overview of contemporary DC-DC converter configurations, categorizing them as either isolated or non-isolated. Every topology is

assessed using essential performance standards. Excellent concentration of power and resilience to failure, bidirectional functionality, and dependability are key characteristics highlighted in setups suitable for HVDC operations. Current advances in bidirectional DC–DC conversion devices with broad voltage converting ranges for mixed energy storage technologies are reviewed in this study [5], with an emphasis on transformer-based isolated topologies and non-isolated topologies based on impedance networks. A thorough assessment approach is employed to evaluate seven isolated and eight non-isolated converter types. The article addresses the primary issues with converter design for hybrid energy storage systems, focusing future research on fault diagnostic techniques, electromagnetic compatibility, and the integration of wide-bandgap devices. Multi-level DC-DC converters are categorized in this evaluation [6] according to voltage category, isolation technique, electrical power transfer direction, power switching techniques, and voltage transformation.

Additionally, it provides an overview of other voltage boosting methods, such as multi-stage strategies and new structural developments. Additionally, multi-stage converters are categorized into important classes, including modular, multi-level, interleaved, cascaded, quadratic gain, and hybrid designs. The present article [7] provides a comprehensive examination of bidirectional DC-DC converters, emphasizing their vital role in photovoltaic power-generating facilities. The paper categorizes converters into two types: non-isolated and isolated. It also discusses the topologies, control methods, and switching techniques of each type. The authors emphasize the crucial importance of a higher level of power and reliability for the stable operation of photovoltaic systems. The study [8] provides the most recent advancements in DC-DC converter technology, with an emphasis on their topologies and control methods.

Although the number of studies on DC-DC converter topologies is increasing, literature still faces several significant limitations. Firstly, the results of many current studies are fragmented and inconsistent due to the lack of a common framework for assessing converter performance. Furthermore, the integration of wide-bandgap semiconductor technology under practical operating circumstances is not given enough attention. Few studies have been done on the scalability of converter designs for large-scale renewable systems and their flexibility in responding to dynamic, bidirectional energy flows. Furthermore, the majority of assessments fail to consider real-world implementation limitations, including industrial cost-performance trade-offs, electromagnetic interference (EMI) mitigation, and heat management. These research gaps underscore the necessity for a thorough assessment that considers the practical constraints and new requirements of current topologies, in addition to classifying them.

II. Bibliometric Analysis

A bibliometric survey conducted on July 7, 2025, using the Web of Science database and the keyword "DC-DC Converters" for the period 2016-2025 identified 7,321 publications. The dataset encompasses various types of scholarly outputs, with research articles (4080) and conference proceedings (3149; 43.0%) constituting the majority. Additionally, 209 review papers (2.9%) were identified, while the remaining entries, comprising early access articles, book chapters, editorial materials, and letters, account for a smaller portion of the total. These results reflect the prominence of original research and conference dissemination in the DC-DC converter domain, emphasizing the field's active development and its close link to ongoing technological innovation. The collected dataset was systematically examined across several bibliometric dimensions, including author productivity, citation patterns, journal distribution, country-wise contributions, keyword occurrences, article characteristics, and abstract content. The analysis was exclusively based on records indexed within the Web of Science Core Collection, ensuring the inclusion of high-quality and peer-reviewed scientific publications in the field.

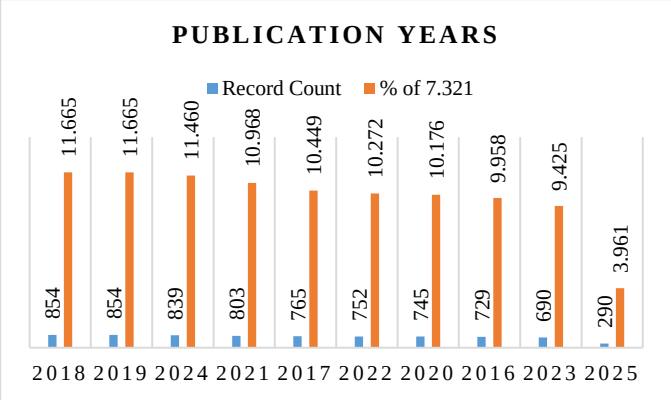


Fig.2. Annual distribution of publications related to DC-DC converter research based on Web of Science data

Fig. 2 illustrates the annual distribution of publications related to DC-DC converters over the past decade (2016–2025). As of the date on which the bibliometric analysis was conducted, a total of 290 documents had been indexed for the year 2025, indicating ongoing research activity within the current year. A total of 7321 publications were identified within the defined time frame. As of the date of the bibliometric analysis, the documents published in 2025 accounted for approximately 3.96% of the overall output, highlighting a modest yet ongoing contribution in the current year.

The co-occurrence keyword mapping is displayed in Fig. 3. The present research employed key keywords and authors' keywords from the Web of Science database to analyze co-occurrence keywords. Considering the weight of the item indicated by the dimension and color associated with every label and circle, the resultant diagram shows the associations among these terms. The program automatically groups similar nodes into color-differentiated clusters. Utilizing the VOSviewer approach, items are arranged according to their degree of relatedness.

Fig. 4 presents a collaboration network map illustrating the relationships among the most prolific authors in the field of DC-DC converters. As illustrated in the figure, Frede Blaabjerg occupies a central position within the collaboration map, having

authored 76 publications in the domain of DC-DC converters, which have collectively received 1887 citations. Frede Blaabjerg is followed by Bo Zhang, who ranks second with 53 publications and a total of 527 citations in the field, indicating his significant yet comparatively lower centrality within the scholarly network. Ranking third in terms of productivity is Dmitri Vinnikov, with 49 publications and 252 citations, reflecting his ongoing contributions to the field despite a comparatively lower citation impact. Fig. 5 illustrates the citation network among authors, highlighting the interconnections and influence patterns within the scholarly community engaged in DC-DC converter research. As observed in the figure, Frede Blaabjerg and Bo Zhang emerge as the most highly cited authors in the domain of DC-DC converters, reflecting their substantial academic influence and impact on the field.

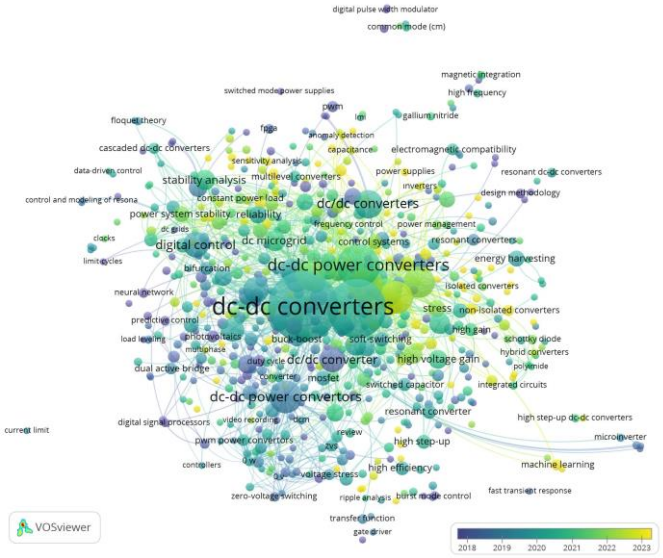


Fig.3. Network diagram study of the co-occurrence of author keywords utilizing VOS viewer

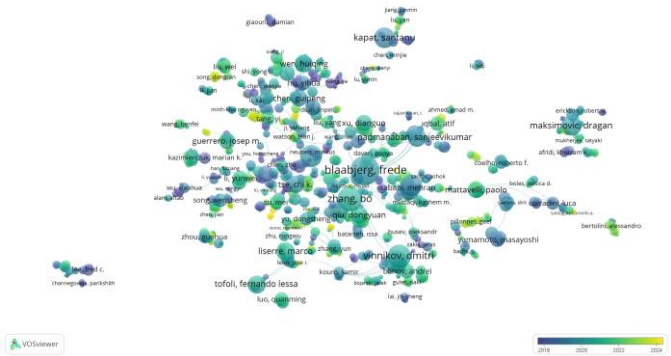


Fig.4. The author's cooperation network

Fig. 6 illustrates the global distribution of co-authorship networks, highlighting the collaborative dynamics among researchers from diverse countries and regions. The diagram illustrates the international collaboration network among countries associated with leading researchers in DC-DC converter studies. Each circle's size indicates the volume of publications originating from that nation, while the lines illustrate joint authorship ties across borders. Notably, China and the United States stand out as key centers due to their significant representation and numerous collaborative connections, underscoring their pivotal positions in the global research landscape. Several European nations, including the United Kingdom, Denmark, Italy, and Spain, also display notable participation, establishing robust linkages both regionally and globally. Countries such as India, Iran, and Turkey have begun to emerge as significant contributors, reflecting an increase in their research output and growing involvement in international partnerships. The color coding, spanning from 2018 through 2022, reveals the evolution of these collaborations over time, with recent cooperative efforts highlighted in warmer hues.

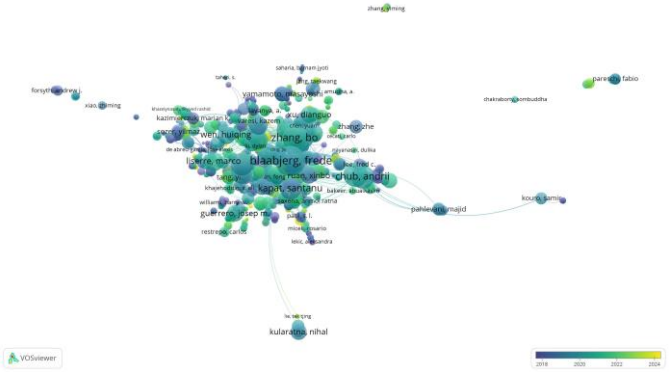


Fig.5. The network of authors' citations

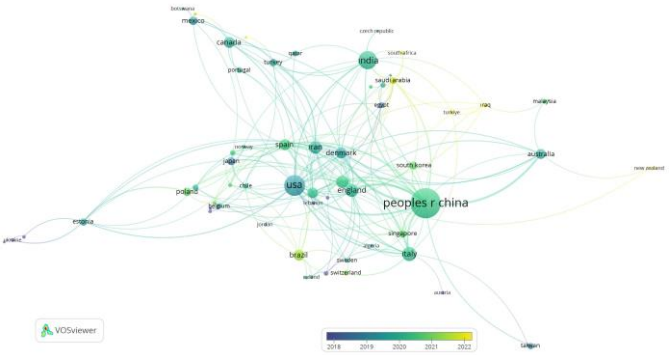


Fig.6. Geographical representation of collaborative authorship patterns across different countries and regions

III. Conventional DC-DC Converters

The architecture of DC-DC converters encompasses a broad spectrum of configurations, each tailored to suit distinct industrial demands. These systems are capable of either elevating or reducing the direct current input voltage and may be engineered to provide multiple voltage outputs or terminals. The fundamental criterion for classifying DC-DC converter systems lies in the presence or absence of galvanic isolation, typically achieved through the integration of a transformer. Based on this distinction, the overarching classification scheme bifurcates DC-DC converters into isolated and non-isolated categories [9]–[13]. In certain DC-DC converters, electrical isolation between the input and output is achieved through the incorporation of a transformer, which functions as a key intermediary component. On the other hand, converter designs

that omit this element establish a continuous conductive path between the primary and secondary sides. A schematic representation summarizing the breadth of DC-DC converter typologies explored in the academic domain is provided in Fig. 7.

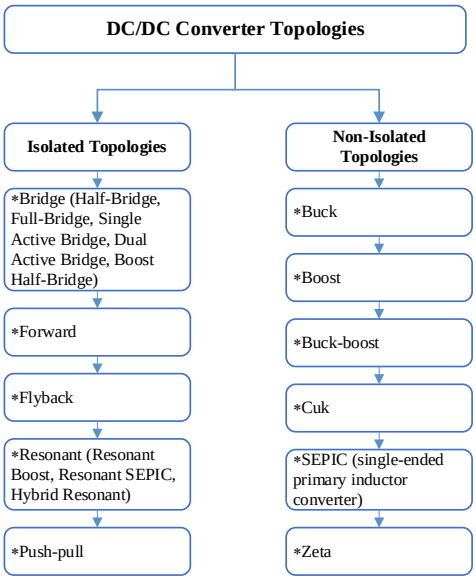


Fig.7. Conventional configurations of DC-DC converters

A. Isolated DC-DC Converter Architectures

In the context of DC-DC converter topologies, the term 'isolation' specifically denotes galvanic separation, signifying that there exists no uninterrupted electrical continuity between the converter's primary (input) and secondary (output) sections. Essentially, isolated DC-DC converter architectures separate the primary terminal and secondary terminal circuits to prevent direct current transfer, typically achieved through the integration of a transformer. This transformer-based separation divides the system into two electrically independent sections. In such isolated configurations, each side generally maintains its distinct ground reference. The energy conversion mechanism typically entails modulating the initial direct current into an alternating waveform, enabling transmission across a high-frequency transformer, after which the signal is rectified to regenerate a stable DC output. Isolated converter systems are often capable of achieving higher voltage conversion ratios compared to non-isolated types. However, challenges such as mitigating the influence exerted by the inherent leakage inductance of the transformer and managing the complexity of transformer construction remain critical design considerations [14].

1) Bridge topologies

Bridge-based isolated DC-DC converters form a highly adaptable class of power conversion systems, distinguished by their use of bridge networks composed of multiple switching elements in tandem with a high-frequency isolation transformer. This category comprises several frequently used designs, including but not limited to the half-bridge, full-bridge, single-active-bridge, dual-active-bridge, and the boost-configured half-bridge variant. Representative schematic diagrams of these bridge-derived topologies are displayed in Fig. 8: Fig. 8(a) features the full-bridge arrangement, and Fig. 8(b) illustrates the half-bridge. At the same time, Fig. 8(c) depicts the dual active bridge. This schematic is adapted from the circuit diagrams in the referenced studies shown below the figure. Among these, the full-bridge circuit in Fig. 8(a) stands out for its ability to generate an output voltage nearly twice that of the half-bridge structure, primarily due to the incorporation of a greater number of active switching devices in its configuration [15]. Consequently, this converter configuration is capable of supporting significantly elevated output power levels. The voltage impressed upon the primary winding of the transformer depends on the conduction sequence of switching devices S1, S4, S2, and S3 and may alternate between positive (+Vs), negative (−Vs), or null values. On the secondary side, the resulting alternating waveform, induced magnetically through transformer action, is processed through a rectification stage to yield a direct current output. The instantaneous value of this secondary voltage fluctuates from zero to a peak amplitude, which is governed by both the transformer's turn ratio and the specific characteristics of the applied pulse-width modulation (PWM) strategy.

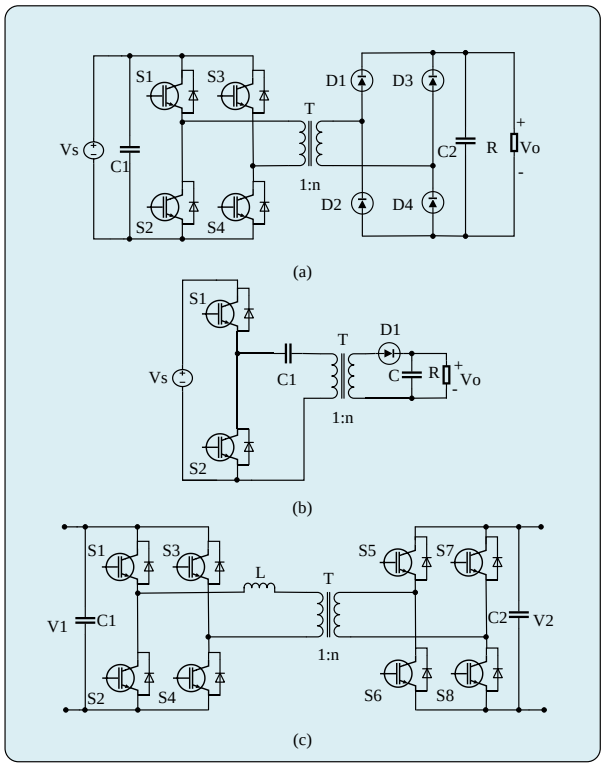


Fig.8. Bridge-type DC-DC converter configurations: (a) topology of the full-bridge converter, (b) topology of the half-bridge converter, (c) topology of the dual active bridge converter [16]–[19]

By contrast, the half-bridge converter configuration presented in Fig. 8(b) utilizes a high-frequency transformer to establish galvanic isolation, thereby enabling the output voltage to assume values that may exceed, fall below, or exhibit reversed polarity relative to the input source. This topology facilitates bidirectional power exchange, thereby enhancing energy transmission efficiency. During operation, the direction and magnitude of power flow are modulated through the active control of semiconductor elements, such as MOSFETs, IGBTs, IGCTs, or FETs, with switches S1 and S2 typically operating in

an alternating, complementary switching sequence. On the other hand, the dual active bridge (DAB) converter depicted in Fig. 8(c) exemplifies a more sophisticated bidirectional architecture, distinguished by its ability to achieve high-efficiency power transfer under varying load and operating conditions. It consists of two inverter stages linked through a coupling inductor and a high-frequency transformer. This innovative concept was first introduced by de Doncker et al. as a means to achieve flexible and efficient power transfer between two DC sources [20]. This topology demonstrates superior efficiency by enabling optimized control, a broad voltage conversion range, and minimal voltage stress during switching transitions, rendering it particularly advantageous for a wide range of industrial applications [21].

The dual active bridge configuration is particularly valued for its straightforward design, ease of scalability, superior efficiency, and high power density. It offers uniform switching behavior, ensures galvanic isolation, and minimizes the presence of passive or idle components, making it well-suited for high-performance power conversion applications [22]–[26]. This topology consists of a pair of H-bridge converter units connected via a high-frequency transformer, with the transformer's leakage inductance playing a dual role: ensuring electrical isolation and acting as a temporary reservoir for energy storage [27]. The single-active-bridge DC-DC converter design is particularly suitable for applications requiring unidirectional power flow, combining economical design with versatile control capabilities. Its basic structure features one actively controlled bridge and one diode-based passive bridge, connected through a transformer operating at high frequency [28]. Moreover, the single active bridge converter can be considered a unidirectional adaptation of the widely recognized DAB topology, which permits power flow in both input-to-output and output-to-input directions and is designed for high-power applications.

A promising strategy to improve galvanically isolated DC-DC converters, particularly for handling extensive variations in input voltage, is the implementation of the boost half-bridge converter architecture [29], [30]. Reference [31] introduces a well-recognized standard design for the half-bridge DC-DC converter with boost functionality. The works referenced in [31], [32] further explore an enhanced version of the traditional half-bridge DC-DC converter topology with boost functionality, detailing modifications and improvements over the conventional design. Due to its advantages such as reduced voltage oscillations associated with diode reverse recovery, zero-voltage switching, simplified input filtering requirements, broad input voltage compatibility, wide operational duty cycle range, elimination of DC magnetization in the power transformer, and absence of circulating currents[31], [33], the conventional half-bridge DC-DC converter with boost functionality finds applicability across numerous sectors. These include DC power distribution networks [32], photovoltaic energy systems [34], fuel cell technologies [35], electric and hybrid vehicle powertrains [36], as well as high-voltage and high-power grid infrastructures [37], among others. Despite its advantages, the boost half-bridge DC-DC converter exhibits several notable limitations. These include excessive DC magnetizing current within the power transformer, a relatively high proportion of electromagnetic components, considerable uncontrolled power dissipation in the transformer, and pronounced voltage stress along with significant switching voltage fluctuations at the secondary rectifier diodes. Furthermore, the converter's performance may be adversely affected by rapid transient disturbances, such as audio-frequency noise, which can induce large instantaneous output currents through the external inductor [31].

With their high power-handling capabilities and galvanic isolation, bridge-based isolated DC-DC converters offer scalable and adaptable solutions suitable for various industrial and renewable energy applications. The complicated transformer design, higher component count, and related switching losses, however, make implementation extremely difficult. Furthermore, in situations of dynamic operation and partial load, the efficiency of these topologies may change significantly. To improve performance and dependability across a range of operating circumstances, future research should focus on developing fault-tolerant control mechanisms, sophisticated modulation techniques, and smaller transformer architectures.

2) Forward DC-DC converter architecture

The forward architecture is a type of isolated DC-DC topology designed to provide a stable and galvanically separated DC output from an input source that may be variable or unregulated. Fundamentally, this configuration is widely applicable across transformer-coupled DC-DC converters, especially those engineered to supply multiple output rails [15]. The secondary side voltage in a forward DC-DC converter is primarily governed by the turn's ratio of the transformer, enabling output levels that can exceed, match, or fall below the input voltage. In terms of operation, it exhibits characteristics akin to those of the flyback converter. This resemblance underpins its extensive adoption in low-power scenarios owing to its advantageous dynamic performance, economical design, and simplicity of implementation. Nevertheless, the forward converter is generally less appropriate for high input voltage conditions, as the primary switch experiences elevated voltage stress, potentially resulting in increased switching losses and reduced overall efficiency [38].

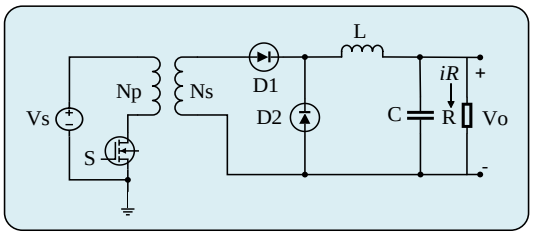


Fig.9. Basic schematic of a forward DC-DC converter configuration [39]

Fig. 9 presents the foundational layout of a forward-mode DC-DC converter. This schematic is adapted from the circuit diagrams in the referenced studies shown below the figure. In contrast to the flyback design, this topology enables current to flow simultaneously through the transformer's primary and secondary coils. At the same time, the switching device—typically a MOSFET—is activated, facilitating the direct delivery of power to the load. Energy is rapidly transferred to the secondary winding through the transformer's magnetic

coupling, with the output voltage determined by the transformer's turns ratio. During this interval, diode D1 is forward-biased, allowing current conduction, whereas diode D2 remains reverse-biased, blocking current. When the switching device is turned off, current ceases in both primary and secondary windings; however, the inductor on the secondary side sustains current flow by forward-biasing diode D2, while diode D1 becomes inactive. This operation, orchestrated by PWM, repeats continuously to maintain stable output voltage regulation.

Overall, the forward converter topology stands out for its simple control mechanism, compact structure, and suitability for low- to moderate-power applications that require multiple output voltages. Its efficiency and dynamic response make it a practical option for regulated DC supplies. However, its performance is limited under high input voltage conditions due to increased voltage stress across the switching device and greater switching losses. Future investigations should focus on optimizing voltage stress handling, exploring soft-switching techniques, and enhancing electromagnetic compatibility to extend its applicability to broader industrial domains.

3) Flyback DC-DC converter architecture

Originating as the isolated variant of the basic buck-boost converter, the flyback DC-DC converter is distinguished by its use of a transformer for galvanic isolation. Due to its straightforward design and inherent isolation properties, this topology is widely used in applications spanning a range of power levels, from low to medium. The incorporation of a high turns-ratio transformer allows substantial voltage amplification. Its widespread adoption in power electronics is attributed to its affordability, simple control mechanisms, compact footprint, and dependable electrical separation between the input and output circuits [40]–[43].

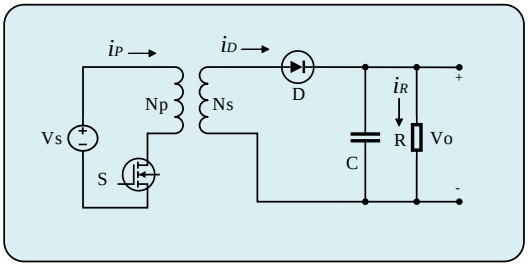


Fig.10. Common structural design of a flyback DC-DC converter [44], [45]

Fig. 10 illustrates the typical arrangement of a flyback DC-DC converter. This schematic is adapted from the circuit diagrams in the referenced studies shown below the figure. A notable benefit of this topology is its reduced component complexity, which facilitates a more straightforward implementation and cost efficiency. The capability to produce multiple output voltages is achieved by incorporating a few supplementary components—namely, an additional transformer winding, along with corresponding diodes and capacitors for each extra output. However, a notable limitation of the flyback design is its relatively limited lateral (cross-regulation) control, especially when compared to transformer-isolated converters based on buck-derived topologies such as forward and full-bridge converters [46]. While traditional flyback DC-DC converters operating at fixed switching frequencies offer simplicity in both design and control, they suffer from significant switching losses under voltage stress conditions and present difficulties in managing efficient power switching through MOSFETs [47]. The flyback converter operates by energizing the transformer's primary coil when the input switch is activated, causing magnetic energy to accumulate in the core. During this period, the voltage induced in the secondary coil is reversed, which keeps diode D1 in a non-conductive state and prevents power from reaching the output. The output capacitor maintains the load voltage while energy is being stored. When the switch turns off, the magnetic field collapses, flipping the polarity across the windings. This change allows diode D1 to conduct, transferring the stored energy to the secondary circuit and onward to the load via the rectifier. This charging and discharging process repeats cyclically, controlled by the switching device, until the core's energy is fully released or the switch closes again.

In essence, the flyback converter remains a cost-effective and compact solution, especially for applications with low to medium power demands and stringent isolation requirements. Its ability to support multiple output voltages and minimal component count enhances its appeal in consumer electronics and auxiliary power supplies. Nevertheless, its limited cross-regulation performance, relatively high switching losses, and suboptimal efficiency at elevated power levels constrain its widespread industrial deployment. Future research directions should address these limitations by exploring resonant switching techniques, digital control schemes, and improved magnetic core designs to enhance overall converter performance and robustness.

4) Resonant DC-DC converter architecture

Resonant DC-DC converter architectures represent a distinctive category within power electronic systems, recognized for their ability to deliver high efficiency and refined control in electrical energy conversion. Unlike conventional switched-mode power supplies, resonant converter architectures are specifically engineered to minimize switching losses, thereby significantly improving overall performance. This section aims to explore the core operating principles, structural elements, and application areas of resonant DC-DC converter technologies. These topologies are well-established solutions, particularly favored in industrial sectors that demand reliable and efficient DC voltage regulation and enhancement [48]. Resonant DC-DC converter topologies are characterized by their ability to handle significant power levels, generate smoother waveform profiles, deliver superior performance, operate at elevated switching frequencies through optimized switching schemes, provide high-quality power conversion, and require significantly fewer electromagnetic components such as transformers and filters [48], [49]. Recent advancements in boost rectifier technologies have contributed to improvements in DC-DC resonant converter designs. An extensive survey covering contemporary resonant converter architectures is provided in reference [50].

Resonant DC-DC converters primarily utilize LC (inductor-capacitor) tank circuits to enhance the switching performance at specific resonant frequencies, thereby significantly reducing energy losses during switching transitions. The primary

categories of resonant converters consist of the series resonant and parallel resonant configurations. Series-resonance-based DC-DC converters are primarily used for voltage step-up applications, utilizing high-frequency operation to minimize switching losses and achieve superior output efficiency. In contrast, parallel resonant converters are preferred in scenarios that require a stable and regulated output voltage, thereby enhancing system stability. Beyond these fundamental types, resonant converter designs also encompass hybrid structures such as mixed resonant, resonant boost, and resonant SEPIC converters. In Fig.11, a combined converter is shown that integrates a boost stage with a series resonant circuit, linking a standard power input to the resonant network to maximize energy conversion efficiency. This schematic is adapted from the circuit diagrams in the referenced studies shown below the figure. When the input voltage remains beneath the upper limit, this topology behaves like a buck converter, employing phase control on the primary winding alongside combined conversion on the secondary side.

In contrast, at lower input voltages, it operates similarly to a boost converter, featuring a full-bridge arrangement on the primary side coupled with a boosting circuit on the secondary side [51]. By substantially reducing switching losses, resonant DC-DC converters achieve superior energy efficiency that frequently surpasses that of conventional converter topologies. Their unique operational principles significantly contribute to lowering high-frequency electromagnetic noise, thereby mitigating electromagnetic interference issues. Such converters ensure reliable and steady performance across varying load demands, rendering them adaptable to a broad spectrum of applications. Despite these advantages, resonant converter designs come with particular challenges. The complexity involved in designing resonant circuits demands advanced technical expertise, which can complicate the development process.

Additionally, the need for specialized components and precise tuning of design parameters often results in increased production costs. Resonant DC-DC converters have found significant applications across various sectors, including high-voltage power systems, electric vehicles, and medical devices. To conclude, resonant DC-DC converters represent a notable leap forward in power electronics, thanks to their superior efficiency and reduced electromagnetic interference. With ongoing technological advancements, the potential to create more advanced and effective converter architectures will grow, steering future investigations toward performance optimization and exploration of novel application areas.

High efficiency, less EMI, and smooth voltage control are all notable benefits of resonant DC-DC converters, particularly when high-frequency switching is involved. Their technical promise is demonstrated by the fact that they are suitable for delicate and high-power applications, such as medical equipment and electric cars. The intricacy of resonant circuit design, the requirement for precise tuning, and the high cost of specialized parts severely hinders adoption. Therefore, future research should focus on integrating sophisticated semiconductors and digital design automation tools to simplify design processes, create adaptive control algorithms, and lower implementation costs.

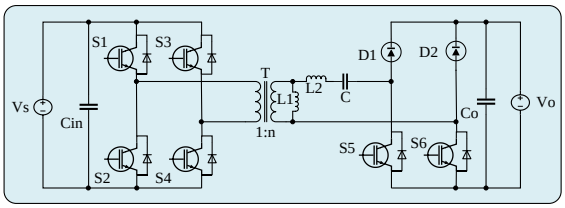


Fig.11. An integrated resonant DC-DC converter topology [51], [52]

5) Push-pull DC-DC converter architecture

The push-pull converter architecture is a key topology in power electronics, highly regarded for its efficiency and adaptability in adjusting DC voltage levels in either direction. This design utilizes a power transformer combined with PWM control and a specialized regulation scheme to manage input voltage transformation. While various forms of push-pull converters exist, this topology is particularly advantageous in applications demanding galvanic isolation and substantial power handling. It is found extensive use in sophisticated setups that incorporate renewable energy sources, such as pumps, solar panels, and battery management systems. Employing push-pull converters enables seamless integration of these diverse power sources with both traditional electrical grids and modern smart grid technologies. The demonstrated applications highlight the practical significance of push-pull converters in managing power flow across a range of electrical networks, including off-grid and remote installations lacking conventional infrastructure [53]. Fig. 12 depicts the fundamental circuit layout of a push-pull DC-DC converter. This schematic is adapted from the circuit diagrams in the referenced studies shown below the figure.

Push-pull DC-DC converter configurations are distinguished by their simple circuit architecture, which enables effective voltage regulation, galvanic isolation, and the achievement of high voltage transformation ratios. Such configurations leverage low-side switching, streamlined gate driver circuits, and enhanced transformer efficiency, making them particularly effective for applications operating at low voltage levels [54]–[56]. Moreover, owing to these beneficial features, push-pull converters are frequently utilized in applications ranging from medium to high power levels [57].

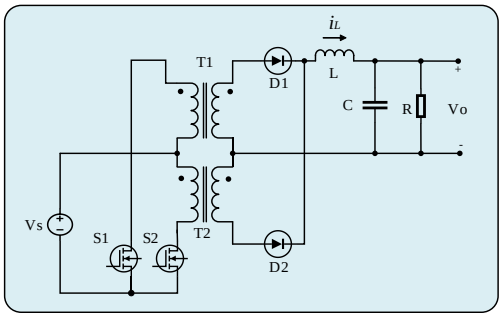


Fig.12. A basic push-pull DC-DC converter design [58]

Push-pull DC-DC converters achieve superior efficiency primarily by ensuring effective transformer coupling and minimizing switching-related losses. The transformer also provides inherent galvanic isolation between the input and output, contributing to enhanced safety and system dependability. This topology is capable of stepping voltages up and down, making it adaptable to a wide range of applications.

The operational principle involves alternately activating each half of the transformer's primary winding using two semiconductor switches, commonly MOSFETs or IGBTs. When the first switch is turned on, the current passes through one portion of the primary winding, inducing voltage on the secondary side. Once this switch is turned off, the second switch is turned on, allowing the current to flow through the opposite half of the primary winding. This alternating activation ensures a continuous flow of energy to the load.

On the secondary coil, the induced voltage undergoes rectification and smoothing through filtering elements to yield a consistent DC output. The transformation of the input voltage at the primary winding depends directly on the switching conditions of the two semiconductor devices, S1 and S2. While S1 is active, the transformer's magnetic flux increases steadily, generating a voltage on the secondary side. At this moment, diode D2 becomes conductive, allowing current to pass, whereas diode D1 stays non-conductive. The output inductor stores energy during this interval and supplies current to the load capacitor. Conversely, when S2 is activated and S1 is off, the transformer flux decreases in the opposite direction, reversing the polarity on the secondary winding. In this phase, diode D1 turns on, while diode D2 turns off, allowing current to flow through D1 and the inductor, thereby maintaining continuous energy delivery to the output. The output capacitor alternately charges and discharges, smoothing the voltage supplied to the load. This repetitive switching pattern persists uninterrupted, guaranteeing consistent power delivery and maintaining voltage stability [58].

The push-pull DC-DC converter is widely recognized for its effective voltage regulation, high efficiency, and inherent galvanic isolation, particularly in medium- to high-power applications. Its symmetrical structure and low-side switching simplify gate drive design and enhance transformer utilization. Despite these strengths, the converter faces limitations related to transformer core saturation, voltage imbalance, and high switch stress under asymmetric loading conditions. Consequently, future studies should focus on developing advanced core reset strategies, implementing real-time current balancing techniques, and exploring control schemes that improve stability under dynamic load variations.

B. Non-isolated DC-DC Converter Topologies

Non-isolated DC-DC conversion topologies constitute an essential class within power electronics, distinguished by their ability to regulate voltage levels without requiring galvanic isolation. Unlike their isolated counterparts, which employ high-frequency transformers to electrically decouple the input and output stages, non-isolated architectures feature a continuous conductive path, allowing direct energy transfer between the source and load terminals. In isolated designs, energy transfer is facilitated through transformer action, ensuring safety and isolation, particularly in high-voltage or sensitive applications. Conversely, non-isolated architectures eliminate the need for a transformer, offering simpler construction, reduced size, and cost-effective operation. These designs are particularly advantageous in systems where electrical isolation is not a mandatory requirement and space or efficiency constraints are critical [9].

Non-isolated DC-DC converter architectures are highly regarded for their superior efficiency, reduced physical dimensions, and cost-effective implementation, making them integral to a broad spectrum of modern electronic applications. The absence of galvanic isolation simplifies their circuit configuration, resulting in enhanced performance and compactness. A wide spectrum of non-isolated converter architectures, including buck, boost, buck-boost, Ćuk, SEPIC, and Zeta configurations, have been engineered to accommodate a range of voltage conversion and current regulation demands across various power processing scenarios. Consequently, these converters play a vital role across multiple domains, ranging from advanced power supply infrastructures to compact, portable electronic systems. Numerous studies in the literature have focused on fault detection and diagnostic methods specific to these converter types [59]. While isolated converters offer advantages in terms of electrical separation and safety, non-isolated designs often present a more pragmatic solution due to their simplicity and efficiency. Despite certain limitations such as constrained voltage gain, extended duty cycles, and occasional reliance on auxiliary components, galvanically non-separated DC-DC converter architectures remain a compelling alternative to their isolated counterparts [9].

1) Buck DC-DC converter architecture

The buck converter stands as a cornerstone configuration within power electronic systems, purpose-built to step down elevated input voltages to a lower, regulated output while achieving commendable efficiency levels. As a member of the non-isolated DC-DC converter family, this topology maintains a direct electrical linkage between source and load, inherently ensuring that the output voltage consistently remains below the input potential [60], [61]. Due to its inherent simplicity and favorable dynamic characteristics, namely its uniform structure and minimum phase behavior, the buck DC-DC converter has seen widespread adoption across various industrial applications. Often designated as a step-down regulator, this configuration is extensively employed across a broad array of electronic architectures due to its simplicity, reliability, and effective voltage reduction capability [62]. The compact structure of buck DC-DC converters is particularly advantageous in miniaturized electronic applications, as it reduces the reliance on bulky passive elements. This topology plays a crucial role in contemporary power electronics, significantly contributing to efficient voltage regulation and power conversion. Its versatility, compactness, and cost-effectiveness have made it a widely adopted solution across various sectors. Despite these advantages, the buck converter's reliance on hard switching mechanisms and reduced efficiency at elevated switching frequencies can result in increased electromagnetic interference, posing challenges for high-frequency applications [63].

Fig. 13 depicts the internal arrangement of a conventional buck-type DC-DC converter. This schematic is adapted from the circuit diagrams in the referenced studies shown below the figure. Upon the activation of the switching element, electrical power is routed from the input source toward the output terminal. Initially, the inductor absorbs energy, resulting in a delayed rise in load current; nevertheless, this is followed by a gradual build-up in both the voltage across the output capacitor

and the current delivered to the load. Throughout this interval, the diode remains in a reverse-bias state, as the voltage at its cathode exceeds that at the anode, thereby inhibiting current conduction.

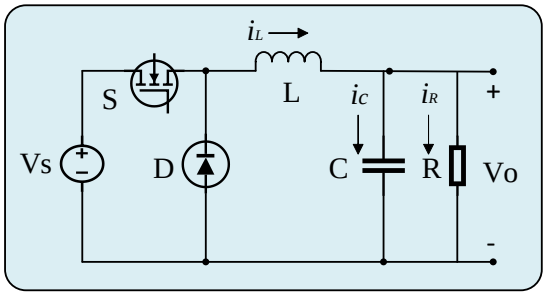


Fig.13. Conventional DC-DC buck converter design [64]–[66].

When the switching device is deactivated, the inductor, having previously accumulated magnetic energy, undergoes a polarity inversion that facilitates sustained current flow despite the switch being in its open state. This energy discharge drives current through the forward-bias diode and continues to energize the load. As the magnetic field within the inductor collapses, the diminishing energy gradually reduces the current until it is depleted. At this point, the output capacitor takes over the responsibility for maintaining voltage regulation by supplying current to the load. This sequence of magnetically and capacitively sustained conduction recurs with each switching interval, thereby ensuring a stable and continuous DC output.

In summary, the buck converter's ease of use, dependability, and effective step-down voltage control make it one of the most basic and popular non-isolated topologies. It is ideal for portable applications and embedded systems due to its compact size and low component count. Its scalability for sophisticated power systems is limited by its hard-switching nature, which increases EMI and reduces efficiency at high frequencies. Future developments should prioritize EMI suppression strategies, soft-switching approaches, and integration with digital control systems to improve their suitability for high-performance and noise-sensitive settings.

2) Boost DC-DC converter architecture

A boost-type DC-DC converter is an advanced energy conversion configuration engineered to elevate a lower direct current input voltage to a significantly higher output level. This topology is widely used in modern electronic and electrical systems, where stringent requirements for voltage amplification and efficient power conditioning exist. Distinguished by its step-up functionality and absence of galvanic isolation, the boost converter operates through controlled energy accumulation in inductive elements during the switch-on period, followed by energy release during the switch-off phase. Its compact form factor, combined with high operational efficiency, renders it indispensable in space-constrained and performance-critical environments, such as portable electronics, electric vehicles, and renewable energy interfaces. By regulating the switching duty cycle through the coordinated operation of passive elements and semiconductor switching devices, the boost converter effectively ensures the desired voltage elevation [67], [68]. Fig. 14 depicts the prototypical topology of a boost DC-DC converter, which relies on the intricate orchestration of switching elements and inductive energy storage to facilitate voltage elevation. This schematic is adapted from the circuit diagrams in the referenced studies shown below the figure. The operational sequence can be delineated into distinct yet interrelated phases: the inductor's energy accumulation during the conduction interval, the subsequent energy delivery to the load when the switch transitions to the off state, and the precise voltage regulation achieved via output filtering and control circuitry. Upon closing the power switch, current flows through the inductor, generating a growing magnetic field that stores energy. When the switch is disengaged, this stored magnetic flux propels current continuity, thereby sustaining power flow to the load and effectuating an output voltage exceeding that of the source. The converter's output voltage magnitude is governed by modulating the duty cycle of the PWM signal applied to the switch; augmenting this duty cycle proportionally raises the output voltage, while reducing it produces the inverse effect.

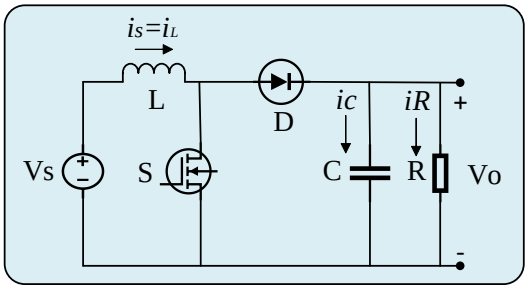


Fig.14. Traditional architectures for DC-DC boost converter [10], [69]

Boost DC-DC converters are critical elements within modern power management frameworks, due to their high efficiency and compact form factors, which effectively address the diverse energy demands of contemporary technologies. As advancements in electrical power systems and electronic devices continue to accelerate, these converter topologies are expected to exhibit enhanced performance metrics and expanded applicability, thereby maintaining their pivotal role in the continually advancing domain of electrical and electronic engineering.

Overall, the boost converter is a highly effective topology for stepping up low input voltages, offering a compact and efficient solution widely used in electric vehicles, battery systems, and renewable energy platforms. Its straightforward control and minimal component requirements enhance design simplicity. Nevertheless, drawbacks such as high output voltage ripple, increased switch stress, and limited efficiency at low load conditions constrain its performance in high-power and precision-demanding applications. Future research should address these limitations by incorporating interleaved structures, digital predictive control, and advanced wide-bandgap devices to improve efficiency, dynamic response, and thermal stability.

3) Buck-Boost DC-DC converter architecture

Buck-boost converter topologies hold a pivotal position within power electronics due to their inherent capability to seamlessly regulate output voltage levels both above and below

the input supply. This dual-mode operation facilitates dynamic voltage adaptation, enhancing overall system flexibility and conversion efficacy. In non-isolated buck-boost implementations, the converter operates in boost mode when the output voltage exceeds the input voltage and switches to buck mode when the output falls below the input level. The transition between these operating states is determined by the modulation index of the PWM control signal that governs the switching device. As a result, precise voltage regulation is achievable, enabling the output to be adjusted responsively with respect to the input voltage magnitude [70], [71]. Buck-boost converter configurations are highly regarded across various industrial sectors for their exceptional versatility, enabling precise modulation of voltage levels that either step up or step down the input voltage while maintaining circuit simplicity through a reduced component count [72], [73]. Fig. 10 depicts the fundamental structural layout of a buck-boost DC-DC converter. This schematic is adapted from the circuit diagrams in the referenced studies shown below the figure.

Similar to other DC-DC converter frameworks, the buck-boost converter's operation is divided into two primary intervals, governed by the switching device's conduction status: the energy accumulation interval and the energy delivery interval. The interplay between these phases underpins the converter's functional dynamics. The core elements of this topology include a semiconductor switch, an inductive coil, a diode, and an energy storage capacitor. The converter toggles between buck and boost modes depending on the switch position. When the switch is engaged, the current flows through the inductor, resulting in the storage of magnetic energy within its core, while the diode remains reverse-biased, preventing current passage. Under steady-state conditions, with a zero average voltage across the inductor and zero net current through the capacitor, the opening of the switch reverses the polarity of the inductor voltage. Consequently, the diode transitions to a forward-bias state, enabling the inductor to discharge its stored magnetic energy into the capacitor and the load. The capacitor replenishes its charge while simultaneously supplying the load. This cyclical process of switching on and off repeats continuously, with the power switch regulated by PWM to control output voltage and current.

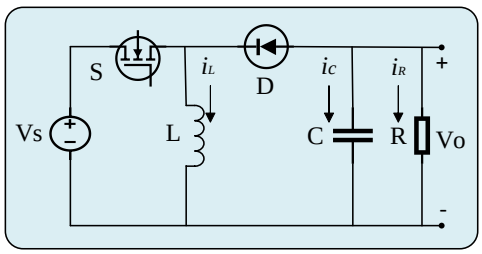


Fig.15. A buck-boost DC-DC converter's basic configuration [74]–[76].

Buck-boost converters offer notable advantages, including operational versatility across extensive voltage ranges and high adaptability to varying application requirements. Their compatibility with diverse input and output voltage levels facilitates integration into a wide range of systems, such as energy management networks, battery storage units, renewable energy platforms, and electric vehicle infrastructures. Furthermore, these converters are favored for their compact form factors and superior energy conversion efficiencies. Anticipated advancements in these topologies, along with integration with emerging technologies, are expected to drive significant enhancements in the functional effectiveness and efficiency of modern power transformation mechanisms.

The buck-boost topology stands out due to its bidirectional voltage handling capability, enabling it to function efficiently under both step-up and step-down scenarios. Despite its widespread applicability and structural simplicity, it tends to suffer from performance degradation at extreme duty ratios, particularly in terms of efficiency and output ripple. To address these shortcomings, novel control strategies and the incorporation of low-loss passive elements are being actively explored. Enhancing transient response and minimizing ripple noise remain key focal points for future development.

4) Cuk and SEPIC DC-DC converter architectures

Within the broad spectrum of DC-DC converter architectures, the Cuk and SEPIC topologies stand out due to their versatile operational capabilities. The non-isolated Cuk converter notably produces an output voltage with polarity inverse to the input, while offering the flexibility to generate output voltages either higher or lower than the input magnitude [77]. The Cuk converter topology is uniquely recognized for its dual capability to both elevate and reduce voltage levels. Key features include a wide conversion ratio, capacitive energy transfer that inherently isolates the input from the output, seamless buck-boost functionality, continuous current conduction on both input and output circuits, the use of low-side switching devices, and the predominant use of capacitive elements for power transmission [78], [79]. A distinguishing attribute of the Cuk converter is its ability to maintain stable output voltage regulation regardless of fluctuations in the input voltage. Furthermore, this topology naturally excels in scenarios that demand a negative polarity output, offering this characteristic inherently without the need for additional circuitry.

Illustrated in Fig. 16(a) is the elementary topology of the Cuk converter. The SEPIC converter, devoid of galvanic isolation, achieves voltage transformation by rapidly toggling its semiconductor switching element. This schematic is adapted from the circuit diagrams in the referenced studies shown below the figure. This arrangement allows for the adjustment of the output voltage to levels either above or below the input voltage, while maintaining the same polarity throughout the process [77]. The SEPIC converter architecture employs an integrated buck-boost approach, distinguished by maintaining a nearly steady input current and producing an output voltage that preserves the input's polarity, in contrast to the polarity-inverting behavior of the Cuk converter. Compared to many conventional DC-DC converter designs, the SEPIC converter typically offers superior efficiency [80]. Fig.16(b) illustrates a relatively simple configuration of a SEPIC DC-DC converter. The switching actions of power semiconductors in these converter topologies are precisely regulated by PWM techniques to maximize operational efficiency. The Cuk converter's behavior is divided into two distinct intervals determined by the status of the switching device. As shown in Fig. 16(a), during the ON interval, current flows through inductor L1, while the diode is reverse-blocked, preventing

current passage. When the switch toggles to the OFF state, the diode becomes conductive, enabling capacitor C1 to charge with a voltage polarity identical to the input, thanks to the closed loop on the diode's left side. Simultaneously, the energy stored in L1 during the ON phase is transferred to the load through inductor L2 and capacitor C2 via the diode.

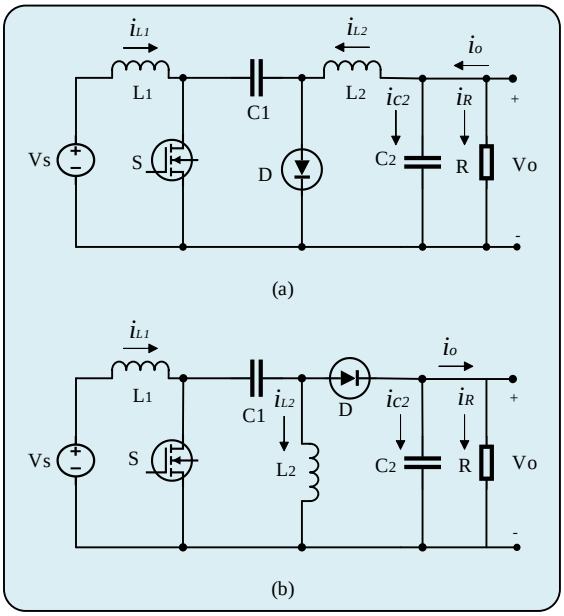


Fig.16. The fundamental Cuk and SEPIC DC-DC converter schemes: (a) Cuk converter topology, (b) SEPIC converter topology [81]–[84].

Conversely, the SEPIC converter, illustrated in Fig. 16(b), shares a comparable operational concept but with nuanced differences. With the switch closed, the input voltage causes current to flow through L1, while capacitor C1, charged in the previous cycle, supplies current to the adjacent circuit segment. The current flowing through L2 induces a voltage that opposes the polarity of the input source. When the switch opens, L1's stored energy reverses polarity and injects power into the circuit, ensuring continuous energy delivery. The combined current from the input source and L1 charges capacitor C2, which subsequently discharges during the next switching cycle. Similarly, inductor L2 changes polarity, contributing to the total load current. Capacitor C2 delivers power to the load, preserving an output voltage polarity consistent with that of the input source. In both converter topologies, the input and output voltages share identical polarity at the load terminals. Control is continuously maintained via PWM signals that switch the power device on and off, thereby regulating energy transfer cyclically.

Renowned for their adaptability and efficiency, the Cuk and SEPIC converter architectures serve as pivotal elements in power electronics, delivering robust and dependable operation across diverse domains, including renewable energy systems, DC microgrid management, electric vehicle charging platforms, and advanced power supply networks. Their ability to handle fluctuating input and output conditions underscores their value in both consumer-grade electronics and industrial-scale operations. Amid the growing worldwide focus on optimizing energy use, these converter configurations are anticipated to serve as fundamental components driving the progression of advanced and highly efficient power management frameworks. The rapid advancement of semiconductor technology presents an opportunity to further refine these topologies in terms of operational efficiency, compactness, and ease of integration. Consequently, future research should aim to improve their functional adaptability, minimize their physical footprint, and enhance their maintainability. With ongoing innovation, Cuk and SEPIC converters are likely to remain fundamental components in shaping the upcoming generation of sustainable and adaptive power structures.

Both the Cuk and SEPIC converter designs offer enhanced input-output continuity and polarity-specific control, making them suitable for a range of power conditioning applications. However, the increased complexity arising from the addition of inductors and capacitors introduces trade-offs in terms of cost, size, and control difficulty. While these architectures offer smoother input currents and stable output profiles, practical deployment often demands refined component selection and efficient layout strategies. Ongoing efforts in miniaturization and digital control optimization may facilitate broader adoption in compact electronic systems.

5) Zeta DC-DC converter architecture

The fundamental role of DC-DC converter architectures lies in their ability to modulate and convert DC voltage levels to desired magnitudes. Among the spectrum of available topologies, the Zeta converter is distinguished by its unique configuration and multifaceted operational capabilities. As a fourth-order, nonlinear, and non-isolated power electronic system, the Zeta converter adeptly facilitates both voltage step-up and step-down functions, adapting seamlessly to varying application demands [85], [86]. Analogous to the Cuk and SEPIC converter families, the Zeta DC-DC converter preserves a unipolar voltage orientation at its input and output nodes. This intrinsic characteristic enables seamless functionality in both voltage elevation (boost) and reduction (buck) modes. Fundamentally, the Zeta topology guarantees congruent polarity between its input and output voltages, thereby broadening its utility across diverse power management applications [87], [88]. Zeta DC-DC converter architectures outperform SEPIC converters due to their integrated continuous feedback control, which facilitates improved voltage stability, enhanced load transient response, reduced output voltage ripple, and simplified rectification processes [89]. Fig.17 depicts the circuit schematic of a conventional Zeta DC-DC converter topology. This schematic is adapted from the circuit diagrams in the referenced studies shown below the figure.

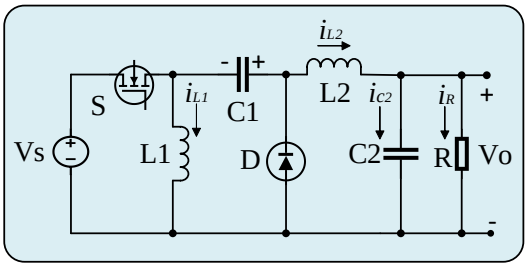


Fig.17. Traditional Zeta DC-DC converter structure [88], [90]–[93]

A conventional Zeta DC-DC converter configuration comprises a power switch, a diode, a pair of inductors, and a capacitor. This configuration enables both energy transfer and voltage regulation to be effectively managed. By incorporating two inductive components, the Zeta converter structure enables dual-mode operation, allowing for both boosting and bucking, making it highly adaptable for various applications. Output voltage regulation is achieved by dynamically adjusting the switching device's operation, enabling seamless toggling between voltage increase and decrease modes. When the switch is engaged, the diode is blocked from conducting, permitting both inductors L1 and L2 to accumulate magnetic energy in alignment with the polarity of the input supply. During this phase, the source voltage Vs is directly imposed across these inductors. Upon deactivation of the switch, the voltage polarity across L1 and L2 inverts, activating the diode to enter conduction mode and allowing continuous power transfer to the output load [90]. In the off state of the switching component, the semiconductor diode governs the current pathway, ensuring effective energy delivery throughout the system. The Zeta converter topology incorporates a pair of inductors that are pivotal for energy retention, enabling the converter to function seamlessly in both voltage elevation and reduction regimes. Additionally, the inclusion of a capacitor enhances the voltage stability by smoothing output voltage fluctuations and mitigating transient oscillations, ensuring a more consistent and regulated output.

The Zeta converter represents a refined solution for applications that demand stable voltage levels under dynamic load and input conditions. Its capability to perform both buck and boost actions with maintained polarity makes it favorable in regulated supply systems. Nevertheless, its relatively complex control characteristics and dependence on dual inductors can limit its attractiveness in cost-sensitive scenarios. Advancements in integrated magnetics and control circuitry may help simplify implementation and expand their utility across more compact and efficient designs.

C. Multi-level DC-DC Converter Architectures

Multi-level DC-DC converters represent a significant advancement in modern power conversion technologies and have emerged as transformative solutions within semiconductor-based power electronic systems. These advanced converter structures are designed to minimize harmonic distortion, improve energy efficiency, and enable the generation of higher output voltage levels. The fundamental concept behind multi-level converter technology involves intentionally combining several distinct voltage steps to achieve the targeted output, thereby reducing voltage stress on switching devices and enhancing the overall system efficiency.

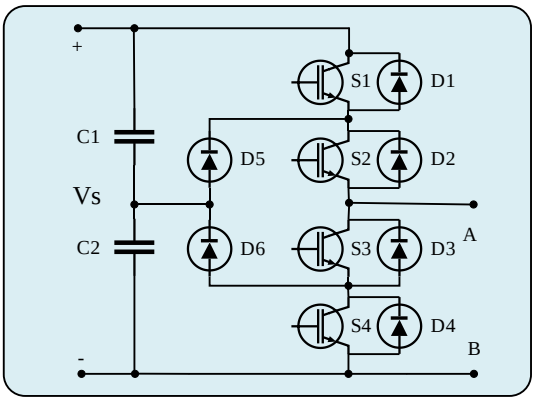


Fig.18. Multi-level DC-DC converter with diode-based voltage clamping [94]

Multi-level DC-DC converters are available in both isolated and non-isolated variants. Within the realm of non-isolated designs, notable configurations encompass diode-clamped, flying-capacitor, and cascaded H-bridge topologies. The diode-clamped approach is widely appreciated for its straightforward architecture and economic feasibility. Nevertheless, its applicability at higher voltage ranges is constrained due to the increasing number of clamping diodes required, which adds complexity and poses challenges for scalability in practical deployments. While flying capacitor converter topologies have demonstrated acceptable performance within certain operating conditions, their efficiency remains constrained. Moreover, the utilization of this architecture at medium voltage levels presents challenges, primarily due to the excessive number of flying capacitors required, which increases system complexity and physical footprint [95]. Historically, flying capacitor converter topologies have been predominantly utilized in applications involving high-voltage, high-power DC-to-AC conversion. These multi-level converter topologies offer significant benefits in efficiency and power handling capabilities when compared to many traditional converter designs [96], [97].

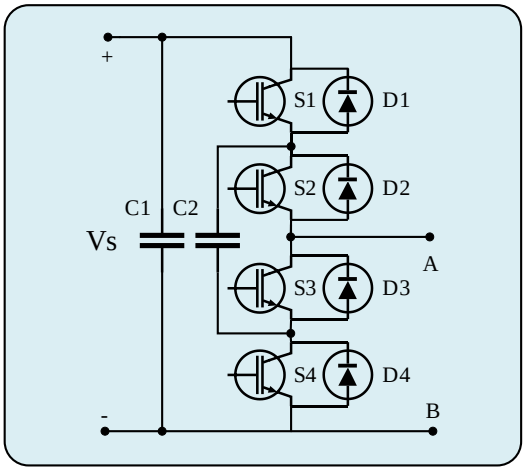


Fig.19. Flying capacitor-based multi-level DC-DC converter architecture [94]

Among multi-level converter methodologies, the cascaded H-bridge configuration stands out as one of the most extensively adopted architectures in practical applications. This configuration is formed by connecting multiple identical H-

bridge units in series. Cascaded H-bridge DC-DC converters offer the benefit of achieving elevated output voltage levels while maintaining high power quality. These results are obtained by integrating several discrete converter cells within the overall structure. Additionally, the inherent modularity of cascaded H-bridge designs provides excellent fault tolerance, enhancing their suitability for various practical applications [98]. Due to its ability to generate high-level output voltages and maintain superior waveform fidelity, the cascaded H-bridge architecture is particularly advantageous for medium- to high-power applications [99]–[101]. Fig. 18, Fig. 19, and Fig. 20, respectively, depict the circuit configurations of the diode-clamped, flying capacitor, and cascaded multi-level DC-DC converter topologies.

In multi-level DC-DC converter designs, an array of capacitors arranged on the secondary side partitions the primary voltage into multiple discrete stepped levels. The series connection of these capacitors establishes defined voltage increments. As illustrated in Fig. 18, the diode-clamped multi-level DC-DC converter employs interconnected capacitors to divide the input voltage into three separate levels. This schematic is adapted from the circuit diagrams in the referenced studies shown below the figure. Within this topology, diodes function as voltage clamps, maintaining the output voltage at specific set points. The architecture comprises power semiconductor switches, such as MOSFETs or IGBTs, alongside diodes and capacitors, with the switching devices orchestrating the current flow to precisely regulate the output voltage. High-speed switching components are essential for optimal operation. Diodes play a crucial role by limiting the voltage at each level, enabling precise regulation of the output voltage steps. Capacitors attached to each voltage level serve to stabilize the voltage, minimizing fluctuations and improving overall system efficiency.

Diode-clamped multi-level converters offer several key benefits, including minimized harmonic distortion, improved operational efficiency, robust handling of high voltage and power demands, and reduced thermal loading. These attributes render them especially well-suited for use in electric vehicle drives, energy storage solutions, renewable energy systems, and power distribution networks. However, challenges remain, including their complex circuit topology, higher manufacturing costs, and difficulties in maintaining voltage balance across capacitor levels. Advancements in semiconductor components and sophisticated control techniques are poised to resolve current obstacles, enabling substantial enhancements in the efficacy and broader implementation of diode-clamped multi-level DC-DC converters within forthcoming power electronic frameworks.

As demonstrated in Figure 19, the flying capacitor multi-level DC-DC converter topology achieves its secondary-side output voltage levels by allocating the voltage across a series-connected capacitor network. This schematic is adapted from the circuit diagrams in the referenced studies shown below the figure. This architecture's distinct advantage lies in its ability to generate multiple voltage levels without relying on additional voltage clamping devices, such as diodes. Rather than relying solely on passive components, this topology incorporates flying capacitors that are dynamically engaged and disengaged within the circuit, allowing for accurate construction of discrete voltage steps. Although flying capacitor multi-level converters present challenges in capacitor voltage balancing and operational complexity, their performance benefits make them highly suitable for applications demanding high voltage, high power, and improved efficiency. These topologies effectively minimize harmonic distortion, rendering them attractive for integration into industrial power systems, electric vehicle powertrains, energy storage solutions, and renewable energy installations. Nonetheless, issues such as maintaining voltage equilibrium across capacitors, increased complexity of the control system, and higher implementation costs remain significant considerations. Progress in semiconductor devices and sophisticated control techniques is projected to improve the feasibility and widespread deployment of flying capacitor multi-level DC-DC converters in next-generation power electronic systems.

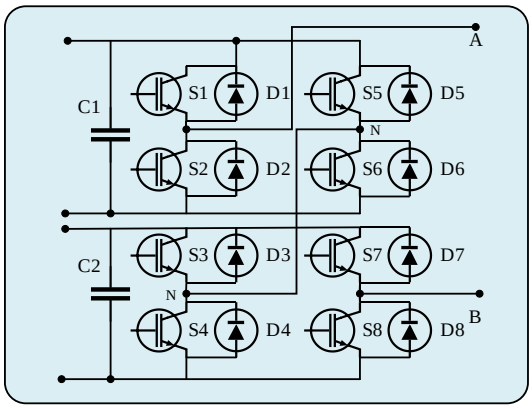


Fig.20. Architecture of cascaded multi-level DC-DC converter system [94]

The cascaded multi-level DC-DC converter topology, depicted in Fig. 20, consists of multiple H-bridge converter units connected in sequence, each supplied by a distinct power source. This schematic is adapted from the circuit diagrams in the referenced studies shown below the figure. This architecture represents an advanced methodology that employs several converter stages in series to achieve elevated output voltage levels while simultaneously reducing harmonic distortion and enhancing overall efficiency. Within cascaded multi-level systems, each converter module contributes a fraction of the total primary voltage by either stepping it up or down. These individual voltage increments are then combined through serial connection to synthesize a high-voltage output. Such converters are well-suited for high-power applications demanding substantial voltage conversion, superior efficiency, reduced harmonic distortion, and operational flexibility. These converter types excel in high-power scenarios requiring significant voltage transformation, enhanced efficiency, minimized harmonic content, and versatile operational capabilities. The modular nature of cascaded converters facilitates the management of high voltages by dividing the voltage transformation across multiple stages, each operating at a lower voltage level. To ensure optimal performance, challenges related to isolation, thermal management, synchronization, and electromagnetic interference mitigation must be effectively addressed during the design process. Progressive breakthroughs in semiconductor manufacturing and advanced control frameworks are anticipated to establish

cascaded multi-level DC-DC converter architectures as integral elements within emerging power conversion technologies, with particular relevance to sectors such as electric transportation, sustainable energy integration, uninterrupted power supply systems, and heavy industrial power infrastructure. Table I presents a comparative analysis of the conventional DC-DC converter topologies discussed in this study, evaluated based on key performance indicators, including efficiency, cost, circuit complexity, application suitability, bidirectional capability, and voltage gain [10], [14], [102], [103].

Multi-level converter systems play a significant role in the advancement of high-voltage and high-efficiency power conversion infrastructure. By layering voltage steps, they successfully reduce switching stress and suppress harmonic content, which are crucial in demanding power applications. Still, their elevated design intricacy, component count, and control synchronization requirements can hinder widespread deployment, especially in cost-critical or space-constrained environments. Future research is expected to center on modularization, reliability, and simplified control logic to unlock their broader potential in sustainable energy platforms.

TABLE I
COMPARATIVE EVALUATION OF DC–DC CONVERTER TOPOLOGIES

Topology	Efficiency	Cost	Complexity	Application Suitability	Bidirectional Capability	Voltage Gain
Buck (Step-Down)	High	Low	Low	Battery chargers, regulated low-voltage loads	Limited	Step-down (Vout < Vin)
Boost (Step-Up)	Medium-High	Low	Low	PV converters, LED drivers	Moderate	Step-up (Vout > Vin)
Buck-Boost	Moderate	Moderate	Moderate	EV powertrains, portable electronics	Yes	Step-up/down
Ćuk	Moderate	Moderate	High	Audio systems, power factor correction	Yes	Negative output, flexible gain
SEPIC (Single-Ended Primary Inductor Converter)	Moderate	High	High	Automotive systems, renewable energy integration	Yes	Step-up/down (positive output)
Zeta	Moderate	High	High	LED lighting, constant current applications	Limited	Step-up/down
Forward	High	Moderate	Moderate	Isolated telecom and embedded power supplies	No	Fixed-ratio step-down
Flyback	Moderate	Low	Moderate	Small SMPS, low-power instrumentation	No	Wide gain range
Push-Pull	High	Moderate	High	Medium power supplies, symmetrical loads	Yes	High gain (with center-tap)
Half-Bridge	High	Medium-High	High	Medium-power converters, motor drives	Yes	Medium-to-high gain
Full-Bridge	High	High	Very High	High-voltage inverters, HVDC systems	Yes	High voltage gain
Single Active Bridge	High	High	High	Power transfer, battery management systems	Yes	Isolated, controllable gain
Dual Active Bridge	High	High	Very High	Bidirectional energy storage, fast EV chargers	Strong	Wide gain range
Boost Half-Bridge	High	High	High	High-voltage DC distribution	Yes	Step-up with isolation
Resonant (LLC, Resonant Boost, Hybrid Resonant, etc.)	Very High	High	Very High	Data centers, EV charging stations	Yes	Wide gain (frequency-dependent)

IV. Discussion and Future Directions

Despite the fact that DC-DC converter technologies have been the subject of extensive study, the majority of these studies focus on optimal operating conditions, often overlooking real-world issues such as environmental unpredictability, long-term reliability, and fault tolerance. Furthermore, it is difficult to make meaningful comparisons and occasionally results in contradictory findings in the literature when studies lack defined assessment parameters, such as efficiency, cost, control complexity, EMI levels, and power density. A uniform multi-criteria assessment system should be established to bridge this gap and enable objective, consistent evaluation of various converter types. A novel categorization strategy based on functional modularity, control flexibility, and integration possibilities with digital and AI-assisted control systems may also provide a more comprehensive and modern viewpoint than the conventional input-output voltage classifications. Future studies should concentrate on intelligent, adaptive control algorithms, modular converter architectures that work with hybrid energy systems, fault-tolerant designs, and ecologically friendly solutions. By moving forward in these areas, it will be easier to get over present obstacles and create converter designs that are more adaptable, dependable, and application-specific.

V. Conclusion

This study has provided a comprehensive analysis of DC-DC converter architectures, systematically classifying them into isolated and non-isolated categories. These converters have become indispensable components in modern power electronic systems due to their ability to regulate voltage with high precision and efficiency. The demand for compact, energy-efficient, and adaptable power conversion solutions is growing across various sectors, including industrial automation,

renewable energy integration, and consumer electronics, driven by the pursuit of higher performance and environmental sustainability. DC-DC converters play a pivotal role in minimizing energy dissipation, reducing thermal stress, and improving overall system reliability. Their applications
span from smart electrical grids and electric vehicles to portable devices and sustainable energy platforms. In renewable energy systems, for example, they stabilize variable outputs from sources like photovoltaic panels and wind turbines, facilitating smooth and reliable power delivery.

Looking ahead, the evolution of DC-DC converter technologies is expected to be shaped by several key research directions. One significant trend is the adoption of wide-bandgap semiconductor devices, such as Gallium Nitride (GaN) and Silicon Carbide (SiC). These materials offer superior electrical characteristics, including higher switching frequencies, lower conduction losses, and better thermal performance, compared to traditional silicon-based devices, enabling the development of next-generation high-efficiency power converters. Another promising avenue lies in integrating artificial intelligence (AI) and machine learning (ML) techniques into converter control strategies. AI-assisted algorithms have the potential to optimize real-time performance, adapt to dynamic load conditions, improve fault detection, and enhance energy management in complex systems. Moreover, the increasing electrification of transportation systems presents new challenges and opportunities for the development of DC-DC converters. Electric vehicles (EVs) demand high power density, bidirectional operation, and robust thermal management. Future research should focus on designing converters tailored for on-board chargers, traction inverters, and battery management systems. Microgrids and distributed energy systems also present a fertile ground for innovation. The need for reliable, flexible, and intelligent power conversion solutions in these decentralized networks highlights the importance of programmable and interoperable converter architectures that can seamlessly interact with various sources and loads.

In conclusion, future developments in DC-DC converter technology are likely to be driven by the convergence of material science innovations, intelligent control systems, and application-specific demands. By advancing switching techniques, incorporating low-loss components, and leveraging AI-driven adaptability, DC-DC converters will continue to be central to the development of energy-efficient and sustainable power systems. Their role in shaping the future of electric mobility, renewable integration, and smart energy infrastructures remains both foundational and transformative.

ACKNOWLEDGMENT

This study is derived from the doctoral dissertation of Abdurrahim Erat.

REFERENCES

[1] J. Chen, M.-K. Nguyen, Z. Yao, C. Wang, L. Gao, and G. Hu, "DC-DC Converters for Transportation Electrification: Topologies, Control, and Future Challenges," *IEEE Electrif. Mag.*, vol. 9, no. 2, pp. 10–22, 2021, doi: 10.1109/MELE.2021.3070934.

[2] T. Sutikno, A. S. Samosir, R. A. Aprilianto, H. S. Purnama, W. Arsadiando, and S. Padmanaban, "Advanced DC–DC converter topologies for solar energy harvesting applications: a review," *Clean Energy*, vol. 7, no. 3, pp. 555–570, Jun. 2023, doi: 10.1093/ce/zkad003.

[3] M. F. Akhtar, S. R. S. Raihan, N. A. Rahim, M. N. Akhtar, and E. Abu Bakar, "Recent Developments in DC-DC Converter Topologies for Light Electric Vehicle Charging: A Critical Review," *Applied Sciences*, vol. 13, no. 3. 2023. doi: 10.3390/app13031676.

[4] Z. W. Khan, H. Minxiao, C. Kai, L. Yang, and A. u. Rehman, "State of the Art DC-DC Converter Topologies for the Multi-Terminal DC Grid Applications: A Review," in *2020 IEEE International Conference on Power Electronics, Smart Grid and Renewable Energy (PESGRE2020)*, 2020, pp. 1–7. doi: 10.1109/PESGRE45664.2020.9070529.

[5] J. Wang, B. Wang, L. Zhang, J. Wang, N. I. Shchurov, and B. V Malozyomov, "Review of bidirectional DC–DC converter topologies for hybrid energy storage system of new energy vehicles," *Green Energy Intell. Transp.*, vol. 1, no. 2, p. 100010, 2022, doi: https://doi.org/10.1016/j.geits.2022.100010.

[6] M. Abolghasemi, I. Soltani, M. Shivaie, and H. Vahedi, "Recent advances of step-up multi-stage DC-DC converters: A review on classifications, structures and grid applications," *Energy Reports*, vol. 13, pp. 3050–3081, 2025, doi: https://doi.org/10.1016/j.egyr.2025.02.025.

[7] A. Tuluhong, Z. Xu, Q. Chang, and T. Song, "Recent Developments in Bidirectional DC-DC Converter Topologies, Control Strategies, and Applications in Photovoltaic Power Generation Systems: A Comparative Review and Analysis," *Electronics*, vol. 14, no. 2. 2025. doi: 10.3390/electronics14020389.

[8] M. Mezouari, M. Megrini, and A. Gaga, "High efficiency DC–DC converter for renewable energy integration and energy storage applications: A review of topologies and control strategies," *Control Eng. Pract.*, vol. 162, p. 106371, 2025, doi: https://doi.org/10.1016/j.conengprac.2025.106371.

[9] F. Mumtaz, N. Z. Yahaya, S. T. Meraj, B. Singh, R. Kannan, and O. Ibrahim, "Review on non-isolated DC-DC converters and their control techniques for renewable energy applications," *Ain Shams Eng. J.*, vol. 12, no. 4, pp. 3747–3763, 2021.

[10] A. Amir, A. Amir, H. S. Che, A. Elkhateb, and N. Abd Rahim, "Comparative analysis of high voltage gain DC-DC converter topologies for photovoltaic systems," *Renew. energy*, vol. 136, pp. 1147–1163, 2019.

[11] S. Dahale, A. Das, N. M. Pindoriya, and S. Rajendran, "An overview of DC-DC converter topologies and controls in DC microgrid," in *2017 7th International Conference on Power Systems (ICPS)*, 2017, pp. 410–415.

[12] A. Lavanya, J. D. Navamani, K. Vijayakumar, and R. Rakesh, "Multi-input DC-DC converter topologies-a review," in *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, 2016, pp. 2230–2233.

[13] M. Parvez, A. T. Pereira, N. Ertugrul, N. H. E. Weste, D. Abbott, and S. F. Al-Sarawi, "Wide bandgap DC–DC converter topologies for power applications," *Proc. IEEE*, vol. 109, no. 7, pp. 1253–1275, 2021.

[14] S. A. Gorji, H. G. Sahebi, M. Ektesabi, and A. B. Rad, "Topologies and control schemes of bidirectional DC–DC power converters: An overview," *IEEE Access*, vol. 7, pp. 117997–118019, 2019.

[15] F. L. Luo and H. Ye, *Advanced dc/dc converters*. crc Press, 2016.

[16] S. Ikeda and F. Kurokawa, "Isolated and wide input ranged boost full bridge DC-DC converter with low loss active snubber," in *2017 IEEE Energy Conversion Congress and Exposition (ECCE)*, 2017, pp. 2213–2218.

[17] F. Krismer and J. W. Kolar, "Efficiency-optimized high-current dual active bridge converter for automotive applications," *IEEE Trans. Ind. Electron.*, vol. 59, no. 7, pp. 2745–2760, 2011.

[18] A. Gnanasaravanan and M. Rajaram, "Dynamic response analysis and output voltage control of asymmetric half bridge DC–DC converter for low voltage applications," *Int. J. Electr. Power Energy Syst.*, vol. 43, no. 1, pp. 774–778, 2012.

[19] A. G. Saravanan and M. Rajaram, "Fuzzy controller for dynamic performance improvement of a half-bridge isolated DC–DC converter," *Neurocomputing*, vol. 140, pp. 283–290, 2014.

[20] R. W. A. A. De Doncker, D. M. Divan, and M. H. Kheraluwala, "A three-phase soft-switched high-power-density DC/DC converter for high-power applications," *IEEE Trans. Ind. Appl.*, vol. 27, no. 1, pp. 63–73, 1991.

[21] S. Shao, H. Chen, X. Wu, J. Zhang, and K. Sheng, "Circulating current and ZVS-on of a dual active bridge DC-DC converter: A review," *Ieee Access*, vol. 7, pp. 50561–50572, 2019.

[22] A. Kumar, A. H. Bhat, and P. Agarwal, "Review and comparative

analysis of dual active bridge isolated DC to DC converter with different control techniques," *Int. J. Ind. Electron. Drives*, vol. 4, no. 2, pp. 69–84, 2018.

[23] Q. Bu, H. Wen, J. Wen, Y. Hu, and Y. Du, "Transient DC bias elimination of dual-active-bridge DC–DC converter with improved triple-phase-shift control," *IEEE Trans. Ind. Electron.*, vol. 67, no. 10, pp. 8587–8598, 2019.

[24] J.-Y. Lee, H.-S. Kim, and J.-H. Jung, "Enhanced dual-active-bridge DC–DC converter for balancing bipolar voltage level of DC distribution system," *IEEE Trans. Ind. Electron.*, vol. 67, no. 12, pp. 10399–10409, 2019.

[25] L. Chen, S. Shao, Q. Xiao, L. Tarisciotti, P. W. Wheeler, and T. Dragičević, "Model predictive control for dual-active-bridge converters supplying pulsed power loads in naval DC micro-grids," *IEEE Trans. Power Electron.*, vol. 35, no. 2, pp. 1957–1966, 2019.

[26] S. Zengin and M. Boztepe, "A novel current modulation method to eliminate low-frequency harmonics in single-stage dual active bridge AC–DC converter," *IEEE Trans. Ind. Electron.*, vol. 67, no. 2, pp. 1048–1058, 2019.

[27] H. Qin and J. W. Kimball, "Generalized Average Modeling of Dual Active Bridge DC–DC Converter," *IEEE Trans. Power Electron.*, vol. 27, no. 4, pp. 2078–2084, 2012, doi: 10.1109/TPEL.2011.2165734.

[28] R. Jha, M. Forato, S. Prakash, H. Dashora, and G. Buja, "An analysis-supported design of a single active bridge (SAB) converter," *Energies*, vol. 15, no. 2, p. 666, 2022.

[29] D. Vinnikov, A. Chub, E. Liivik, F. Blaabjerg, and Y. Siwakoti, "Boost half-bridge DC-DC converter with reconfigurable rectifier for ultra-wide input voltage range applications," in *2018 IEEE Applied Power Electronics Conference and Exposition (APEC)*, 2018, pp. 1528–1532.

[30] H.-S. Lee, B. Kang, W.-S. Kim, and S.-J. Yoon, "Reduction of input voltage/current ripples of boost half-bridge DC-DC converter for photovoltaic micro-inverter," *Sol. Energy*, vol. 188, pp. 1084–1101, 2019.

[31] C.-E. Kim, G.-W. Moon, and S.-K. Han, "Voltage Doubler Rectified Boost-Integrated Half Bridge (VDRBHB) Converter for Digital Car Audio Amplifiers," *IEEE Trans. Power Electron.*, vol. 22, no. 6, pp. 2321–2330, 2007, doi: 10.1109/TPEL.2007.904222.

[32] H.-S. Lee, H.-J. Choe, and J.-J. Yun, "Improved Boost Half-Bridge DC–DC Converter for DC Distribution Networks," *J. Electr. Eng. Technol.*, vol. 17, no. 5, pp. 2889–2898, 2022, doi: 10.1007/s42835-022-01107-1.

[33] C. Yoon, J. Kim, and S. Choi, "Multiphase DC–DC Converters Using a Boost-Half-Bridge Cell for High-Voltage and High-Power Applications," *IEEE Trans. Power Electron.*, vol. 26, no. 2, pp. 381–388, 2011, doi: 10.1109/TPEL.2010.2060498.

[34] H.-S. Lee and J.-J. Yun, "Quasi-Resonant Voltage Doubler With Snubber Capacitor for Boost Half-Bridge DC–DC Converter in Photovoltaic Micro-Inverter," *IEEE Trans. Power Electron.*, vol. 34, no. 9, pp. 8377–8388, 2019, doi: 10.1109/TPEL.2018.2883535.

[35] K. Fathy, H. W. Lee, T. Mishima, and M. Nakaoka, "Boost-half bridge single power stage PWM DC-DC converter for small scale fuel cell stack," in *2006 IEEE International Power and Energy Conference*, 2006, pp. 426–431.

[36] V. A. G. Cunha, A. O. C. Neto, G. B. Lima, and L. C. G. Freitas, "A Bridgeless Boost Half Bridge DC-DC Converter for Electrical and Hybrid Vehicle Applications," in *2019 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America)*, 2019, pp. 1–6. doi: 10.1109/ISGT-LA.2019.8895388.

[37] C. Yoon, J. Kim, and S. Choi, "Multiphase DC–DC converters using a boost-half-bridge cell for high-voltage and high-power applications," *IEEE Trans. Power Electron.*, vol. 26, no. 2, pp. 381–388, 2010.

[38] B. Su, T. Yang, Z. Lu, and D. Xu, "Soft-switching dual forward DC/DC converters employing secondary side control," in *2009 IEEE Energy Conversion Congress and Exposition*, 2009, pp. 1855–1859.

[39] K.-H. Cheng, C.-F. Hsu, C.-M. Lin, T.-T. Lee, and C. Li, "Fuzzy–neural sliding-mode control for DC–DC converters using asymmetric Gaussian membership functions," *IEEE Trans. Ind. Electron.*, vol. 54, no. 3, pp. 1528–1536, 2007.

[40] R. Kanthimathi and J. Kamala, "Analysis of different flyback converter topologies," in *2015 international conference on industrial instrumentation and control (ICIC)*, 2015, pp. 1248–1252.

[41] M. C. Taneri, N. Genc, and A. Mamizadeh, "Analyzing and comparing of variable and constant switching frequency flyback DC-DC Converter," in *2019 4th International Conference on Power Electronics and their Applications (ICPEA)*, 2019, pp. 1–5.

[42] J.-W. Yang and H.-L. Do, "Soft-switching dual-flyback DC–DC converter with improved efficiency and reduced output ripple current," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 3587–3594, 2017.

[43] Z. Zhang, M. Liao, D. Jiang, X. Yang, and S. Li, "High step-up isolated forward-flyback DC/DC converter based on resonance with pulse frequency modulation," *J. Power Electron.*, vol. 21, no. 2, pp. 483–493, 2021, doi: 10.1007/s43236-020-00186-5.

[44] V. Parkash, P. Kumar, P. Sharma, and G. Sapra, "Design and implementation of flyback converter as high voltage power supply for nanofibers production," *Mater. Today Proc.*, vol. 45, pp. 5285–5291, 2021.

[45] Y. Wu, Y. Huangfu, R. Ma, A. Ravey, and D. Chrenko, "A strong robust DC-DC converter of all-digital high-order sliding mode control for fuel cell power applications," *J. Power Sources*, vol. 413, pp. 222–232, 2019.

[46] R. W. Erickson, "DC–DC power converters," *Wiley Encycl. Electr. Electron. Eng.*, 2001.

[47] K.-H. Chen and T.-J. Liang, "Design of Quasi-resonant flyback converter control IC with DCM and CCM operation," in *2014 International Power Electronics Conference (IPEC-Hiroshima 2014-ECCE ASIA)*, 2014, pp. 2750–2753.

[48] M. Salem *et al.*, "Three-phase series resonant DC-DC boost converter with double LLC resonant tanks and variable frequency control," *IEEE Access*, vol. 8, pp. 22386–22399, 2020.

[49] M. Salem, A. Jusoh, N. R. N. Idris, H. S. Das, and I. Alhamrouni, "Resonant power converters with respect to passive storage (LC) elements and control techniques–An overview," *Renew. Sustain. Energy Rev.*, vol. 91, pp. 504–520, 2018.

[50] F. Alaql and I. Batarseh, "Review and comparison of resonant DC-DC converters for wide-output voltage range applications," in *2020 IEEE Energy Conversion Congress and Exposition (ECCE)*, 2020, pp. 1197–1203.

[51] X. Zhao *et al.*, "A high-efficiency hybrid series resonant DC-DC converter with boost converter as secondary for photovoltaic applications," in *2015 IEEE Energy Conversion Congress and Exposition (ECCE)*, 2015, pp. 5462–5467. doi: 10.1109/ECCE.2015.7310428.

[52] X. Zhao, L. Zhang, R. Born, and J.-S. Lai, "A High-Efficiency Hybrid Resonant Converter With Wide-Input Regulation for Photovoltaic Applications," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 3684–3695, 2017, doi: 10.1109/TIE.2017.2652340.

[53] G. I. Vacheva, K. Genev, and N. L. Hinov, "Modeling and Simulation of DC-DC Push-Pull Converter," in *2022 57th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)*, 2022, pp. 1–4. doi: 10.1109/ICEST55168.2022.9828584.

[54] Q. Wu, Q. Wang, J. Xu, and L. Xiao, "A wide load range ZVS push–pull DC/DC converter with active clamped," *IEEE Trans. Power Electron.*, vol. 32, no. 4, pp. 2865–2875, 2016.

[55] S. Musumeci and S. Di Mauro, "Low voltage single fuel cell interface by Push-Pull converter: A case of study," in *2017 6th International Conference on Clean Electrical Power (ICCEP)*, 2017, pp. 541–548.

[56] S. Bal, A. K. Rathore, and D. Srinivasan, "Naturally clamped snubberless soft-switching bidirectional current-fed three-phase push–pull DC/DC converter for DC microgrid application," *IEEE Trans. Ind. Appl.*, vol. 52, no. 2, pp. 1577–1587, 2015.

[57] R. Kalpana, "Configurations of modular push-pull buck dc-dc converters for 12KW telecom SMPS and its design," in *2016 Biennial International Conference on Power and Energy Systems: Towards Sustainable Energy (PESTSE)*, 2016, pp. 1–7.

[58] H. Zenk, "Comparison of the Performance of Photovoltaic Power Generation-Consumption System with Push-Pull Converter under the Effect of Five Different Types of Controllers," *Int. J. Photoenergy*, vol. 2019, no. 1, p. 3810970, 2019.

[59] H. Givi, E. Farjah, and T. Ghanbari, "A comprehensive monitoring system for online fault diagnosis and aging detection of non-isolated DC–DC converters' components," *IEEE Trans. Power Electron.*, vol. 34, no. 7, pp. 6858–6875, 2018.

[60] S. N. Singh, "Selection of non-isolated DC-DC converters for solar photovoltaic system," *Renew. Sustain. Energy Rev.*, vol. 76, pp. 1230–1247, 2017.

[61] T. K. Nizami and C. Mahanta, "An intelligent adaptive control of DC–DC buck converters," *J. Franklin Inst.*, vol. 353, no. 12, pp. 2588–2613, 2016.

[62] B. B. Naik and A. J. Mehta, "Sliding mode controller with modified sliding function for DC-DC Buck Converter," *ISA Trans.*, vol. 70, pp. 279–287, 2017.

[63] C. Nan, R. Ayyanar, and Y. Xi, "A 2.2-MHz active-clamp buck converter for automotive applications," *IEEE Trans. Power Electron.*, vol. 33, no. 1, pp. 460–472, 2017.
[64] J. Wang, J. Rong, and L. Yu, "Dynamic prescribed performance sliding mode control for DC–DC buck converter system with mismatched time-varying disturbances," *ISA Trans.*, vol. 129, pp. 546–557, 2022.
[65] W. Chen, Z. Ge, Y. Cheng, H. Du, Q. Du, and M. Yu, "Current-constrained finite-time control algorithm for DC-DC buck converter," *J. Franklin Inst.*, vol. 358, no. 18, pp. 9467–9482, 2021.
[66] T. K. Nizami, A. Chakravarty, and C. Mahanta, "Time bound online uncertainty estimation based adaptive control design for DC–DC buck converters with experimental validation," *IFAC J. Syst. Control*, vol. 15, p. 100127, 2021.
[67] İ. Yazici, "Robust voltage-mode controller for DC–DC boost converter," *IET Power Electron.*, vol. 8, no. 3, pp. 342–349, 2015.
[68] F. L. Tofoli, D. de C. Pereira, W. Josias de Paula, and D. de S. Oliveira Junior, "Survey on non-isolated high-voltage step-up dc–dc topologies based on the boost converter," *IET power Electron.*, vol. 8, no. 10, pp. 2044–2057, 2015.
[69] T. K. Nizami and A. Chakravarty, "Neural network integrated adaptive backstepping control of DC-DC boost converter," *IFAC-PapersOnLine*, vol. 53, no. 1, pp. 549–554, 2020.
[70] S. Vadi, F. B. Gurbuz, S. Sagiroglu, and R. Bayindir, "Optimization of pi based buck-boost converter by particle swarm optimization algorithm," in *2021 9th International Conference on Smart Grid (icSmartGrid)*, 2021, pp. 295–301.
[71] K. Prag, M. Woolway, and T. Celik, "Data-driven model predictive control of DC-to-DC buck-boost converter," *IEEE Access*, vol. 9, pp. 101902–101915, 2021.
[72] N. Rana and S. Banerjee, "Interleaved tri-state buck-boost converter with fast transient response and lower ripple," in *2019 IEEE Transportation Electrification Conference (ITEC-India)*, 2019, pp. 1–5.
[73] N. Rana, S. Banerjee, S. K. Giri, A. Trivedi, and S. S. Williamson, "Modeling, analysis and implementation of an improved interleaved buck-boost converter," *IEEE Trans. Circuits Syst. II Express Briefs*, vol. 68, no. 7, pp. 2588–2592, 2021.
[74] E. Babaei, M. E. S. Mahmoodieh, and H. M. Mahery, "Operational modes and output-voltage-ripple analysis and design considerations of buck–boost DC–DC converters," *IEEE Trans. Ind. Electron.*, vol. 59, no. 1, pp. 381–391, 2011.
[75] K. Wang, D. Liu, and L. Wang, "The implementation of synergetic control for a DC-DC buck-boost converter," *Procedia Comput. Sci.*, vol. 199, pp. 900–907, 2022.
[76] M. Martinez-Lopez, J. Moreno-Valenzuela, and W. He, "A robust nonlinear PI-type controller for the DC–DC buck–boost power converter," *ISA Trans.*, vol. 129, pp. 687–700, 2022.
[77] S. Kumar, R. Kumar, and N. Singh, "Performance of closed loop SEPIC converter with DC-DC converter for solar energy system," in *2017 4th International Conference on Power, Control & Embedded Systems (ICPCES)*, 2017, pp. 1–6.
[78] L. Kathi, A. Ayachit, D. K. Saini, A. Chadha, and M. K. Kazimierczuk, "Open-loop small-signal modeling of Cuk DC-DC converter in CCM by circuit-averaging technique," in *2018 IEEE Texas Power and Energy Conference (TPEC)*, 2018, pp. 1–6.
[79] B. P. Mokal and K. Vadirajacharya, "Extensive modeling of DC-DC Cuk converter operating in continuous conduction mode," in *2017 International Conference on Circuit, Power and Computing Technologies (ICCPCT)*, 2017, pp. 1–5.
[80] M. Verma and S. S. Kumar, "Hardware design of sepic converter and its analysis," in *2018 International Conference on Current Trends towards Converging Technologies (ICCTCT)*, 2018, pp. 1–4.
[81] S. Sivakumar, M. J. Sathik, P. S. Manoj, and G. Sundararajan, "An assessment on performance of DC–DC converters for renewable energy applications," *Renew. Sustain. Energy Rev.*, vol. 58, pp. 1475–1485, 2016.
[82] M. B. Ferrera, S. P. Litran, E. D. Aranda, and J. M. A. Marquez, "A converter for bipolar DC link based on SEPIC-Cuk combination," *IEEE Trans. Power Electron.*, vol. 30, no. 12, pp. 6483–6487, 2015.
[83] J. Marjani, A. Imani, A. Hekmati, and E. Afjei, "A new dual output DC-DC converter based on SEPIC and Cuk converters," in *2016 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM)*, 2016, pp. 946–950.
[84] A. H. R. Rosa, L. M. F. Morais, G. O. Fortes, and S. I. S. Júnior, "Practical considerations of nonlinear control techniques applied to static power converters: A survey and comparative study," *Int. J. Electr. Power Energy Syst.*, vol. 127, p. 106545, 2021.
[85] R. K. Pachauri and Y. K. Chauhan, "Modeling and simulation analysis of PV fed Cuk, Sepic, Zeta and Luo DC-DC converter," in *2016 IEEE 1st international conference on power electronics, intelligent control and energy systems (ICPEICES)*, 2016, pp. 1–6.
[86] K. Manikandan, A. Sivabalan, R. Sundar, and P. Surya, "A study of landsman, sepic and zeta converter by particle swarm optimization technique," in *2020 6th international conference on advanced computing and communication systems (ICACCS)*, 2020, pp. 1035–1038.
[87] M. Kaouane, A. Boukhelifa, and A. Cheriti, "Implementation of incremental-conductance MPPT algorithm in a photovoltaic conversion system based on DC-DC ZETA converter," in *2016 8th International Conference on Modelling, Identification and Control (ICMIC)*, 2016, pp. 612–617.
[88] J. S. Alagesan, J. Gnanavadivel, N. S. Kumar, and K. S. K. Veni, "Design and Simulation of Fuzzybased DC-DC Interleaved Zeta Converter for Photovoltaic Applications," in *2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI)*, 2018, pp. 704–709.
[89] M. M. Nishat, M. R. K. Shagor, H. Akter, S. A. Mim, and F. Faisal, "An optimal design of PID controller for DC-DC zeta converter using particle swarm optimization," in *2020 23rd International Conference on Computer and Information Technology (ICCIT)*, 2020, pp. 1–6.
[90] A. Raj, S. R. Arya, and J. Gupta, "Solar PV array-based DC–DC converter with MPPT for low power applications," *Renew. Energy Focus*, vol. 34, pp. 109–119, 2020.
[91] H. Sarkawi, Y. Ohta, and P. Rapisarda, "On the switching control of the DC–DC zeta converter operating in continuous conduction mode," *IET Control Theory Appl.*, vol. 15, no. 9, pp. 1185–1198, 2021.
[92] N. Vosoughi, M. Abbasi, E. Abbasi, and M. Sabahi, "A Zeta-based switched-capacitor DC-DC converter topology," *Int. J. Circuit Theory Appl.*, vol. 47, no. 8, pp. 1302–1322, 2019.
[93] R. F. Rajakumari and M. S. Ramkumar, "Design considerations and performance analysis based on ripple factors and switching loss for converter techniques," *Mater. Today Proc.*, vol. 37, pp. 2681–2686, 2021.
[94] N. Ghasemi, F. Zare, C. Langton, and A. Ghosh, "A new unequal DC link voltage configuration for a single phase multilevel converter to reduce low order harmonics," in *Proceedings of the 2011 14th European Conference on Power Electronics and Applications*, 2011, pp. 1–9.
[95] S. Du, B. Wu, N. R. Zargari, and Z. Cheng, "A flying-capacitor modular multilevel converter for medium-voltage motor drive," *IEEE Trans. Power Electron.*, vol. 32, no. 3, pp. 2081–2089, 2016.
[96] S. Qin, Y. Lei, Z. Ye, D. Chou, and R. C. N. Pilawa-Podgurski, "A high-power-density power factor correction front end based on seven-level flying capacitor multilevel converter," *IEEE J. Emerg. Sel. Top. Power Electron.*, vol. 7, no. 3, pp. 1883–1898, 2018.
[97] C. B. Barth *et al.*, "Design and control of a GaN-based, 13-level, flying capacitor multilevel inverter," *IEEE J. Emerg. Sel. Top. Power Electron.*, vol. 8, no. 3, pp. 2179–2191, 2019.
[98] A. Marquez *et al.*, "Discontinuous-PWM method for multilevel N-cell cascaded H-bridge converters," *IEEE Trans. Ind. Electron.*, vol. 68, no. 9, pp. 7996–8005, 2020.
[99] A. M. Alcaide *et al.*, "Variable-angle PS-PWM technique for multilevel cascaded H-bridge converters with large number of power cells," *IEEE Trans. Ind. Electron.*, vol. 68, no. 8, pp. 6773–6783, 2020.
[100] Y. Koyama, Y. Nakazawa, H. Mochikawa, A. Kuzumaki, K. Sano, and N. Okada, "A transformerless 6.6-kV STATCOM based on a hybrid cascade multilevel converter using SiC devices," *IEEE Trans. Power Electron.*, vol. 33, no. 9, pp. 7411–7423, 2017.
[101] H. D. Tafti, A. I. Maswood, G. Konstantinou, C. D. Townsend, P. Acuna, and J. Pou, "Flexible control of photovoltaic grid-connected cascaded H-bridge converters during unbalanced voltage sags," *IEEE Trans. Ind. Electron.*, vol. 65, no. 8, pp. 6229–6238, 2017.
[102] F. A. Abbas, T. A. Abdul-Jabbar, A. A. Obed, A. Kersten, M. Kuder, and T. Weyh, "A Comprehensive Review and Analytical Comparison of Non-Isolated DC-DC Converters for Fuel Cell Applications," *Energies*, vol. 16, no. 8. 2023. doi: 10.3390/en16083493.
[103] S. Pourjafar, H. Afshari, P. Mohseni, O. Husev, O. Matiushkin, and N. Shabbir, "Comprehensive Comparison of Isolated High Step-up DC-DC Converters for Low Power Application," *IEEE Open J. Power Electron.*, vol. 5, pp. 1149–1161, 2024, doi: 10.1109/OJPEL.2024.3433554.

BIOGRAPHIES

Abdurrahim Erat received the B.S. and M.S. degrees in Electrical and Electronics Engineering from the University of Niğde Ömer Halisdemir, Niğde, in 2013 and the University of Necmettin Erbakan, Konya, in 2018, respectively. He is currently in the dissertation phase of the PhD program at Gaziantep University, Department of Electrical and Electronics Engineering. He has been working as a Lecturer in the Department of Electrical and Energy at Şırnak Vocational School, Şırnak University since 2019.

Ahmet Mete Vural received the B.S. and M.S. degrees in Electrical and Electronics Engineering from Gaziantep University, Gaziantep, in 1999 and 2001, respectively. He also received the Ph.D. degrees in Electrical and Electronics Engineering from the Çukurova University, Adana, in 2012. He works as a full-time Professor in Electrical and Electronics Engineering Department in Gaziantep University, Gaziantep, Türkiye. He leads the Power Systems Research Group. His research interests include application of power electronics to power systems, multi-level converters, FACTS devices, power quality, microgrid, smartgrid, renewable energy, and energy storage.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 3, September 2025
Research Article

Modeling of an Electro-Hydraulic System for Wave Generation in a Wave Channel

Batin Demircan, Tugce Yaren, Ersin Akyuz, Sabri Bicakci

***Abstract*—This paper presents a black-box mathematical model of the electrohydraulic system controlling the wave generation structure in a wave channel under various operating conditions. For the position control of the system, NI-CRIO 9074 hardware and LabVIEW software were used. Open-loop position control experiments of the hydraulic cylinder were conducted using stimulus signals with different initial positions (0, 120, and 240 mm) and amplitudes. Data were recorded for different sampling times: 1 ms, 2 ms, 5 ms, and 10 ms. The recorded data were processed using the System Identification Toolbox (SIT) in MATLAB, and system models were developed in both continuous and discrete time domains using transfer function, state space, and AutoRegressive with eXogenous input (ARX) models. These models were compared and analyzed based on their fit rates to the training and test data. Among the system models with high compliance rates, the top three models were selected for further comparison using an additional test dataset. Based on this evaluation, the transfer function model (120 mm initial condition and 1 ms sampling time) type was identified as the best-performing model. This model was successfully integrated into the real-time control study, achieving effective controller performance.**

Index Terms—**Wave channel, Electro-hydraulic system, Hydraulic control, System identification, Black-box.**

Batın Demircan, Department of Electronics and Automation, Balıkesir Vocational School, Balıkesir University, Balıkesir, Turkey,
Institue of Science, Department of Electrical and Electronics Engineering, Balıkesir University, Balıkesir, Turkey,
Renewable Energy Research, Application and Development Center, Balıkesir University, Balıkesir, Turkey,
(e-mail: batin.demircan@balikesir.edu.tr).
https://orcid.org/0000-0002-0765-458X

Tugce Yaren, Department of Electrical and Electronics Engineering, Faculty of Engineering, Balıkesir University, Balıkesir, Turkey,
Renewable Energy Research, Application and Development Center, Balıkesir University, Balıkesir, Turkey,
(e-mail: tugce.yaren@balikesir.edu.tr).
https://orcid.org/0000-0001-9937-3111

Ersin Akyuz, Department of Electrical and Electronics Engineering, Faculty of Engineering, Balıkesir University, Balıkesir, Turkey,
Renewable Energy Research, Application and Development Center, Balıkesir University, Balıkesir, Turkey,
(e-mail: eakyuz@balikesir.edu.tr).
https://orcid.org/0000-0001-9786-3221

Sabri Bicakci, Department of Electrical and Electronics Engineering, Faculty of Engineering, Balıkesir University, Balıkesir, Turkey,
Renewable Energy Research, Application and Development Center, Balıkesir University, Balıkesir, Turkey,
(e-mail: sbicakci@balikesir.edu.tr).
https://orcid.org/0000-0002-2334-8515
Manuscript received Jan 18, 2025; accepted Apr 22, 2025.
DOI: 10.17694/bajece.1622216

I. INTRODUCTION

WAVE CHANNEL SYSTEMS are critically important in many engineering and research fields, such as interaction with coastal structures, wave energy, and flood and tsunami modeling [1-3]. These systems enhance the accuracy of experimental studies by generating water waves in specific forms. Wave generation is typically carried out using flap, plunger, or piston-type mechanical systems, and these mechanisms are driven by technologies such as electric motors, servo systems, pneumatic actuators, and electrohydraulic systems (EHS) [4, 5].

EHS, with its high power-to-weight ratio and rapid response capabilities, stands out by enabling the generation of complex waveforms in wave channel applications. However, their complex structures, arising from nonlinear dynamics, time-varying parameters, and susceptibility to external disturbances, necessitate accurate dynamics modeling and the development of effective control strategies [6].

System identification techniques play a critical role in the modeling processes of EHSs. These techniques enable the development of models that accurately represent the system's dynamics using various mathematical approaches such as transfer function (TF), state space (SS), and AutoRegressive with eXogenous input (ARX) models with external inputs, leveraging experimental data [7-10]. Each modeling method offers different advantages in analyzing system dynamics and designing effective control strategies. While transfer functions and state-space models are more traditional approaches, the ARX model offers a data-driven alternative that can be particularly useful in scenarios where system dynamics are complex and cannot be easily captured by physical models.

The choice of modeling technique depends on the specific requirements of the EHS application, as well as the desired accuracy, computational efficiency, and the complexity of the system dynamics. Each method contributes to a deeper understanding of system behavior and facilitates the development of effective control strategies. In the literature, numerous studies have been conducted using various methods and approaches to better understand the complex dynamics of EHSs and to develop accurate system models. Rahmat et al. used the MATLAB System Identification Toolbox (SIT) to derive an ARX model of the EHS and designed a PID controller tuned with the Ziegler-Nichols method based on the obtained model [8]. Similarly, Das et al. utilized artificial neural networks to model the nonlinear dynamics of an electrohydraulic actuator (EHA) system and validated the model with experimental data [11]. Salleh et al. also used SIT to identify an ARX model of the EHA system and reported that, during the model validation phase, the predicted model demonstrated high accuracy in representing the actual system behavior, achieving a best fit of 94.88% [12].

The Recursive Least Squares (RLS) method is one of the approaches used for system identification of EHSs. Ghazali et al. modeled the EHA system with time-varying parameters using the system identification method and derived a linear discrete-time model. The model parameters were estimated using the RLS algorithm [13]. Furthermore, the use of nonlinear state-space models, as explored by Yan et al., enables a more comprehensive representation of system behavior under varying operating conditions and significantly enhances the accuracy of the identified models [14].

Black-box modeling techniques have gained increasing importance in the modeling of EHSs due to their capability to represent complex system dynamics without requiring detailed knowledge of underlying physical processes. These approaches rely solely on input-output data to develop mathematical models that can predict system behavior under various operating conditions [15]. Jin and Wang highlighted that black-box methods, such as artificial neural networks and fuzzy systems, can effectively capture the dynamics of hydraulic cylinders by fitting models to experimental data, thereby addressing the inherent modeling uncertainties present in hydraulic systems [16]. This capability is particularly important in EHS applications, where inherent nonlinearities and varying operational conditions often complicate traditional modeling approaches.

Mitov et al. successfully achieved control performance by analyzing the robustness of an electro-hydraulic steering control system using a black-box model [17]. The implementation of advanced control strategies, such as Model Predictive Control (MPC) and Linear Quadratic Regulator (LQR), largely depends on accurate system identification. By comparing the performance of these controllers, Mitov et al. demonstrated that the effectiveness of control strategies is significantly influenced by the quality of the underlying system model obtained through identification techniques [18].

Recent advances in electrohydraulic systems have highlighted the critical need for accurate and efficient modeling approaches, particularly under varying operating conditions. While traditional white-box models rely heavily on detailed physical parameters, recent studies have demonstrated the growing importance of black-box and data-driven techniques for capturing system dynamics with limited sensor data [19, 20]. For instance, Schwarz and Lohmann [21] proposed a robust identification method based solely on pressure signals, enabling effective modeling even in sensor-constrained mobile hydraulic systems. Similarly, energy-focused investigations by He et al. [22] emphasized the role of pump-controlled architectures in improving system efficiency, while Han et al. [23] introduced deep learning methods to predict complex nonlinearities such as friction in hydraulic actuators. In a related effort, Kim et al. [24] demonstrated that neural network inverse models could outperform traditional adaptive control in force control of hydraulic actuators, highlighting the limitations of physical models under uncertain dynamics. Likewise, Jiang et al. [25] proposed a neural network-based adaptive disturbance rejection strategy that integrates model reference control with nonlinear hydraulic models, showing increased robustness under high nonlinearities. Dindorf [26] further contributed to this domain by introducing a discrete incremental hydraulic positioning system using binary valves, which achieves accurate sensorless step positioning based on dynamic modeling and simulation. Building upon these developments, this study presents a black-box modeling framework for an electrohydraulic wave-generation system, emphasizing model accuracy, experimental validation, and integration into real-time position control.

In this study, the dynamics of the EHS, illustrated in the schematic diagram in Fig. 1, were examined using a black-box modeling approach applied to the wave generation mechanism located in a 24x1x1 m laboratory-scale wave channel at the Department of Civil Engineering, Faculty of Engineering, Balıkesir University. Experiments were conducted by applying stimulus signals with varying initial positions (0 mm, 120 mm, and 240 mm) and amplitude values, and the corresponding system responses were recorded. The collected data were processed using the SIT, and models were obtained in transfer function (TF), state-space (SS), and ARX structures. The prediction performance of these models was evaluated under different sampling times (1 ms, 2 ms, 5 ms, and 10 ms) for each initial condition. The model fit (FIT) ratios of the TF, SS, and ARX models were assessed for all initial positions. Furthermore, the models developed using training datasets were tested on validation datasets at different time scales.

This study presents a comprehensive analysis of the dynamics of a wave-generating mechanisms under varying initial conditions and sampling times, contributing to the modeling process of EHSs. The findings lay a significant foundation for improving the control performance of electrohydraulic systems in wave channel applications. The main contributions of this study are outlined as follows:

- Experimental modeling of an electrohydraulic flap-type wave generator system under different initial positions and amplitude levels.
- Development and comparison of TF, SS, and ARX models using the SIT, based on a black-box modeling approach.
- Evaluation of model accuracy across various sampling times to support the design of effective control strategies.

The structure of this paper is organized as follows: the section "Materials and Methods" details the EHS, the methodology for conducting experiments, and the development of transfer function, state-space, and ARX models using the SIT. The section "Results" discusses the validation of the obtained models under various operating conditions and their comparative analysis. The section "Conclusion" provides the final conclusions and future work.

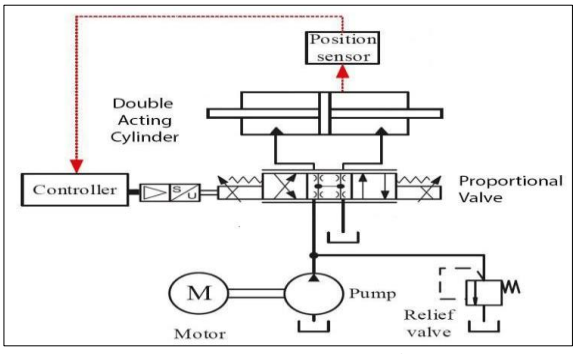


Fig. 1. Schematic diagram of EHS

II. Material and Methods

A. System Overview and Modeling

The EHS used in this study consists of a hydraulic cylinder and a hydraulic pump unit, as shown in Fig. 2a. The hydraulic cylinder is a double-acting, double-rod servo cylinder, while the hydraulic pump unit is equipped with a proportional

directional control valve, a hydraulic power supply, and safety components. This system is central to the experimental setup, as shown in Fig. 2b, and plays a key role in generating and controlling the wave dynamics, as shown in Fig. 2c.

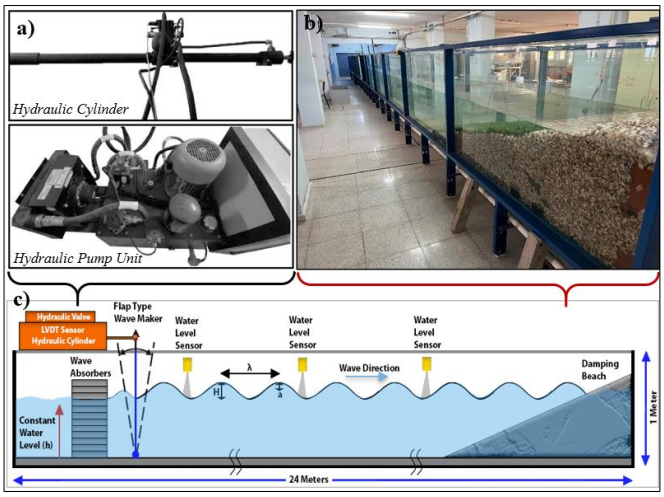


Fig. 2. (a) Hydraulic cylinder and pump unit (EHS), (b) physical wave channel system, (c) general structure of the wave channel.

The NI-CRIO 9074 packaged controller (PAC) used for EHS position control is operated by the user via LabVIEW software running on a Windows platform. Table 1 presents the hardware specifications of the NI-CRIO 9074 PAC [27].

TABLE I
NI-CRIO 9074 PAC SPECIFICATIONS

FEATURE	SPEC
Processor	400MHz Freescale MPC2500
Memory	128 MB
Flash Memory	256 MB
I/O Module Slots	8 hot-swappable C Series module

To provide a clearer understanding of the architecture within the experimental setup, a detailed block diagram of the system is shown in Fig. 3.The diagram illustrates the interactions between the hardware components—including the NI-CRIO-9074 controller, LabVIEW interface, hydraulic valve, servo cylinder, and the flap-type wave maker—along with their electrical, mechanical, and hydraulic connections. This structure enabled real-time data acquisition and control during the experiments.

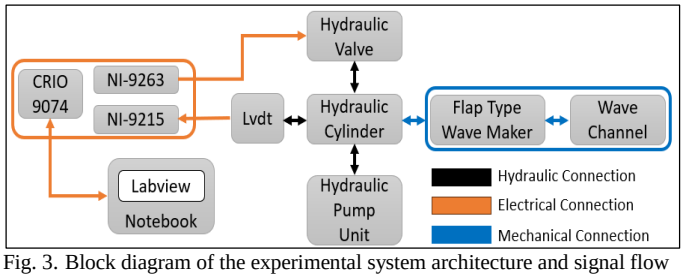


Fig. 3. Block diagram of the experimental system architecture and signal flow

The hydraulic servo cylinder used in the wave channel is custom-made, designed to handle a frequency of 0.625 Hz at a 300 mm stroke value and a +/- 150 amplitude. The general structure of the servo cylinder is shown in Fig. 4. Since the servo cylinder is double-ended and double-acting, the chamber volumes are equal when the cylinder is at the 350 mm position.

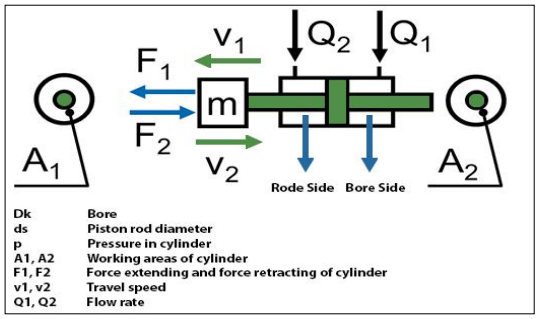


Fig. 4. Hydraulic servo cylinder.

The servo cylinder consists of two main parts: the bore side and the rod side. Because the cylinder is symmetrical, the areas on both sides are equal. The area and force generated by the servo cylinder are calculated using the following equations:

$$A = \frac{\pi}{4}(D^2 - d^2) \quad (1)$$

$$F = P.A \quad (2)$$

where D is the piston diameter, d is the rod diameter, P is the hydraulic pressure, A is the piston area, and F is the generated hydraulic force. The cylinder volume, the time required for piston travel, and the volumetric flow rate are calculated based on piston area and stroke length (L), as follows:

$$V = A.L \quad (3)$$

$$t = \frac{L}{v} \quad (4)$$

$$Q = A.v \quad (5)$$

where v is the average piston velocity, V is the cylinder volume. The design criteria for the cylinder include a piston diameter of 50mm, a rod diameter of 28mm, a stroke length of 700mm, an oil flow rate of 48 lpm, and a fixed operating pressure of 30 bar. The results of these calculations are provided in Table 2.

TABLE II
CALCULATED HYDRAULIC SERVO CYLINDER PARAMETERS

SPEC	VALUE
Area (cm^2)	13.4774
Volume (l)	0.9434
Force (kN)	4.0432
Time (sec)	1.1792
Velocity (m/s)	0.5935
Outflow (lpm)	63.930

TABLE III
HYDRAULIC PROPORTIONAL DIRECTIONAL VALVE PARAMETERS

SPEC	VALUE
Nominal Voltage	24 VDC
Absorbed Power	70 W
Maximum Current	2.6 A
Current Signal	-10/+10 V
Duty Cycle	4-20 mA
Max operating pressure -P-A-B	100%
Max operating pressure -T port	350 bar
Nominal flow with Δ10bar P-T	210 bar

The proportional directional valve (PDV) used in the EHS is the Duplomatic Oleodinamica DSE3J model, which receives control signals in the range of ±10 V. It governs the flow of

hydraulic fluid to the servo cylinder, thereby enabling position control within the 0–700 mm range. The NI-CRIO 9074 PAC communicates with this valve via analog output modules. The fluid used has a density of 0.869 kg/L and a kinematic viscosity of 46 cSt. The specifications of the PDV used in the system are summarized in Table 3.

Although all hydraulic components in the wave channel system can be modeled mathematically, the nonlinear characteristics of the hydraulic system—such as pressure-dependent behavior, temperature sensitivity, and fluid dynamics—make the derivation of a complete analytical model highly challenging. Therefore, only essential mathematical formulations for the hydraulic servo cylinder and its associated flow dynamics are considered. The PDV ports—P (connected to the pump), T (connected to the tank), and A/B (connected to the two sides of the cylinder)—allow bidirectional control of the cylinder position, as illustrated in Fig. 5.

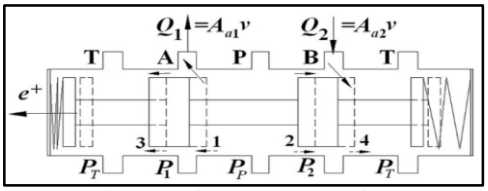


Fig. 5. Schematic of the PDV flow at neutral and for positive excitation.

The motion of the hydraulic piston can be generally described by Newton's second law, which states that the net force acting on a moving mass equals the sum of inertial, damping, and spring forces. This relationship is expressed in the classical form:

$$F = m\ddot{x} + b\dot{x} + kx \tag{6}$$

However, in electro-hydraulic systems, the driving force is not applied directly but is instead generated by hydraulic pressure acting on the piston surfaces. By considering the effective pressure forces on both sides of the piston, along with the return stiffness and friction force F_f, the dynamic equation of motion can be reformulated as:

$$\dot{y} = v\,,\ \dot{v} = \frac{P_1 A_{a1} - P_2 A_{a2} - k_y - F_f}{m_a} \tag{7}$$

where P denotes the pressures acting on the left and right chambers of the piston, k is the spring stiffness coefficient, y represents the piston displacement, and m_a is the equivalent moving mass of the system.

To fully describe the behavior of the electro-hydraulic system, it is essential to model not only the piston dynamics but also the flow dynamics through the PDV, since the control input directly affects the pressure applied to the cylinder chambers. The valve dynamics determine how the control signal e influences the hydraulic flow rates entering and leaving the piston chambers.

The relationship between the control input, pressure differences, and flow rates is governed by nonlinear orifice equations. For positive excitation ($e \geq 0$), piston extension occurs, and flow is directed through valve ports P–A and B–T. The corresponding flow rates are expressed as:

$$Q_1 = A_{a1}v = \left(C_1 e\sqrt{(P_p - P_1)} + C_1\sqrt{(P_p - P_1)}\right) - C_1\sqrt{(P_1 - P_T)} \tag{8}$$

$$Q_2 = A_{a2}v = \left(C_4 e\sqrt{(P_2 - P_T)} + C_1\sqrt{(P_2 - P_T)}\right) - C_1\sqrt{(P_p - P_2)} \tag{9}$$

where the terms Q_1 and Q_2 denote the volumetric flow rates through the valve ports P–A and B–T, respectively. P_p is the pump pressure, and P_T is the tank pressure. C_1 and C_4 are dimensionless discharge coefficients representing the flow characteristics of the valve orifices. Similarly, in the case of a negative command (e < 0), the retraction of the piston can be modeled with the P−B and A−T flows.

Frictional forces are another critical component of the model and the frictional force in the system can be expressed using a static modeling approach, as:

$$F_f = F_0, \qquad |v| \leq v_0\,, \tag{10}$$

$$F_f = F_c + (F_b - F_c)\exp\{-\frac{(v - v_b)^2}{(v_s - v_b)^2}\} + a_v(v - v_0), |v| > v_0 \tag{11}$$

where F_0 represents the adhesion force; v_0 denotes the slip velocity; F_c indicates the Coulomb friction; v_s corresponds to the Stribeck velocity; and a_v represents the viscous friction coefficient. Additionally, F_b signifies the maximum limit lubrication friction, and represents the associated velocity.

The complete hydraulic power unit includes components such as safety valves, filters, and a fixed-displacement pump. The PDV controls the flow to the hydraulic cylinder based on command voltages from the NI-CRIO control system, which interacts with an analog input/output module and an LVDT sensor for real-time position feedback. Due to the inherent nonlinearity of the EHS and its sensitivity to environmental and operational variables deriving an exact white-box model is challenging. Therefore, in this study, a black-box modeling approach is adopted. System identification techniques using input-output data are employed to develop empirical models (TF, SS, ARX), which capture the dominant system dynamics under various initial conditions and sampling intervals.

B. Data Acquisition

A flap-type wave generator was utilized in a 24 m × 1 m × 1 m wave channel to produce various wave forms, and the laboratory-scale physical setup of the wave channel system is shown in Fig. 6.

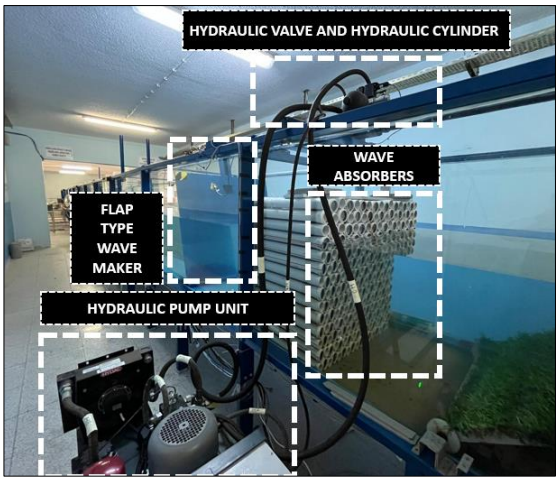


Fig. 6. Wave channel.

For the development and identification of the EHS model used in the position control of the flap-type wave generator,

both LabVIEW and MATLAB software were utilized. As illustrated in Fig. 7, the configuration includes the NI-CRIO 9074 and analog input-output modules serve as the PAC hardware. The position of the hydraulic cylinder is controlled by a proportional valve, and feedback is obtained from a 4–20 mA output LVDT sensor. The hydraulic power unit provides the required flow and pressure for system operation.

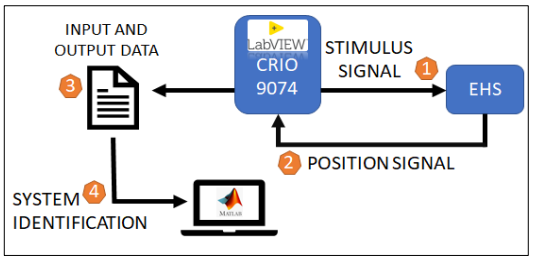


Fig. 7. EHS system identification data collection and mode generation structure.

To excite the system dynamics, stimulus signals composed of three different sinusoidal components were applied at initial positions of 0 mm, 120 mm, and 240 mm. These signals, given in Equations 12−14, were used to measure and record the position data of the hydraulic cylinder in response to these signals in real time.

$$y = 1.5(sin2\pi(0.05t)) + 0.9(sin2\pi(0.2t)) + 2.2(sin2\pi(t)) \quad (12)$$

$$y = 1.1(sin2\pi(0.05t)) + 0.9(sin2\pi(0.2t)) + 2.2(sin2\pi(t)) \quad (13)$$

$$y = 1.1(sin2\pi(0.05t)) + 0.5(sin2\pi(0.2t)) + 0.8(sin2\pi(t)) \quad (14)$$

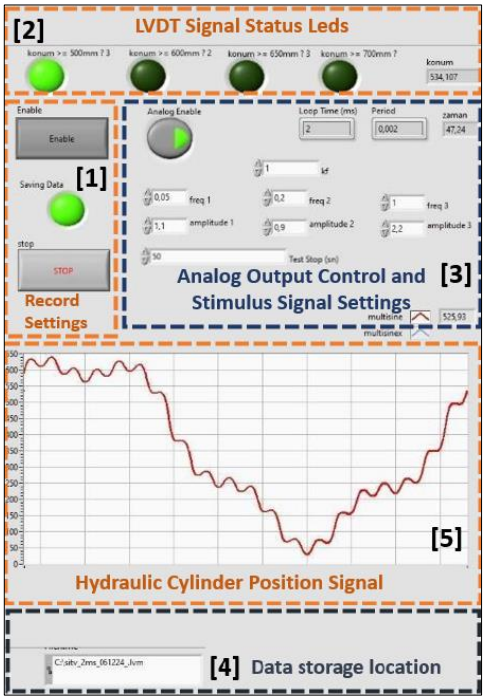


Fig. 8. EHS control software developed in Labview software.

The experimental interface, developed in LabVIEW and shown in Fig. 8, includes the following functional components: (1) start/stop control for testing, (2) LED indicators showing the real-time LVDT position, (3) configuration of stimulus signal parameters and control of the analog output, (4) file path selection for saving acquired data, and (5) real-time visualization of the cylinder position.

At each initial position, the stimulus signal was applied at four different cycle times (1 ms, 2 ms, 5 ms, and 10 ms), and the resulting position data were recorded. The collected input-output data were then transferred to MATLAB for system identification. TF, SS, and ARX models were generated using the SIT. An overview of the identification process is provided in Fig. 9.

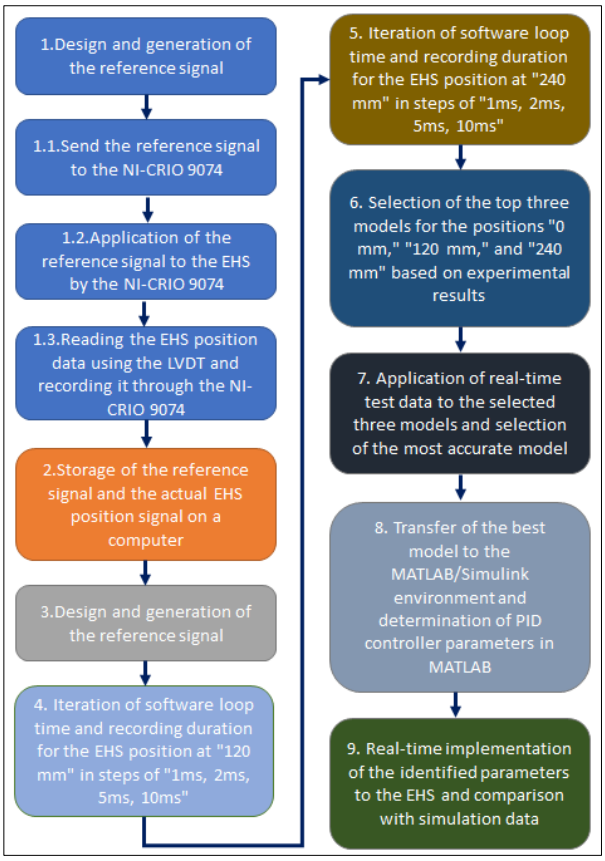


Fig. 9. EHS identification process

C. *System Identification*

System identification methods are categorized into three main approaches based on the level of system knowledge: *White-box* model method is used when the physical structure and mathematical model of the system are fully known. It is typically defined based on physical principles. *Grey-box* model approach combines theoretical and experimental methods and is employed when some parts of the system are known while the remaining parts are completed through data or estimation. *Black-box* model method is utilized when the internal structure of the system is entirely unknown. Mathematical models are developed solely based on input-output data. Instead of understanding the system dynamics, the black-box model focuses on making predictions through data-driven approaches, which is particularly advantageous for complex, nonlinear, or systems that are challenging to model physically.

The SIT enables the development of mathematical models based on input-output data of systems. It includes features for creating transfer function models, state-space models, autoregressive models (ARX, ARMAX), and nonlinear models such as Hammerstein-Wiener.

In this study, the selected models and their characteristics were chosen to analyze the dynamic behavior of the cylinder system from different perspectives and to obtain a model that

best represents the system's features. The transfer function model with three poles and two zeros directly represents the input-output relationship of the system and is useful for understanding the fundamental properties of dynamic systems. Both continuous and discrete-time representations allow accurate examination of the real system in digital control environments (discrete-time) and theoretical analysis processes (continuous-time).

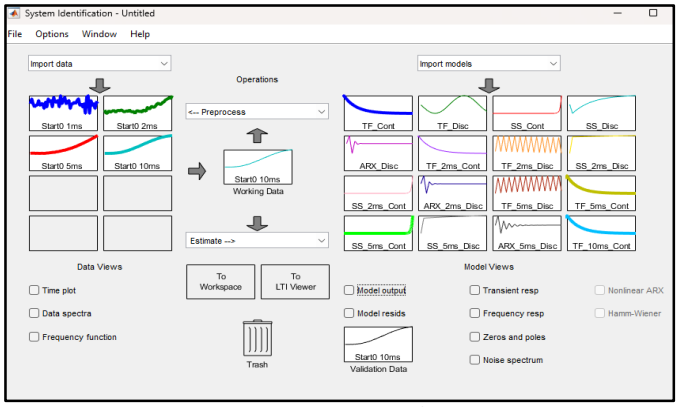


Fig. 9. Matlab SIT interface.

Second- and third-order state-space models offer a more flexible and detailed dynamic representation by directly addressing the system's state variables. Low-order state-space models are capable of representing the fundamental characteristics of the system while maintaining manageable complexity. Continuous-time and discrete-time state-space models provide a suitable framework for both the theoretical analysis of the physical system and for numerical simulations and control design. The ARX (4,4,1) and ARX (3,3,1) models are valid only for discrete-time systems and were selected due to their ability to quickly and practically model autoregressive structures. ARX models are a suitable option for rapidly predicting the linear relationship between high-frequency data and input-output. In particular, higher-order models like ARX (4,4,1) can capture more complex relationships between input and output, while lower-order models like ARX (3,3,1) provide the ability to generalize with a simpler structure. In selecting these model structures, the goal was to evaluate the linear and nonlinear dynamic characteristics of the system, compare the prediction performance of different structures, and optimize model accuracy from both theoretical and practical perspectives. This approach not only adds academic depth to the system identification process but also enhances the applicability of the obtained models in various applications.

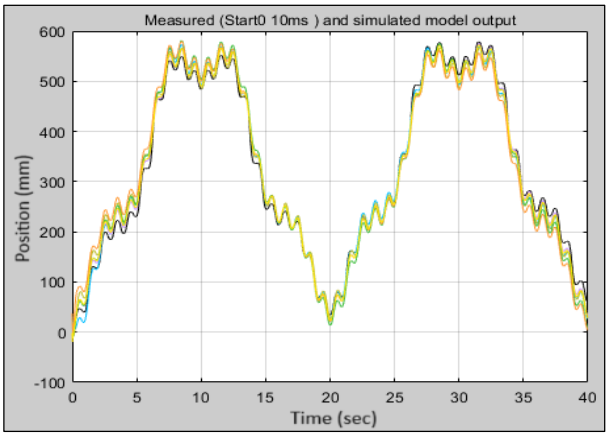


Fig. 10. Model estimation results for the training data with a 0 mm initial condition and a 10 ms sampling time (a).

In the system identification process, a distinction is made between training data and test data to evaluate the accuracy and generalization capability of the model. Training data is used to estimate the model parameters and enables the system to learn its input-output behavior. This data is selected to represent the fundamental dynamics of the system. Test data, on the other hand, is used to assess the model's performance and generalization ability. This dataset, which is not used during training, is crucial for analyzing the model's accuracy on different data sets beyond the training data. Experimental data collected from the system under 0-120-240 mm initial conditions, with four different sampling times, were sequentially imported into the SIT interface as training data, as shown in Fig. 9.

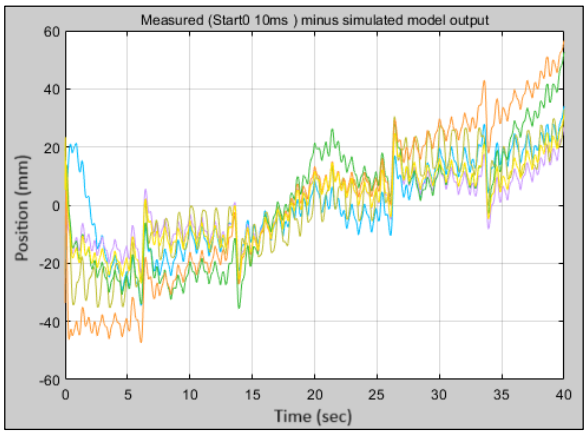


Fig. 11. Model estimation results for the training data with a 0 mm initial condition and a 10 ms sampling time (b).

In this study, the performance of the predicted models was evaluated using the best-fit ratio between the measured system output and the simulated model output. Fig. 10 and Fig.11 presents graphs illustrating the fit of the predicted models to the measured system output, using input-output data with a 10 ms sampling time as an example. Fig. 10 illustrates how accurately the predicted models replicate the system output, whereas Fig. 11 presents the error signal between the measured and simulated outputs. Fig. 10 illustrates a comparison between the measured system output and the outputs of various simulated models using input-output data sampled at 10 ms. The ARX (4,4,1) model, shown in purple, and the ARX (3,3,1) model, shown in yellow, achieved the highest fit ratios of 93.79% and 92.49%, respectively, demonstrating superior performance in capturing the system dynamics. The continuous-time transfer function model, represented in blue, yielded a fit ratio of 91.29%, while the discrete-time transfer function model, shown in green, achieved 90.19%. Both the continuous-time and discrete-time state-space models, depicted in light and dark green, exhibited fit ratios of 88.13%. The third-order discrete-time state-space model, shown in orange, had the lowest performance with a fit ratio of 83.94%. As illustrated in the graph, the ARX models—particularly those shown in purple and yellow—most accurately replicated the system behavior.

Fig. 11 illustrates the error signals between the measured system output and the outputs of various simulated models. The ARX (4,4,1) (purple, 93.79%) and ARX (3,3,1) (yellow, 92.49%) models demonstrated the best performance, exhibiting the lowest error signals throughout the time interval. These models remained closest to zero error, indicating a highly

accurate replication of the system dynamics. The continuous-time transfer function model (blue, 91.29%) and the discrete-time transfer function model (green, 90.19%) followed with moderate error magnitudes. The continuous-time and discrete-time state-space models (light green and dark green, both 88.13%) showed slightly higher error variations. The third-order discrete-time state-space model (orange, 83.94%) yielded the largest error signal, particularly in the latter stages of the simulation. The color-coded traces in Figure 11 clearly reveal that the ARX models outperformed the others in terms of both accuracy and consistency across the entire simulation period.

The ARX (4,4,1) and ARX (3,3,1) models demonstrated the best performance with fit ratios of 93.79% and 92.49%, respectively. The fit ratios for the transfer function (both continuous and discrete-time) and state-space models were calculated as 91.29%, 90.19%, 88.13%, and 83.94%, respectively. As shown in the graphs, the ARX models, in particular, successfully represented the system's dynamics with a low error signal.

III. RESULTS

A. *Identification of EHS Model*

The model prediction performance of the cylinder system for different initial conditions (0 mm, 120 mm, 240 mm) was analyzed at different sampling times (1 ms, 2 ms, 5 ms, 10 ms) and is presented in Tables 4, 5, and 6, respectively. For each initial condition, the FIT of different model structures, such as the TF, SS, and ARX models, were evaluated. Additionally, the models obtained from the training dataset were tested against validation datasets at different time scales.

According to the results obtained for the 0 mm initial condition presented in Table 3, the ARX models provided the highest fit ratios (89.47%–94.92%), particularly in discrete-time systems, and successfully represented the system's dynamic behavior. In this context, the ARX (4,4,1) model delivered the most reliable results, showing consistent performance across all sampling times (ranging from 92.5% to 94.92%). Among the TF models, the continuous-time 3-pole, 2-zero structure exhibited good performance at low sampling times (85.77%–91.29%), while the discrete-time versions, with low accuracy rates (e.g., -21.98%), failed to sufficiently represent the system's dynamics.

SS models, particularly second-order continuous-time models (88.7%–92.02%), provided satisfactory results and demonstrated a high generalization capacity during validation processes. However, the performance of third-order state-space models was lower, which is associated with the negative impact of increased model complexity on generalization capacity. Continuous-time models generally achieved higher accuracy at different sampling times, reflecting the flexibility and universal characteristics of continuous models.

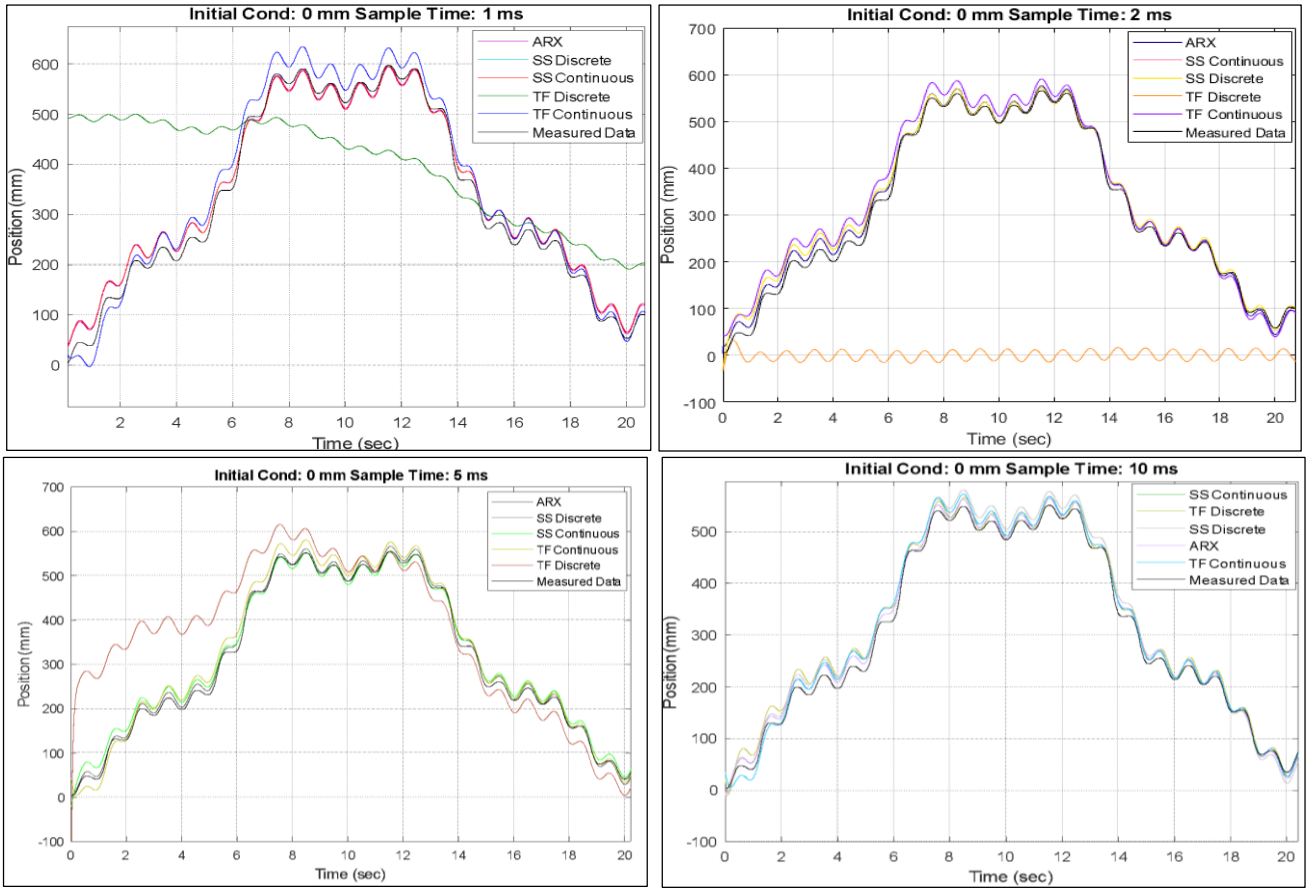


Fig.12. System identification results for 0 mm initial condition with 1, 2, 5, and 10 ms sample times.

The validation FIT results indicate that model performance was also evaluated with test data at different sampling times. Specifically, at low sampling times (1 ms), models created using training data maintained high fit ratios during validation, even when tested with high sampling time test data (e.g., 10 ms). For instance, the SS (Continuous) second-order model, trained with 1 ms data and tested with 10 ms data, achieved a 95.72% validation fit. This demonstrates the generalization ability and applicability of continuous models across different sampling times. In conclusion, while the ARX models

demonstrated superior performance in discrete-time systems, continuous-time transfer function and state-space models allowed for a broader representation of linear systems. These findings highlight the necessity of carefully evaluating the performance and generalization capacities of the chosen model structures in relation to specific applications during the modeling process. The performance of the predicted models at 1 ms, 2 ms, 5 ms, and 10 ms sampling times for the 0 mm initial condition was compared to the measured data and presented in Fig. 12. This figure clearly illustrates how the predicted models represent the system dynamics and the impact of sampling time on model performance.

Table 5 presents the prediction and validation performance of different model structures at 1 ms, 2 ms, 5 ms, and 10 ms sampling times for the 120 mm initial condition. Continuous-time models, particularly the TF Continuous and second-order SS Continuous structures, provided high fit ratios across all sampling times.

The TF Continuous model exhibited stable performance with fit ratios ranging from 94.92% to 96.61%, demonstrating consistent results from low to high sampling times. The SS Continuous models, especially the second-order structure, provided successful results (e.g., 90.62% fit at 10 ms), while the increased complexity of the third-order model negatively impacted validation performance.

Among the discrete-time models, the ARX (4,4,1) model delivered the highest fit ratios and best represented the system dynamics with results ranging from 94.06% to 96.94%. The ARX models demonstrated consistently high performance, not only with training data but also in validation tests. For instance, the ARX (4,4,1) model trained with 5 ms data achieved a fit ratio of 92.67% when validated with 10 ms data. In contrast, TF Discrete models failed to represent the system dynamics, yielding negative fit ratios. This is due to the insufficient representation of information caused by the low sampling frequency in discrete-time TF models.

Validation data analysis shows that continuous-time models maintained their generalization capacity across different sampling times. For example, the TF Continuous model trained with 1 ms data achieved a fit ratio of 89.33% when validated with 10 ms data. Similarly, SS Continuous models exhibited high validation fit at low sampling times, but a decrease in validation performance was observed with the third-order structure.

Table 6 presents the model prediction and validation results conducted under the initial condition of 240 mm. Among the continuous-time models, the TF Continuous and SS Continuous models have stood out. Especially the 3rd degree SS Continuous model has successfully represented the system's dynamics with high accuracy rates. For example, this model, trained with a 5 ms sampling time, achieved a 96.14% fit rate and reached up to 97.12% in validation tests.

The TF Continuous model also achieved satisfactory results in terms of generalization ability, but its accuracy rates were lower compared to the SS Continuous model. Among the discrete-time models, the ARX (4,4,1) model has provided the highest fit rates and has stood out as the structure that best represents the system dynamics with results reaching 97.77%. ARX models have also maintained their generalization capacity in validation tests and have demonstrated stable performance at all sampling times. For example, the ARX (4,4,1) model trained with a 5 ms sampling time achieved a 97.6% fit rate in the 10 ms validation test. In contrast, the TF Discrete models have shown low performance and generally yielded inadequate results in validation tests. At low sampling times (1 ms and 2 ms), both continuous and discrete-time models achieved higher accuracy rates, indicating that low sampling times better capture the detailed dynamic behaviors of the system.

Continuous-time models have reliably represented system dynamics, especially at longer sampling times, while maintaining their generalization capacity. ARX models, on the other hand, have provided reliable results under all test conditions with high accuracy and generalization capacity in discrete-time systems. These analyses reveal that the ARX (4,4,1) model, the 3rd degree SS Continuous model, and the TF Continuous model are the structures that best represent the system dynamics. While the ARX model stands out as a strong option for discrete-time control applications, continuous-time models allow for a broader range of analysis and control of the system. These models provide a suitable foundation for further testing with different types of signals.

The ARX (4,4,1) model has provided the best results for discrete-time systems under both 0 mm and 240 mm initial conditions. This model is a strong option for digital control applications. The TF Continuous model has provided the best result under the initial condition of 120 mm and is recommended for continuous-time analysis and control applications. For ARX models, a sampling time of 10 ms is ideal in terms of both prediction and validation performance. For continuous-time TF models, lower sampling times (e.g., 1 ms) provide higher accuracy.

Among the models evaluated under different initial conditions and sampling intervals, the best-performing structures were selected for further comparison. First, the 3-pole, 2-zero continuous-time transfer function model obtained at a 120 mm initial condition and a 1 ms sampling interval (as shown in Equation 15) was preferred due to its high fit rate (96.31%) under low sampling intervals and its stable performance during the validation phase.

$$G_1 = \frac{86.48\,s^2 + 845.1s + 9902}{s^3 + 13.99s^2 + 128.9s + 0.7653} \tag{15}$$

Secondly, the ARX(4,4,1) model shown in Equation 16—obtained under a 240 mm initial condition and a 10 ms sampling interval—was chosen for its high prediction accuracy (97.77%) and strong generalization capability.

$$G_2 = \begin{cases} A(z) = 1 - 1.079z^{-1} - 0.2649z^{-2} + 0.07386z^{-3} - 0.4173z^{-4} \\ B(z) = 45.6z^{-1} - 134z^{-2} + 132.2z^{-3} - 43.8z^{-4} \end{cases} \tag{16}$$

Finally, the second-order continuous-time state-space model shown in Equation 17, obtained under the 0 mm initial condition and a 5 ms sampling interval, was preferred due to its low complexity and consistent performance (93.19%) demonstrated during the validation phase.

$$G_3 = \begin{cases} A = \begin{bmatrix} 0.003973 & -0.08541 \\ 0.1342 & -8.391 \end{bmatrix}, \; B = \begin{bmatrix} 0.001494 \\ -0.04018 \end{bmatrix} \\ C = [3.326e+4 \quad -1.065], \quad D = [0] \end{cases} \tag{17}$$

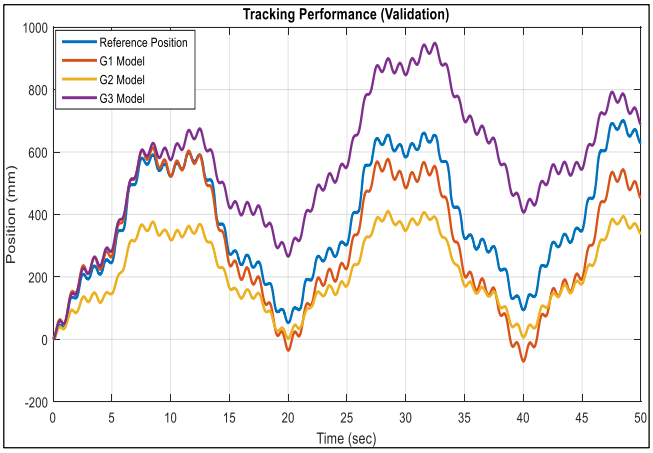


Fig. 13. Tracking performance during the validation phase for different models (G_1, G_2, G_3).

The performance of the system models was further evaluated using three commonly used error metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Integral of Time-weighted Absolute Error (ITAE), as shown in Equation 18.

Fig. 14 indicates that the G_1 model exhibits markedly reduced RMSE and MAE values in comparison to G_2 and G_3, suggesting superior tracking accuracy and robustness. Moreover, the ITAE values, which highlight long-term cumulative error, demonstrate that G_1 exhibits enhanced performance with time. (The ITAE values were scaled by a factor of 10^9 for visualization purposes.)

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(e(t))^2},\ ITAE = \int t|e(t)|dt$$
$$MAE = \frac{1}{N}\sum_{i=1}^{N}|e(t)| \quad (18)$$

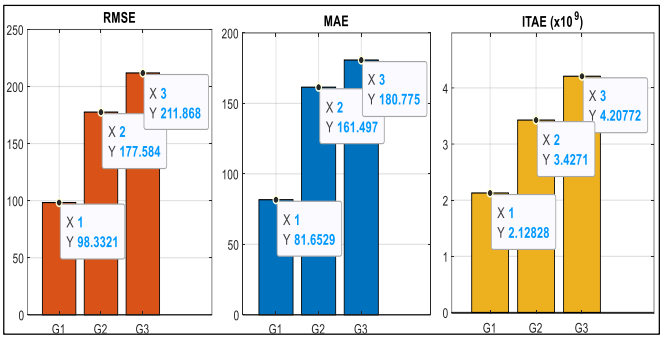


Fig. 14. Tracking error metrics comparison.

TABLE IV
MODEL ESTIMATION AND VALIDATION RESULTS WITH A 0 MM INITIAL CONDITION

Sample Time	Model Type	Model Structure	FIT (%)	Validation FIT 1 ms → 2 ms	Validation FIT 1 ms → 5 ms	Validation FIT 1 ms → 10 ms
1 ms	TF (*Continuous*)	3 pole, 2 zero	85.77	85.71	86.11	87.5
	SS (*Continuous*)	Second-order	88.7	93.83	96.12	95.72
	SS (*Continuous*)	Third-order	NaN	NaN	NaN	NaN
	TF (*Discrete*)	3 pole, 2 zero	-21.98	x	x	x
	SS (*Discrete*)	Second-order	88.7	x	x	x
	SS (*Discrete*)	Third-order	NaN	x	x	x
	ARX (*Discrete*)	(3,3,1)	89.47	x	x	x
	ARX (*Discrete*)	(4,4,1)	92.5	x	x	x
Sample Time	**Model Type**	**Model Structure**	**FIT (%)**	**Validation FIT 2 ms → 1 ms**	**Validation FIT 2 ms → 5 ms**	**Validation FIT 2 ms → 10 ms**
2 ms	TF (*Continuous*)	3 pole, 2 zero	85.3	81.07	87.39	89.49
	SS (*Continuous*)	Second-order	89.95	84.0	92.41	93.81
	SS (*Continuous*)	Third-order	89.96	86.16	91.81	93.36
	TF (*Discrete*)	3 pole, 2 zero	-122.9	x	x	x
	SS (*Discrete*)	Second-order	89.95	x	x	x
	SS (*Discrete*)	Third-order	89.96	x	x	x
	ARX (*Discrete*)	(3,3,1)	93.21	x	x	x
	ARX (*Discrete*)	(4,4,1)	93.4	x	x	x
Sample Time	**Model Type**	**Model Structure**	**FIT (%)**	**Validation FIT 5 ms → 1 ms**	**Validation FIT 5 ms → 2 ms**	**Validation FIT 5 ms → 10 ms**
5 ms	TF (*Continuous*)	3 pole, 2 zero	89.31	84.76	88.23	91.06
	SS (*Continuous*)	Second-order	92.02	83.48	89.64	93.19
	SS (*Continuous*)	Third-order	64.14	59.48	62.12	66.53
	TF (*Discrete*)	3 pole, 2 zero	18.66	x	x	x
	SS (*Discrete*)	Second-order	92.02	x	x	x
	SS (*Discrete*)	Third-order	64.14	x	x	x
	ARX (*Discrete*)	(3,3,1)	95.5	x	x	x
	ARX (*Discrete*)	(4,4,1)	94.92	x	x	x
Sample Time	**Model Type**	**Model Structure**	**FIT (%)**	**Validation FIT 10 ms → 1 ms**	**Validation FIT 10 ms → 2 ms**	**Validation FIT 10 ms → 5 ms**
10 ms	TF (*Continuous*)	3 pole, 2 zero	91.29	84.84	88.43	89.53
	SS (*Continuous*)	Second-order	88.13	81.55	84.31	86.16
	SS (*Continuous*)	Third-order	83.94	75.45	79.47	81.68
	TF (*Discrete*)	3 pole, 2 zero	90.19	x	x	x
	SS (*Discrete*)	Second-order	88.13	x	x	x
	SS (*Discrete*)	Third-order	83.94	x	x	x
	ARX (*Discrete*)	(3,3,1)	93.79	x	x	x
	ARX (*Discrete*)	(4,4,1)	92.49	x	x	x

TABLE V
MODEL ESTIMATION AND VALIDATION RESULTS WITH A 120 MM INITIAL CONDITION

Sample Time	Model Type	Model Structure	FIT (%)	Validation FIT 1 ms → 2 ms	Validation FIT 1 ms → 5 ms	Validation FIT 1 ms → 10 ms
1 ms	TF (*Continuous*)	3 pole, 2 zero	96.31	90.7	89.63	89.33
	SS (*Continuous*)	Second-order	NaN	NaN	NaN	NaN
	SS (*Continuous*)	Third-order	75.36	74.96	76.6	76.89
	TF (*Discrete*)	3 pole, 2 zero	-126	x	x	x
	ARX (*Discrete*)	(3,3,1)	90.68	x	x	x
	ARX (*Discrete*)	(4,4,1)	94.06	x	x	x
Sample Time	**Model Type**	**Model Structure**	**FIT (%)**	**Validation FIT 2 ms → 1 ms**	**Validation FIT 2 ms → 5 ms**	**Validation FIT 2 ms → 10 ms**
2 ms	TF (*Continuous*)	3 pole, 2 zero	95.16	91.62	95.86	95.57
	SS (*Continuous*)	Second-order	NaN	NaN	NaN	NaN
	SS (*Continuous*)	Third-order	67.32	71.6	65.36	64.85
	TF (*Discrete*)	3 pole, 2 zero	-138.7	x	x	x
	ARX (*Discrete*)	(3,3,1)	94.88	x	x	x
	ARX (*Discrete*)	(4,4,1)	95.21	x	x	x
Sample Time	**Model Type**	**Model Structure**	**FIT (%)**	**Validation FIT 5 ms → 1 ms**	**Validation FIT 5 ms → 2 ms**	**Validation FIT 5 ms → 10 ms**
5 ms	TF (*Continuous*)	3 pole, 2 zero	94.92	90.94	93.27	95.04
	SS (*Continuous*)	Second-order	93.73	86.85	92.35	93.55
	SS (*Continuous*)	Third-order	82.41	83.16	82.72	82.18
	TF (*Discrete*)	3 pole, 2 zero	-48.64	x	x	x
	ARX (*Discrete*)	(3,3,1)	96.99	x	x	x
	ARX (*Discrete*)	(4,4,1)	96.64	x	x	x
Sample Time	**Model Type**	**Model Structure**	**FIT (%)**	**Validation FIT 10 ms → 1 ms**	**Validation FIT 10 ms → 2 ms**	**Validation FIT 10 ms → 5 ms**
10 ms	TF (*Continuous*)	3 pole, 2 zero	96.61	93.26	94.46	96.58
	SS (*Continuous*)	Second-order	90.62	93.47	92.02	91.05
	SS (*Continuous*)	Third-order	90.05	90.65	91.55	90.55
	TF (*Discrete*)	3 pole, 2 zero	40.02	x	x	x
	ARX (*Discrete*)	(3,3,1)	95.36	x	x	x
	ARX (*Discrete*)	(4,4,1)	92.67	x	x	x

TABLE VI
MODEL ESTIMATION AND VALIDATION RESULTS WITH A 240 MM INITIAL CONDITION

Sample Time	Model Type	Model Structure	FIT (%)	Validation FIT 1 ms → 2 ms	Validation FIT 1 ms → 5 ms	Validation FIT 1 ms → 10 ms
1 ms	TF (*Continuous*)	3 pole, 2 zero	77	74.67	75.04	76.36
	SS (*Continuous*)	Second-order	89.03	94.55	95.15	93.06
	SS (*Continuous*)	Third-order	95.63	91.16	89.7	88.3
	TF (*Discrete*)	3 pole, 2 zero	76.36	x	x	x
	ARX (*Discrete*)	(3,3,1)	89.69	x	x	x
	ARX (*Discrete*)	(4,4,1)	91.87	x	x	x
Sample Time	**Model Type**	**Model Structure**	**FIT (%)**	**Validation FIT 2 ms → 1 ms**	**Validation FIT 2 ms → 5 ms**	**Validation FIT 2 ms → 10 ms**
2 ms	TF (*Continuous*)	3 pole, 2 zero	87.69	84.1	88.43	89.68
	SS (*Continuous*)	Second-order	90.82	83.33	93.51	93.53
	SS (*Continuous*)	Third-order	93.91	88.12	96.52	96.69
	TF (*Discrete*)	3 pole, 2 zero	80.4	x	x	x
	ARX (*Discrete*)	(3,3,1)	93.65	x	x	x
	ARX (*Discrete*)	(4,4,1)	95.1	x	x	x
Sample Time	**Model Type**	**Model Structure**	**FIT (%)**	**Validation FIT 5 ms → 1 ms**	**Validation FIT 5 ms → 2 ms**	**Validation FIT 5 ms → 10 ms**
5 ms	TF (*Continuous*)	3 pole, 2 zero	87.55	82.84	86.1	88.95
	SS (*Continuous*)	Second-order	93.34	83.39	90.81	93.17
	SS (*Continuous*)	Third-order	96.14	86.9	93.03	97.12
	TF (*Discrete*)	3 pole, 2 zero	68.05	x	x	x
	ARX (*Discrete*)	(3,3,1)	97.65	x	x	x
	ARX (*Discrete*)	(4,4,1)	97.6	x	x	x
Sample Time	**Model Type**	**Model Structure**	**FIT (%)**	**Validation FIT 10 ms → 1 ms**	**Validation FIT 10 ms → 2 ms**	**Validation FIT 10 ms → 5 ms**
10 ms	TF (*Continuous*)	3 pole, 2 zero	89.57	82.72	86.82	88.12
	SS (*Continuous*)	Second-order	95.48	86.17	91.14	93.8
	SS (*Continuous*)	Third-order	94.81	89.59	95.34	96.74
	TF (*Discrete*)	3 pole, 2 zero	11.67	x	x	x
	ARX (*Discrete*)	(3,3,1)	97.57	x	x	x
	ARX (*Discrete*)	(4,4,1)	97.77	x	x	x

 ISSN: 2147-284X http://dergipark.gov.tr/bajece

B. Real-Time Control Implementation

Among the models identified using different structures and sampling times, the one referred to as G_1, which exhibited the lowest tracking error during addition test, was selected for control application. This model provided the most accurate dynamic representation of the EHS and was thus integrated into the closed-loop control design. Based on the G_1 system model, the sampling time and controller configuration were determined for real time application.

For the control of the EHS, a Proportional-Integral (PI) controller was employed. A control interface was developed in LabVIEW to apply voltage signals in the range of –10 V to +10 V to the proportional directional control valve. Various signal types (e.g., sine, square, and step) could be generated for reference tracking. The measured position of the hydraulic cylinder was continuously compared with the desired position command, and the resulting position error was used to compute the control input. The most appropriate cycle time for the control loop was selected based on the previously identified system dynamics.

The use of PI controllers in hydraulic systems offers significant advantages in terms of maintaining equilibrium and enhancing system response. By combining proportional and integral components, the PI controller delivers both immediate reactions to system deviations and corrections based on historical errors. The *P* term in PI control refers to the control mechanism that responds rapidly in proportion to the error between the desired and measured positions. In this study, this response was executed via a proportional directional control valve, which quickly intervenes in system deviations to make instantaneous corrections. The *I* term in PI control, on the other hand, represents the accumulation of error. It attempts to compensate for constant imbalances in the system (such as friction, leakage, or unmodeled dynamics) based on errors accumulated over time. If a persistent error exists, the integral component integrates this error and sends additional corrective signals to the proportional directional control valve, thereby eliminating permanent discrepancies in the system. By combining these two terms, the PI controller ensures that the EHS operates more precisely, stably, and reliably.

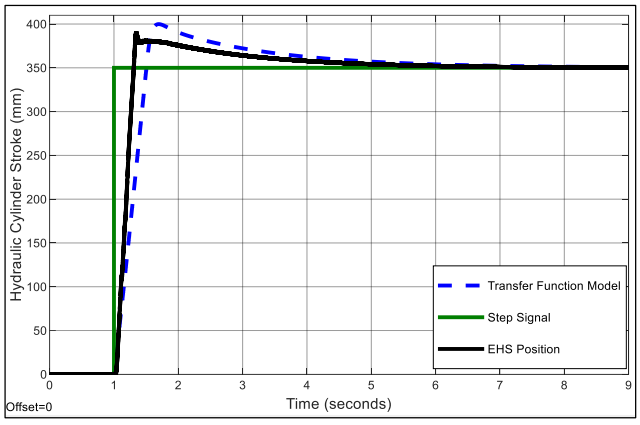


Fig. 15. Transfer function model and EHS position graphs (Green: step signal; Black: EHS position; Blue: transfer function model position).

In this study, since the model of the EHS provided the transfer function with the highest accuracy and fidelity, the determination of the PI parameters was carried out within a MATLAB/Simulink simulation environment and subsequently applied to the EHS. The PI controller parameters for the control of the EHS were determined based on the model developed in the Simulink environment. Initially, the proportional gain K_p and integral gain K_i were both set to 0, and the system response was observed. Subsequently, the K_p and K_i values were gradually increased to optimize the behavior of the closed-loop system. During the simulation process, the accuracy and stability of the EHS position control were analyzed, and the optimal values were determined as K_p = 0.28 and K_i = 0.14. These parameters were then integrated into the LabVIEW software and implemented in the system. As a result, it was verified that the PI controller successfully adapted to the dynamic characteristics of the EHS and effectively achieved the desired position control.

A LabVIEW control loop speed of 1 ms was employed when implementing the PI controller coefficients on the EHS and generating the desired step signal. Figure 15 presents the system's closed-loop response under PI control, showing the actual EHS position, the applied step reference signal, and the theoretical response predicted by the identified transfer function model. This model, derived from the system identification process, was used in the MATLAB/Simulink environment for tuning the controller parameters. The close agreement between the real system response and the model output verifies both the effectiveness of the PI controller and the accuracy of the identified model used in the control design.

IV. Conclusion and Discussion

This study comprehensively examined the dynamics of an EHS in a laboratory-scale wave channel using black-box modeling and advanced control techniques. Utilizing the SIT in MATLAB/Simulink, various model structures—including continuous-time transfer functions, state-space models, and discrete-time ARX models—were developed and evaluated under different initial conditions and sampling times. According to comparative simulation results, a third-order continuous-time transfer function was identified as the most suitable model structure. Since responses were obtained from the transfer function model with a 1-millisecond cycle time, the control program in LabVIEW was also set to a 1-ms cycle time for the controller.

Additionally, based on the most accurate model obtained, a PI controller was implemented for EHS position control using LabVIEW. The controller parameters were optimized in the MATLAB/Simulink environment, with the proportional gain K_p set to 0.28 and the integral gain K_i set to 0.14. The integration of this PI controller into the system provided significant improvements in tracking performance and system stability by rapidly responding to instantaneous deviations and accounting for accumulated errors. Overall, the findings of this study highlight the critical importance of accurate system identification in enhancing the performance of electro-hydraulic systems in wave generation. The selected modeling approaches and control strategies improve the predictability and reliability of such systems, providing a robust foundation for future research.

Although the developed models and the PI control implementation yielded promising results, the study focused on a specific system configuration and controlled operating conditions. Factors such as external disturbances, strong nonlinearities beyond the tested range, and long-term operational behavior were not within the scope of this work. Future research may explore advanced nonlinear identification techniques—such as neural networks or hybrid

grey-box models—to better capture complex dynamics, and validate the approach under broader operating scenarios to improve generalizability.

REFERENCES

[1] M. Özbulut, "Düzenli Dalgalar Üreten Bir Sayısal Dalga Tankının SPH Yöntemi ile Modellenmesi," Uludağ University Journal of The Faculty of Engineering, pp. 551–570, Sep. 2019, doi: 10.17482/uumfd.509229.

[2] M. A. Drzewiecki and J. Guzinski, "Application of the ISE Optimized Proportional Control of the Wave Maker in a Towing Tank," IEEE Access, vol. 10, pp. 42151–42162, 2022, doi: 10.1109/ACCESS.2022.3168047.

[3] S. Mahjouri, R. Shabani, G. Rezazadeh, and P. Badiei, "Active Control of A Piston-Type Absorbing Wavemaker with Fully Reflective Structure," China Ocean Engineering, vol. 34, no. 5, pp. 730–737, Sep. 2020, doi: 10.1007/s13344-020-0066-9.

[4] Y. Liu, Y. Zheng, R. Song, J. Chen, and H. Jin, "Wave generation characteristic analysis of piston and flap type wave maker with rotary-valve-control vibrator," Journal of Vibration and Control, vol. 26, no. 15–16, pp. 1297–1308, Aug. 2020, doi: 10.1177/1077546319895664.

[5] B. Nouioui and M. Doğan, "Irregular Wavemaker (Piston Type) in a Numerical and Physical Wave Tank," Usak University Journal of Engineering Sciences, vol. 5, no. 2, pp. 95–116, Dec. 2022, doi: 10.47137/uujes.1180866.

[6] M. Y. Coşkun and M. İtik, "Reinforcement learning based position control of an electro-hydraulic system," Niğde Ömer Halisdemir University Journal of Engineering Sciences, vol. 12, no. 1, pp. 280–288, 2023, doi: 10.28948/ngmuh.1163241.

[7] H. U. Akova and T. Balkan, "Elektro-Hidrolik Yük Simülatörü İçin Geribesleme Doğrusallaştırma Yöntemi ile Kuvvet Kontrolü Tasarımı," in *Proceedings of the VIII. Ulusal Hidrolik Pnömatik Kongresi*, Istanbul, Türkiye, 2012, pp. 295–306.

[8] Rahmat, "Modeling and Controller Design of an Electro-Hydraulic Actuator System," Am J Appl Sci, vol. 7, no. 8, pp. 1100–1108, Aug. 2010, doi: 10.3844/ajassp.2010.1100.1108.

[9] M. F. Rahmat, S. Rozali, A. Wahab, and Zulfatman, "Application Of Draw Wire Sensor In Position Tracking Of Electro Hydraulic Actuator System," *International Journal on Smart Sensing and Intelligent Systems*, vol. 3, no. 4, pp. 736-755, 2010.

[10] Y. Lin, Y. Shi, and R. Burton, "Modeling and Robust Discrete-Time Sliding-Mode Control Design for a Fluid Power Electrohydraulic Actuator (EHA) System," IEEE/ASME Transactions on Mechatronics, vol. 18, no. 1, pp. 1–10, Feb. 2013, doi: 10.1109/TMECH.2011.2160959.

[11] J. Das, S. Kr. Mishra, R. Saha, S. Mookherjee, and D. Sanyal, "Nonlinear modeling of an electrohydraulic actuation system via experiments and its characterization by means of neural network," Journal of the Brazilian Society of Mechanical Sciences and Engineering, vol. 40, no. 2, p. 58, Feb. 2018, doi: 10.1007/s40430-018-0979-x.

[12] S. Salleh, M. F. Rahmat, S. M. Othman, and H. Z. Abidin, "Application of Draw Wire Sensor in the Tracking Control of an Electro-Hydraulic Actuator System," *Journal of Technology*, vol. 73, no. 6, pp. 51–57, 2015.

[13] R. Ghazali, M. Sam, M. F. Rahmat, K. Jusoff, Zulfatman, and A. Hashim, "Self-Tuning Control of an Electro-Hydraulic Actuator System," *International Journal on Smart Sensing and Intelligent Systems*, vol. 4, no. 2, pp. 189-204, 2011.

[14] J. Yan, B. Li, H.-F. Ling, H.-S. Chen, and M.-J. Zhang, "Nonlinear State Space Modeling and System Identification for Electrohydraulic Control," Math Probl Eng, vol. 2013, pp. 1–9, 2013, doi: 10.1155/2013/973903.

[15] S. Kizir, T. Yaren, and E. Kelekçi, Matlab Simulink Destekli Gerçek Zamanlı Kontrol: Teori ve Mühendislik Uygulamaları. Ankara, Türkiye: Seçkin Yayıncılık, 2019.

[16] L. Jin and Q. Wang, "Accurate model identification of the inertial mass dynamic of hydraulic cylinder with model uncertainty," Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering, vol. 233, no. 5, pp. 501–510, May 2019, doi: 10.1177/0959651818802113.

[17] A. Mitov, T. Slavov, and J. Kralev, "Robustness Analysis of an Electrohydraulic Steering Control System Based on the Estimated Uncertainty Model," Information, vol. 12, no. 12, p. 512, Dec. 2021, doi: 10.3390/info12120512.

[18] A. Mitov, J. Kralev, T. Slavov, and I. Angelov, "Comparison of Model Predictive Control (MPC) and Linear-Quadratic Gaussian (LQG) Algorithm for Electrohydraulic Steering Control System," E3S Web of Conferences, vol. 207, p. 04001, Nov. 2020, doi: 10.1051/e3sconf/202020704001.

[19] B. Yu, Q. Zhu, J. Yao, J. Zhang, Z. Huang, Z. Jin, and X. Wang, "Design, mathematical modeling and force control for electro-hydraulic servo system with pump-valve compound drive," *IEEE Access*, vol. 8, pp. 171988–172003, 2020, doi: 10.1109/ACCESS.2020.3012091.

[20] J. Zhao, D. Song, B. Zhu, Z. Chen, and Y. Sun, "Nonlinear Backstepping Control of Electro-Hydraulic Brake System Based on Bond Graph Model," *IEEE Access*, vol. 8, pp. 19100–19112, Jan. 2020, doi: 10.1109/ACCESS.2020.2968513.

[21] J. Schwarz and B. Lohmann, "Robust identification and control of mobile hydraulic systems using a decentralized valve structure," Control Engineering Practice, vol. 151, Art. no. 106030, 2024. Available: https://doi.org/10.1016/j.conengprac.2024.106030.

[22] A. He, L. Wei, Q. Lu, and P. He, "An Investigation of Energy Consumption Characteristics of the Pump-Control System for Electric Excavator Arms," Applied Sciences, vol. 14, no. 23, Art. no. 10791, Nov. 2024. doi: 10.3390/app142310791.

[23] S. Han, G. Orzechowski, J.-G. Kim, and A. Mikkola, "Data-driven friction force prediction model for hydraulic actuators using deep neural networks," Mechanism and Machine Theory, vol. 192, Art. no. 105545, 2024. doi: 10.1016/j.mechmachtheory.2023.105545.

[24] S.-W. Kim, B. Cho, S. Shin, J.-H. Oh, J. Hwangbo, and H.-W. Park, "Force Control of a Hydraulic Actuator With a Neural Network Inverse Model," IEEE Robotics and Automation Letters, vol. 6, no. 2, pp. 2814–2821, Apr. 2021, doi: 10.1109/LRA.2021.3062353.

[25] S. Jiang, H. Wang, and G. Zhao, "Research on neural network model reference adaptive disturbance rejection control of digital hydraulic cylinder," Advances in Mechanical Engineering, vol. 14, no. 12, pp. 1–15, 2022, doi: 10.1177/16878132221140706.

[26] R. Dindorf, "Dynamic Modeling and Simulation of a Discrete Incremental Hydraulic Positioning System Controlled by Binary Valves," Applied Sciences, vol. 14, no. 7, Art. no. 2973, Apr. 2024. doi: 10.3390/app14072973.

[27] National Instruments Corp., "cRIO-9074 - NI." Accessed: Jul. 20, 2024. [Online]. Available: https://www.ni.com/en-tr/support/model.crio-9074.html

BIOGRAPHIES

Batin Demircan graduated with a B.Sc. degree in electrical and electronics engineering from the Kirikkale University in 2015, M.Sc. degree in electrical and electronics engineering from the Balikesir University in 2019. Since 2022, he has been an Lecturer and PhD.Candidate with the Electronics and Automation Department, Balikesir Vocational School, Balikesir University, Balikesir. His research interest include control systems, renewable energy systems and IoT distributed systems.

Tugce Yaren graduated with a B.Sc. degree in Mechatronics Engineering in 2014, received her M.Sc. degree in 2018, and completed her Ph.D. in 2024, all from Kocaeli University, Kocaeli, Turkey. From 2016 to 2024, she worked as a Research Assistant in the Department of Mechatronics Engineering at Kocaeli University. Since 2024, she has been serving as an Assistant Professor in the Department of Electrical and Electronics Engineering at Balıkesir University. Her research interests include the modeling of nonlinear systems, real-time control, system dynamics and control (e.g., optimal control, model predictive control), and robotics.

Ersin Akyuz is an Associate Professor at the Department of Electrical and Electronics Engineering, Faculty of Engineering, Balikesir University, Balikesir, Turkey. He graduated from the Istanbul Technical University in 1996. He received his MSc in Gazi University, Department of Electric-Electronic in 2003. He received his PhD in Balikesir University Mechanical Engineering in 2010. His research interests are renewable energy, hydrogen energy, and IoT distributed systems.

Sabri Bicakci was born in Balikesir, Turkey in 1982. He received the B.S. degree in electronics engineering from the Erciyes University, Kayseri, in 2005, M.S. degree in electrical and electronics engineering from the Balikesir University, Balikesir, in 2009 and the Ph. D. degree in mechanical engineering from Balikesir University, Balikesir, in 2012. From 2005 to 2015, he was a Research Assistant with Electrical and Electronics Engineering Department, Balikesir University. Since 2015, he has been an Assistant Professor with the Mechatronics Engineering Department, Balikesir University. His research interest includes robotics and control systems.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 3, September 2025
Research Article

Renewable Energy Transition in Türkiye: The Impact of Solar and Wind Based Capacity Growth on Market Prices and Cost Dynamics

Tamer Emre

***Abstract*— This study investigates the impact of solar and wind-based energy capacity expansion on electricity market prices and cost structures in Türkiye from 2020 to 2024. The rapid integration of renewable energy resources has significantly transformed Türkiye's electricity market, a key example due to the country's ambitious renewable energy goals. Using high-frequency data from the Turkish Electricity Market (EPİAŞ) and applying econometric methods such as time-series analysis and regression techniques, we demonstrate that the substantial growth in renewable energy capacity has resulted in a pronounced "duck curve" effect. This has led to significant midday price reductions, with solar based energy generation increasing over tenfold, contributing to a 20% drop in midday prices. Wind energy, though variable, also contributed to lower marginal prices during periods of high generation. Despite these benefits, the intermittency of renewable resources remains a challenge, highlighting the continued need for conventional energy resources such as imported coal and natural gas for grid stability. Market interventions, including price caps and renewable energy support mechanisms (YEKDEM), have mitigated extreme price volatility. However, they emphasize the need for improved grid flexibility, energy storage solutions, and more accurate forecasting. This study provides empirical evidence of the duck curve effect in Türkiye's electricity market, offering policy recommendations for optimizing renewable energy integration, enhancing forecasting precision, and reducing system costs. Our findings emphasize the dual role of renewables in driving price reductions while posing challenges for grid management, underscoring the importance of innovative market mechanisms and technological advancements for a sustainable energy transition.**

Index Terms—**Duck-curve, marginal price calculation PTF, renewable energy, solar capacity, Turkish energy market, wind capacity.**

I. Introduction

ENERGY SYSTEMS are one of the fundamental pillars of modern societies, playing a critical role in economic growth and sustainable development (UN, 2025). The increasing global demand for energy, combined with the limited reserves of fossil fuels and their environmental impacts, has heightened the interest in renewable energy resources for power generation (IEA, 2024). Among renewable energy resources, solar and wind-based power have gained prominence in recent years due to reduced costs and efficiency improvements driven by technological advancements (EMBER, 2024). This transformation is seen as a key strategy for ensuring energy security, reducing environmental impacts, and enhancing energy independence (MENR, 2017).

Türkiye, with its strategic location bridging Europe and Asia, is increasingly impacted by the global shift towards renewable energy due to its rising energy demands fueled by industrialization, urbanization, and population growth. These dynamics have not only escalated energy consumption but have also necessitated the adoption of sustainable energy generation strategies. In response, Türkiye's energy policies have prominently featured renewable energy investments aimed at enhancing energy security and minimizing external dependencies. The comparative lower operational costs and environmental advantages of solar and wind-based energy over traditional fossil fuel-based power plants have further propelled the adoption of these renewable resources.

With a population exceeding 80 million and a rapidly evolving socio-economic and technological landscape, Türkiye stands out in the energy economy due to its geopolitical position at the heart of regional energy resources. Despite the reflection of cultural norms and daily life practices in electricity consumption habits, longitudinal comparisons offer insights into the direction of these changing patterns. The renewable energy transition, which has been gaining momentum since the early 2000s when natural gas was the predominant price-setting resource, has significantly altered the dynamics of electricity markets designed during that era. Although non-market interventions were thought to prevent market disruptions during crises, such approaches proved insufficient during the energy crisis that began in 2021 and waned by 2023. During this turbulent period, Türkiye implemented unique measures such as price caps and reasonable profit methods to stabilize energy prices, despite rising costs. Situated in a geography prone to nearly all types of natural disasters, Türkiye, much like its resilient population, has rapidly adapted to challenging conditions. In these adaptive processes, it remains an open question whether the increased share of renewable energy has led to lower energy costs and whether the current market mechanisms support a more affordable and sustainable energy policy.

Tamer Emre, is the Director of Strategic Planning & Corporate Governance of Akenerji Elektrik Üretim A.Ş. Miralay Şefik Bey Sk., Akhan, No:15, Kat:4, 34437, Gümüşssuyu / Beyoğlu / İstanbul / Türkiye,(e-mail: tamer.emre@akenerji.com.tr), IEEE Senior Member
https://orcid.org/0000-0001-9354-8128

Manuscript received Jan 09, 2025; accepted Jul 12, 2025.
DOI: 10.17694/bajece.1636476

Studies on countries like Germany and the Netherlands indicate that while renewables lower market prices, they can also contribute to price volatility (Liebensteiner, 2025) (Wang, 2022). In Türkiye, the transition to renewable energy has been particularly impactful due to the country's strategic location and rising energy needs. Industrialization, urbanization, and population growth have continuously increased energy demand, necessitating the development of sustainable solutions.

This study hypothesizes that increasing wind and solar based energy capacities in Türkiye have significantly influenced electricity market prices by reducing peak-hour costs, creating a duck-curve effect, and introducing market volatility that requires further regulatory adjustments.

Türkiye's energy policies aim to ensure energy security and reduce dependence on imports by fostering renewable energy investments. Compared to conventional fossil-fuel power plants, solar and wind-based energy offer lower operational costs and environmental benefits, making them a preferred choice for expansion.

II. Renewable Energy Development in Türkiye

A. *The Role of Renewable Energy in Market Transition*

The integration of renewable energy resources plays a vital role in the transition to a low-carbon economy. Solar and wind-based energy contribute to reducing carbon emissions and lowering long-term energy costs. However, studies suggest that while renewables enhance sustainability, they can lead to increased price fluctuations in electricity markets (Pollitt.M., 2024). Similar trends are observed in Türkiye, where improving energy storage technologies and demand-side management mechanisms are crucial for mitigating such effects. Figure 1 illustrates the trends in installed energy capacities for solar based, wind based, thermal, and hydroelectric energy resources in Türkiye between 2020 and 2024, along with the average energy prices. The figure reveals a consistent upward trajectory in renewable energy capacities, particularly for solar and wind-based energy, underscoring the country's transition towards sustainable energy generation. The Market Clearing Price (MCP), which is PTF in Türkiye, depicted by the green line, shows a continuous decline after 2022, largely due to the tenfold increase in energy generation driven by the Global Energy Crisis. This reduction in prices can be attributed to the growing share of cost-efficient renewable energy resources in the energy mix. The figure offers valuable insights into the dynamic relationship between energy diversification and its economic implications over the observed period.

Türkiye has taken significant steps to diversify its energy portfolio through regulations and support mechanisms that promote the use of renewable energy resources. Particularly since the mid-2010s, initiatives such as the Renewable Energy Resources Support Scheme (YEKDEM) have accelerated investments in solar and wind-based energy. Technological advancements in efficiency and cost reductions have further enhanced the scalability of these resources, contributing to an increasing share of renewables in the country's energy generation. Large-scale tenders for Renewable Energy Resource Areas (YEKA), which are pre-permitted and auctioned, have been recognized as best practice examples globally. The Konya Karapınar 1 GW Solar based Power Plant, one of the largest solar based fields in Europe, is a prime example of this approach.

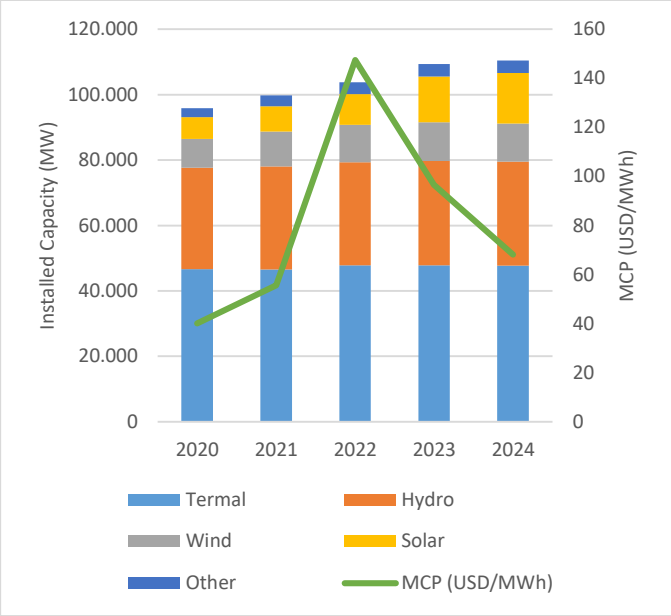


Fig.1. Trends in installed energy capacity and Market Clearing Prices (MCP) in Türkiye (2020–2024) (EMRA, 2020) (EMRA, 2021) (EMRA, 2022) (EMRA, 2023) (EMRA, 2024) (EXIST, 2025)

Figure 2 illustrates the growth trends of solar and wind-based energy from renewable resources in Türkiye. Notably, solar based energy experienced a rapid increase, particularly after 2020, while wind energy demonstrated a steady and continuous upward trajectory. According to the National Energy Plan of Türkiye, the targets for 2035 include 50 GW of wind energy capacity and 70 GW of solar based energy capacity (MENR, 2024).

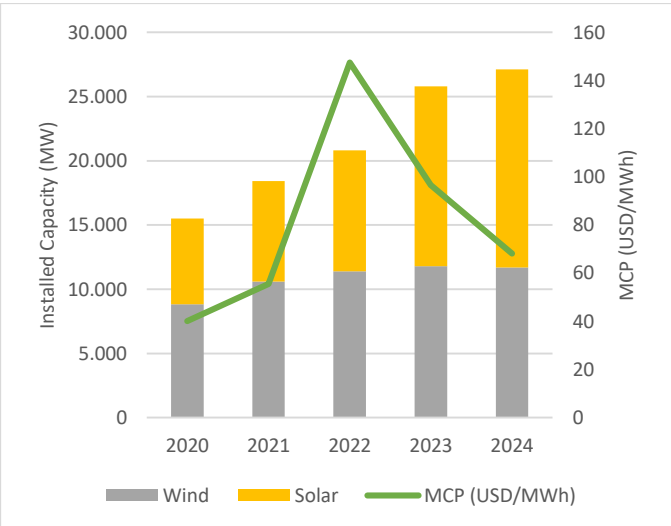


Fig.2. Change in renewable energy capacity in Türkiye

B. Technological Advancements in Solar and Wind based Energy

The last decade has witnessed remarkable advancements in solar based photovoltaics and wind turbine technologies. Lower costs and improved efficiencies have driven increased investments. However, renewables are subject to the cannibalization effect, where increased generation from solar and wind resources reduces wholesale electricity prices and subsequently diminishes their own profitability (Lopez Prol, 2020).

Similarly, wind energy technologies have seen increased generation capacity due to advancements in turbine design, the use of larger rotor diameters, and improvements in weather forecasting models. Wind energy, particularly during nighttime hours, contributes to electricity generation, offering the potential to enhance supply security through the establishment of hybrid energy systems when combined with solar based power. However, wind energy generation remains subject to seasonal variability and short-term weather conditions.

C. Impact on Market Prices and Policy Implications

Electricity market dynamics are significantly influenced by supply-demand balance. The expansion of renewables has led to lower electricity prices at certain hours while creating price instability. Studies show that solar and wind-based energy lower day-ahead market prices but can introduce imbalances in the system (Chava, 2024).

Furthermore, the integration of renewable energy resources into energy markets requires the development of various market mechanisms to ensure security of supply and price stability. The intermittent nature of renewable energy generation can increase system balancing costs and pose challenges to energy security. Therefore, the development of mechanisms such as energy storage solutions, smart grid systems, and demand-side management is critical to successfully implementing the renewable energy transition.

Since Türkiye's 2035 Road Map Plan has been announced as 120 GW - renewable energy capacity from wind and solar resources, to achieve this target, a variety of policies and strategies have been developed. Incentives for renewable energy investments, market regulations, and the promotion of private sector participation in renewable energy projects have contributed to the accelerated growth of this sector. However, in order to ensure the sustainability of the energy transition, further regulatory frameworks are needed to improve the effectiveness of market mechanisms. In particular, the integration of energy storage systems, the strengthening of grid infrastructure, and the development of renewable energy forecasting algorithms will allow Türkiye to more efficiently harness its renewable energy potential. Additionally, in line with the goals of reducing carbon emissions and decreasing dependence on fossil fuels, more innovative solutions in the energy sector need to be implemented.

III. Data Collection and Mathematical Framework

In this study, various data resources were utilized to examine the impact of the increase in renewable energy capacity on electricity prices in Türkiye. The data were obtained from official institutions such as TEİAŞ, EPİAŞ, EPDK, and TÜİK, and were supplemented with reports from international energy agencies and market analyses. Time series analysis, regression models, and optimization techniques were employed to assess the effects of solar and wind-based energy on market prices. As a result of these analyses, it was found that the increase in solar and wind-based energy generation between 2020 and 2024 in Türkiye led to a decrease in electricity prices, and the duck curve effect became more pronounced. Furthermore, it was concluded that new regulations are required to support the integration of renewable energy into market mechanisms.

The data obtained from the EPIAŞ Transparency Platform (2020, 2021, 2022, 2023, and 2024 data) were considered to represent a five-year period that accurately reflects the dynamics of the market. It was evaluated that a longer period would have a higher likelihood of containing misleading effects.

Various data cleaning and processing techniques were employed to ensure the reliability of the obtained data and their suitability for analysis:

Cleaning of Missing and Erroneous Data: Linear interpolation and average values from previous years were used for missing data. Erroneous or outlier values were identified using the IQR (Interquartile Range) method and excluded from the dataset.

Data Merging and Visualization: Data from different resources were merged on a common time scale. The data were processed and figures were created using Python (Pandas, NumPy) and MATLAB for time series analysis.

A. Methodology

Among the data used, consumption data, which is the primary driver of the price formation mechanism, was first analyzed. Hourly generation data for each resource-based power plant was also examined. In the price formation mechanism, data for independent purchases and independent sales quantities, as well as the curve gap between these two sets of data, were utilized.

$$\text{Curve gap} = \text{price independent sales} - \text{price independent purchases} \quad (1)$$

The curve gap is a fundamental component of the supply-demand curve, and an increase in this gap leads to a rise in prices. The increase in the share of independent sales (PIS) for power plants benefiting from renewable energy support mechanisms enables the curve to intersect at a lower price. This relationship can be observed through regression analysis. Using the established optimization model, it is possible to obtain the necessary parameters for the desired target. It is also known that national investment plans are based on the results of such analyses.

In markets where the theoretical maximum price is infinite, the actual value is derived by multiplying or approaching the theoretical maximum price in markets with price caps. This

ratio is referred to as the "saturation index." The correlation between the saturation index and the wind capacity generation curve will reveal the effect of this resource type on prices. Such an analysis, as applied in Türkiye's electricity market, cannot be implemented in European markets. Following the 2021 energy crisis, measures were introduced to set a new price cap once the saturation index reaches 70% of the theoretical maximum (which was 60% in previous version) (ACER, 2023).

$$\text{Saturation index} = \text{hourly marginal price (MCP)} / \text{theoretical maximum price} \quad (2)$$

To observe the effect of solar based capacity on prices, the correlation between generation levels and prices must be examined. To identify the structure and sensitivity of the duck curve in pricing, the maximum solar based generation and illumination values at noon (12:00 PM) were utilized. Even if the generation at this time does not meet the demand, the midday consumption pattern still generates a curve, providing a baseline price relative to the average price for this consumption. The downward trends of this baseline value with annual generation and a corresponding index reveal the duck-curve structure in the price curve.

B. *Mathematical Background*

In this study, various mathematical and statistical methods have been utilized to examine capacity increases, price changes, and market dynamics:

1) Time Series Analysis

The ARIMA (AutoRegressive Integrated Moving Average) model has been employed to investigate the temporal fluctuations in electricity prices and generation volumes. The Fourier transform has been applied to analyze the hourly variations in solar and wind-based energy generation.

A multiple regression analysis has been conducted to measure the impact of capacity increases in solar and wind-based energy on prices. Independent variables include solar and wind-based capacity (in megawatts, MW), and the prices of imported coal and natural gas (in USD per MWh). The dependent variable is the PTF (Market Clearing Price, in Turkish Lira per MWh). Regression model:

$$PTF=\beta0+\beta1\times SolarMW+\beta2\times WindMW+\beta3\times CoalUSD/MWh+\beta4\times GasUSD/MWh+\epsilon \quad (3)$$

2) Duck-Curve and Daily Price Change Model

Kernel Density Estimation (KDE) has been utilized to model the price drops caused by high solar based generation during midday hours. To demonstrate the duck-curve effect, the relationship between solar based energy generation at noon and market prices has been analyzed.

3) Optimization Models

Linear programming has been applied to analyze the impact of renewable energy generation on base-load power plants. The objective function is to minimize the total system costs. Constraints include demand-response limits, generation capacity limits, and emission restrictions.

These mathematical approaches allow for an assessment of the impact of renewable energy resources on the system from both economic and technical perspectives.

4) Limitations

This study utilized data from EPIAŞ, which lacks region-specific granularity. The methodology for calculating market prices in Türkiye aligns with the common optimization equation used in Europe, yet it retains unique characteristics. In this research, specifics concerning the optimization calculations were not examined, and interpretations were solely based on the observed outcomes. Therefore, results from different countries have been reviewed in the literature, but the analyses have focused on Türkiye.

IV. ANALYSIS

When examining the data from the Turkish electricity market, it cannot be said that consumption has continuously increased during the period between 2020 and 2025. According to Figure-3, the significant rise in consumption observed in 2023 is primarily due to the earthquake on February 6, 2023, which nearly erased three provinces from the map. Consumption, which had been showing a cautious increase until 2023, is notably above the normal trend in 2024. This can be attributed to a significantly higher air conditioning load during an unusually hot summer period and the increased use of devices in the advancing technological world contributing to this rise in consumption.

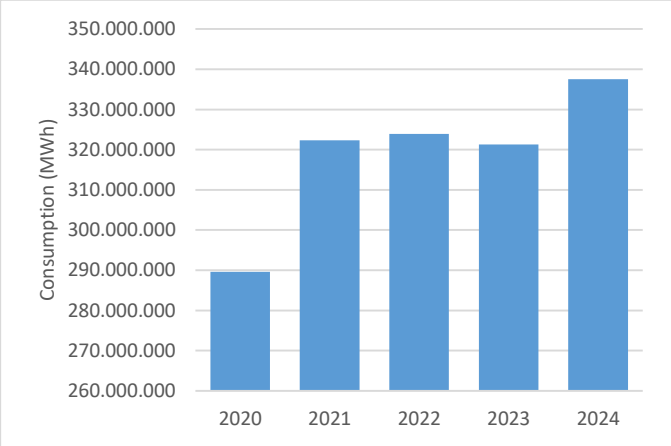


Fig.3. Change in consumption for 5 years period (EPİAŞ Data)

In recent years, heightened awareness of climate change has prompted a strategic shift from traditional portfolio balancing to a focus on diversifying energy resources through a green transformation agenda, prioritizing renewable energy resources. This shift is particularly evident in the increased investments in wind and solar based energy sectors, driven by technological advancements and declining costs. Figure 4 and 5 depict the evolution of energy generation between 2020 and 2025, incorporating the impact of additional capacity introduced during this period.

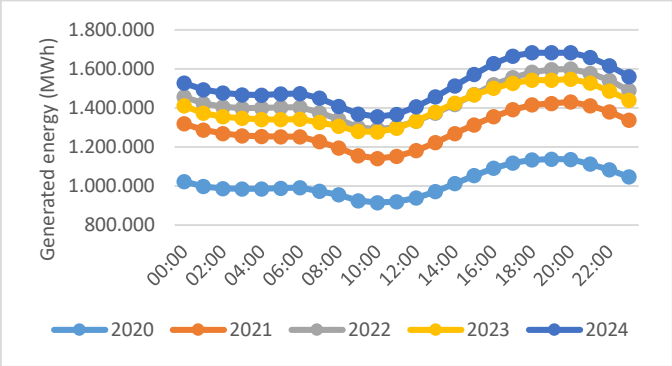


Fig.4. Annual generation volumes of wind power plants (EPİAŞ data)

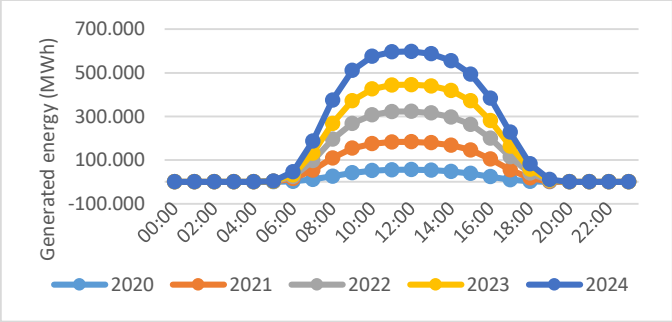


Fig.5. Annual generation volumes of solar power plants (EPİAŞ data)

As investments in these resources increase, there is also an observable rise in generation volumes. For instance, the continual increase in wind generation over the years, except for 2022, indicates that the year-to-year variability in wind conditions remains below the level of investment. In 2022, a less favorable year for wind saw generation occasionally drop below the levels of the previous year.

Regarding the regression analysis of the relationship between curve gap and saturation index for the year 2024, which aims to exclude the effects of the 2021-2023 energy crisis, the r-squared value is 0.61. Despite other drivers, this value indicates a strong correlation.

Solar based generation exhibits a consistent year-over-year increase. The illumination that starts after 6 AM continues until 7 PM, peaking at noon. The generation at noon between 2020 and 2024 has shown an increase of more than tenfold.

As for the dynamics of conventional energy resources without additional investments, what are the trends? In Türkiye, market pricing is predominantly influenced by two conventional resources: imported coal and natural gas. These resources are distinguishable by their operational flexibility in terms of activation and deactivation times. Given that the accuracy of wind forecasts hovers around 15%, these fuels continue to play a critical role as base-load suppliers in the energy market. An analysis of Figure-6 reveals that, except for the supply crisis in 2021, there has been a consistent upward trend in the utilization of imported coal, solidifying its essential role as a base-load resource.

Figure-7 indicates that during the crisis period marked by shortages in imported coal supply, the generation from natural gas resources was exceptionally high, particularly noting a decline in consumption in 2023 and 2024. From a market perspective, the underutilization of natural gas resources, which can be rapidly ramped up and down, implies that the marginal price did not reach its highest possible level and that the required energy was instead met through alternative resources (in this case, additional wind and solar based capacity). The fact that these resources are compensated through YEKDEM via a non-market method creates a zero-priced supply in terms of market dynamics.

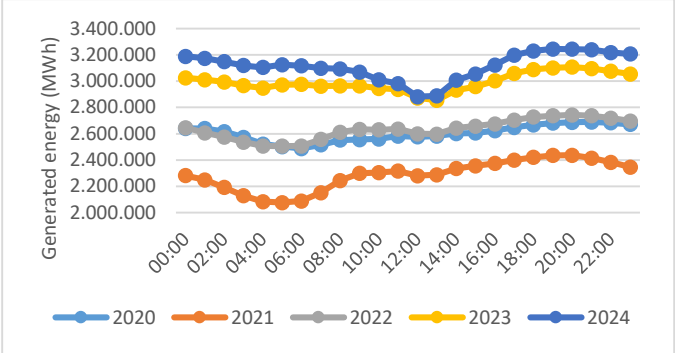


Fig.6. Annual generation volumes of imported coal power plants (EPİAŞ data)

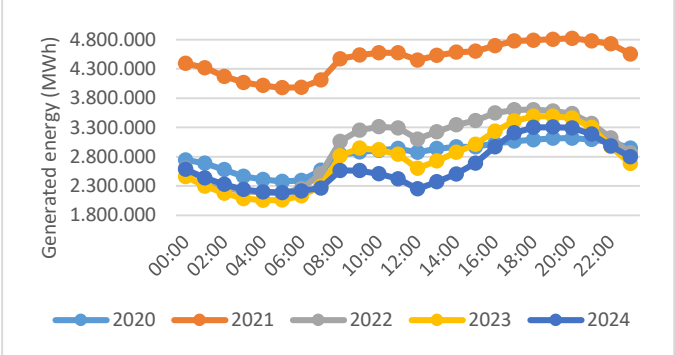


Fig.7. Annual generation volumes of natural gas power plants (EPİAŞ data)

Figure-8 displays the trajectory of price-independent sales offers over the years. According to the figure, the increases in 2022 and 2023 remained below the levels of 2021. This could be due to a decrease in generation from various resources (such as wind and hydroelectric power). In 2024, an increase in price-independent capacity is observed, influenced by both wind and solar based capacities. Another important aspect to note is the surplus of capacity in 2021 and 2022 that opted not to participate in the support mechanism. This capacity is not zero-cost in terms of market dynamics and does impact the marginal price.

Figure-9 shows the trajectory of the reference price over the years. During 2022 and 2023, the global energy crisis led to increased costs, which, in turn, caused marginal prices to rise; however, these increases were kept within certain limits through market regulations. During this period, the market was supported by non-market mechanisms in line with the literature. Assuming a traditional load curve for 2020 and 2021, a limited price drop during midday hours can be observed. In contrast, a significant decline in marginal prices at midday is noted for the years 2022, 2023, and 2024. This decline is attributed to the relaxation of supply and demand dynamics in the market. The

alleviation is markedly demonstrated by the consistent influence of wind energy on prices throughout the day.

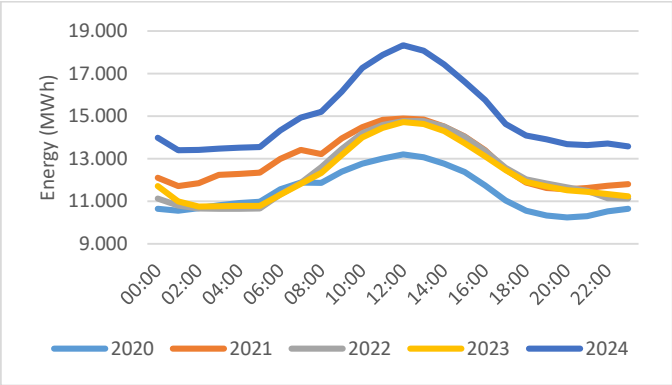


Fig.8. Price-independent sales offers (EPİAŞ data)

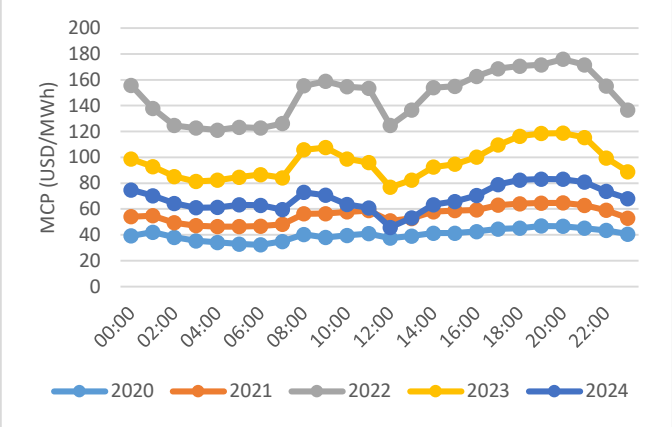


Fig.9. Annual changes in MCP (EPİAŞ data)

Figure-10 employs a different approach, displaying the ratio of solar based generation at its peak generation hour, noon, to the daily average generation. This comparison allows for a clearer observation of how high solar based energy generation at noon influences the downward trend in prices. The figure shows that in 2020, generation at noon was 94% of the daily average, but this percentage subsequently decreases to 92%, 85%, 80%, and 67% over the years. This trend clearly indicates the formation of a duck-curve pattern in the price curve.

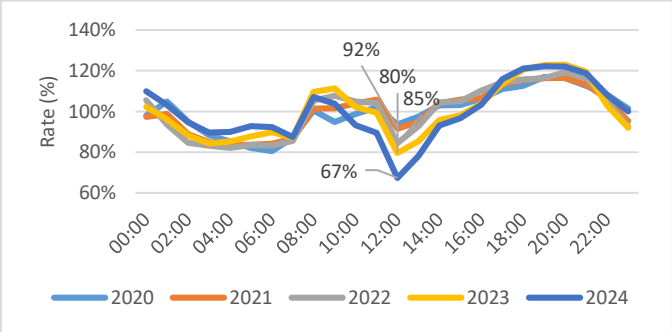


Fig.10. Ratio of noon-prices to daily average

The integration of renewable energy resources into the system often results in imbalances due to the inadequacy of forecast data. Figure-11 displays the quantities of 0 and 1 coded YAL (upward regulation) and YAT (downward regulation) instructions, which reflect the costs in the balancing markets as the system operator faces an increasing share of renewable energy resources and a decreasing generation of conventional fuels over the years. When evaluated independently of price, a regular decrease in these quantities is observed following the crisis period. The reason for not including the costs of the capacity mechanism in these calculations is that this mechanism inherently requires reserves to be held under all circumstances.

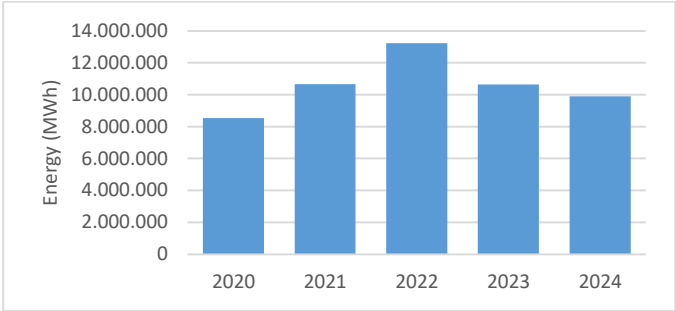


Fig.11. Annual changes of upward and downward regulation volumes (EPİAŞ data)

Figure-12 displays YEKDEM prices. Despite significant fluctuations during the crisis period, these prices have remained steadily close to and below market prices. In this context, it can be said that the capacity affecting price formation through price-independent sales, being compensated at prices below the formed marginal prices, contributes to reducing the overall costs.

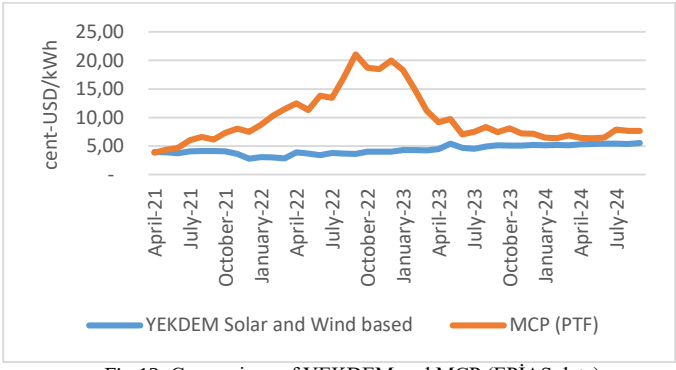


Fig.12. Comparison of YEKDEM and MCP (EPİAŞ data)

V. Results and Discussion

Solar and wind-based energy have significantly reduced peak electricity prices. Between 2020 and 2024, noon-time electricity generation from solar based increased tenfold, leading to an evident duck-curve effect. Moreover, the simultaneous observation of the duck-curve effect in both market prices and load profiles implies the growing role of distributed generation in meeting local demand, particularly through rooftop solar systems and behind-the-meter consumption.

Intermittency challenges persist, necessitating the continued use of natural gas for grid stability. Market mechanisms such as price caps have helped mitigate volatility, but long-term

stability requires storage solutions and improved forecasting models.

VI. Policiy Recomendations

To ensure a sustainable transition to renewables, this study proposes:

-Investment in Energy Storage: Battery storage and pumped hydro should be expanded to counteract price fluctuations.

-Improved Forecasting Models: AI-driven predictions should enhance grid balancing strategies.

-Market Design Adjustments: Revisiting the YEKDEM subsidy mechanism to align incentives with long-term market stability.

-International Benchmarking: Comparing Türkiye's market interventions with successful global examples, such as California's and Germany's capacity market designs.

VII. Conclusion

This study has thoroughly examined the impact of Türkiye's renewable energy transition on the electricity market. Recent capacity expansions in wind and solar-based power have significantly increased their contribution to total electricity generation, with solar-based energy generation during midday rising more than tenfold. This surge has resulted in a substantial decline in electricity prices during peak solar-based hours, visibly altering the market's supply-demand balance. Graphical analyses, particularly from 2022 to 2024, reveal that marginal prices have consistently dropped during midday, confirming the emergence of a pronounced duck curve in the Turkish electricity market. Furthermore, the analysis indicates that when wind generation is high, the depth of the price curve is reduced, reflecting a lower saturation index and decreased marginal prices. Both renewable energy resources have demonstrated a price-suppressing effect on the market.

However, the unpredictability of renewable energy remains a challenge. The relatively low forecasting accuracy of wind generation, in particular, underscores the continued necessity for baseload power resources. Imported coal and natural gas-based power plants remain critical in ensuring system reliability. During periods of crisis, increased natural gas-based generation has helped meet the system's flexibility requirements. Meanwhile, regulatory interventions and price control mechanisms have helped mitigate extreme price fluctuations; however, long-term sustainability requires further structural reforms.

In conclusion, renewable energy investments in Türkiye have substantially influenced the electricity market, especially through their impact on price mechanisms. Nonetheless, improving system reliability and enhancing renewable energy forecasting accuracy remain essential. Future developments such as smart grid systems, energy storage solutions, and refined market mechanisms will enable more efficient utilization of renewable energy resources, ultimately reducing costs and enhancing market stability.

References

[1] ACER. (2023, 01 10). Harmonised maximum and minimum clearing prices for single day-ahead coupling. Guideline on capacity allocation and congestion management (CACM Regulation) for EU 2015/1222. ACER.

[2] Chava, J. (2024). Impacts of energy price cap system on electricity and gas prices in the era of energy crisis. California State University, Dissertation Graduate Thesis. San Bernardino: California State University.

[3] EMBER. (2024). Türkiye Elektrik Görünümü. EMBER.

[4] ACER. (2023, 01 10). Harmonised maximum and minimum clearing prices for single day-ahead coupling. Guideline on capacity allocation and congestion management (CACM Regulation) for EU 2015/1222. ACER.

[5] Chava, J. (2024). Impacts of energy price cap system on electricity and gas prices in the era of energy crisis. California State University, Dissertation Graduate Thesis. San Bernardino: California State University.

[6] EMBER. (2024). Türkiye Elektrik Görünümü. EMBER.

[7] EMRA. (2020). Market Annual Report. Ankara: Energy Market Regulatory Authority (EMRA).

[8] EMRA. (2021). Market Annual Report. Ankara: Energy MArkets Regulatory Authority.

[9] EMRA. (2022). Market Annual Report. Ankara: Energy Markets Regulatory Authority.

[10] EMRA. (2023). Market Annual Report. Ankara: Energy Markets Regulatory Authority.

[11] EMRA. (2024). Monthly Report. Ankara: Energy Markets Regulatory Authority.

[12] EXIST. (2025, 01). Transparency Platform. EXIST: https://seffaflik.epias.com.tr/home

[13] IEA. (2024). World Energy Outlook. Paris: International Energy Agency.

[14] Liebensteiner, M. (2025). High electricity price despite expansion in renewables: How market trends shape Germany's power market in the coming years. Energy Policy, 198, 114448.

[15] Lopez Prol, J. (2020). The cannibalization effect of wind and solar in the California wholesale electricity market. Energy Economics, 85, 104552, https://doi.org/10.1016/j.eneco.2019.104552.

[16] MENR. (2017). National Energy Efficiency Implementation Plan 2017-2023. Ankara: Ministry of Energy and Natural Resources.

[17] MENR. (2024). 2035 Road Map in Renewable Energy Launch Presentation. (s. 7). İstanbul: MENR.

[18] Pollitt.M., G. d.-H. (2024). Recommendations for a future-proof electricity market design in Europe in light of the 2021-2023 energy crisis. Energy Policy, 188, 114051, https://doi.org/10.1016/j.enpol.2024.114051.

[19] UN. (2025, 01). The Sustainable Development Goals. United Natıons Official Page: https://turkiye.un.org/en/sdgs adresinden alındı

[20] Wang, D. (2022). Electricity price instability over time: Time series analysis and forecasting. Sustainability, 14,9081 https://doi.org/10.3390/su14159081.

BIOGRAPHIES

Tamer Emre Ankara, Türkiye, in 1978. He completed his B.Sc. in Electrical and Electronics Engineering at Gazi University in 2001, holds M.Sc. degrees in Electrical and Electronics Engineering (Power Systems) from Gazi University in 2004, and Business Administration (Production Management and Quantitative Methods) from Hacettepe University in 2014 and his PhD in Energy Systems Engineering from Gazi University in 2022.

Dr. Emre began his career in the energy sector in 2004 as a Foreign Trade Specialist at GAMA Holding, where he was responsible for procuring equipment for its affiliated companies. From 2005 to 2014, he held engineering and mid-level management positions at TEDAŞ and TEİAŞ, which are public enterprises. During this period, he was part of a small team that undertook the establishment activities of the Turkish

Electricity Market. He continued his career as the Coordinator of Electricity Market Services at PwC Turkey from 2014 to 2019. Between November 2019 and May 2021, he served as the Central Coordination Director at Kazancı Holding, where he was responsible for coordinating between the Management and the Electric Distribution Companies. Since May 2021, he has been working at EPİAŞ for about two and a half years as the Director of Energy Market Operations, managing operations of electricity and gas markets, environmental markets, and eligable consumer markets, as well as financial settlement operations. He has overseen the launch of the Futures Electricity Market, Futures Gas Market, and Renewable Energy Guarantees of Origin Markets, and has been involved in the design of numerous innovative projects. Since May 2024, he has been serving as the Director of Strategic Planning and Corporate Governance at Akenerji. Throughout his over 20 years of professional life in the energy field, Dr. Emre has also continued his academic work, currently teaching "Energy Economics and Environment" to undergraduate students at Gazi University's Faculty of Technology. He speaks English, Russian, and Japanese. He is married and a father of one daughter.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 3, September 2025

Research Article

Development of a New Computational System Supported by Artificial Intelligence for Detection of Real-Time Retinal Diseases

Hasan Memis and Emrullah Acar

Abstract— **Retinal diseases such as choroidal neovascularization (CNV), diabetic macular edema (DME), and drusen are among the leading causes of vision loss worldwide, requiring early and accurate diagnosis to prevent irreversible damage. Optical Coherence Tomography (OCT) provides high-resolution imaging of retinal structures, making it a valuable tool in ophthalmological diagnosis. This study presents a novel artificial intelligence (AI)-supported computer-aided diagnostic system for the real-time classification of retinal diseases using OCT images. The proposed system integrates a DenseNet-201 deep learning model with a hash-based data integrity mechanism and a user-friendly interface for clinical deployment. The DenseNet-201 model achieved superior performance with an accuracy of 94.42%, an F1- score of 0.9442, and an AUC of 1.00, outperforming other widely used models such as GoogleNet, ResNet50, and EfficientNetB0. Unlike existing systems, our approach includes automatic image validation, eliminates data redundancy through hashing, and is optimized for practical use via the Gradio interface. These features address major limitations in prior studies, such as a lack of real-time capability, data inconsistency, and insufficient clinical integration. The system not only improves diagnostic accuracy but also reduces clinician workload, ensuring faster and more reliable decision-making in the detection of retinal diseases. This work demonstrates the feasibility of deploying AI-powered diagnostic tools in real-world ophthalmic settings and lays the groundwork for future development of integrated, scalable healthcare solutions.**

Index Terms—**Retinal Diseases, Optical Coherence Tomography (OCT), Artificial Intelligence (AI), Deep Learning, Computer-Aided Diagnosis, Clinical Decision Support System, Real-Time Classification, Medical Image Analysis, DenseNet-201, GoogleNet, ResNet50, EfficientNetB0, Gradio Interface**

I. INTRODUCTION

IN MANY parts of the world, especially in Turkey, changes in eating habits, a decline in physical and social activities, and excessive engagement with technology have not only reduced overall quality of life but also contributed to serious health problems [1]. One of the most significant issues is the partial or total loss of vision. Additionally, with the rapid advancement of technology, screens have become deeply integrated into our daily lives, leading to prolonged screen exposure throughout the day. This prolonged exposure results in excessive eye fatigue and may contribute to future vision-related complications [2]. The primary objective of this study is to explore how the benefits of modern technology—particularly artificial intelligence—can be leveraged to support healthcare, rather than merely highlighting health warnings or precautions. Since the introduction of AI into everyday life, it has found applications across numerous fields, with a particularly significant impact in healthcare. Therefore, this study aims to provide a comprehensive and precise evaluation while also optimizing diagnostic efficiency and reducing time consumption.

Among all human functions, vision is undoubtedly one of the most critical, as it provides meaning to our interaction with the environment. With vision, the human environment gains meaning. In addition, thanks to the ability to see, muscle functions are fulfilled without restriction. We can explain this by the fact that a person who has lost their eyesight moves very slowly. Therefore, problems that may occur in eyesight will affect our quality of life significantly negatively. It will limit our activities considerably and will cause us to barely maintain our vital functions [3].

The human eye is one of the most complex and essential organs, responsible for facilitating communication with the external world through vision. It operates via a series of intricate biological processes and is composed of various protective and sensory structures. Because vision is highly sensitive, even minor damage or dysfunction in any of its anatomical components can lead to severe visual impairment. To prevent such ocular health issues, it is essential to understand the eye's anatomy and its associated functions. For this purpose, several key terms related to the diseases in question, which are the subject of this study, need to be examined simply [4].

The eye is the most complex sensory organ responsible for enabling vision and facilitating interaction with the external world. Its primary anatomical components include the cornea, iris, lens, sclera (outer fibrous layer), choroid (vascular layer), retina, vitreous gel, and optic nerves. Briefly, the process of vision begins when light rays from an object enter the eye through the cornea, pass through the iris—which adjusts in size according to light intensity—then pass through the lens, and finally converge onto the retina at the back of the eye as an inverted image. These rays falling on the yellow spot on the

Hasan Memiş, is with Department of Electrical Engineering of Batman University, Batman, Turkey, (e-mail: hasanmms21@gmail.com).
https://orcid.org/0009-0006-4366-4722

Emrullah Acar, is with Department of Electrical Engineering of Batman University, Batman, Turkey, (e-mail: emrullah.acar@batman.edu.tr).
https://orcid.org/0000-0002-1897-9830

Manuscript received June 5, 2025; accepted Jul 19, 2025.
DOI: 10.17694/bajece.1715185

retina stimulate the photoreceptors in this layer and send optical signals to the brain to be evaluated, and as a result, a meaningful image is formed from these signals [5][6].

The retina is a highly differentiated neuroectodermal tissue located at the innermost and posterior part of the eyeball, where light is focused. It consists of two distinct regions: the central retina (macula) and the peripheral retina. Photoreceptor cell bodies are situated in the outer nuclear layer, while bipolar, horizontal, and amacrine cell bodies are located in the inner nuclear layer [7]. Signals generated by the photoreceptors are transmitted to the brain via the optic nerves. There are two main types of photoreceptors in the retina: rods and cones, which absorb light and enable sharp, color-rich vision. The most common retinal diseases include choroidal neovascularization (CNV), diabetic macular edema (DME), and drusen. CNV refers to the formation of abnormal new blood vessels beneath the retina. These vessels often leak fluid, leading to dark spots in the visual field, distorted lines, and sudden vision loss. It is most frequently observed in the advanced stages of age-related macular degeneration [8]. DME is a complication of diabetic retinopathy characterized by fluid accumulation in the macula, resulting in blurred vision. Since the macula is responsible for detailed central vision, swelling in this area can cause significant visual impairment [9][10]. Drusen are yellowish deposits that accumulate under the retinal pigment epithelium and are considered an early indicator of age-related macular degeneration. Their presence may cause symptoms such as central blurring, faded color perception, and scotomas in the visual field [11].

Optical coherence tomography (OCT) is an advanced imaging technique that enables high-resolution visualization of biological tissue layers. It generates cross-sectional images by calculating the intensity and time delay of infrared light waves emitted from the device and reflected from various tissue structures. The light beam from the source is split by a semi-permeable mirror called a beamsplitter; one part is directed to a reference mirror, while the other part is directed into the patient's eye. The light reflected from the retinal layers and the reference mirror is then combined in an interferometer. The resulting time delay is calculated and used to construct a detailed tomographic image [12].

Some of the studies using OCT images with artificial intelligence are as follows: Li et al. (2019) applied deep transfer learning with VGG-16 on 109,312 OCT images and achieved 98.6% accuracy [13]. Serener et al. (2019) classified dry AMD using ResNet with 99.5% accuracy [14]. Motozawa et al. (2019) used CNN and transfer learning to distinguish normal and AMD subtypes with 99% accuracy [15]. Sun et al. (2020) employed ResNet-50 with an attention mechanism and reached 98.17% accuracy on the Duke dataset [16]. Acar et al. (2021) developed an automatic cataract detection system using VGGNet and DenseNet on fundus images, achieving up to 97.94% accuracy and highlighting the potential of deep learning in teleophthalmology [17]. Shi et al. (2021) developed the Med- XAI-Net model for geographic atrophy detection and achieved 91.5% accuracy [18]. Rajagopalan et al. (2021) trained a CNN model on 12,000 OCT images and achieved 97.01% accuracy [19]. Sahin (2022) compared a proposed CNN model with MobileNet50 and reported 94% accuracy [20]. Elsharkawy et al. (2022) used ResNet-50 on the Kermany dataset and reached 96.21% accuracy [21]. He et al. (2022) achieved 99.78% accuracy on the UCSD dataset using ResNet-50 and the LOF algorithm [22]. Aykat et al. (2023) developed a novel hybrid model combining EfficientNetV2S and Xception (EffXceptNet) for classifying CNV, DME, and drusen in OCT images, achieving 99.90% accuracy and outperforming conventional CNN architectures [23]. Baharlouei et al. (2023) applied a CNN model based on wavelet scattering transform and achieved 97.1% accuracy on the Heidelberg and Duke datasets [24]. Kulyabin et al. (2024) used VGG16 on the OCTDL dataset and reported 85.9% accuracy [25]. Gencer et al. (2025) developed a hybrid model integrating EfficientNetB0 and Xception with SE blocks, achieving 99.58% and 99.18% accuracy on the UCSD and Duke datasets, respectively [26].

Existing AI-based methods, although successful in controlled settings, often suffer from limitations such as a lack of real-time functionality, redundancy in data processing, high computational cost, and an insufficient user interface design for clinical deployment. Most prior models also do not incorporate effective mechanisms to prevent data duplication or ensure image type validation, which are essential for consistent clinical performance. The proposed method overcomes these limitations by combining a high-performing DenseNet-201 deep learning model with a hash-based data integrity mechanism and a streamlined user interface via Gradio, enabling clinicians to diagnose retinal diseases such as CNV, DME, and drusen with high precision (accuracy: **94.42%**, F1-score: **0.9442**, AUC: **1.00**) in real time. Furthermore, the model's ability to focus solely on OCT images, validate them automatically, and avoid saving duplicates enhances both accuracy and efficiency, making the system a robust and scalable solution for clinical environments.

A key distinction from previous studies in visual data processing is that the system assigns a unique hash value to each tomographic image. This prevents data repetition and redundancy, reduces unnecessary system load, and enables faster access to images. Moreover, restricting the system to accept only OCT (Optical Coherence Tomography) images improves the efficiency of the analysis process by ensuring the model focuses solely on relevant medical data. Storing patient information alongside the images enhances the integrity of medical records. The proposed model will utilize an optimized convolutional neural network (CNN) architecture that is capable of performing both diagnosis and segmentation tasks with a reduced number of deep layers. The model's performance will be evaluated comparatively using pre-trained architectures (GoogleNet, ResNet, EfficientNet, DenseNet) and expert assessments. Wi th these components combined, the study aims to develop an AI- assisted retinal disease diagnosis system with high accuracy based on OCT images.

II. MATERIAL AND METHOD

a. Dataset Collection

The OCT-2017 dataset consists of retinal images obtained using a medical imaging technique called Optical Coherence Tomography OCT. The dataset is published on the Kaggle platform. The dataset contains images from four different retinal disease categories: DME (Diabetic Macular Edema), Drusen, CNV (Choroidal Neovascular Membrane), and Normal. In this framework:

- CNV: Choroidal Neovascular Membrane category includes 37,205 training images and 242 test images.
- DME: Diabetic Macular Edema category includes 11,350

training images and 242 test images.
- DRUSEN: There are 8,616 training images and 242 test images in the Drusen category.
- NORMAL: There are 26,318 training images and 242 test images in the Normal category [27].

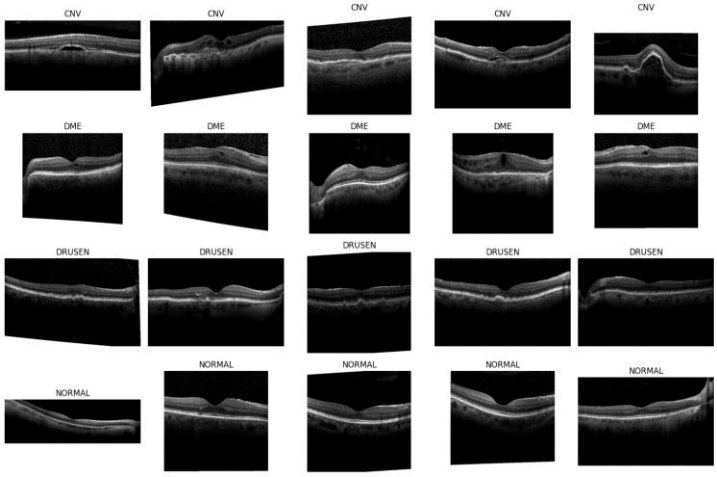


Fig 2.1. Examples from the OCT dataset [27]

b. Basic Working Principle

The basic working principle of the system describes the process of a client (doctor, specialized personnel or other service) submitting a decision support request via Gradio, as illustrated by a sequence diagram in Figure 2.1. The process starts with the client sending a request to Gradio. If the uploaded data is in image format, the request is forwarded to the Image Validator layer. If the image format is not supported, the uploaded image is not a colour or OCT image, and it provides feedback to the user with an unsupported image warning. If it is a supported image, the validator checks the accuracy of the OCT retinal image. Then, the AI model makes a prediction based on the input image. During the prediction process, the AI model makes a prediction using the data. Finally, the prediction result and the image are saved in the database and sent to the client via Gradio. However, if the image sent is not valid, i.e., has an unsupported format or structure, the validator responds to the client with an Unsupported Media Type (415) error. Similarly, if the incoming token is invalid, in this way, a doctor or other expert or service can send a request to the decision support system and receive an estimate. It is also worth mentioning that if the uploaded image has already been evaluated, it is re-evaluated, and the result is retrieved again. However, the image is not saved in the database again to avoid data redundancy, as it was previously saved with the given hash.

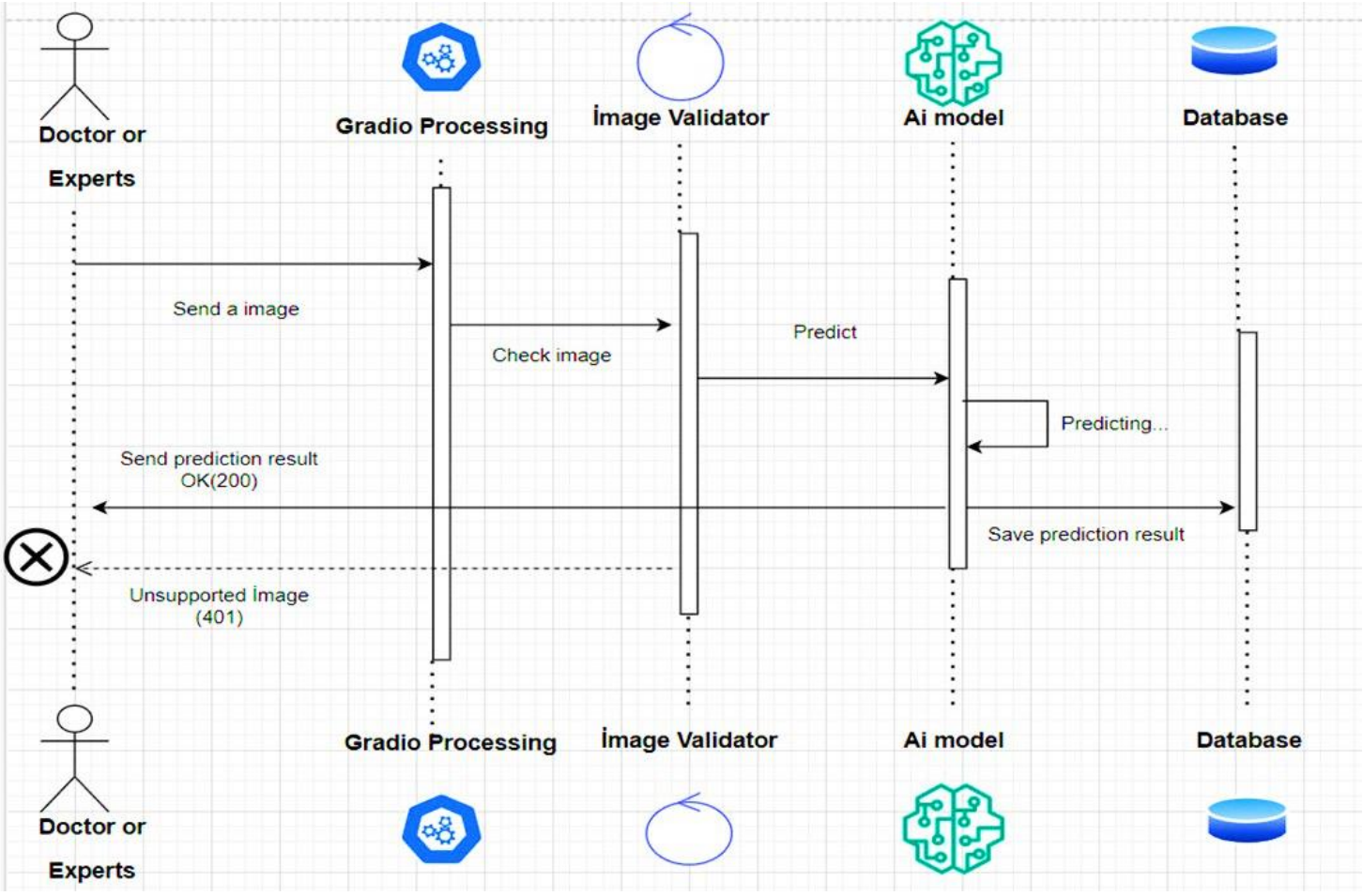


Fig. 2.2. System Basic Working Principle

c. Deep Learning

Artificial intelligence involves automating processes that require human reasoning and logic, such as mathematical computations, and human-specific capabilities, including event interpretation, experience-based benefits, learning, and inference. It can be said that artificial intelligence, which is a modelling of the human brain, is an effort to bring human and superhuman abilities to machines. Machine learning is the self-learning of operations

based on the codes written by users, without waiting for commands from the user or without the need to write new code [28]. Deep learning is a subset of machine learning that consists of a large number of hidden layers, and the output of one hidden layer is the input of the next hidden layer. It is a subfield of machine learning based on artificial neural networks that tries to mimic the way the human brain works. It aims to create systems that can automatically learn complex tasks by training with large data sets [29].

1. Convolutional Neural Networks (CNN)

Convolutional Neural Networks (CNN) are a multilayer forward artificial neural network, which is one of the deep learning architectures that needs fewer parameters and less training in large network models. It is used in image processing analysis, object recognition, and natural language processing. At first, it was used for object recognition because it could not meet the requirements due to hardware deficiencies, but later, with the development of graphical processing units, it has been frequently preferred in areas such as image processing and recognition, sound processing, computer vision, face recognition, text processing and classification, and its popularity has increased. The crowding and complexity of data, along with the development of technology, have underscored the importance of convolutional neural networks and increased their use [30]. Convolutional neural networks perform better on complex data than fully connected layers, as they progress hierarchically by adding and superimposing what they have learnt in the previous layer. CNN architectures consist of input layer (Input Layer), convolution layer (Convolution Layer), flattened linear unit layer (RELU), pooling layer (Pooling Layer), fully connected layer (Fully Connected Layer), dropout layer (Dropout Layer) and classification layer (Classification Layer) [31]. The number of layers depends on the work of the person designing the model and the study. In convolutional neural networks, much less preprocessing is needed to create the filters to be trained compared to traditional techniques due to the automatic self-training of different features [32].

In this study, deep learning architectures such as GoogleNet, ResNet, EfficientNet, and DenseNet are used, and they are compared with each other, and the comparison results are briefly mentioned. As can be seen from the results, the most efficient result is obtained with DenseNet, so the graphs and evaluations are based on this model.

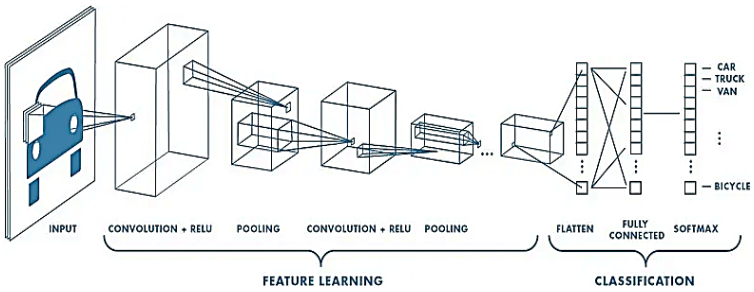


Fig. 2.3. DenseNet Architecture [33]

a) GoogleNet

Instead of stacking convolution and pooling layers directly on top of each other, GoogleNet is structured with Inception modules connected in parallel. This approach aims to increase learning capacity with non-linear activation functions while reducing computational and memory costs. In the 22-layer architecture, 1×1 convolution layers are used to manage the number of parameters and computational complexity [34].

b) ResNet

ResNet was developed to solve the vanishing gradient problem seen in deep networks and is differentiated by its residual block structure. This structure provides a deeper and more efficient learning by learning the input and output differences instead of direct input-output mapping. There are versions with different depths (ResNet18, ResNet34, ResNet50, ResNet101, ResNet152) [35].

c) EfficientNet

EfficientNet is an architecture that scales the depth, width and input resolution of the network in a balanced way. Convolutional layers, bottleneck layers and inverse residual blocks are used to increase efficiency. It aims to achieve high accuracy with fewer parameters than traditional CNN architectures [36].

d) DenseNet

The images were trained with the DenseNet201 artificial intelligence model for image classification and retinal disease detection. DenseNet-201 has an input size of 224x224x3, where 224 is the width, 224 is the height, and 3 is the depth (number of color channels). DenseNet201 is a deep learning model where each layer receives features from all previous layers as input, thus improving information flow and gradient propagation [37]. This dense connection structure makes the model more efficient and provides high performance with fewer parameters. In the tests performed with the DenseNet201 model in this study, successful results were obtained. The main reason why the model is preferred is that it has more layers than other DenseNet models. Since it has 201 layers, the model can be trained in many more layers, and much more accurate results can be obtained. In classical convolutional neural network models, the layers are connected in such a way that the output of one layer depends on the input of the other layer, but in the DenseNet architecture, the outputs of all previous layers can be given as input for any layer, offering much more effective results in terms of performance and accuracy [38].

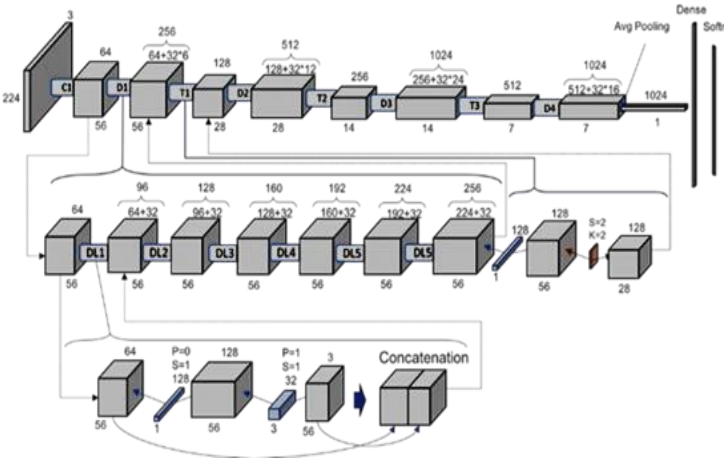


Fig. 2.4. DenseNet Architecture [39]

III. RESULTS AND DISCUSSION

During the model development studies, various deep learning models, including GoogleNet, ResNet (ResNet50), EfficientNet (EfficientNetB0), and DenseNet (Densenet201), were applied. As can be seen from the detailed comparison and analyses below, the most efficient results were obtained with the DenseNet-201 model. Therefore, the graphs and analyses in this study are mostly related to the DenseNet-201 model.

Train accuracy shows how accurately the model predicts the training data. In other words, it is the success of the model on the data it sees during learning. Validation accuracy indicates the model's success on unseen data (validation). This evaluates the generalization ability of the model. If train accuracy is high but validation accuracy is low, there may be overfitting. If both are low, there may be underfitting. If both are high and close to each other, the model has learnt well and can generalise. Based on this information, it is understood that the DenseNet-201 model gives by far the most successful result among the other models for the graph in Figure 3.1. The lowest accuracy is observed in EfficientNetB0 and GoogleNet models. In addition, since Densenet201, the model with the highest F1 score in the F1 score comparison graph in Figure 3.2, gives very balanced and successful results in terms of both precision and recall, this model stands out as the model that can discriminate between classes best. Since EfficientNetB0 has the lowest F1 score with a ratio of 0.83, this model seems to be less successful compared to the others. In this parameter and many similar parameters, it is clearly seen that the Densenet201 model shows the best result among the models tested in this study.

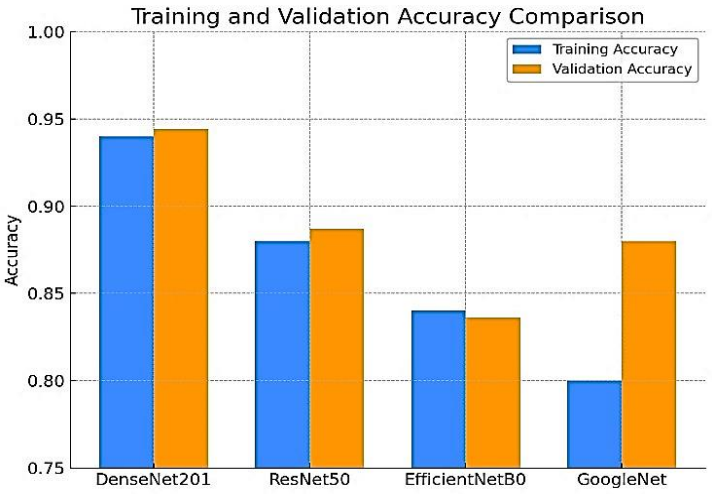


Fig. 3.1. Comparison of Train accuracy and Validation accuracy of the models

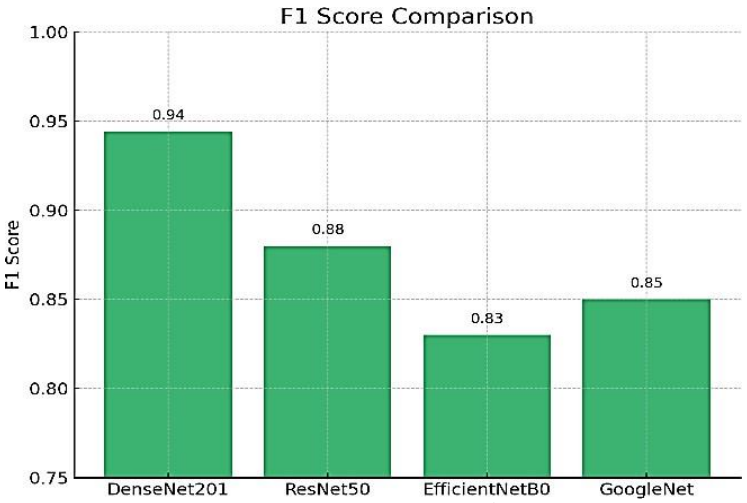


Fig. 3.2. F1 Score Comparison of Models

The two most important factors that determine the complexity of deep learning models are the number of layers and the number of parameters. A high number of layers allows the model to learn more complex patterns, which is especially advantageous in tasks such as image processing, where edges at the lower level, shapes at the middle level and objects at the upper level can be recognized. A higher number of parameters allows the model to learn more information; however, this requires more data, more memory and more computational power. In this context, the DenseNet201 model stands out among other models with 201 layers and approximately 20 million parameters. In comparison, GoogleNet has 22 layers and 7 million parameters, ResNet50 has 50 layers and 25.6 million parameters, and EfficientNetB0 has 82 layers and 5.3 million parameters. This technical evaluation shows that DenseNet201 is the best-performing model in this study. However, it is not correct to say that DenseNet201 is the best model in all cases; many factors, such as the size and complexity of the dataset, data imbalance, and hardware, affect the model's performance. Nevertheless, in this study, it was observed that the best results were obtained with DenseNet201. Therefore, the technical analyses, results and graphical presentations of the study were based on this model.

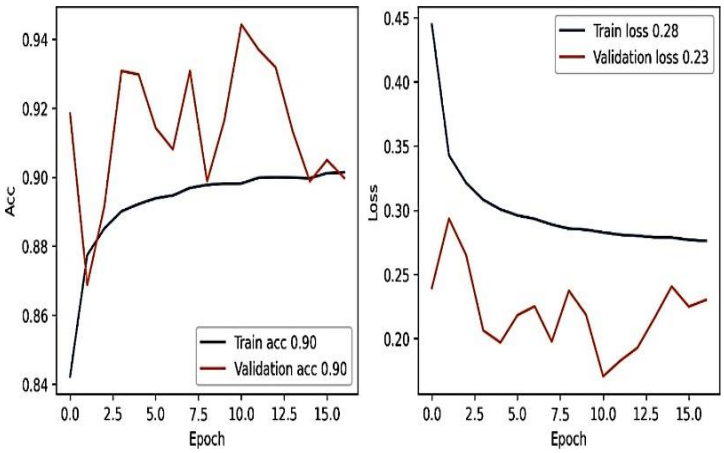


Fig. 3.3. Model accuracy and loss curves of DenseNet

Figure 3.3 shows the accuracy and loss curves obtained during the validation process of the model training for the DenseNet-201 model.

In this study, the training and validation loss curves of the DenseNet201 model, as shown in Figure 3.3, were analyzed to evaluate the model's learning behavior over time. The training loss curve exhibited a consistent downward trend, indicating that the model effectively learned from the training data and progressively improved its performance. In contrast, the validation loss curve showed more fluctuations, suggesting potential overfitting risk and a relatively unstable generalization performance on unseen data. To address this issue, an early stopping mechanism was employed. Although the training cycle was initially set for 128 epochs, the training process was automatically halted around the 15th epoch due to a lack of significant improvement in the validation loss. This decision was based on a predefined "patience" parameter, which monitors the model's performance on the validation set and stops the training if no further improvement is observed for a specific number of epochs. This technique effectively prevented overfitting, preserved the model's generalization capability, and ensured optimal training efficiency.

The observed trends in Figure 3.3 confirm this mechanism: while training accuracy continued to improve, validation accuracy plateaued and then stabilized. Similarly, validation loss ceased to decrease meaningfully, indicating that the model had reached a critical threshold in learning. By terminating training at this point, unnecessary training time and computational resources were avoided. Furthermore, early stopping contributed to maintaining the model's robustness, reliability, and clinical applicability—key considerations in a medical imaging context where unseen data performance is paramount.

Overall, the use of early stopping played a crucial role in balancing training performance with generalization ability, optimizing resource usage, and enhancing the clinical readiness of the proposed AI system.

In order to evaluate the classification performance, in other words, the success of the study, some metric values are used as shown in Table 3.1. The mathematical ratio of true positive (TP), true negative (TN), false positive (FP) and false negative (FN) to each other provides some metric results to show the success of the study. Knowing some metrics such as accuracy, sensitivity, recall, specificity, precision, false positive rate, and F1 score will contribute significantly to the judgment of the study's performance.

TABLE I
PERFORMANCE RESULTS OF DENSENET

Metrics	**Values**
Accuracy	0.9442
Sensitivity	0.9442
Specificity	0.9814
Precision	0.9475
False positive rate	0.0185
F1 score	0.9442

Table 3.1 shows the performance results and success measures of the classification model. Accuracy, or the rate at which the model makes correct predictions, is 94.42%. This indicates that the majority of instances are correctly classified. Sensitivity, or the ability of the model to correctly recognize true positives, is also 94.42%. It shows the percentage of actual positives that the model correctly predicts as positive. Specificity, which refers to the model's ability to correctly recognize true negatives, is 98.14%. It indicates the percentage of actual negatives that the model correctly predicts as negative. Precision is the ratio of the number of instances that the model predicts as positive to the number of positive cases, and it is 94.75%. It shows the proportion of samples classified as positive that are indeed positive. The false positive rate, which is the rate at which the model incorrectly classifies true negatives as positives, is 1.85%. It shows the percentage of negative samples that the model incorrectly identifies as positive. The F1 score, which balances sensitivity and precision, is 94.42%. It indicates how well the classification model balances recall and precision. Overall, these results demonstrate that the model performs well. High accuracy, sensitivity, specificity, and precision suggest that the model can accurately classify both positive and negative cases.

In this study, the classification performance of the proposed DenseNet201-based model was evaluated using Receiver Operating Characteristic (ROC) curves. ROC curves are widely used to assess the performance of classification models by illustrating the trade-off between the true positive rate (TPR) and the false positive rate (FPR) across various threshold settings. In Figure 3.4, ROC curves for each class (CNV, DME, DRUSEN, NORMAL) are presented on a single graph using the one-vs-rest approach, allowing a comprehensive and comparative analysis of the model's discriminative power across all classes.

The results indicate that all classes achieved an AUC (Area Under the Curve) value of 1.00, which suggests that the model distinguishes each class from the others with perfect accuracy.

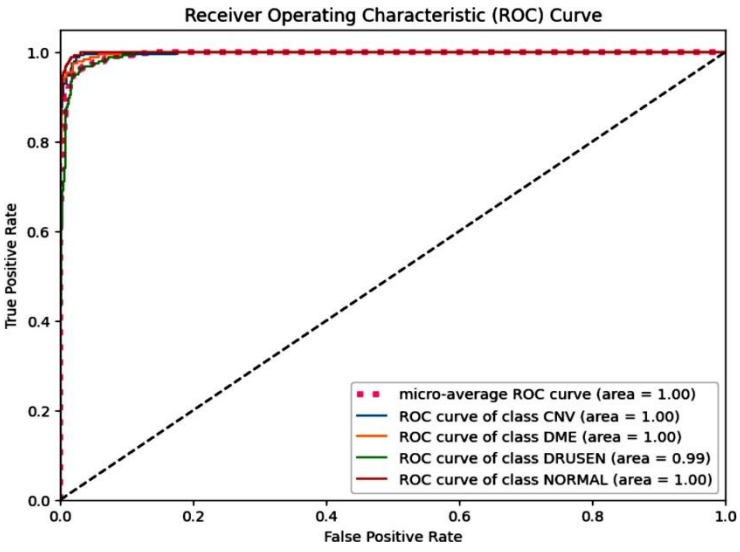


Fig. 3.4. ROC curve of the model

The ROC curves cluster closely around the upper-left corner of the graph, signifying both high sensitivity and a low false positive rate—hallmarks of exceptional classification performance.

In multi-class classification problems, visualizing ROC curves for all classes in a single graph is a common and effective practice in the academic literature. For instance, Rajagopalan et al. (2021) utilized this approach to evaluate the overall diagnostic strength of their deep learning models by comparing the ROC curves of each class collectively [19]. This strategy enables direct visual comparison of class-level performance, particularly beneficial for assessing critical clinical classes such as CNV and DME. Although presenting separate ROC curves for each class may offer visual clarity, it limits holistic comparison and may lead to misinterpretation due to inconsistent axis scales. Therefore, displaying all ROC curves within a single unified graph ensures scale consistency and facilitates direct, comparative interpretation of model performance across classes. This makes it a valuable method for evaluating the overall diagnostic reliability of the model, especially in clinical contexts.

IV. CONCLUSION

In this study, an artificial intelligence-based computer-aided diagnostic system was developed for the automated detection of retinal diseases using Optical Coherence Tomography (OCT) images. Various state-of-the-art deep learning models, including GoogleNet, ResNet, EfficientNet, and DenseNet, were implemented and comparatively evaluated. Among these models, DenseNet-201 achieved the highest classification performance with an accuracy of 94.42%, an F1 score of 94.42%, and an AUC of 1.00. These results indicate that DenseNet-201 is highly effective in distinguishing between different retinal disease categories and normal cases.

The system was designed not only to achieve high diagnostic accuracy but also to optimize data management through hashing mechanisms and to ensure usability through a user-friendly interface. Moreover, the system was developed with real-world clinical applicability in mind, aiming to support healthcare professionals in early and accurate diagnosis of retinal diseases. By reducing the workload of clinicians and accelerating the diagnostic process, the system has the potential to improve patient outcomes and contribute to more efficient clinical workflows.

The proposed system aims to address the shortcomings observed in previous studies in the literature. Although some studies report high accuracy rates, they often fall short in terms of ease of use, real-time processing, and clinical applicability. Key distinguishing features of this study include a hash-based data recording mechanism that enhances data security, acceptance of only OCT (Optical Coherence Tomography) images to ensure analysis efficiency, intuitive and training-free user interfaces, and integration potential with national health software such as e- Nabız. Another notable advantage is the ability of the system to be embedded into OCT devices via embedded software, enabling direct output generation from the device. Patient- specific OCT images are securely stored in the MongoDB database with a unique hash value assigned at the time of upload. If the same image is uploaded again, only the prediction result is displayed without saving a duplicate, thereby reducing data redundancy and improving system performance. The Gradio-based interface is designed to reduce the workload of healthcare professionals and save time. Furthermore, the developed patient interface allows individuals to communicate directly with their doctors via email.

Despite these strengths, the system has several practical limitations. If the system is to be installed individually on the computers located in relevant departments of the Ministry of Health, the process may lead to excessive time consumption and face limitations due to insufficient hardware capabilities, as not every institutional computer may support the operational requirements of the model. Alternatively, deploying it via centralized servers could increase operational costs. Additionally, OCT devices may not be available in every healthcare unit, limiting their widespread use. Most importantly, since patients' medical images are classified as sensitive personal health data, using such images for analysis requires special authorization from the Ministry of Health. This administrative requirement is often a lengthy and complex process, posing a significant barrier to real-world implementation. Therefore, while the current version of the system cannot yet fully achieve large-scale clinical integration, it lays a strong conceptual foundation for future development and adoption.

Looking ahead, several directions for future improvement are evident. Expanding the dataset with more diverse and real-world clinical images will be crucial to enhancing the model's generalization capability. Integration of hybrid approaches, such as combining deep learning with traditional image processing techniques or clinical metadata, may further boost diagnostic performance. Efforts should also be made to optimize the system for real-time performance and to ensure seamless integration with existing healthcare platforms and electronic medical record systems. Collaborations with hospitals and healthcare providers will be essential to test and refine the system in actual clinical practice. Ultimately, embedding the model directly into OCT devices or national healthcare platforms could significantly increase its accessibility and impact. Furthermore, more effective use and widespread access may be possible by integrating the model directly into OCT devices or health system software and the e-Nabız platform. However, adequate technical infrastructure, financial support and official permissions are required to realize these goals.

In conclusion, this study demonstrates the feasibility and potential of an AI-supported OCT image analysis system for retinal disease detection. While further work is needed to fully realize its clinical utility, the results obtained provide a strong foundation for future advancements and broader adoption in ophthalmological practice.

V. REFERENCES

[1] World Health Organization. (2019). World report on vision. Geneva: WHO.

[2] Resnikoff, S., Lansingh, V. C., Washburn, L., Felch, W. C., & Gauthier, T. M. (2020). Vision loss and its impact on quality of life. *Ophthalmic Epidemiology*, 27(2), 85–90.

[3] Lamoureux, E. L., & Fenwick, E. K. (2016). Health-related quality of life and visual impairment. Current Opinion in Ophthalmology, 27(3), 238–243.

[4] Forrester, J. V., Dick, A. D., McMenamin, P. G., Roberts, F., & Pearlman, E. (2015). The Eye: Basic Sciences in Practice. Elsevier

Health Sciences.
[5] Berger, John, Ways of Seeing, Penguin Books, UK 2008, p.7-33.
[6] Kolb, H. (2005). Simple Anatomy of the Retina. Webvision.
[7] Curcio, C.A., Sloan, K.R., Kalina, R.E.,Hendrickson, A.E. (1990) Human photoreceptor topography. The Journal of Comparative Neurology, 292 (4), 497-523.
[8] Dandekar SS, Jenkins SA, Peto T, et al.: Autofluoresence imaging of choroidal neovascularization due to age-related macular degeneration. Arch Ophthalmol. 2005;123:1507-1513.
[9] Bhende, M, Shetty, S, Parthasarathy, M. K., & Ramya, S. (2018). Optical coherence tomography: A guide to interpretation of common macular diseases. Indian journal of ophthalmology, 66(1), 20.
[10] Erdoğan, Alper, Epiretinal Membran Cerrahisinde Prognozu Etkileyen Faktörler, Uzmanlık tezi, s.8.
[11] Deutman AF, Hansen LMAA: Dominantly inherited drusen of Bruch's membrane. Br J Ophthalmol 1970; 34:373-382.
[12] Aydın,Ali,Bilge,A.Hamdi, Optik Koherens Tomografinin Glokomda Yeri,s.78.
[13] Li, F., Chen, H., Liu, Z., Zhang, X.-D., Jiang, M.-S., Wu, Z.-Z., and Zhou, K.-Q. (2019). Deep learning-based automated detection of retinal diseases using optical coherence tomography images. Biomedical Optics Express, 10(12), 6204–6226. https://doi.org/10.1364/BOE.10.006204
[14] Serener, A., and Serte, S. (2019). Dry and wet age-related macular degeneration classification using OCT images and deep learning. 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), 1–5. https://ieeexplore.ieee.org/document/8741768
[15] Motozawa, N., An, G., Takagi, S., Kitahata, S., Mandai, M., Hirami, Y., Yokota, H., Akiba, M., Tsujikawa, A., Takahashi, M., & Kurimoto, Y. (2019). Optical coherence tomography- based deep-learning models for classifying normal and age- related macular degeneration and exudative and non-exudative age-related macular degeneration changes. Ophthalmology and Therapy, 8(4), 527–539. https://doi.org/10.1007/s40123-019- 00207-y
[16] Sun, W., Zhang, H., & Yao, Z. (2020). Automatic diagnosis of macular diseases from OCT volume based on its two- dimensional feature map and convolutional neural network with attention mechanism. Journal of Biomedical Optics, 25(9), 096004. https://doi.org/10.1117/1.JBO.25.9.096004
[17] Acar, E., Türk, Ö., Ertuğrul, Ö. F., and Aldemir, E. (2021). Employing deep learning architectures for image-based automatic cataract diagnosis. Turkish Journal of Electrical Engineering & Computer Sciences, 29(5), 2649–2662. https://doi.org/10.3906/elk-2103-77
[18] Shi, X., Keenan, T. D. L., Chen, Q., De Silva, T., Thavikulwat, A. T., Broadhead, G., Bhandari, S., Cukras, C., Chew, E. Y., and Lu, Z. (2021). Improving interpretability in machine diagnosis: Detection of geographic atrophy in OCT scans. Ophthalmology Science, 1(3), 100038. https://doi.org/10.1016/j.xops.2021.100038
[19] Rajagopalan, N., Venkateswaran, N., Josephraj, A. N., and Srithaladevi, E. (2021). Diagnosis of retinal disorders from Optical Coherence Tomography images using CNN. PLoS ONE, 16(7), e0254180. https://doi.org/10.1371/journal.pone.0254180
[20] Şahin, M. E. (2022). A deep learning-based technique for diagnosing retinal disease by using optical coherence tomography (OCT) images. Turkish Journal of Science & Technology,17(2),417–426. https://doi.org/10.55525/tjst.1128395
[21] Elsharkawy, M., Sharafeldeen, A., Soliman, A., Khalifa, F., Ghazal, M., El-Daydamony, E., Atwan, A., Sandhu, H. S., & El- Baz, A. (2022). A novel computer-aided diagnostic system for early detection of diabetic retinopathy using 3D-OCT higher- order spatial appearance model. *Diagnostics, 12*(2), 532. https://doi.org/10.3390/diagnostics12020532
[22] He, T., Zhou, Q., and Zou, Y. (2022). Automatic detection of age-related macular degeneration based on deep learning and local outlier factor algorithm. Diagnostics, 12(2), 532. https://doi.org/10.3390/diagnostics12020532
[23] Aykat, Ş., & Şenan, S. (2023). Advanced detection of retinal diseases via novel hybrid deep learning approach. Traitement du Signal, 40(6), 2367–2382. https://doi.org/10.18280/ts.400604
[24] Baharlouei, Z., Rabbani, H., & Plonka, G. (2023). Wavelet scattering transform application in classification of retinal abnormalities using OCT images. Scientific Reports, 13, 19013. https://doi.org/10.1038/s41598-023-46200-1
[25] Kulyabin, M., Zhdanov, A., Nikiforova, A., Stepichev, A., Kuznetsova, A., Ronkin, M., Borisov, V., Bogachev, A., Korotkich, S., Constable, P. A., & Maier, A. (2024). OCTDL: Optical coherence tomography dataset for image-based deep learning methods. Scientific Data, 11(365). https://doi.org/10.1038/s41597-024-03182-7
[26] Gencer, G., & Gencer, K. (2025). Advanced retinal disease detection from OCT images using a hybrid squeeze and excitation enhanced model. PLOS ONE, 20(2), e0318657. https://doi.org/10.1371/journal.pone.0318657
[27] Mooney, P. T. (2018). Kermany 2018 OCT Dataset. https://www.kaggle.com/datasets/paultimothymooney/kermany2018
[28] C. Cortes, X. Gonzalvo, V. Kuznetsov, M. Mohri, andS. Yang. Adanet: Adaptive structural learning of artificial neural networks. arXiv preprint arXiv:1607.01097, 2016.
[29] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In CVPR, 2016
[30] Urmamen H E. "Optik Koherans Tomografi Görüntüleri İle Retinal Hastalıkların Evrişimsel Sinir Ağı Kullanılarak Teşhis Edilmesi". Yüksek Lisans Tez Çalışması, 2023, s.22
[31] Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In CVPR, 2015
[32] Özdemir E, Türkoğlu İ. Yazılım Güvenlik Açıklarının Evrişimsel Sinir Ağları (CNN) ile Sınıflandırılması, 2022
[33] He, X. ve Dong, F. (2023). A deep learning-based mathematical modeling strategy for classifying musical genres in musical industry. Nonlinear Engineering, 12, 20220302. https://doi.org/10.1515/nleng-2022-0302
[34] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1–9.
[35] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778.
[36] Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In Proceedings of the 36th International Conference on Machine Learning (ICML), PMLR, 6105–6114.
[37] 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 (CVPR), pp. 4700-4708.
[38] Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., & Chen, T. (2018). Recent Advances in Convolutional Neural Networks. Pattern Recognition, 77, 354-377
[39] Gajera, H. ve Trada, P. (2020). An AI Approach for Disease Detection in Plants. International Research Journal of Engineering and Technology (IRJET), 7(7), 5287–5290.

VI.BIOGRAPHIES

HASAN MEMİŞ, born in Diyarbakır in 1988, completed his high school education in Diyarbakır. In 2011, he was admitted to the Electrical and Electronics Engineering Department at Gaziantep University and graduated in 2017. In 2023, he started to study his master's degree programme in Electrical

and Electronics Engineering at Batman University. He researches topics such as artificial intelligence and image processing. He has also been teaching as an electrical and electronic engineering teacher at the Ministry of National Education in Şırnak since 2019.

EMRULLAH ACAR received the B.S. degree in electrical and electronics engineering from Çukurova University, Adana, Turkey, in 2009, the M.S. degree in biomedical engineering from Istanbul Technical University, Istanbul, Turkey, in 2010, and the M.S. and Ph.D. degrees in electrical– electronics engineering from Dicle University, Diyarbakır, Turkey, i n 2012 and 2017 respectively. He is currently an Associate Professor of electrical–electronics engineering with Batman University and also the head of the electronics division. His research interests include digital image processing, machine learning, and remote sensing applications. Dr. Acar received awards and honors include The Scientific and Technological Research Council of Turkey (TUBITAK) Grant, Erasmus Mobility Grant, Sweden, Internship Grant, Germany and Visiting Researcher, Czech Republic.

Research Article

Bibliometric Analysis of Publication in the Field of Artificial Intelligence and Cyber Security

Songul Karakus and Sevinc Ay

***Abstract*— With the rapid changes in technology, artificial intelligence and cybersecurity have become two important areas inseparable from each other in today's digital world. The aim of this article is to analyze academic research on artificial intelligence and cybersecurity using bibliometric analysis method by providing an overview of artificial intelligence and cybersecurity. The findings obtained from the bibliometric analysis conducted between 2002-2025 (available at the time of the search) using the Web of Science database revealed that there were a total of 1296 publications and 4023 authors related to artificial intelligence and cybersecurity. The most frequently used keywords were examined by the authors using the VOSviewer software, and as a result of this examination, cybersecurity was found to be repeated 344 times, and artificial intelligence was found to be repeated 238 times. It has been also seen that the data obtained included publications from 97 countries and the country with the highest number of publications has China with 282 publications. The relationships between artificial intelligence and cyber security technologies have been also visualized with VOSviewer software. While artificial intelligence has been suggested in the theme analysis, methods such as intrusion detection system and feature selection have been identified as niche themes. As a result, this study shows the current status of studies on artificial intelligence and cyber security. It is also thought to guide future studies in this field.**

Index Terms— **Artificial intelligence, Bibliometric analysis, Cyber security, R Studio, VOSwiever program.**

I. Introduction

ARTIFICIAL INTELLIGENCE can be defined as computer systems or computer software that can imitate human intelligence functions such as learning, problem solving, visual perception, decision making based on past experiences, etc. Artificial intelligence technologies include almost all areas including health [1], education [2], finance and banking systems [3], automotive [4], e-commerce and marketing [5, 6], agriculture [7], law [8], human resources [9], energy [10], environment [11], art [12], tourism [13], search engines [14], translation [15], storytelling [16], games [17] and cybersecurity [18].

Cyber security can be defined as various measures, strategies or technologies that aim to protect digital systems, computers, devices, networks, data and information systems against all kinds of cyber attacks, threats, disruptions, theft and abuse. With the widespread adoption of digital technologies alongside the internet, cybersecurity has become critically important not only in protecting our personal data in the digital realm but also in safeguarding highly sensitive information at the government level [19]. According to the relevant literature, cyber security has three main objectives: confidentiality, integrity and availability. They have been briefly called the CIA trio. Confidentiality is the preservation of personal data and confidential matters to which access authority must be granted. Integrity means that data cannot be changed or any action can be taken without authorization. Finally, usability can be expressed as using the relevant system as expected [20].

In recent years, artificial intelligence and cybersecurity have begun to attract more and more attention in today's digital world. While artificial intelligence is a support element in detecting threats and improving defense strategies in the field of cyber security, it can also be a threat element when used in cyber attacks by malicious people. Therefore, we can talk about a two-way impact of artificial intelligence in the field of cyber security.

In this study, a bibliometric analysis of the studies conducted in the fields of artificial intelligence and cyber security was conducted to find out the general trend. Bibliometrics is defined as the statistical and mathematical analysis of publications prepared on a subject or field using data such as year, citation, author, country, subject, source and university. The bibliometric analysis method is a method that examines and analyzes studies in scientific literature and explains them with numerical data. This analysis can be used to determine the direction in which the studies are trending, author and institution interactions, distribution across countries, academic networks, etc. in order to reveal the necessary information from scientific maps and networks [21, 22, 23]. In addition, with this method, relationship densities, power values and connections between data, maps or networks are determined [24]. Connections reveal the potential for similarities, relationships or partnerships that are obtained as a result of the analysis of the resulting data with different elements. With the bibliometric

Songul Karakus, is with Department of Computer Engineering University of Bitlis Eren, Bitlis, Turkey,(e-mail: skarakus@beu.edu.tr).
https://orcid.org/0000-0003-1999-0203

Sevinc Ay, is with Department of Distance Education Center University of Firat, Elazig, Turkey, (e-mail: say@firat.edu.tr).
https://orcid.org/0009-0001-6309-0889

Manuscript received Nov 24, 2024; accepted Mar 07, 2025.
DOI: 10.17694/bajece.1585201

analysis method, analyses such as citation analysis, authors and collaborations, keyword and trend analysis, number of publications, etc. can be performed [25]. Bibliometric analysis is a very powerful method for evaluating scientific studies, ensuring collaborations and determining the current situation.

Studies on artificial intelligence and cyber security are very important in identifying possible threats in this area and taking security measures or making existing security measures more effective. Particularly, the development of artificial intelligence in areas such as machine learning, deep learning, natural language processing, and the internet of things significantly affects cyber security applications. Artificial intelligence-based methods are used in cyber security for threat detection and anomaly detection [26, 27], malware detection [28], network intrusion detection [29], response to cyber attacks [30], vulnerability detection and penetration testing [31], social engineering and phishing attacks [32], and phishing/spam detection [32, 33].

Some studies on artificial intelligence and cyber security are as follows. In the study conducted by Aytan and Barışçı [35], artificial intelligence-based attack detection and analysis were performed in cyber defense. For this purpose, they tried to detect denial of service attacks and information scanning attacks using the "KDD Cup'99" data set and the machine learning methods in the Weka program. As a result, they achieved a success rate of 99%. Özdemir [36] investigated the impact of artificial intelligence and machine learning on cyber threat intelligence and examined how they can be integrated into applications related to cyber threat intelligence with case studies. Li et al. [37] conducted a study on how to improve real-time false alarm detection in network attacks by combining machine learning methods with deep learning to detect real alarms more accurately. They achieved successful results in the study. Liu et al. [38] conducted a review study focusing on methods that use machine learning techniques to model and detect hybrid cyber attacks in the industrial Internet of Things. They also discussed the difficulties and research trainings of analyzing industrial internet of things and cyber attacks and made some suggestions. Li [39] conducted a compilation study on cyber security and artificial intelligence. In this study, machine learning and deep learning methods are first summarized in terms of combating cyber attacks. Then, the attacks that artificial intelligence may be exposed to were classified.

The remainder of the study is organized as follows. In the second part of the study, the method of the research is mentioned. In the 3rd chapter, the findings are discussed, and in the last chapter, a general evaluation is made and suggestions are made.

II. METHOD OF THE RESEARCH

The aim of this study is to conduct a bibliometric analysis of studies involving the concepts of artificial intelligence and cyber security between 2002 and 2025 (available at the time of the search) in the light of different parameters. It is aimed that the bibliometric analysis results will guide and contribute to new studies in the field of artificial intelligence and cyber security. Within the scope of the study, the terms "artificial intelligence" and "cyber security" were searched on the Web of Science database using the conjunction "and" and selecting all categories.

Within the scope of the study, the contents obtained from the Web of Science database were examined within the framework of the following questions:

- What is the distribution of studies indexed in the Web of Science database on artificial intelligence and cyber security by category?
- What is the distribution of studies indexed in the Web of Science database on artificial intelligence and cyber security by year?
- Which authors are prominent in the studies indexed in the Web of Science database on artificial intelligence and cybersecurity, and what are the connections between them?
- What are the keywords used in studies indexed in the Web of Science database on artificial intelligence and cybersecurity, and what are the connections between them?
- What is the distribution of studies indexed in the Web of Science database on artificial intelligence and cybersecurity by country and the strength of the connections between these countries?
- What is the distribution of studies indexed in the Web of Science database on artificial intelligence and cybersecurity according to the most cited publications and the strength of the connection between these publications?

In studies where the keywords artificial intelligence and cyber security are used together, analysis by year is important in determining the increasing trend in these areas in recent years. In addition, determining the study areas on artificial intelligence and cyber security issues provides information to researchers who want to work interdisciplinary on this subject. Having a country analysis of the studies will reveal the relations between the countries and will also provide an idea about the prominent language and university connections in this regard. The author analysis conducted between the studies is important in terms of determining the authors who are the school in this field and revealing the cooperation between the publications. While the analyses will guide the studies to be conducted in this field, it is thought that the literature analysis will guide those who will work in this field.

Content indexed in Web of Science was used as the database. The keywords "artificial intelligence" and "cyber security" were used in the search dated 15.10.2024. In the search conducted using the AND conjunction, the All Fields field in Web of Science was selected and 1296 publications were reached. It was seen that these publications have many different categories, the oldest being 2002 and the newest being 2025, including 657 Proceeding Papers, 560 articles, 75 review articles, 24 early Access, 9 book chapters, 5 editorial material, 4 Retracted Publications, 2 data papers, 1 book, 1 book review, and 1 letter. The languages used in the studies were English, German, Chinese, Spanish and Turkish.

For the purpose of bibliometric analysis carried out on studies obtained from the Web of Science database and using the keywords "artificial intelligence" and "cyber security" together,

the Voswiever (version 1.6.20) program and the R studio programs "bibliometrix" package and the biblioshiny application were used. VOSviewer and R Studio are among the most widely used software in bibliometric analysis, offering important functions such as network analysis, clustering and visualization of citation relationships.

III. FINDINGS

Within the scope of the study, bibliometric analysis methods were applied to the studies indexed in the Web of Science database on artificial intelligence and cyber security in order to find answers to the research questions listed above. The categories of the data obtained as a result of the analysis are given in Figure 1.

Fig.1. Categories of publications containing the concepts of artificial intelligence and cybersecurity in the Web of Science database [40]

Among the categories with the most publications in the fields of artificial intelligence and cyber security, the first five places were taken by Computer Science Artificial Intelligence with 35.76%, Computer Science Information Systems with 30.31%, Electrical and Electronics Engineering with 27.70%, Computer Science Theory Methods with 26.094% and Computer Science Interdisciplinary Applications with 22.66%, respectively. Among the fields studied, there are also publications belonging to different categories such as Physics, Chemistry, Mathematics, Environmental Sciences and Telecommunications. General information about the data obtained from 1296 studies subject to bibliometric analysis is given in Table 1.

.

TABLE I
GENERAL INFORMATION ABOUT THE STUDIES

Time period	2002:2025
Sources (Magazines, Books, etc.)	719
Total number of publications	1296
Annual Growth Rate %	2,97
Average age of studies	2.71
Average number of citations per study	11.79
References	49394
Keywords (ID)	874
Authors	4023
Number of authors of single-authored studies	105
Number of single-authored studies	114
Co-authors per study	3.89
International co-authorships %	30.71

When Table 1 is examined, it is seen that 1296 publications published between 2002 and 2025 are included in 719 different sources. The total number of references for publications was determined as 49394. Other data obtained from the table include the information that 114 of these publications were

single-authored and that an average of 2.97% of new publications were produced annually.

The data obtained through the bibliometric analysis applied within the scope of the study were examined in terms of distribution and link strength according to the years of publication, fields of publication, countries with the most publications, universities with the most publications, most used keywords, most cited authors and most cited publications. Firstly, the distribution according to the years of publication was examined. The findings obtained in the study are given under subheadings.

Publications in the Web of Science database using the terms artificial intelligence and cyber security together were made between 2002 and 2025. Among the publication years, the highest number was reached in 2024 with 224. Figure 2 shows the distribution of publications by year.

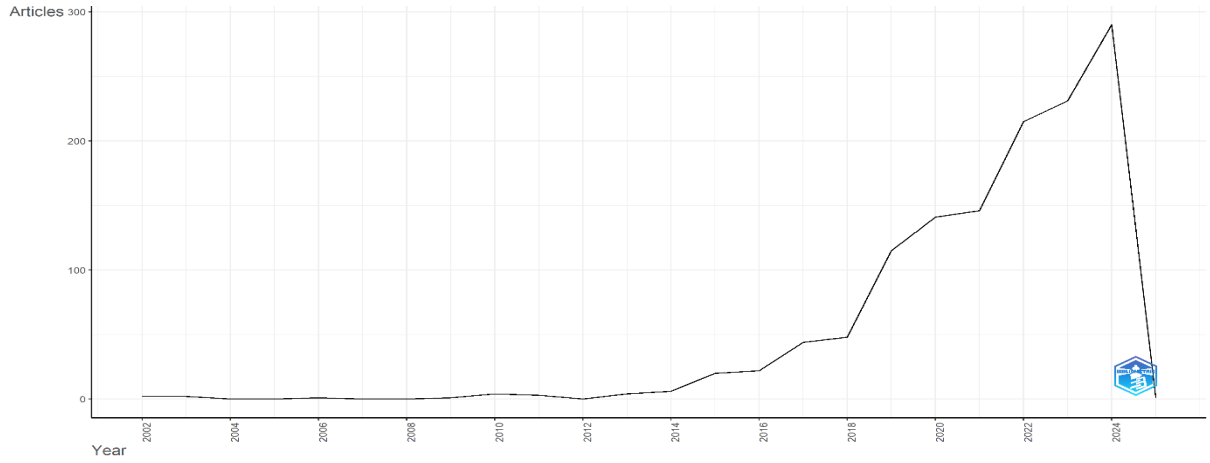


Fig.2. Distribution of publications by year

As Figure 2 shows, very little work has been done in the field between 2002 and 2014. In total, 23 studies were carried out between these years. It is observed that this area, where studies were very few until 2014, gained rapid momentum especially after 2019.

A. Author Profile and Network Analysis

A total of 4023 authors' works were obtained from the dataset obtained from the Web of Science database between the years 2002-2025. The five most influential authors working in the fields of artificial intelligence and cybersecurity over the years are Yang Zhang with 19 publications, Jinjin Li with 14 publications, Jingshi Yang with 14 publications, Yicheng Liu with 13 publications, and Yisen Wang with 11 publications, respectively.

The most relevant authors are presented in Figure 3.

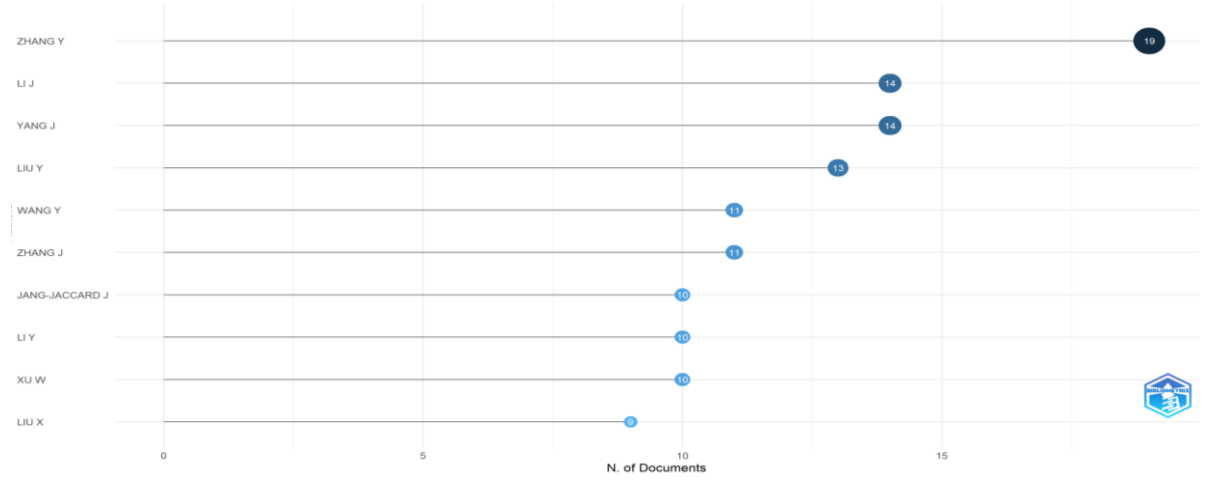


Fig.3. The most relevant authors and the number of publications they produced

When we look at the productivity status of the authors over the years, it is seen that Yang Zhang, Jinjin Li, Jingshi Yang, Yicheng Liu, Yisen Wang took the first places between 2015 and 2025, when the intensity of the studies increased. Figure 4 shows the productivity status of the authors.

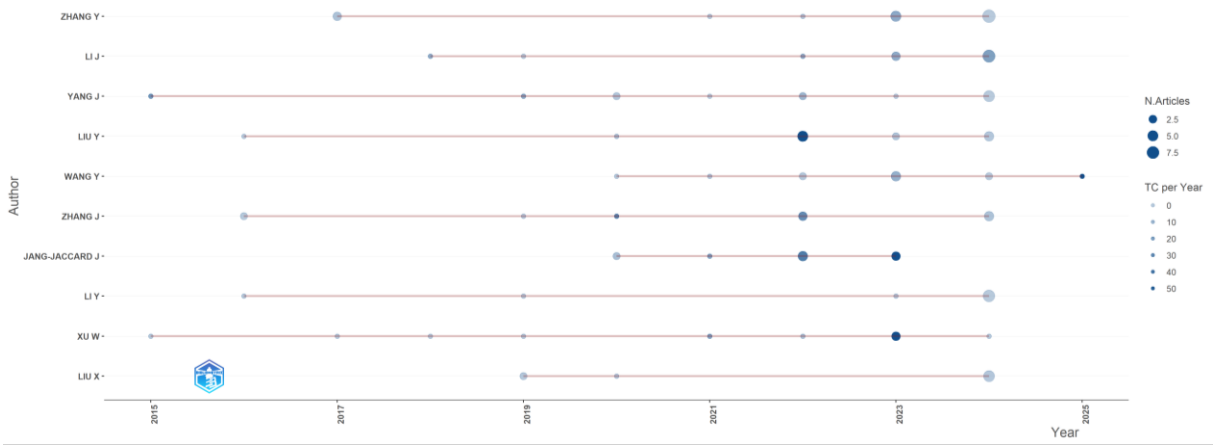


Fig.4. Productivity status of writers by year

B. *Keyword Analysis*

In publications obtained from the Web of Science database in the fields of artificial intelligence and cyber security, the most frequently used keywords by authors were cyber security with 344 repetitions, artificial intelligence with 238 repetitions, machine learning with 214 repetitions, deep learning with 133 repetitions, and internet of things with 50 repetitions. Another keyword analysis conducted within the scope of the study is to determine the usage density of author keywords that changes over the years. As a result of this analysis, when the studies conducted in recent years are examined, it is seen that the concepts of deep learning, artificial intelligence, internet of things, blockchain, and intrusion detection come to the fore, especially in 2022 and beyond. Figure 5 shows the network map of the change in author keywords in the publications in the dataset over the years

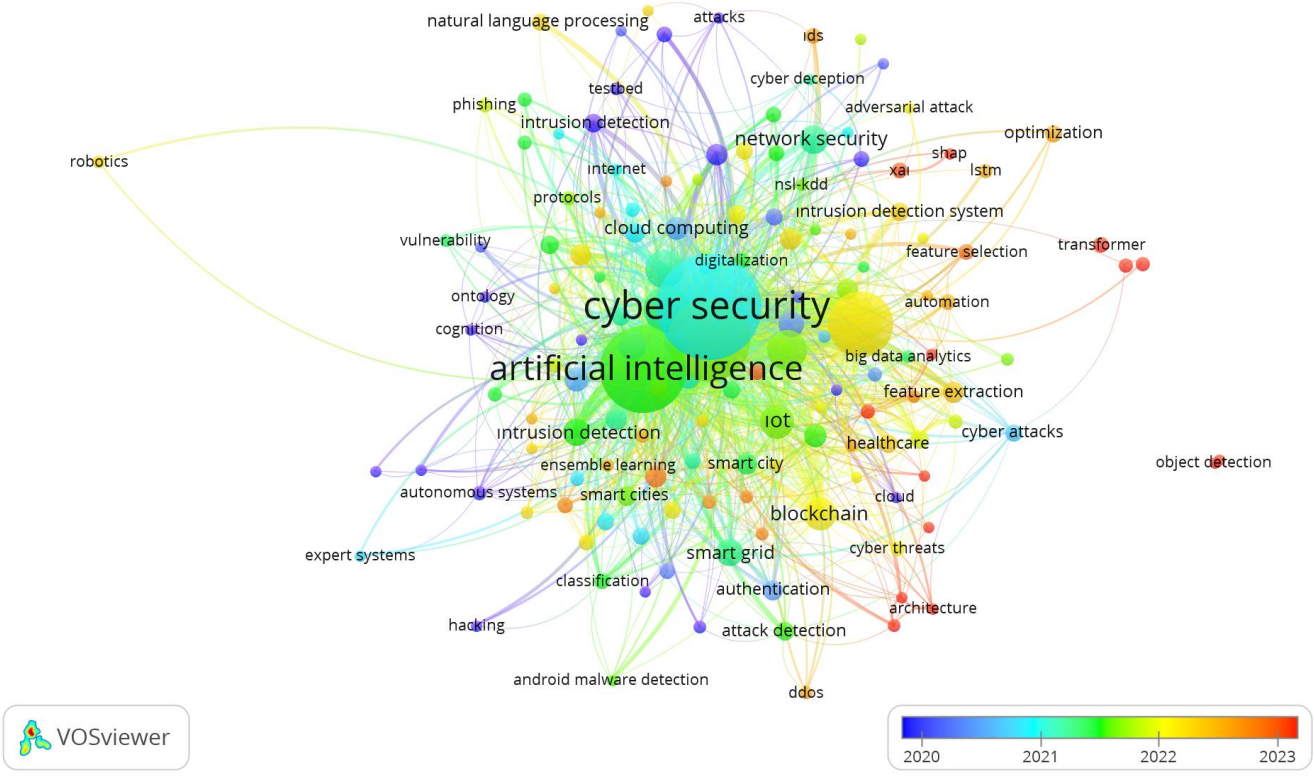


Fig.5. Most frequently used keywords and their links

In the visualization of the most frequently used keywords in the word cloud format obtained from the studies obtained from the Web of Science database, it is seen that the first five words are artificial intelligence, cyber security, attacks, intrusion detection and classification. Figure 6 shows the World cloud

obtained from the most frequently used keywords and prepared by the R-Studio program.

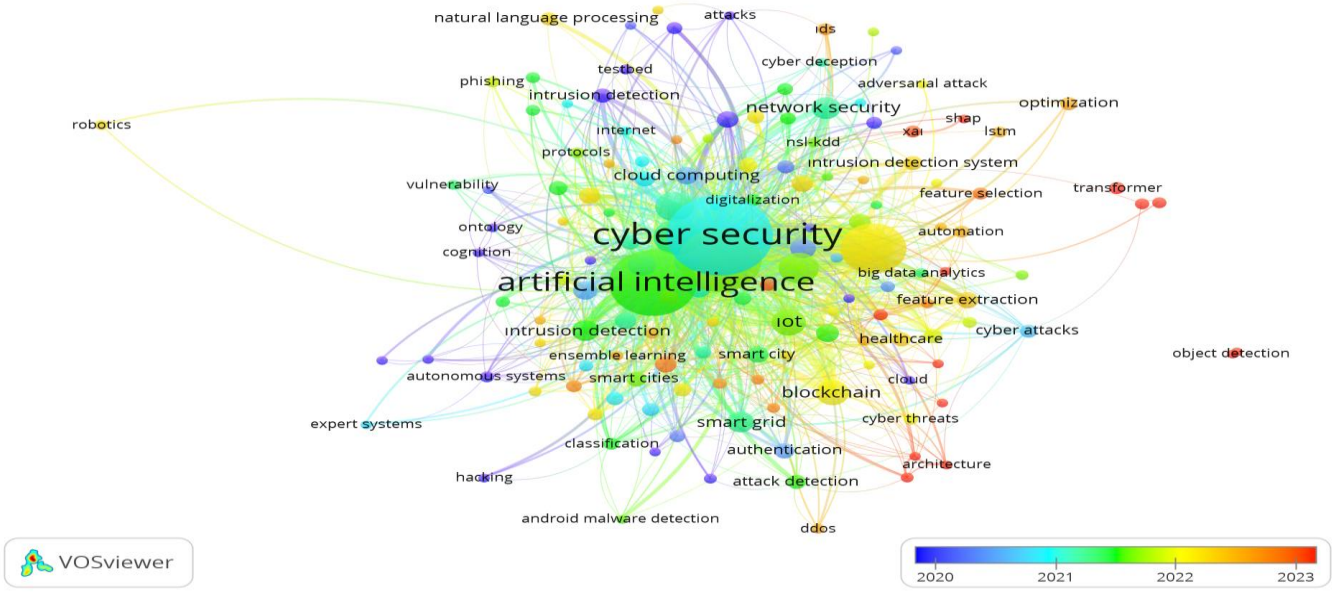


Fig.6. Keyword cloud

C. *Network Analysis of Countries*

It was determined that the data obtained from the Web of Science database included publications from 97 countries. Among the 62 countries that passed the threshold by meeting the conditions of at least 1 publication and 1 citation, it was determined that the countries with the most publications were China (282 publications), the USA (217 publications), India (125 publications), the UK (102 publications) and Saudi Arabia (96 publications), respectively. The top five most cited countries are the USA (3846 citations), the UK (3260 citations), India (2950 citations), China (2564 citations) and Australia (1832 citations). The top five countries in terms of total connection power are the United States, China, Saudi Arabia, India and Australia. Table 2 shows the publication numbers for each country.

TABLE II
NUMBER OF PUBLICATIONS, NUMBER OF CITATIONS AND TOTAL LINK STRENGTH OF COUNTRIES

Country	Number of Publications	Number of Citations	Total Link Strength
People's Republic of China	282	2564	18955
USA	217	3846	19063
India	125	2950	17445
England	102	3260	12990
Saudi Arabia	96	1241	18092
Australia	83	1832	13446
South Korea	66	799	7944
Italy	57	873	6499
Israel	50	496	3583
United Arab Emirates	41	447	8605
Canada	41	426	5743
Germany	41	375	4812
Pakistan	35	481	10653
Greece	35	297	3559
Spain	33	260	3811
France	31	305	2799
Malaysia	26	225	3995
Singapore	25	909	2135
Romania	24	62	387

According to the findings obtained from the data set, the countries are listed according to the number of publications in Table 2, which shows the number of publications, the number of citations and the strength of the connection between them. In order to make a bibliographic analysis of the publications according to their countries, the relationship between 62 countries that have a relationship between them was examined within the scope of the criteria of at least 1 work published and 1 citation. Figure 7 shows the analysis showing the bibliographic coupling of countries.

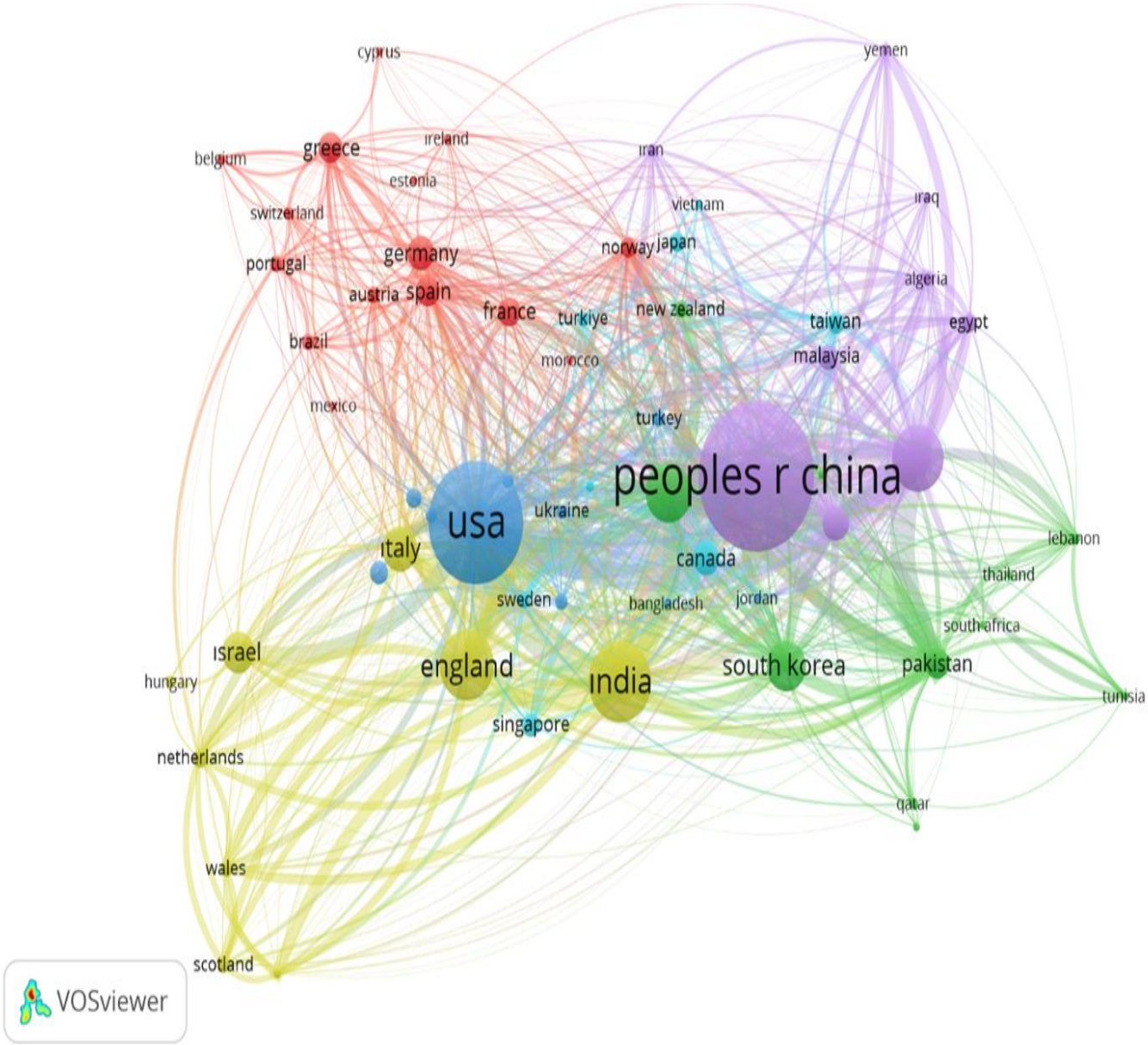


Fig.7. Network map showing bibliographic matching of countries

D. Citation Analysis

Author and source citation analysis was examined in order to perform citation analysis on the data set obtained from the Web of Science database. In order to perform author citation analysis, the criterion of having at least 1 publication and 1 citation was applied. Out of 4023 authors, 2964 authors passed the threshold value. Among the authors, Niti Upadhyay (470 citations), Zhon Lin Wang (405 citations), Jin Yang (341 citations), Julian Jan-Jaccard (312 citations) and Fadi Al-Turjman (296 citations) are at the top of the list. Figure 8 shows the network map of the most cited authors based on the findings from the dataset.

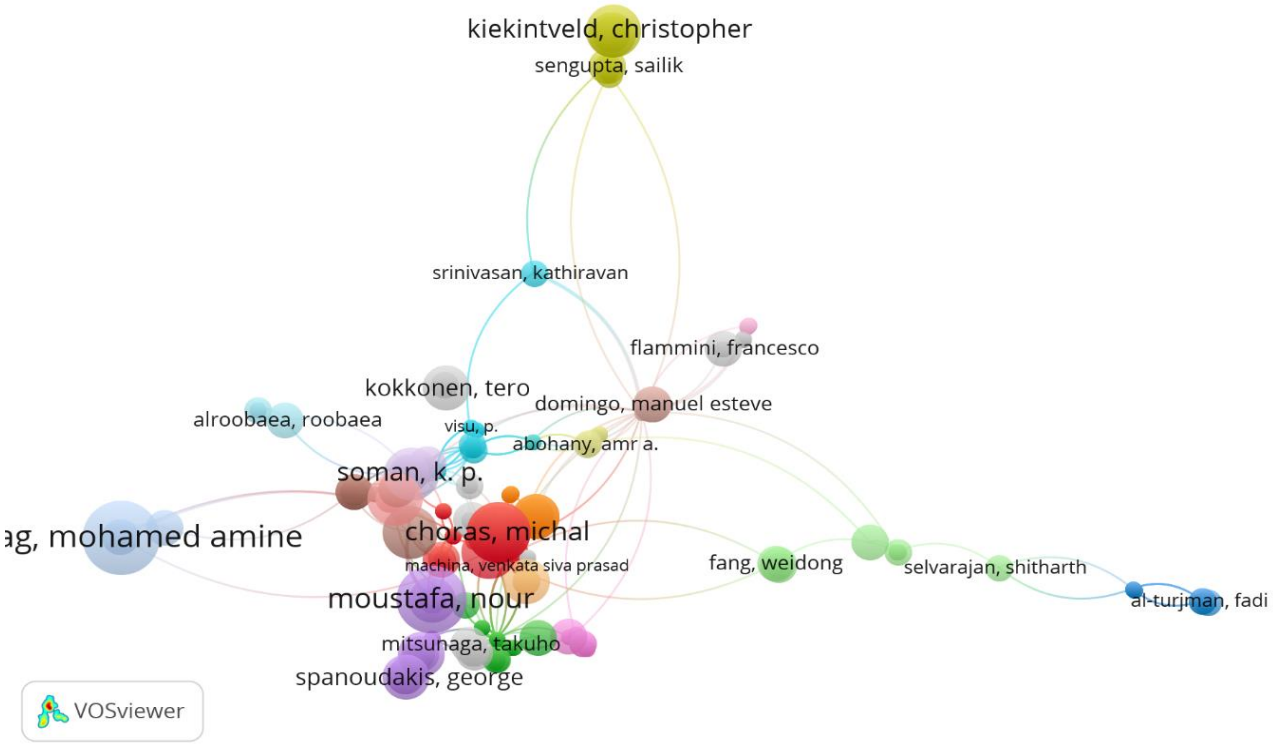


Fig.8. Network analysis showing authors' citation ties

Another analysis most frequently used in studies is the analysis of sources. When the analysis of the resources was examined, it was determined that the production was highest after 2018 and reached its highest values in 2024. Table 3 shows the data for the top ten sources with the most publications.

TABLE III
PUBLICATION NUMBERS OF SOURCES

Sources	Articles
Proceedings of the 3rd International Conference on Cybersecurity, Artificial Intelligence and Digital Economy 2024, CSAIDE 2024	115
IEEE Access	59
Electronic	24
Sensors	23
2023 IEEE Cybersecurity and Resilience, CSR International Conference	12
Computers and Security	12
Sustainability	12
2023 25th International Conference on Advanced Communication Technologies, ICACT	11
Applied Sciences-BASEL	11
Computers and Electrical Engineering	8

As a result of examining the sources in terms of bibliographic analysis, bibliographic match was revealed. Bibliographic matching on sources is defined as the citation of a common work by two independent sources as a result of the citation analysis. While examining the citation analysis of the sources, the criterion of having at least 1 publication and 1 citation was determined and 501 sources that passed the threshold value were listed. Figure 9 shows the network map showing the citation analysis of the sources.

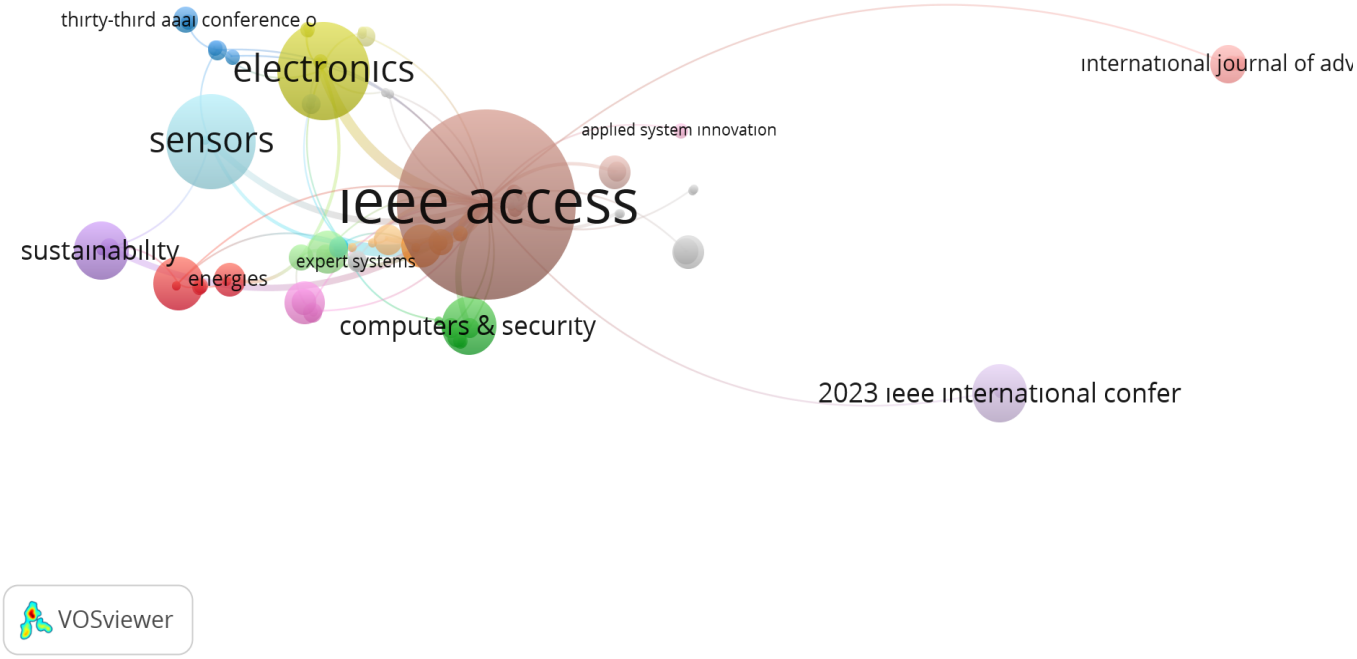


Fig.9. Citation network map of sources

E. Featured Trend Topics in The Field (Themes Analysis)

New fields of study and trending topics in the fields of cyber security and artificial intelligence have differentiated over the years. These topics have changed according to the prominent technological developments of the period. There are subject areas that have about to disappeared, as well as subject areas that feed a different subject in the new period and form the basis for its development. Identifying prominent themes and topics is important for researchers who want to work in this field to show current trends in the field. Min Cluster Frequency was selected as "5" for the Thematic map. Clustering Algorithm was selected as "Louvain". The selected algorithm is an algorithm that focuses on densities and prevents distributed clustering. Figure 10 presents a thematic map that provides a comprehensive analysis of academic studies in the field.

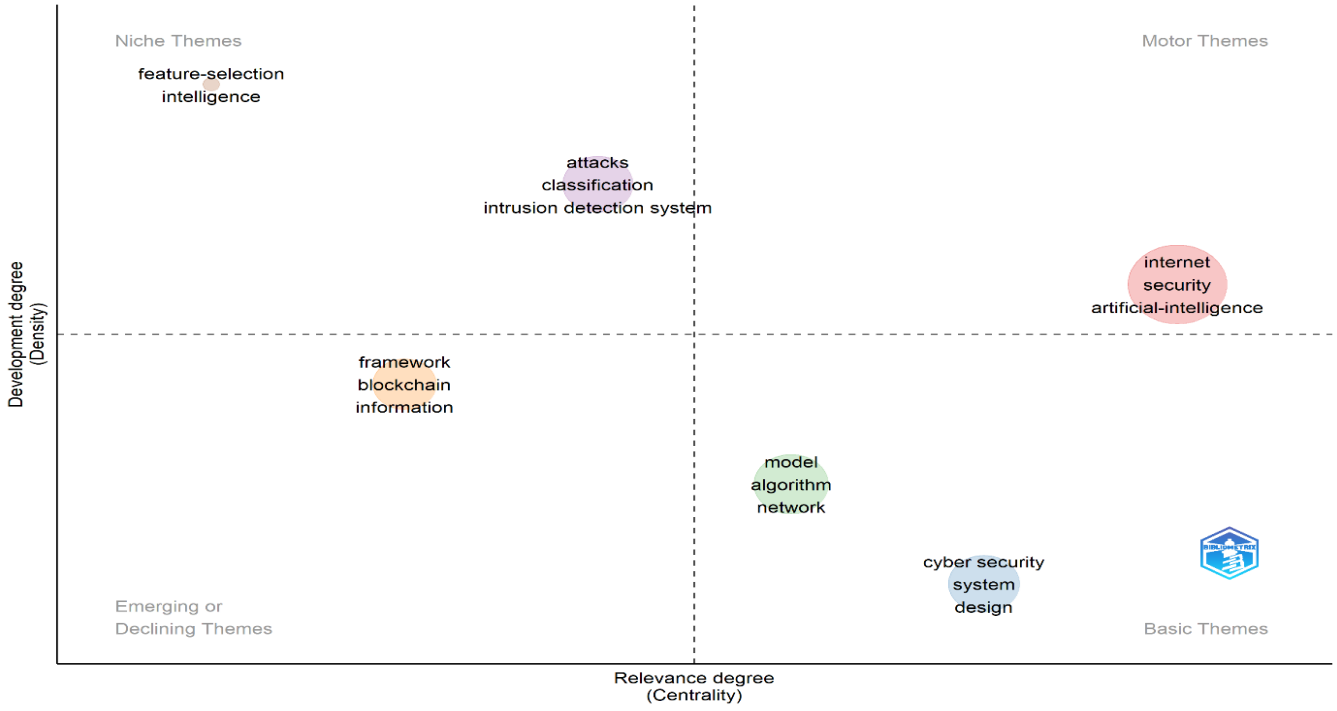


Fig.10. Agenda for future research (thematic map)

The X axis of the diagram represents network cluster centrality, which measures the degree to which clusters interact with each other and indicates the importance of a study theme. The Y-axis represents density, which is a measure of the internal connectivity of network clusters and a measure of theme growth.

Niche themes are defined as themes that cover a more specific area rather than a broad and general area While these themes were not directly related to the research area, they were developed and included in the field. Niche themes that stand out in the research area are categorized into two groups. The first set of prominent themes includes attacks, classification, intrusion detection system. The other cluster, which has not yet developed and attracted academic interest, consists of feature-selection and intelligence.

The motor themes have high centrality and high density. It includes important subject areas that shape the research field and contribute to its development. In other words, it means important and central topics in the field. The main cluster, which includes internet, security, artificial-intelligence, represents the leading motor themes in the field.

Emerging and declining themes have low centrality and low density. It represents themes that are likely to trend in the future and themes that are slowly fading into oblivion. In the diagram, the themes of framework, blockchain and information are in this group.

The basic themes have high centrality and low density. They represent interconnected topics that are more prominent in interdisciplinary studies. It is represented by two clusters in the diagram. The first cluster includes model, algorithm and network topics. The other cluster consists of cyber security, system and design topics, which are the leading and most fundamental topics in the research field.

Thematic Evolution is another analysis made while examining the issues in the study area. Themes that emerged and disappeared on selected important breakpoint dates are presented. Figure 11 shows thematic evolution, which shows past studies in this field and sheds light on future studies.

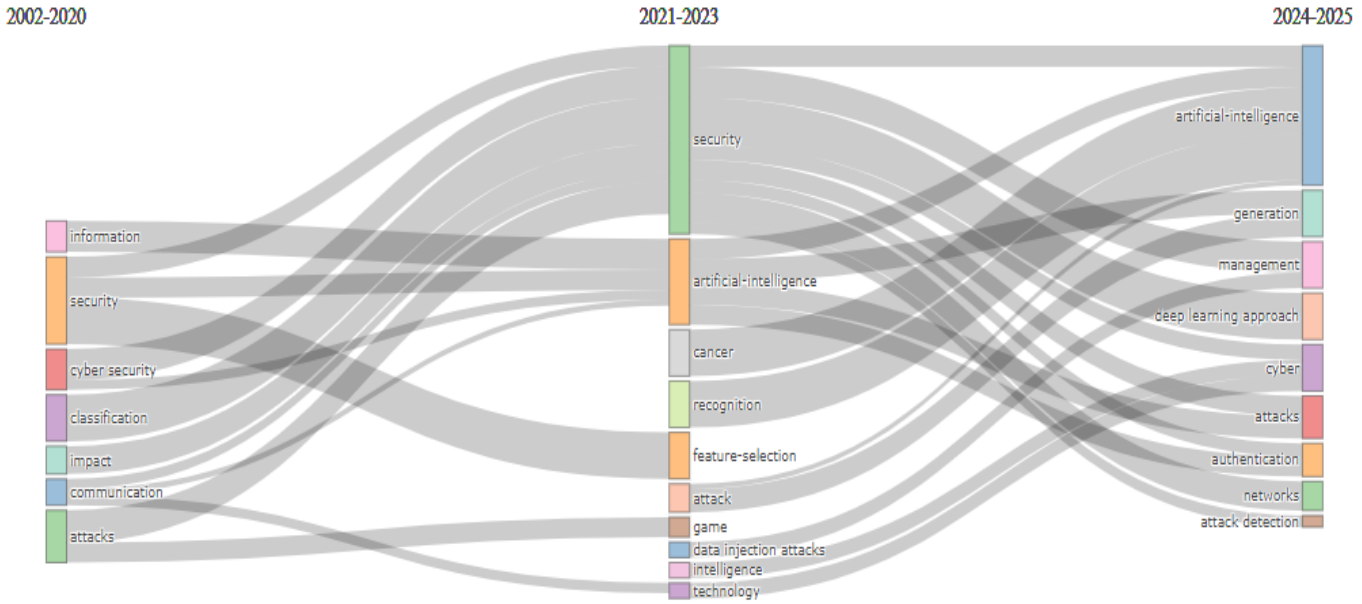


Fig.11. Thematic evolution of prominent topics in the research field

Among the parameters selected for the analysis, Min Cluster Frequency was set as 5, Clustering Algorithm as “Walktrap”, Min Weight Index as “0.1”, Number of Labels as “3”. The year 2020 was chosen as the breaking point in the diagram. This date was chosen because it is the year when the researches increased the most. After 2020, it is seen that the information field completely disappeared and fed the artificial intelligence subject area in the following period. Classification, impact, communication topics were combined under the security topic in the next period. The subject areas that are expected to remain important in the future are artificial intelligence, deep learning aproach, authentication, attack detection.

IV. CONCLUSION AND RECOMMENDATIONS

In today's digital world, artificial intelligence and cybersecurity have become increasingly integrated and complementary to each other. More effective solutions can be offered when combating threats in cyber security, By integrating these two areas. In this study, a bibliometric analysis was conducted on studies in the field of artificial intelligence and cyber security.

The Web of Science database was searched using the keywords "artificial intelligence" and "cyber security" using the conjunction "and" and selecting "all fields". A data set was created with the data obtained as a result of this search. The results of the bibliometric analysis conducted according to keywords, authors, countries and citations within the scope of the study are summarized below. The distribution by years of 1296 publications made between 2002 and 2025 (available at the time of the search) was examined. It was stated that the publications in the field of study increased after 2019 and the year 2024 stood out with 224 publications among the publication years. Studies have been prepared mostly in English. Among 4023 authors, the most influential authors were determined to be Yang Zhang, Jinjin Li, Jingshi Yang, Yicheng Liu, and Yisen Wang. According to the keyword analysis, the most frequently used keywords are cyber security, artificial intelligence, machine learning, deep learning and internet of things. Data obtained from the Web of Science database showed that studies were conducted in 97 countries. China (282 publications), USA (217 publications), India (125 publications), UK (102 publications) and Saudi Arabia (96 publications) are ranked as the top five countries. Another criterion taken into consideration when analyzing on a country basis is the number of citations. The most cited countries are the USA, China and India. The most commonly used sources in the studies were Proceedings of the International Conference on Cybersecurity, Artificial Intelligence and Digital Economy, IEEE Access, International Conference on Electronics, Sensors, Cybersecurity and Resilience.

The prominent study topics in the fields of cyber security and artificial intelligence were analyzed in the themes analysis section. In the topics examined in four groups as niche themes, motor themes, emerging and declining themes and basic themes, artificial-intelligence came to the fore in motor themes, while intrusion detections system, feature-selection were identified as niche themes. Cyber security was found as basic themes. When the study topics were analyzed according to years, it was determined that artificial intelligence, deep learning aproach, authentication, which was selected as the breaking point in 2020, are promising subject areas.

It is thought that the study will guide new researchers who will work in this field on the use of artificial intelligence in cyber security. In order to increase the effectiveness of the study, Voswiever (version 1.6.20) and R Studio programs were used together and the analysis results obtained from both programs were presented. The scope of the study can be expanded to include deep learning and cryptology fields and contribute to the literature. In addition, the study was conducted covering only the Web of Science database. The scope can be expanded by adding data from Scopus and Google Scholar databases in future studies.

REFERENCES

[1] H. Hoşgör, H. Güngördü. "Sağlıkta Yapay Zekanın Kullanım Alanları Üzerine Nitel Bir Araştırma.", European Journal of Science and Technology, Vol.35, 2022, pp. 395-407.

[2] F. Coşkun, H.D. Gülleroğlu. "Yapay zekânın tarih içindeki gelişimi ve eğitimde kullanılması.", Ankara University Journal of Faculty of Educational Sciences (JFES), Vol.54, No.3, 2021, pp.947-966.

[3] E. Gümüş, B. Medetoğlu, S. Tutar. "Finans ve bankacılık sisteminde yapay zekâ kullanımı: kullanıcılar üzerine bir uygulama.", Journal of Bucak Business Administration Faculty, Vol.3, No.1, 2020. pp.28-53.

[4] B. Kesici, M.S. Yıldız. "Kalite kontrol faaliyetlerinde yapay zekâ kullanımı ve bir otomotiv yan sanayisinde uygulanması.", Yalova Journal of Social Sciences, Vol.6, No.12, 2016, pp.307-323.

[5] E.Ö. Çizer, "Pazarlamada Yapay Zeka Uygulamaları." Individual, Society, Politics in the Digitalizing World, Congress Abstracts, İstanbul, Turkey, 26-27 May 2022.

[6] H. Güven, E.T.A. Güven. "Yapay Zekâ Uygulamalarının E-Ticarette Kullanımı.", International Journal of Management and Administration, Vol.7, No.13, 2023, pp.69-94.

[7] İ Terzi, M.M. Özgüven, Z Altaş, T. Uygun. "Tarımda Yapay Zeka Kullanımı.", International Erciyes Agriculture, Animal & Food Sciences Conference, Kayseri, Turkey, 24-27 April 2019.

[8] Z. Öğretmen Kotil. "Hukukta Yapay Zekâ Uygulamaları.", Artificial Intelligence Working Group, 2022, pp.1-12.

[9] C. Tiftik. "Insan kaynakları yönetiminde yapay zekâ teknolojileri ve uygulamaları.", IBAD Journal of Social Sciences, Vol.9, 2021. pp.374-390.

[10] R. Dirik, E. Taşkesen, Ö. Dirik. "Yenilenebilir Enerji Kaynaklarında Yapay Zekâ Kullanımı.", In International Conference on Recent Academic Studies, Vol.1, 2023, pp.28-35.

[11] N.S. Partigöç. "Afet risk yönetiminde yapay zekâ kullanımının rolü.", Journal of Information Technologies (JIT), Vol.15, No.4, 2022, pp.401-411.

[12] A. Tekin. "Yapay zeka kullanımının sanata etkileri.", Journal of Urban Academy, Vol.11 No.4, 2018, pp.692-702.

[13] M.Y. Başer, A. Olcay. "Akıllı turizmde yapay zekâ teknolojisi." Gaziantep University Journal of Social Sciences, Vol. 21, No.3, 2022. pp. 1795-1817.

[14] T.M. Erbir, O. Sevli. "Arama Motorlarında Yapay Zekâ Teknolojilerinin Kullanımı." 10th International Nardin Artuklu Scientific Researches Conference , Turkey, 2023.

[15] X. Liu, C. Li. (2023). *Yapay zeka ve çeviri*. Routledge Çeviri Teknolojisi Ansiklopedisi Routledge, 2023, p. 280-302.

[16] B. Anadolu. "Dijital hikâye anlatıcılığı bağlamında yapay zekânın sinemaya etkisi: Sunspring ve It's No Game filmlerinin analizi.", Journal of Erciyes Communication, Vol.1, 2019, pp.39-56.

[17] A. Efe. "Oyun Teorisi Perspektifinden Bilgisayar Oyunlarında Yapay Zekâ Kullanımı Üzerine Bir Değerlendirme.", Istanbul Gelisim University Journal of Social Sciences (IGUJSS), Vol. 11, No.1, 2024, pp.422-441.

[18] M.M. Mıjwıl, E. Sadıkoğlu, E. Cengiz, H. Candan. "Siber Güvenlikte Yapay Zekanın Rolü ve Önemi: Bir Derleme.", Journal of Data Science, Vol.5, No.2, 2022, pp.97-105.

[19] R. Von Solms, J. Van Niekerk. "From information security to cyber security.", Computers & Security, Vol.38, 2013, pp.97-102.

[20] A. Tekke, A. Lale. "Sosyal Medyada Etik, Bilgi Manipülasyonu ve Siber Güvenlik", Journal of Academic Inquiries (AID), Vol.16, No.2, 2021, pp. 44-62.

[21] N. De Bellis. "Bibliometrics and citation analysis: from the science citation index to cybermetrics.", scarecrow press, 2009.

[22] B. Godin. "On the origins of bibliometrics.", Scientometrics, Vol.68, No.1, 2006, pp.109-133.

[23] C. Chen. "Science mapping: A systematic review of the literature", Journal of Data and Information Science, Vol.2, 2017, pp. 1-40.

[24] N. Jan Eck, L. Waltman. "Software survey: VOSviewer, a computer program for bibliometric mapping.", Scientometrics, Vol.84, No.2, 2010, pp.523-538.

[25] Z.A. Polat, M. Alkan. "Jeodezi, jeoinformasyon ve arazi yönetimi dergisi'nin bibliyometrik analizi.", TMMOB Chamber of Surveying and Cadastre Engineers, Vol.15, 2015, pp.25-28.

[26] F. Ullah, H. Naeem, S. Jabbar, S. Khalid, M.A. Latif, F. Al-Turjman, L. Mostarda. "Cyber security threats detection in internet of things using deep learning approach.", IEEE access, Vol.7, 2019, pp.124379-124389.

[27] B.J. Radford, L.M. Apolonio, A.J. Trias, J.A. Simpson, *Network traffic anomaly detection using recurrent neural networks*, arXiv preprint arXiv:1803.10769, 2018.

[28] M.J.H. Faruk, H. Shahriar, M. Valero, F.L. Barsha, S. Sobhan, M.A. Khan, M. Whitman, A. Cuzzocrea, D. Lo, A. Rahman, F. Wu, "Malware detection and prevention using artificial intelligence techniques.", In 2021 IEEE international conference on big data (big data), 15-18 December 2021 (pp. 5369-5377). IEEE.

[29] M. Stampar, K. Fertalj. “Artificial intelligence in network intrusion detection.” In 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), IEEE. May 2015.
[30] F.M. Teichmann, S.R. Boticiu. “Adequate responses to cyber-attacks.”, International Cybersecurity Law Review, Vol.5, No.2, 2024, pp.337-345.
[31] S. More, A. Rohela. “Vulnerability assessment and penetration testing through artificial intelligence.”, International Journal of Recent Trends in Engineering Research, Vol.4, No.1, 2018, pp.217-224.
[32] M. Schmitt, I. Flechais. “Digital deception: generative artificial intelligence in social engineering and phishing.”, Artificial Intelligence Review, Vol.57, No.12, 2024, pp.1-23.
[33] İ. Yurtseven, S. Bagriyanik, S. Ayvaz. “A review of spam detection in social media.”, In 2021 6th International Conference on Computer Science and Engineering (UBMK), IEEE, September 2021.
[34] D.J. Dsouza, A.P. Rodrigues, R. Fernandes. *Multi-modal Comparative Analysis on Execution of Phishing Detection using Artificial Intelligence*. IEEE Access, 2024.
[35] B. Aytan, N. Barışçı. “Siber savunma alanında yapay zekâ tabanlı saldırı tespiti ve analizi.” In Proceeding of the 2nd International Symposium on Innovative Approaches in Scientific Studies, Samsun, December 2018.
[36] B. Özdemir. “Siber tehdit istihbaratında yapay zeka ve makine öğrenmesi.”, The Journal of Defence and Security Research, Vol.1, No.1, 2024, pp.75-96.
[37] S. Li, D. Qin, X. Wu, J. Li, B. Li, W. Han. “False alert detection based on deep learning and machine learning.”, International Journal on Semantic Web and Information Systems (IJSWIS), Vol.18, No.1, 2022, pp.1-21.
[38] Y. Liu, S. Li, X. Wang, L. Xu. “A review of hybrid cyber threats modelling and detection using artificial intelligence in IIoT.”, Computer Modeling in Engineering & Sciences, Vol.140, No.2, 2024.
[39] J.H. Li. “Cyber security meets artificial intelligence: a survey.” Frontiers of Information Technology & Electronic Engineering, Vol.19, No.12, 2018, pp.1462-1474.
[40] Web of Science. https://www.webofscience.com/wos/ woscc/analyze-results/b2364eca-c198-4339-b048-dc965c916cdd-4986caa2 (Access Date: 15. 10. 2024

BIOGRAPHIES

Songul Karakus received her undergraduate degree from Fırat University, Department of Computer Education and Instructional Technologies in 2010. She completed her master's degree at Fırat University, Department of Computer Education and Instructional Technologies. She started her doctoral studies at Fırat University Software Engineering Department. She completed her doctorate education in 2020. She is currently working as a faculty member at Bitlis Eren University Computer Engineering Department. Research regions; information security, machine learning, cyber security, software engineering.

Sevinc Ay received her undergraduate degree from Fırat University, Department of Software Engineering in 2016. She completed her master's degree at Fırat University, Department of Software Engineering. She started her doctoral studies at Fırat University Software Engineering Department. She completed her doctorate education in 2023. She is currently working as a Dr. Lecturer at Firat University Distance Education Center. Research regions; deep learning, machine learning, computer vision, image proccessing, video processing, object detection, object tracking, data science.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 3, September 2025

Research Article

Deepfake Detection for Digital Image Security using Deep Learning Methods

Huseyin Alperen Dagdogen, Resul Das, Ibrahim Turkoglu

***Abstract*—With the rapid increase and proliferation of digital technologies, manipulated content is also increasing in parallel. With the widespread use of deepfake technology, the detection of manipulated content and deceptive content reduces the risks of manipulated data. This situation leads to serious security consequences at social, political, and personal levels with the creation of fake news, misleading videos, and audio recordings. This technology can also cause serious problems such as malicious use and violation of privacy. Therefore, it is vital to develop preventive measures such as deepfake detection and to use this technology correctly and ethically. The detection of images created using deepfake techniques aims to detect manipulations in media files such as video and audio using artificial intelligence and machine learning techniques. Deepfake detection is usually carried out using deep learning algorithms and models.**

In this study, a hybrid model consisting of transformer-based networks and Convolutional Neural Networks (CNNs) is used to classify fake and real images. When the results of the study were examined, it was seen that the hybrid model used gave more successful results compared to the literature. The applications were carried out on the Casia-WebFace dataset. According to the results obtained, the proposed artificial intelligence method plays an important role in the classification process of images produced using DeepFake techniques. 98.82% accuracy rate was achieved for the Casia-WebFace dataset. These results show that the proposed artificial intelligence model is effective and successful in predicting deepfake techniques.

***Index Terms*—DeepFake, Vision Transformer, Convolutional Neural Networks, Forgery, Deep Learning, Image Security.**

I. Introduction

WITH the advancement of technology, the digital world is becoming increasingly complex, and its boundaries are becoming blurred. The rise of deepfake technology further deepens this complexity. Deepfake appears as a phenomenon that emerges from the combination of manipulated images and sounds created using artificial intelligence and deep learning methods. However, the impacts of this new technology do not stop at the entertainment world; it also presents serious ethical, social, and security dimensions. DeepFake was initially used in the film industry to enhance effects, but over time it has become applicable in a wider area [1]. It appears in a wide range of areas, from works of art to political debates, from news to social media content. However, every aspect of this technology needs to be examined because the boundary between reality and fiction is becoming increasingly blurred with the manipulation of images and sounds. However, the spread of deepfake technology also brings with it some important concerns. In recent years, this technology, which has also been used in areas such as object tracking and detection [2] and determining attack types in large data sets, has led to important discussions in terms of reliability and ethical use [3]. The reliability of manipulated content in the fields of image processing and review and the potential for malicious use of this technology have become serious concerns. In this context, it has become very important to develop a conscious approach to the ethical use and security of deepfake technology. Additionally, the rapid advancement of technology has led to the emergence of a persistent security gap in today's world. These security vulnerabilities affect people in every aspect of life, whether social, political, or economic, and the consequences are severe. Particularly in recent years, the development of deepfake technology and the creation of fake images, videos, and audio files have made these issues a part of our lives. Initially, DeepFake technology was a technique aimed at replacing a targeted point on a video, image, or audio with another type [4]. Recently, this technology has advanced to the point where it can generate both fake and unique content in conjunction with generative AI. To prevent the complexity that has emerged with the proliferation of DeepFake technology, the detection of fake images is of extreme importance. This article aims to detect images produced by DeepFake technology using a new technology known as Transformer Networks, intending to prevent the mentioned issues. The results obtained from the experiments conducted in the study show that successful results can be achieved when the detection of fake images is done on a pixel basis, and additionally, the trained model can make a clear distinction between fake and real images.

A. Related Works

Deepfake detection studies include examining techniques developed to distinguish and detect fake content from reality. These studies focus on developing solutions to detect and block deepfake videos using technologies such as image processing, artificial intelligence, deep learning, and similar methods. Various studies in the literature have used deep learning models to determine the authenticity of videos and images in datasets.

Huseyin Alperen Dagdogen is with the Department of Software Engineering, Technology Faculty, Firat University, Elazig, 23119 Türkiye e-mail: hadagdogen@firat.edu.tr

Resul Das is with the Department of Software Engineering, Technology Faculty, Firat University, Elazig, 23119 Türkiye e-mail: rdas@firat.edu.tr

Ibrahim Turkoglu is with the Department of Software Engineering, Technology Faculty, Firat University, Elazig, 23119 Türkiye e-mail: iturkoglu@firat.edu.tr

Manuscript received Jan 21, 2025; accepted Feb 15, 2025.
DOI: 10.17694/bajece.1624564

TABLE I: Review of Studies in The Literature

References	Year	Model	Measured Metrics	Cons
[5]	2020	Xception - MobileNet	Accuracy	Does not use inter-frame correlations Performance results at low resolution not examined
[6]	2018	VGG Face	Accuracy AUROC	Parameters such as learning rate, batch size or dropout are not optimized Limited model comparisons have been made It used only a CNN-based approach
[7]	2020	MesoNet - CNN	Accuracy	The model used is a shallow model, not a deep one Has a limited number of parameters A lightweight and low computational power model
[8]	2019	SCNNet	Accuracy AUROC	Not compared with methods based on time series analysis The performance of the model at different threshold values has not been examined
[9]	2018	InceptionV3	Accuracy	Too little data Low resistance to adversarial attacks
[10]	2020	Deepfake Stack	Precision Recall F1 Score Accuracy AUROC	Too little data Low resistance to adversarial attacks
[11]	2022	CNN	Precision Recall F1 Score Accuracy	The model is based solely on CNN, no new generation model is used Insufficient dataset Lack of temporal analysis
[12]	2024	Inception-ResNet-V1	Accuracy	It is not stated how durable the model is in different conditions Imbalances in training data
[13]	2024	Xception	Accuracy ROC	Computational cost Lack of explainability Vulnerability to new deepfake techniques

In the study numbered [5], deep learning technologies called Xception and MobileNet were used to automatically detect DeepFake videos. Four datasets using four different and popular DeepFake technologies created by FaceForensics++ [14] were used as training and evaluation datasets. The results obtained in the study show variable accuracy between 91% and 98% depending on the level of development of deepfake technologies in the industry. In addition, the researchers developed a voting mechanism that can detect fake videos using a combination of four methods instead of just one method. However, since the application was performed on a high-resolution dataset such as FaceForensics++, generalization to low-resolution datasets is limited. When the studies in the literature are examined, it is seen that they are generally divided into four separate categories [15]. These techniques are generally deep learning-based techniques, classical machine learning-based methods, statistical techniques, and blockchain-based techniques. The studies also evaluated the performance of the detection capabilities of various methods according to different datasets and concluded that deep learning-based methods are more successful than other methods in detecting deepfakes [15]. Most of the studies in the literature generally focus on the production and detection of Dee. Most of the time, in these studies, algorithms such as Generative Adversial Networks (GANs) [16], autoencoders, Variational Autoencoders (VAEs) [17], Recurrent Neural Networks (RNNs) [18] and CNNs [19] are used for deep learning and generative artificial intelligence applications. Many of these algorithms mentioned in the literature are generally used for deepfake production. It is stated that algorithms such as CNNs, Long Short-Term Memory (LSTM), Capsule Networks, GANs, and Autoencoders are used for the detection of produced deepfake images. Especially in recent years, with the development of generative artificial intelligence, GANs models play an important role both in image generation and in the detection of fake images. In the study numbered [6], Deep Convolutional Generative Adversial Network (DC-GAN) model was selected to generate 64x64 images and Progressive Growing Generative Adversial Network (PG-GAN) models were selected for 256x256 and 1024x1024 images. The aim here is to help adapt the training data to various image sizes and qualities. In addition, VGG16 architecture was used for the classification process and the results were observed to be 80%. However, since the obtained success rate is low, it is not at the level to be used in forensic studies. Non-deep learning methods require lower computational power and data requirements. However, these methods cannot reach the accuracy and consistency rates of deep learning-based methods. [7] presents a deep learning-based deepfake detection method called MesoNet. MesoNet is a two-step process. A process known as feature extraction extracts facial features in the first stage. The second stage uses a classification process to classify these features as fake or real. The study also introduces a new activation function called Pish to improve the accuracy and consistency of MesoNet. With Pish, MesoNet achieves an accuracy rate of 86.89% on the FaceForensics++ dataset. Only the combination of Swish and Mish performs better (over 87%), followed by the baseline combination (ReLU + Leaky ReLU) with 86.76%. The paper “Detecting Deepfake-Forged Contents with Separable Convolutional Neural Network (SCNNet) and Image

Segmentation" by C.M.Yu, C.T.Chang, and Y.W.Ti published in 2019 presents a deep learning-based deepfake content detection method called SCNNet. The accuracy rates for acSCNNet are comparable to other deep learning-based techniques used to detect fake content. However, SCNNet can use less data and computational power due to its separability technique. Therefore, SCNNet is suitable for applications that do not need a lot of data or processing power.

Deepfake detections are not only performed on images but also on data such as video, audio, and signals. When examining the study numbered [20], it is observed that the published articles are generally categorized under the four mentioned headings, and the models have been tested on these types of data. Most studies have used similar datasets. Kothandaraman et al. aimed to classify fake and real faces using the Inception-ResNet-V1 hybrid model in their study [12]. The proposed model is pre-trained with the VGGFace2 dataset, thus performing face recognition with high accuracy rates. In the study, face extraction was performed from the dataset using the MTCNN (Multi-Task Cascaded Convolutional Networks) algorithm. In this way, the input data was made more regular and the model's learning rate contributed to the success. However, how the model would react to factors such as different lighting conditions, angles, and rotations was not tested and specified. Similarly, Joshi et al. used the Xception model to classify images and videos produced with deepfake technology in their study [13]. The Xception architecture, which uses depthwise separable convolutions rather than CNNs architectures, provides more effective results. A test accuracy rate of 93.01% was obtained in the study.

B. Main Contrubitions

The proposed hybrid model in this study provides significant contributions to the literature on the detection and classification of images produced by deepfake techniques. In particular, our approach provides new strategies that provide both higher prediction accuracy and improved interpretability. The main contributions of the study can be summarized as follows:

- The proposed model combines the strengths and advantages of both architectures by combining the traditional CNNs architecture and the new generation Vision Transformer (ViT) architecture. While CNNs capture features, ViT learns long-term dependencies. By hybridizing these two powerful architectures, challenging tasks can be detected with higher accuracy and generalization.
- The proposed hybrid model uses CNNs to optimize the high computational cost and create an efficient structure. Thus, it requires lower processing power than pure Transformer models and offers a stronger generalization ability than CNNs models alone.
- The proposed model achieved better accuracy than traditional CNN-based methods such as VGG16, SCNNet, and CNN-MesoNet. This demonstrates that transformer architecture can significantly improve the performance when combined with CNNs in image classification.

C. Paper Organization

This study consists of 4 main sections including the introduction. After the introduction, the second section presents the proposed approach and implementation steps, and the third section presents the findings and discussion. The last section concludes the paper. Figure 1 shows the relevant organization chart.

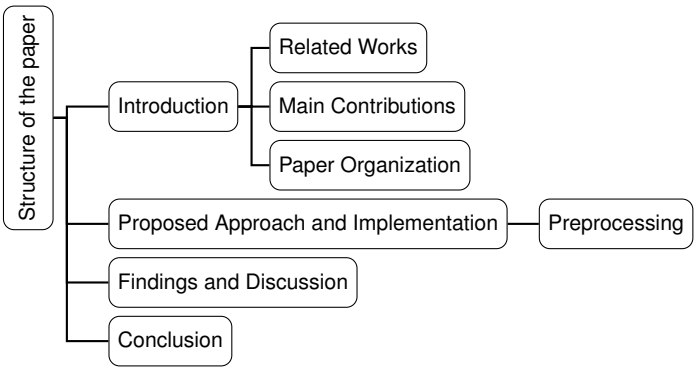


Fig. 1: Organization of the paper

II. Proposed Approach and Implementation

Convolutional neural networks have long been at the forefront of computer vision research. However, a noteworthy new development has surfaced recently in image processing. By modifying the transformer design, which has proven effective in Natural Language Processing (NLP), to analyze pixel-based data directly, ViT challenges conventional CNNs techniques. In essence, the ViT processes images using larger units known as "patches" rather than the traditional pixel-based method [21]. Using this method, the image is divided into patches, each of which is flattened before being fed into transformer blocks. ViT most distinctive characteristic is that, rather than using the convolutional and pooling layers seen in conventional CNNs, it uses transformer blocks that are rich with attention mechanisms. Figure 2 displays the model's architecture that was suggested for the investigation.

The study's suggested hybrid model combines the ViT model with convolutional neural networks. While ViT was used for the classification process, CNNs was used for feature extraction. Once the attributes of the input photos have been extracted, the ViT model has been utilized to classify the images and determine whether or not a certain image is fake.

The first step in the model is processing the images through several normalizing and convolutional layers. To reduce dimensionality, convolutional layers first extract feature maps from the images and subsequently identify patterns. After that, feature vectors are produced by flattening the feature maps. Activation, batch normalization, pooling, and convolutional layers make up the network architecture used for feature extraction. In the convolutional networks, the images begin at 512×512 dimensions and gradually shrink to 32×32 dimensions. The 512×512 pictures that pass through the convolutional layer first travel via the batch normalization layer before undergoing 2×2 maximum pooling operations.

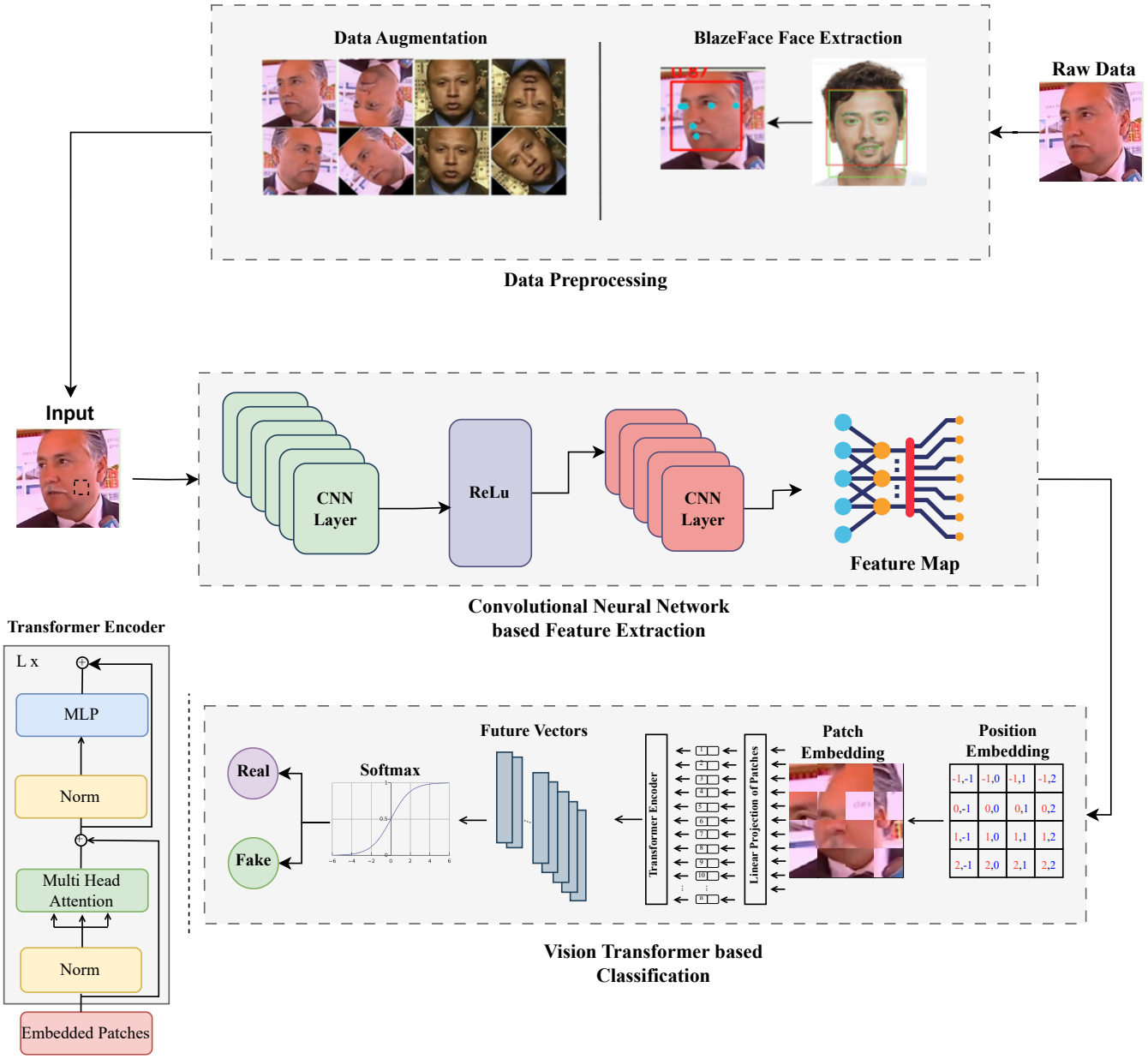


Fig. 2: The proposed hybrid model for deepfake detection

Features are captured sequentially by going through blocks of sizes 256×256, 128×128,64×64, and 32×32. For every size type, these operations are performed four times, using a stride value of 1.

A sequence of Transformer layers is applied to feature vectors. An attention mechanism pre-trained for better efficiency. The categorization task then made use of these relationships. The ViT network was used to train the captured features, and classification was carried out. Instead of processing the images' pixels directly, they use attention and self-attention mechanisms, which is different from standard CNN-based methods. Equation 1 displays the formula for the self-attention mechanism.

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q} \cdot \mathbf{K}^{\intercal}}{\sqrt{d_k}}\right) \cdot \mathbf{V} \qquad (1)$$

As can be seen from the equation 1, Q stands for the query, K for the key, and V for the value matrices. The size of the key matrix is denoted by d_k. The self-attention mechanism's fundamental computational processes are illustrated in this formula. The six feedforward and attention layers that comprise the Transformer layers in the suggested model each have 1024-dimensional hidden units. Eight heads, each with 128 dimensions, are used in each attention layer. A linear layer has been incorporated into the photos after they were separated into patches of a particular size. The pictures have been split into 196 16×16 patches. Every patch has been transformed

Algorithm 1 A general interval ViT classification algorithm and CNNs feature extraction.

Function ViT()

function CNNFeatureExtract(data)
  for each image in data **do**
    Preprocess the image *e.g., resize, normalize*
    Pass the image through the CNN
    Extract features from the CNN's intermediate layer
    Store the features
  end for
end function
function VisionTransformerClassification(features, labels, validation_data, max_epochs, patience, initial_lr, factor, cooldown, min_lr)
  Initialize Vision Transformer model with initial_lr
  best_model ← Vision Transformer model
  best_loss ← ∞
  no_improvement ← 0
  current_lr ← initial_lr
  for epoch = 1 to max_epochs **do**
    Train the Vision Transformer using features and labels with current_lr
    Compute validation loss on validation_data
    if validation_loss ¡ best_loss **then**
      best_loss ← validation_loss
      best_model ← Vision Transformer model
      no_improvement ← 0
    else
      no_improvement ← no_improvement + 1
      if no_improvement ≥ patience **then**
        break *Early stopping*
      end if
    end if
    if no_improvement ≠ 0 and no_improvement % cooldown == 0 **then**
      current_lr ← current_lr * factor
      if current_lr ¡ min_lr **then**
        current_lr ← min_lr
      end if
      Update Vision Transformer model with current_lr
    end if
  end for
  return best_model
end function
Input: Training data $\mathcal{D} = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$
Extract features using CNNFeatureExtract($\mathcal{D}$)
Output: Extracted features $\mathcal{F} = \{f_1, f_2, \ldots, f_n\}$
Validation data $\mathcal{V} = \{(x_{v1}, y_{v1}), (x_{v2}, y_{v2}), \ldots, (x_{vm}, y_{vm})\}$
TrainedModel ← VisionTransformerClassification($\mathcal{F}$, $\mathcal{D}$, $\mathcal{V}$, max_epochs, patience, initial_lr, factor, cooldown, min_lr)
return TrainedModel

into 49 2048-size feature vectors.

Furthermore, the vectors used for position embedding have a size of 1024. Moreover, the model incorporates the positional interactions between patches by embedding positions for every patch. Classification is done using the output from the transformer layers. For classification purposes, the outputs from the transformer layers were routed to an output layer with softmax activation for two classes after passing via an intermediary layer of size 2048. The output was run through a classification head to identify the class to which the provided image belongs. Algorithm 1 displays the algorithm used in the pertinent investigation.

A. Preprocessing

In the study, data preprocessing was performed in two stages. These are face extraction and data augmentation processes. BlazeFace [22] model was used for face extraction. The BlazeFace Model developed by Google is lightweight and performs fast face detection. This model, which is optimized to run on devices with low hardware features such as embedded systems and mobile devices, performs face detection using deep learning techniques, especially since it is a low-dimensional model. BlazeFace model, which has a CNN-based architecture, extracts the features of the image with a lightweight backbone architecture and creates feature maps to define regions on the face. The extracted feature maps are subjected to a regression process in the next stage. In the regression process, whether the image is a face or not is classified from the feature maps. If the image is a face, key points are determined. The model enables fast processing for real-time use and offers high-accuracy face detection. A convolutional block of the BlazeFace model is shown in Figure 3.

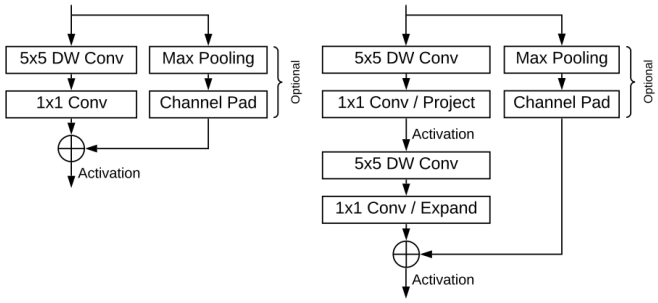


Fig. 3: BlazeBlock and Double BlazeBlock [22]

CNNs and deep learning were used to create face identification algorithm BlazeFace. A customized architecture that can identify objects quickly and accurately is the foundation of the operating principle [22]. The BlazeFace model's foundation, the Backbone Network, usually uses several backbone networks. These networks are employed for feature learning and data processing. These face detection-specific networks, such as MobileNetV1 [23] or MobileNetV2 [24], are often quick and lightweight models.

Traditional data augmentation methods have been applied in data augmentation procedures. Literature uses various data

augmentation approaches. These include methods such as data slicing, SMOTE, synthetic data generation, and augmentation. The data augmentation method used in this study helps to diversify the existing data set and provides a wider perspective to the model. Augmentation techniques in image processing include cropping, rotation, reflection, and brightness change. In this study, rotation was applied to the existing data set to perform data augmentation procedures. Different perspectives were obtained with these images rotated at certain angles. It is especially used in direction-independent object recognition operations. Thanks to the rotation of the data, the memorization of the training data of the model is prevented. Overfitting is prevented and the generalization ability of the model is improved with augmentation techniques such as rotation in small data sets. In this study, the images in the existing data set were rotated by 90 and 180 degrees, respectively, and recorded. An essential dataset for advancing facial recognition systems is the Casia-WebFace dataset [25]. Researchers utilize this dataset, a sizable database with various facial traits, to create and evaluate facial recognition systems. Facial photos of persons of various racial, age, and gender identities are included in Casia-WebFace. Additionally, it contains images captured in a variety of settings, including varying lighting, expressions, stances, and viewpoints. This kind is quite useful for assessing the performance of facial recognition systems in practical settings.

The Casia-WebFace collection contains 494,414 face images of 10,575 real identities collected from the web. These photographs cover a variety of facial features and were taken in a variety of environments. The effectiveness of face recognition systems in different environments is evaluated using this richness [26]. In the study, 80% of the dataset is used for training, 10% for validation, and 10% for testing. Figure 4 shows sample photographs from the Casia dataset and data augmentation examples.

Fig. 4: Casia WebFace Dataset Image Samples

III. Findings and Discussion

In the study, DeepFake classification was performed using CNNs and ViT architectures. The features extracted with CNNs were classified using ViT, and successful results were obtained. The proposed model has performed a faster and more accurate detection process. The differentiation of manipulated images has been achieved with a high success rate. In the preprocessing of the images, the BlazeFace face extraction

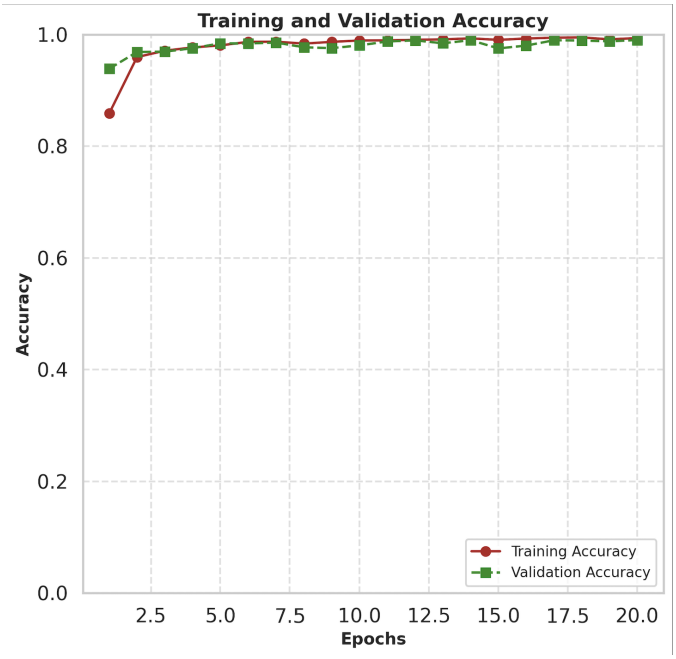


Fig. 5: Train accuracy graph

network was used, and existing data was augmented using data augmentation techniques. The data was trained using the EarlyStopping and ReduceLROnPlateau functions. The patience value for the EarlyStopping function was set to 5, and the patience value for the ReduceLROnPlateau function was set to 10. After all parameters were determined, the training was completed at 20 epochs, and the success rate for the CasiaWebFace dataset was observed to be 98.82%. The training accuracy graph of the study is shown in Figure 5, and the training loss graph is shown in Figure 6. Figure 5 shows that the graph shows that ViT can be an effective tool for DeepFake detection. The model can learn the training data quickly and the accuracy levels are quite high. The rapid increase in accuracy in the first 20 periods shows that the model learns the training data quickly and successfully learns the entire structure from simple features to more complex features. The proposed model uses the strengths of both models thanks to the hybridization of CNNs and ViT architectures. When CNNs, which have a successful structure in capturing features and relationships, are combined with the strong feature extraction ability of ViT capacity to learn global context and long-range connections, a high increase in the performance rate is achieved. The results obtained show that the study carried out by classifying the features extracted by CNNs with ViT can learn complex features that are important for distinguishing deepfakes. The accuracy fixed at approximately 98.82% shows that the model has completely learned the training data.

The parameters of the hybrid model proposed in the study are shown in Table II. In order for the attention mechanism to work more efficiently, the Feed Forward Network depth was determined as 6 and the number of neurons in the hidden layers was selected as 1024. The model, which consists of 8 attention heads, uses a multi-head attention mechanism. EarlyStopping was used to prevent overfitting and to reduce

TABLE II: Hyperparameter Description of Proposed Model

Model	Hyperparameters	Description	Value
ViT	Feed Forward Network	Artificial neural network without feedback or loop	6
	Dimensional Hidden Units	Number of neurons in the hidden layer	1024
	Number of Head	The number of topics an attention mechanism is divided into	8
	Patience Value for Early Stopping	To avoid overfitting and reduce unnecessary computational costs	5
	Patience Value for ReduceLROnPlateau	Allows the model to reach a better minimum	10
	Initial Learning Rate	The initial learning rate value used during training of the model	3e-4
	Learning factor	Used in learning rate schedulers in ReduceLROnPlateau	0.1
	Optimizer	is the algorithm that updates the weights to minimize the loss function	AdamW
CNNs	Input Layer	Layer representing the image data given to the model	512x512
	Kernel Size	Specifies the size of the filter used in the convolution layer	7x7
	Stride	Specifies how many pixels the kernel will shift on the input data	1
	Activation Function	Mathematical function that determines the output of neurons	ReLU
	Dropout	Regularization technique used to prevent overfitting	0.5

TABLE III: Proposed Model Accuracy Graph

Reference	Method (Classification)	DataSets	Accuracy	F1 Score	Recall	Precision
[5]	Xception - MobileNet	FaceForensics++	91%-98%	-	-	-
[7]	CNN- MesoNet	FaceForensics++	87%	-	-	-
[8]	SCNNET	FaceForensics++, Celeb-Df	95%	-	-	-
[6]	VGG16	CelebA-HQ	80%	-	-	-
[9]	Conv-LSTM	Hoha	97.01%	-	-	-
[10]	DeepFake Stack	FaceForensics++	99.65%	1.00	0.99	0.99
[11]	CNN	CasiaV2	92.23%	0.97	0.85	0.91
[12]	Inception-ResNet-V1	VGGFace2, DeepFake Dataset	97.5%	-	-	-
[13]	Xception	Kaggle deepfake_faces	93.01%	-	-	-
Proposed Model	CNN-ViT	Casia-WebFace	98.23%	0.97	0.98	0.97

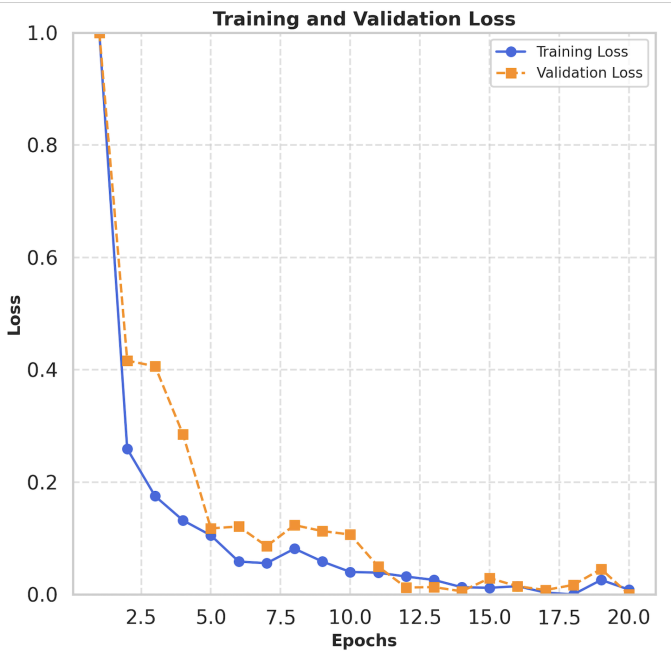


Fig. 6: Traning Loss Graph

unnecessary computational load. The EarlyStopping patience value was selected as 5. In addition, the ReduceLROnPlateau patience value, which is the learning rate planner, was selected as 10. In this way, the most optimal training hyperparameters of the model were selected with both EarlyStopping and ReduceLROnPlateau. Initially, the learning rate was selected as 3e-4 and the learning factor was determined as 0.1. In this way, the learning rate of the model was adjusted in a controlled manner during training.

As seen in the graph in Figure 6, the rapid decrease in the training loss in the first few epochs and the slowdown afterwards indicate that the learning process of the model is successful. As a result of 20 epochs of training with EarlyStopping, the training loss reached 0.2. The rapidly decreasing training loss shows that the model is learning fast. The slower decrease in the training loss after that indicates that the model starts to make smaller improvements. As seen in Figure 6, it is important to note that the training loss does not drop to 0, however, the low training loss obtained shows that the model is likely to perform well in general. Thus, when all results are taken into account, it shows that the proposed hybrid model for deepfake detection has a successful training process. The comparison of the results is shown in Table III. Although there are studies with higher success rates, the proposed model has an original structure thanks to a hybrid approach that demonstrates the strong feature extraction of the CNNs architecture and the classification power of the ViT architecture. Transformer models are particularly good at contextual learning on large datasets and therefore can perform better in image-based studies that contain more data. The proposed hybrid model combines traditional and new-generation deep learning techniques by hybridizing CNN and ViT architectures, providing a powerful representation of learning. While working on small datasets, the reduction of ViT's large data requirement and the preservation of its ability

to capture local features make the proposed approach unique and advantageous. One of the biggest advantages of the hybrid model is that it exhibits an effective approach to detecting time-varying inconsistencies. While CNNs and RNNs are successful in feature extraction and learning relationships between consecutive data points, Transformer-based models perform a more effective learning process since they capture spatial and temporal correlations better. In addition, the proposed hybrid model offers a fast and flexible learning process by providing scalability, parallelization, and efficient computation on large-scale image sets.

As shown in Table III, although CNN-based architectures show strong performance, their integration with new generation models such as GAN like ViT increases their performance rates. The proposed model CNNs-ViT provides a competitive accuracy rate compared to traditional CNNs models, indicating that ViT integration provides advantages. In model selection, high accuracy rates, used datasets and general model architecture should be taken into consideration.

IV. Conclusion

The analysis of the findings of the study shows that the hybrid classification model based on ViT and CNNs can work well to detect images generated by deepfake techniques. The model, which learns the training data quickly, performs the classification of fake images with high accuracy rates. In this article, local features of the images are extracted with CNNs, and maps of the extracted features are created. These maps, which are represented at high levels, are classified by ViT. The casia-WebFace dataset is used for training and testing in the study and 98.23% accuracy, 0.97 F1 Score, 0.98 Recall, 0.97 Precision values are obtained. Images generated by deepfake techniques, which have become a part of our lives today, create a major security gap. Many models have been proposed in the literature for the detection of fake images, but the model proposed in this study is among the models with high accuracy rates. In this way, the proposed model has great importance in terms of image security. As a result of the hybridization of CNNs and ViT architectures, the model explains the sensitivity. Especially with the increase in fake face images, these security vulnerabilities have become a bigger problem, and thus the model can become a new criterion for identifying these fake face images. The success of the proposed model is of great importance, especially when considering the security threats caused by deepfake technologies. With the increase in fake face images, security vulnerabilities such as identity theft, fake news, and fraud are increasing. In this context, the model's ability to detect fake images with high accuracy rates will contribute to ensuring security in the digital environment.

As a result, our hybrid model based on ViT and CNNs is considered as an important step towards increasing digital security by providing a reliable and effective method for detecting fake face images. The CNNs and ViT model opens the door to new research by showing that self-attention-based transformative models can be effective in deepfake detection applications. The proposed model can be tested with larger datasets in future studies to increase its generalization ability and to be integrated into real-time applications. The model, which is an important step for increasing digital security, will be tested with different datasets and trained with more realistic deepfake images, which will provide a more robust and comprehensive solution in terms of security.

Abbreviations

GANs Generative Adversial Networks
VAEs Variational Autoencoders
RNNs Recurrent Neural Networks
CNNs Convolutional Neural Networks
LSTM Long Short-Term Memory
DC-GAN Deep Convolutional Generative Adversial Network
PG-GAN Progressive Growing Generative Adversial Network
SCNNet Separable Convolutional Neural Network
NLP Natural Language Processing
ViT Vision Transformer

References

[1] F. Zengin, "Akilli makine çaği sinemasina giriş: Sinema sanatinda yapay zekâ teknolojilerinin kullanimi," *İletişim Çalışmaları Dergisi*, vol. 6, no. 2, pp. 151–177, 2020.

[2] R. Daş, B. Polat, and G. Tuna, "Derin Öğrenme ile resim ve videolarda nesnelerin tanınması ve takibi," *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, vol. 31, no. 2, p. 571–581, 2019.

[3] H. Ahmetoğlu and R. Daş, "Derin Öğrenme ile büyük veri kumelerinden saldırı türlerinin sınıflandırılması," in *2019 International Artificial Intelligence and Data Processing Symposium (IDAP)*, 2019, pp. 1–9.

[4] D. Afchar, V. Nozick, J. Yamagishi, and I. Echizen, "Mesonet: a compact facial video forgery detection network," *2018 10th Ieee International Workshop on Information Forensics and Security (Wifs)*, 2018. [Online]. Available: ⟨GotoISI⟩://WOS:000461290400003

[5] D. Pan, L. Sun, R. Wang, X. Zhang, and R. O. Sinnott, "Deepfake detection through deep learning," in *2020 IEEE/ACM International Conference on Big Data Computing, Applications and Technologies (BDCAT)*, 2020, pp. 134–143. [Online]. Available: 10.1109/BDCAT50828.2020.00001

[6] N.-T. Do, I.-S. Na, and S.-H. Kim, "Forensics face detection from gans using convolutional neural network," *ISITC*, vol. 2018, pp. 376–379, 2018.

[7] P. Kawa and P. Syga, "A note on deepfake detection with low-resources," *CoRR*, vol. abs/2006.05183, 2020. [Online]. Available: https://arxiv.org/abs/2006.05183

[8] C. Yu, C. Chang, and Y. Ti, "Detecting deepfake-forged contents with separable convolutional neural network and image segmentation," *CoRR*, vol. abs/1912.12184, 2019. [Online]. Available: http://arxiv.org/abs/1912.12184

[9] D. Güera and E. J. Delp, "Deepfake video detection using recurrent neural networks," in *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 2018, pp. 1–6. [Online]. Available: 10.1109/AVSS.2018.8639163

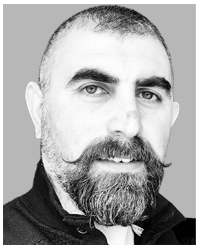
[10] M. S. Rana and A. H. Sung, "Deepfakestack: A deep ensemble-based learning technique for deepfake detection," in *2020 7th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2020 6th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom)*, 2020, pp. 70–75. [Online]. Available: 10.1109/CSCloud-EdgeCom49738.2020.00021

[11] S. S. Ali, I. I. Ganapathi, N.-S. Vu, S. D. Ali, N. Saxena, and N. Werghi, "Image forgery detection using deep learning by recompressing images," *Electronics*, vol. 11, no. 3, 2022. [Online]. Available: https://www.mdpi.com/2079-9292/11/3/403

[12] K. D, S. S. Narayanan, M. I. M, A. Yekopalli, and S. K. S, "Deep fake image classification engine using inception-resnet-v1 network," in *2024 International Conference on Computing and Data Science (ICCDS)*, 2024, pp. 1–5. [Online]. Available: 10.1109/ICCDS60734.2024.10560424

[13] P. Joshi and N. V, "Deep fake image detection using xception architecture," in *2024 5th International Conference on Recent Trends in Computer Science and Technology (ICRTCST)*, 2024, pp. 533–537. [Online]. Available: 10.1109/ICRTCST61793.2024.10578398
[14] A. Rössler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner, "FaceForensics++: Learning to detect manipulated facial images," in *International Conference on Computer Vision (ICCV)*, 2019.
[15] M. S. Rana, M. N. Nobi, B. Murali, and A. H. Sung, "Deepfake detection: A systematic literature review," *IEEE Access*, vol. 10, pp. 25 494–25 513, 2022. [Online]. Available: 10.1109/ACCESS.2022.3154404
[16] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative adversarial networks: An overview," *IEEE signal processing magazine*, vol. 35, no. 1, pp. 53–65, 2018.
[17] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*, 2013.
[18] L. R. Medsker and L. Jain, "Recurrent neural networks," *Design and Applications*, vol. 5, no. 64-67, p. 2, 2001.
[19] T. T. Nguyena, Q. Viet, H. Nguyenb, D. T. Nguyena, D. T. Nguyena, T. Huynh-Thec *et al.*, "Deep learning for deepfakes creation and detection: A survey," *SSRN Electron. J*, vol. 223, p. 103525, 2022.
[20] A. Heidari, N. Jafari Navimipour, H. Dag, and M. Unal, "Deepfake detection using deep learning methods: A systematic and comprehensive review," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, p. e1520, 2023.
[21] S. Khan, M. Naseer, M. Hayat, S. W. Zamir, F. S. Khan, and M. Shah, "Transformers in vision: A survey," *ACM computing surveys (CSUR)*, vol. 54, no. 10s, pp. 1–41, 2022.
[22] V. Bazarevsky, Y. Kartynnik, A. Vakunov, K. Raveendran, and M. Grundmann, "Blazeface: Sub-millisecond neural face detection on mobile gpus," *arXiv preprint arXiv:1907.05047*, 2019.
[23] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.
[24] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510–4520.
[25] D. Yi, Z. Lei, S. Liao, and S. Z. Li, "Learning face representation from scratch," *CoRR*, vol. abs/1411.7923, 2014. [Online]. Available: http://arxiv.org/abs/1411.7923
[26] J. Dong, W. Wang, and T. Tan, "CASIA image tampering detection evaluation database," in *2013 IEEE China Summit and International Conference on Signal and Information Processing*. IEEE, Jul. 2013. [Online]. Available: https://doi.org/10.1109/chinasip.2013.6625374

Huseyin Alperen Dagdogen is a research assistant at Firat University's Software Engineering department. He graduated with B.Sc. and M.Sc. degrees from the Software Engineering Department of Fırat University in 2021 and 2024, respectively. After working for several private sector companies in 2021 and 2022, he was a research assistant at the Software Engineering Department of the Faculty of Engineering and Natural Sciences at Malatya Turgut Ozal University in 2023. After working at Malatya Turgut Ozal University for about 1 year, he started serving as a research assistant at the Software Engineering Department of the Faculty of Technology of Firat University in 2024. His current research areas include machine learning, deep learning, network systems, and graph systems.



Resul Das is a full professor in the Department of Software Engineering at the Faculty of Technology, Firat University. He received his B.Sc. and M.Sc. degrees in Computer Science from Firat University in 1999 and 2002, respectively, and completed his Ph.D. in Electrical and Electronics Engineering in 2008. Between 2000 and 2011, he served as a lecturer in the Department of Informatics and concurrently worked as a network and system administrator at the University's IT Center. Since 2002, he has been an instructor in the Cisco Networking Academy Program, delivering CCNA and CCNP courses. From September 2017 to June 2018, he conducted research as a visiting professor at the University of Alberta, Edmonton, Canada, under the TUBITAK-BIDEB 2219 Postdoctoral Research Fellowship program. He also held the position of Head of the Department of Software Engineering from March 2020 to April 2023.

Professor Das has served in editorial roles for several prestigious academic journals. He was an Associate Editor for IEEE Access and the Turkish Journal of Electrical Engineering and Computer Science. Currently, he is an Associate Editor for several Elsevier journals, including Internet of Things, Alexandria Engineering Journal, and Telematics and Informatics Reports, as well as the IEEE Open Journal of the Communications Society (OJ-COMS) and the International Journal of Grid and Utility Computing (Inderscience).

Globally recognized for his academic contributions, Prof. Das has been consistently listed among the top 2% of the "World's Most Influential Scientists," compiled by Stanford University researchers, for five consecutive years (2019-2024). His research interests include computer networks, cybersecurity, IoT and systems engineering, data science and visualization, and software quality assurance and testing.

Ibrahim Turkoglu graduated with B.S., M.S., and Ph.D. degrees from the Department of Electrical and Electronics Engineering at the Firat University in 1994, 1996, and 2002 respectively. He completed his doctoral and master's thesis on artificial intelligence and pattern recognition. In particular, he developed two artificial intelligence techniques in his doctoral thesis. He has published over two hundred papers in international conference proceedings and refereed journals and has been actively serving as a reviewer for international journals and conferences. He has also been involved in many national and international projects. His current research areas include artificial intelligence, pattern recognition, bioinformatics, signal/image processing, and embedded systems. He is a lecturer for many courses such as Artificial Intelligence, Embedded Systems, Pattern Recognition, and Deep Learning in the Software Engineering Department.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 3, September 2025

Research Article

Improving Face Detection Performance of Compressed MPEG Videos by Using Frame-Independent Scene Change Detection Method

Mehmet Ozdem

***Abstract*—With the spread of computer vision applications, the performance of such applications also became prominent, especially when real-time and near real-time use cases are considered. If not all, many object detection algorithms follow a frame-based search approach, where all frames of the MPEG stream are analyzed sequentially. This drastically increases the computation time and the hardware requirements for such systems. This paper proposes employing a new scene-change detection method to improve object and face detection performance by eliminating the need to analyze every video frame. The method provides a frame-independent approach and does not require decoding and re-encoding of MPEG video. The paper also reports the performance test outcomes to exhibit the proposed approach's value. The findings show that a scene-change detection method enhances efficiency and decreases computational demands. Focusing on frames that show scene changes shows notable advancements in object detection performance.**

***Index Terms*—Face detection, Object detection, Frame-by-frame analysis**

I. Introduction

Using computer vision techniques has become increasingly popular in the last years in various verticals such as entertainment, security, health, and even e-commerce. Human face detection is also considered one of the dominant problems used across many domains. Being a dynamic object and having high variability in its appearance human face detection and recognition is known to be a difficult problem [1]. Thus, these applications lead to high computation time and increased hardware requirements.

In video content, a scene change typically refers to a transition from one distinct visual or thematic segment to another. This can include cuts, fades, dissolves, or any other transitions that signify a shift in content. Identifying these transitions involves analyzing changes in video frames. Detecting scene changes or scene cuts before object detection helps to eliminate the complexity of frame-by-frame analysis thus drastically improving performance. Yet, the overhead of existing scene change detection techniques also contributes to the computational time of the overall process.

The main contributions of this paper lie in the introduction of a novel scene-change detection method that significantly enhances the performance of object detection algorithms, particularly in real-time applications. By minimizing the need for exhaustive frame-by-frame analysis, the proposed approach optimizes computational efficiency and resource allocation, resulting in improved accuracy for both object and face detection. The method is frame-independent, avoiding the complexities associated with decoding and re-encoding MPEG video, thus streamlining the processing pipeline. The results demonstrate that focusing on frames with detected scene changes leads to notable gains in detection performance while reducing computational requirements. Additionally, the research paves the way for future integration of this method into existing computer vision systems and invites further exploration of its effectiveness across various video coding standards and diverse datasets.

This study proposes a two-phased scene-change detection algorithm where the overhead of scene detection in such applications are significantly reduced.

A. Problem Statement

One of the key challenges faced by object detection algorithms is the performance bottleneck caused by the sequential analysis of all frames in an MPEG video stream [2]. This approach leads to increased computation time and higher hardware requirements, which can hinder real-time and near-real-time applications. Yet another challenge is the performance overhead of the current scene-change detection approaches itself [3]. Detecting scene change is solely performed by visually comparing sequential frames. This challenge is more obvious, especially for videos coded with fixed GOP. There is room for improvement in this area.

II. Related Works

Scene change detection is a crucial aspect of video processing, particularly for tasks such as video indexing, summarization, and retrieval. In MPEG (Moving Picture Experts Group) videos, detecting scene changes helps in segmenting the video into distinct scenes, which is useful for a variety of applications including content analysis, editing, and navigation.

There are a significant amount of research studies reporting the performance and accuracy of scene-cut detection methods. Most of these methods can be applied to compressed videos thus need for decoding and reencoding is eliminated. One

Mehmet Ozdem is with the Turk Telekom, Ankara, Türkiye e-mail: mehmet.ozdem@turktelekom.com.tr

Manuscript received Nov 02, 2024; accepted Nov 28, 2024.

DOI: 10.17694/bajece.1577997

of the important studies on this area, conducted by Sethi and Patel, suggests statistical hypothesis testing to locate the scene cut points in the video [4]. The paper reports very high accuracy (almost 100%) while performance figures are not reported. The model is based on frame-by-frame comparison which promises high accuracy but the performance is still a question.

Another important contribution in this area is the study conducted by Huang and Liao, where scene changes are classified into abrupt scene changes and gradual scene changes [5]. The study examines the intensity of the image to capture potential fade-out and fade-in scenarios to detect gradual scene changes. The proposed model also distinguishes normal changes that happen in the scene and actual scene breaks. The accuracy of the method is reported to be very high concerning previous studies in this domain however the performance benchmarks are not reported.

Fundamentals of scene-change detection were published by Shahraray in 1995 [6]. Study introduces the concepts such as content-based temporal sampling and frame retention which are still being used and most of the studies conducted in this domain follow this terminology. In addition, there are AI-based solutions that take into account frames in videos and images [7].

III. Background

As part of the study, a two-phased scene-change detection algorithm is employed to identify changes between consecutive frames. This algorithm compares frames based on their size (in bytes) and their visual content to detect scene transitions. Once a potential scene change is detected, the corresponding frame is marked as an index frame for further analysis.

A high-level illustration of the proposed scene-change detection solution is given in Figure 1. The frame selector

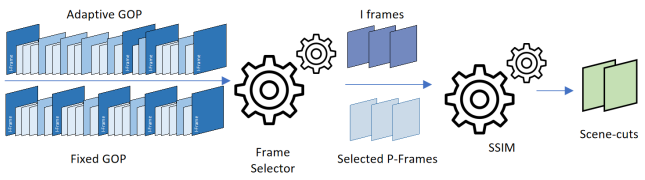


Fig. 1. Building Blocks of Proposed Approach

module identifies if the video is coded with fixed GOP or with adaptive GOP. In case adaptive GOP is used, all I-frames are selected as potential scene-cut frames. In case fixed GOP is used, P-frames close to I-frames in size (bytes) are selected as potential scene-cut frames. All frames marked as potential scene-cuts are then analyzed with SSIM to find the dissimilarity value with the previous I-frame using the SSIM module. MPEG videos are encoded using a series of frames, including I-frames (intra-coded frames), P-frames (predicted frames), and B-frames (bidirectionally predicted frames). Scene change detection generally starts with analyzing these frames:

I-frames serve as key frames containing the complete image data. They are crucial for detecting scene changes because they represent the entire frame independently. P-frames and B-frames depend on preceding or surrounding frames for their data. Scene changes can be harder to detect in these frames due to their reliance on previous frames.

Detection methods often involve comparing I-frames to identify significant differences between them. When the difference between consecutive I-frames exceeds a certain threshold, it suggests a possible scene change.

Advanced Video Coding (AVC) is still by far the most widely used video coding standard for compression and distribution of video content [8]. AVC, also referred to as MPEG-4 Part 10, expresses the video content as a series of frames of various types. Typically, a high-definition (HD) video is composed of 25 frames per second and an ultra-high definition (UHD) video is composed of 60 frames per second.

AVC compression reduces the temporal redundancy across frames using predicted pictures (P-frames) which are coded concerning past full pictures or index frames (I-frame). P-frames contains only the delta concerning the previous I-frame thus, smaller in size. Bi-directional pictures (B-frames) are like P-frames but coded concerning past or future I-frames or P-frames. I-frames or full pictures by nature are heavier in size but still required to be inserted in constant or variable intervals. It is very common to insert an I-frame every 1-2 seconds to enable easier random access, to avoid propagation of errors for longer durations, and to keep the size of frame-to-frame differences at a reasonable level for better compression [9].

This brings us to the concept of a group of pictures (GOP). GOP is the distance between two I-frames expressed in the number of frames. Meaning; for a 25 fps (frames per second) video an I-frame is inserted every 2 seconds GOP is selected to be 50 frames. The notion of GOP is illustrated in Figure 2.

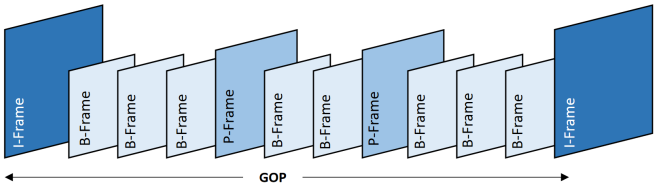


Fig. 2. MPEG Group of Pictures (GOP)

If I-frames are inserted into the video at constant intervals like an I-frame every second or every two seconds, regardless of the scene changes or scene cuts, this is called fixed GOP. A Fixed GOP structure is a static approach where the arrangement of I-frames (intra-coded frames), P-frames (predicted frames), and B-frames (bidirectionally predicted frames) is predetermined and remains constant throughout the video. It offers an easier coding process however introduces the risk of heavier reference frames (B-frames and P-frames) in case there is a scene change within a GOP.

Adaptive GOP, on the other hand, allows for dynamic adjustment of the GOP structure based on the content of the video. The encoder can modify the GOP size and frame types according to scene complexity, motion, and other factors. If not all, most of the decent encoders are capable of coding with adaptive GOP where I-frames are inserted when necessary

[10]. Adaptive GOP promises not only better compression but also improves the chance of identification of scene changes. A comparison of fixed GOP and adaptive GOP approaches is illustrated in Figure 3.

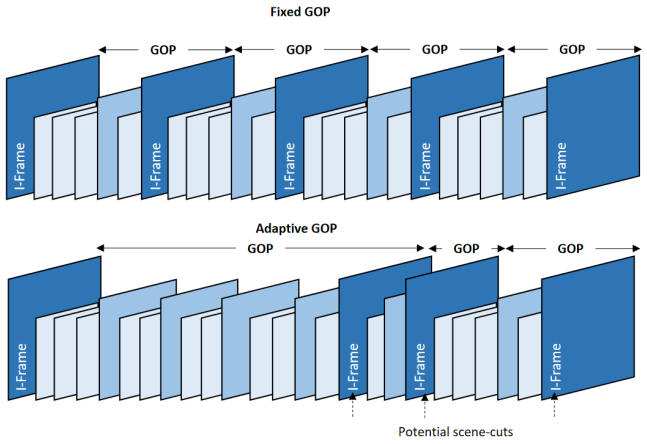


Fig. 3. Fixed GOP vs. Adaptive GOP

Fixed GOP and Adaptive GOP each offer distinct advantages depending on the application. Fixed GOP provides simplicity and predictability, making it suitable for scenarios with consistent content and bitrate requirements. In contrast, Adaptive GOP excels in handling varied content by dynamically adjusting the GOP structure to optimize both video quality and compression efficiency.

As seen, scene change detection is straightforward if the video is coded using adaptive GOP since the assumption of "no scene cut between I-frames" will hold higher chances. Scene detection problem for videos coded with fixed GOP is still a challenge.

In this study, to reduce the complexity of object detection on MPEG videos, a frame selection approach is proposed as opposed to sequential analysis of every single video frame. Frame selection is done based on the detection of scene changes and the sampling of frames for each scene.

The approach is applied to a face detection application. While it is obvious that there will be a performance gain on object detection, the performance drawback of scene-change detection and the increased risk of false-positive results (missing the key object on non-selected frames) are also studied and reported.

IV. Materials and Methods

To address these challenges, we propose a new scene-change detection method that alleviates the need to analyze each frame individually. By identifying scene changes, our method allows for a more efficient allocation of computational resources, thereby improving the performance of object detection algorithms, including face detection. Importantly, our approach is frame-independent, meaning it does not rely on frame-by-frame analysis, and it does not require the decoding and re-encoding of the MPEG video.

The proposed method aims to improve face detection performance by applying a frame-independent scene-change detection approach on compressed MPEG videos. It eliminates the need to analyze every video frame, thereby reducing computation time and hardware requirements. The proposed method aims to improve face detection performance by applying a frame-independent scene-change detection approach on compressed MPEG videos. It eliminates the need to analyze every video frame, thereby reducing computation time and hardware requirements.

As part of the study, a scene-change detection algorithm is employed to identify changes between consecutive frames. This algorithm compares frames based on their size (in bytes) and their visual content to detect scene transitions. Once a scene change is detected, the frame corresponding to the scene change is marked as an index frame for further analysis.

As explained in the previous chapter, detecting scene change for videos coded using fixed GOP is more complex compared to videos coded using adaptive GOP. Therefore, this study is more focused on videos having fixed GOP size.

Noting that P-frames contain only the changes concerning previous index frames a significant increase in size gives us the impression that a scene-change may have occurred. Therefore, the initial phase of the two-phase approach proposed in this study is to detect frame size anomalies in the MPEG frame structure.

Detecting frame size anomalies in MPEG videos involves identifying unusual variations in the size of P-frames (predicted frames) compared to expected values based on typical frame sizes or content characteristics. While there are multiple approaches for anomaly detection a three-step method is preferred in our case:

1) Frame Extraction
2) Statistical Analysis over mean size and standard deviation
3) Anomaly Detection over Z-score method and percentile method

Frame extraction may be done using a variety of tools while open source ffmpeg tool is used in our study. A section from the output of the ffmpeg command where I-frames and P-frames are filtered with their packet sizes is presented in Figure 4. Statistical analysis involves the calculation of average

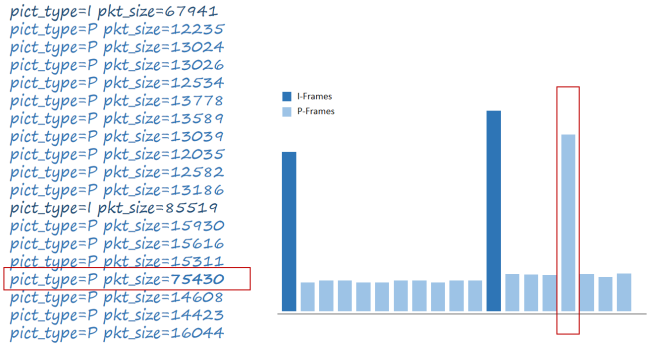


Fig. 4. Video Frames with Type and Size Info

sizes of P-frames and computing the standard deviation to understand the typical variation in frame sizes. Figure 5 presents the implementation of these calculations with Python.

```
frame_sizes = np.array(frame_sizes)

# Compute statistics
mean_size = np.mean(frame_sizes)
std_dev = np.std(frame_sizes)
```

Fig. 5. Implementation of Statistical Analysis

```
# Calculate Z-scores
z_scores = (frame_sizes - mean_size) / std_dev
z_anomalies = np.where(np.abs(z_scores) > z_threshold)[0]

# Calculate percentile thresholds
lower_percentile = np.percentile(frame_sizes, 100 - percentile_threshold)
upper_percentile = np.percentile(frame_sizes, percentile_threshold)

percentile_anomalies = np.where((frame_sizes < lower_percentile) |
                                (frame_sizes > upper_percentile))[0]

# Combine anomalies
anomalies = set(z_anomalies).union(set(percentile_anomalies))
```

Fig. 6. Implementation of anomaly detection

The anomaly detection step employs z-score calculation and percentile threshold together to identify abnormal p-frames. Z-score for each P-frame size is calculated to measure how many standard deviations a frame size is from the mean. As different z-score thresholds are tested, 3 is observed to be the most accurate for the video types used in our case. On top of the z-score, percentile thresholds are also used with a percentile threshold of 95. Figure 6 presents the implementation of these calculations in Python. This approach helps to apply visual comparison only on specific frames rather than traversal of the entire video frame-by-frame.

The scene-change detection algorithm compares identified frames to check significant changes based on visual content. Identified frames are compared with previous frames using structural similarity index (SSIM) calculation [11]. This technique measures the dissimilarity between frames based on pixel intensity differences. A threshold is applied to the dissimilarity measure to ensure that a scene change has occurred. If the dissimilarity value exceeds a predefined threshold, it indicates a significant change between frames, and a scene change is detected.

SSIM is mostly used in the broadcasting industry for measuring the perceived quality of cinematic pictures and digital television content [12]. It simply measures the similarity of the given two images. In this study, SSIM is used to obtain structural dissimilarity of two consecutive video frames. SSIM uses three components to measure the similarity: luminance (l), contrast (c), and structure (s). The formulas to measure these components are as follows.

$$l(x,y) = \frac{2\mu_x\mu_y}{\mu_x^2 + \mu_y^2} \tag{1}$$

$$c(x,y) = \frac{2\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2} \tag{2}$$

$$s(x,y) = \frac{\sigma_{xy}}{\sigma_x + \sigma_y} \tag{3}$$

Here, in formula 1, 2, and 3; x represents the pixel sample mean of x and y represents the pixel sample mean of y. x2 represents the variance of x and y2 represents the variance of y. xy represents the covariance of x and y. In formula 4, the overall SSIM value is the weighted product of these components. For simplicity, in this study, we set α, β, and γ values to 1 so all components are equally weighted.

$$SSIM(x,y) = (l(x,y))^{\alpha} \cdot (c(x,y))^{\beta} \cdot (s(x,y))^{\gamma} \tag{4}$$

Finally, the following formula 5 is used to obtain the dissimilarity of given frames, expressed as a percentage. The higher dissimilarity value is the potential scene-change indicator.

$$DSSIM(x,y) = (1 - SSIM(x,y) \times 100) \tag{5}$$

During tests, 90% is used as the threshold since the assumption of dissimilarity above 90% refers to almost completely different images thus a scene cut. Upon detecting a scene change, the frame corresponding to the scene change is selected as the index frame for face detection and subsequent tracking. This frame represents the new scene and serves as a reference for face detection in the scene.

V. Performance Measurement Results

In the scope of this study, performance tests are conducted to evaluate the effectiveness of the proposed approach. These tests are executed on sample videos having various types and durations.

Expected performance directly depends on the type of content. High-motion video content (sports events, action videos) and relatively static content (news programs) have an impact both on scene change detection performance and on face recognition performance. In total 43 test videos containing multiple samples to represent these two content groups were used during tests. Characteristics of sample videos are summarized in Table I. Performance tests are executed

TABLE I
Sample Video Content for Testing

Type	Avg. Duration	Codec and Resolution	Avg. Bitrate	Number of Samples
Action Movie	00:05:10.01	AVC (MPEG-4 Part 10) 30fps, 1080p	2673 Kbps	25
Sports Content	00:04:00.41	AVC (MPEG-4 Part 10) 30fps, 1080p	2430 Kbps	10
News Program	00:11:42.11	AVC (MPEG-4 Part 10) 30fps, 720p	1141 Kbps	8

for two scenarios. Scenario A; where the source video file is provided to the Face Detector module with all frames, scenario B; where the source video file is first provided to the Scene-change Detector module and the Face Detector module is fed only with the selected frames where the scene changes are detected. These scenarios are expressed as block diagrams in Figure 7. Another factor that impacts test performance is the computational capacity of the test environment. Server hosting the Face Detector module is equipped with 2 processors (CPU) both having 2 CPU cores with 3.0 GHz base frequency. The server is also equipped with 64GB of memory.

Scene-change detector module is hosted on a workstation (PC) with a single processor having 10 CPU cores with a maximum of 4.70 GHz base frequency. The workstation is

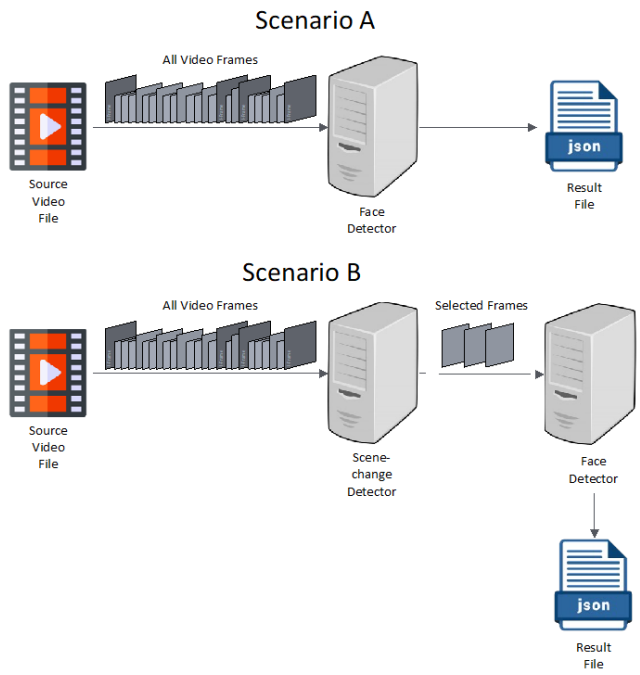


Fig. 7. Test Scenarios

equipped with 32GB of memory. Although it is typical to use Graphical Processing Units (GPU) for such applications, the test in this study is executed on CPU modules for both scenarios. It is assumed that this will have the same impact on both scenarios and thus will not have a significant effect on our benchmarking.

Performance test results are summarized in Table II and Figure 8 for both scenarios. As seen from the test results, despite the overhead of scene-change detection, the scene-change approach significantly improves performance for all content types. The gain is even higher for news programs.

The performance gain can also be seen in the number of frames passed to the Face Detector module concerning the total number of video frames. As expressed in Table III and Figure 9, the significant number of frames is eliminated thus the amount of face detection cycles is drastically reduced. These figures also align with the results listed in Table II.

TABLE II
PERFORMANCE TEST RESULTS

Content Type	Average Processing Time for Scenario A	Average Processing Time for Scenario B
Action Movie	997 seconds	393 seconds
Sports Content	894 seconds	307 seconds
News Program	791 seconds	280 seconds

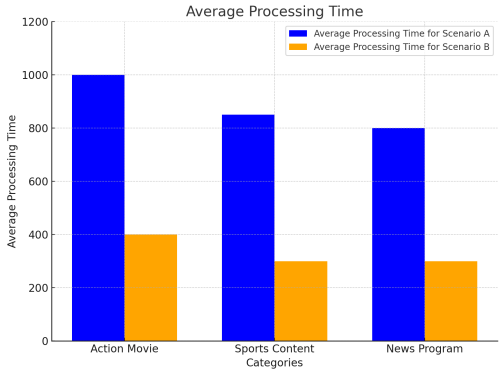


Fig. 8. Average Processing Times per Content Type

TABLE III
NUMBER OF PROCESSED FRAMES BY FACE DETECTOR

Content Type	Average Processed Number of Frames for Scenario A	Processed Number of Frames for Scenario B
Action Movie	9300 frames	1944 frames
Sports Content	7208 frames	1184 frames
News Program	21084 frames	3051 frames

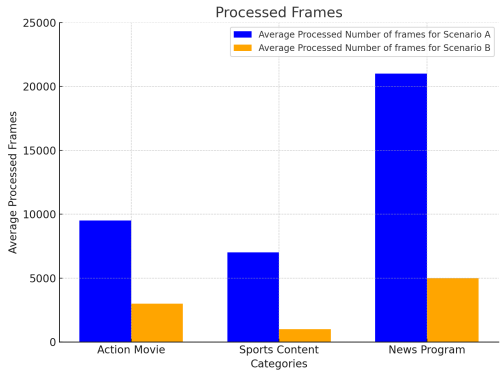


Fig. 9. Processed Frames per Content Type

The objective of the tests was not only to measure the performance gain but also to observe the effect on face detection accuracy. Table IV and Figure 8 summarize the number of face detections per movie type for both scenarios. It is observed that the proposed approach has almost no impact on the number of faces detected.

TABLE IV
NUMBER OF CORRECTLY PREDICTED DATA IN THE TEST RESULT

Content Type	Average Number of Faces Detected for Scenario A	Average Number of Faces Detected for Scenario B
Action Movie	216	211
Sports Content	87	87
News Program	307	300

To summarize; the proposed approach achieves a significantly faster execution time compared to using only the face detector. For instance; for a news program on average the face detection process is completed in 280 seconds, which is around one-third the time taken by the face detector alone (791 seconds). This indicates the efficiency of the proposed method in processing the video frames and performing face detection. Despite the reduced processing time, the approach method maintains a high level of accuracy in face detection.

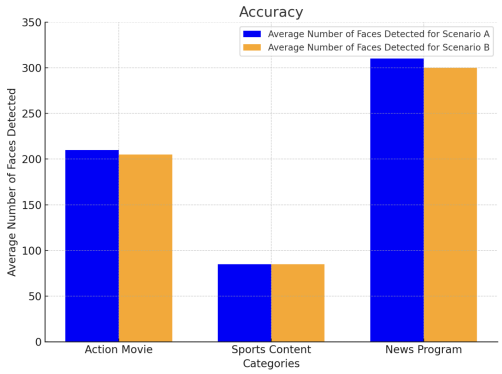


Fig. 10. Accuracy of Proposed Solution

Results highlight the effectiveness of the proposed method in achieving faster face detection while maintaining a high level of accuracy. The optimized processing performance is considered to be beneficial for real-time and near-real-time applications where timely detection is crucial.

VI. Conclusion

The proposed scene-change detection method offers a promising solution for improving the performance of object detection algorithms, particularly in real-time and near-real-time scenarios. By reducing the reliance on frame-by-frame analysis and optimizing computational resource allocation, the approach enhances efficiency while maintaining the accuracy of object detection, including face detection. Future research could explore the integration of the proposed approach into existing computer vision systems and explore its applicability across different domains.

The results demonstrate the improved efficiency and reduced computational requirements achieved by employing a scene-change detection approach. By selectively analyzing frames based on scene changes, significant improvements in object detection performance are observed. It is important to note that the performance of the proposed approach may vary depending on the characteristics of the video, such as scene complexity, motion, and lighting conditions. Further experimentation and evaluation with diverse video datasets are being planned to assess the generalizability and robustness of the method.

Further study is also planned to test the method on different video coding standards such as HEVC, VP9, and AV1. Although AVC is still the dominant standard, these standards are getting more popular and will probably dethrone AVC soon.

References

[1] E. Hjelmås and B. Low, "Face detection: A survey," *Computer Vision and Image Understanding*, vol. 83, no. 3, September 2001.

[2] L. Jiao, R. Zhang, F. Liu, S. Yang, B. Hou, L. Li, and X. Tang, "New generation deep learning for video object detection: A survey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 8, August 2022.

[3] L. Baraldi, C. Grana, and R. Cucchiara, "Measuring scene detection performance," in *Pattern Recognition and Image Analysis: Proceedings of 7th Iberian Conference, IbPRIA 2015*, Santiago de Compostela, Spain, June 2015.

[4] I. Sethi and N. Patel, "A statistical approach to scene change detection," in *Proceedings of SPIE - The International Society for Optical Engineering*, June 1998.

[5] C.-L. Huang and B.-Y. Liao, "A robust scene-change detection method for video segmentation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 12, December 2001.

[6] B. Shahraray, "Scene change detection and content-based sampling of video sequences," in *Proceedings Volume 2419, Digital Video Compression: Algorithms and Technologies*, April 1995.

[7] R. Daş, B. Polat, and G. Tuna, "Recognizing and tracking objects in images and videos with deep learning," *Fırat University, Mühendislik Bilimleri Dergisi*, vol. 31, no. 2, pp. 571–581, 2019.

[8] Bitmovin, "Video developer report 2018," September 2019.

[9] A. Huszak and S. Imre, "Analysing gop structure and packet loss effects on error propagation in mpeg-4 video streams," in *Proceedings of 4th International Symposium on Communications Control and Signal Processing (ISCCSP)*, March 2010.

[10] B. Zatt, M. Porto, J. Scharcanski, and S. Bampi, "Gop structure adaptive to the video content for efficient h.264/avc encoding," in *Proceedings of IEEE International Conference on Image Processing*, September 2010.

[11] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, April 2004.

[12] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Transactions on Image Processing*, May 2011.

Biographies

Mehmet Özdem received his BS degree from Başkent University, Department of Electrical and Electronics Engineering in 2002, his MS degree from Middle East Technical University, Institute of Informatics in 2007, and his PhD degree from Gazi University, Institute of Science, Department of Electrical and Electronics Engineering in 2021. He has more than 20 years of experience in the Telecommunications industry. He worked in many groups (investment, planning, operation, quality assurance, and architecture teams). He currently works as the Network Architecture & Quality Assurance Director and Head of R&D Centers, following responsibilities in Turk Telekom. Simultaneously, he is Vice President of ITU (International Telecommunication Union) SG12 and QSDG organizations. He is also a board member of the Wireless Broadband Alliance. He lectures part-time at universities and has international certificates such as CCIE, PMP, and ITIL.

Publication Ethics

The journal publishes original papers in the extensive field of Electrical-electronics and Computer engineering. To that end, it is essential that all who participate in producing the journal conduct themselves as authors, reviewers, editors, and publishers in accord with the highest level of professional ethics and standards. Plagiarism or self-plagiarism constitutes unethical scientific behavior and is never acceptable.
By submitting a manuscript to this journal, each author explicitly confirms that the manuscript meets the highest ethical standards for authors and coauthors
The undersigned hereby assign(s) to *Balkan Journal of Electrical & Computer Engineering* (BAJECE) copyright ownership in the above Paper, effective if and when the Paper is accepted for publication by BAJECE and to the extent transferable under applicable national law. This assignment gives BAJECE the right to register copyright to the Paper in its name as claimant and to publish the Paper in any print or electronic medium.

Authors, or their employers in the case of works made for hire, retain the following rights:

1. All proprietary rights other than copyright, including patent rights.
2. The right to make and distribute copies of the Paper for internal purposes.
3. The right to use the material for lecture or classroom purposes.
4. The right to prepare derivative publications based on the Paper, including books or book chapters, journal papers, and magazine articles, provided that publication of a derivative work occurs subsequent to the official date of publication by BAJECE.
5. The right to post an author-prepared version or an of ficial version (preferred version) of the published paper on an i nternal or external server controlled exclusively by t he author/employer, pr ovided that (a) such posting is noncommercial in nature and the paper is made available to users without charge; (b) a copyright notice and full citation appear with the paper, and (c) a link to BAJECE's official online version of the abstract is provided using the DOI (Document Object Identifier) link.

Balkan Journal of Electrical & Computer Engineering
BAJECE
An International Peer Reviewed, Indexed and Open Access Journal

ISSN: 2147- 284X
Year: September 2025
Volume: 13
Issue: 3

CONTENTS

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING
(An International Peer Reviewed, Indexed and Open Access Journal**)**

Contact
Batman University
Department of Electrical-Electronics Engineering
Bati Raman Campus Batman-Turkey

Web: https://dergipark.org.tr/en/pub/bajece
http://www.bajece.com
e-mail: bajece@hotmail.com