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 : 1
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 : 1
Year: March 2025

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
https://www.bajece.com
e-mail: bajece@hotmail.com

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 1, March 2025
Research Article

Ensemble-Based Deep Transfer Learning for Robust Gastrointestinal Endoscopy Image Classification

Sehmus Aslan

Abstract— **Gastrointestinal (GI) diseases remain a significant global health challenge, particularly in low-income settings where diagnostic resources are often scarce. Endoscopic examination is essential for detecting and monitoring these diseases, yet the manual analysis of the resulting images is time-consuming, prone to observer variability, and demanding of clinical expertise. Recent advances in computer-aided diagnosis (CAD) using deep convolutional neural networks (CNNs) have shown promise in automating endoscopic image classification, but limited annotated data and the subtlety of GI findings continue to pose challenges. To address these constraints, this study proposes a two-level stacking ensemble framework that combines three pre-trained CNN architectures—ResNet50, DenseNet201, and MobileNetV3Large—with four classical machine-learning meta-classifiers (Logistic Regression, Random Forest, Support Vector Machine, and K-Nearest Neighbors). The KvasirV2 dataset, comprising 8,000 GI endoscopic images across eight classes, was used to train and evaluate the models. Results indicate that the stacking ensemble achieved a top accuracy of 94.33%, surpassing individual CNN baselines by 1–2%. Notably, this multi-level ensemble approach demonstrated improved diagnostic consistency for challenging classes like early-stage esophagitis and normal Z-line, suggesting that synergizing diverse CNN feature extractors can mitigate the limitations of single-network methods. These findings underscore the potential of ensemble-based transfer learning to enhance clinical decision support, reduce observer variability, and facilitate earlier, more accurate detection of GI diseases.**

Index Terms—: **Ensemble Learning, Transfer learning, Gastrointestinal Endoscopy, Deep Convolutional Neural Networks (CNNs), Computer-Aided Diagnosis (CAD)**

I. INTRODUCTION

GASTROINTESTINAL(GI) DISEASES pose a major global health concern, ranking as the seventh leading cause of death in low-income countries in 2021, according to the World Health Organization (WHO) [1]. The diversity of gastrointestinal disorders, ranging from mild inflammatory conditions to life-threatening cancers, underscores the need for accurate and timely diagnosis to prevent complications and improve patient outcomes.

A primary method for diagnosing GI diseases is endoscopic examination, which involves the use of a flexible tube with an attached camera to visualize the GI tract in real time. While this technique facilitates direct observation of the esophagus, stomach, and intestines, and permits biopsy or treatment during the procedure, it also places substantial demands on clinical resources. Gastroenterologists spend considerable time interpreting large numbers of images or videos, which may result in increased workload and heightened fatigue. Moreover, visual assessment is inherently susceptible to inter- and intra-observer variability, as clinicians may reach different conclusions depending on their expertise, training, or even geographic practice environment.

In response to these diagnostic challenges, computer-aided diagnosis (CAD) systems have gained momentum as valuable clinical support tools. By harnessing the power of artificial intelligence (AI), these systems can potentially standardize diagnostic criteria, detect abnormalities more consistently, and expedite the process of disease identification. Specifically, deep convolutional neural networks (CNNs) have shown promise for classifying endoscopic images, as they autonomously extract relevant features from raw data, thereby reducing the need for manual feature engineering. However, one persistent obstacle in developing robust deep learning models in medicine is the limited availability of annotated data, given that privacy regulations often constrain the sharing of medical images among institutions.

To address this limitation, transfer learning has emerged as a strategic approach. Rather than training a deep neural network from scratch, researchers leverage pre-trained models from large-scale datasets—such as ImageNet—and adapt them to medical image classification tasks [2], [3]. This process not only reduces computational burden but also allows the model to inherit feature representations from millions of natural images, improving performance on smaller and more specialized medical datasets. Overall, the confluence of increased GI disease prevalence, the growing volume of endoscopic data, and advancements in deep learning highlights an urgent need for integrating AI-driven diagnostic tools into clinical practice. By doing so, healthcare systems can potentially detect pathologies earlier, reduce clinician workload, and enhance patient care. A summary of the studies in the literature is as follows:

Sehmus Aslan, is with Department of Business Administration of Mardin Artuklu University, Mardin, Turkey,(e-mail: sehmusaslan@artuklu.edu.tr).
https://orcid.org/0000-0003-1886-3421

Manuscript received Jan 31, 2025; accepted Feb 20, 2025.
DOI: 10.17694/bajece.1630294

Gebreslassie et al. [4] compared DenseNet121 and ResNet50 on a subset of 2,000 images from the Kvasir v2 dataset, employing a split ratio of 0.6, 0.3, and 0.1 for training, testing, and validation, respectively. Their findings indicated that ResNet50 attained the highest accuracy of 87.8%, thereby demonstrating the potential utility of transfer learning for gastrointestinal (GI) endoscopic image classification. Poudel et al. [5] employed ResNet50 with scaled-dilation convolutions to classify 4,000 endoscopic images into eight categories, achieving an accuracy of 88%. Their methodology incorporated a batch size of 16, a learning rate of 0.001, and stochastic gradient descent, illustrating the importance of meticulous hyperparameter selection to mitigate overfitting in convolutional neural networks (CNNs). Lonseko et al. [6] adopted an attention-guided CNN incorporating spatial attention and encoder-decoder layers on the Kvasir dataset, achieving 93.19% accuracy and an F1 score of 92.8%. They addressed data imbalance via augmentation strategies, underscoring the relevance of data diversity in model training. Musha et al. [7] investigated 16 pre-trained models, including MobileNetV2, on 2,000 Kvasir v2 images focused on dyed lifted polyps and resection margins. MobileNetV2 performed best, reaching an accuracy of 82.25% under a learning rate of 0.001 with Adam. Auzine et al. [8] explored InceptionV3, InceptionResNetV2, and VGG16 on 9,852 images from the Endoscopic Artifact Detection and Kvasir v2 datasets, reporting 77.65% accuracy with InceptionV3. In a similar vein, Gupta et al. [9] proposed a hybrid architecture combining EfficientNetB7 and ResNet50 to classify 8,000 Kvasir v2 images, achieving 88.19% accuracy. These studies collectively highlighted the continued success of transfer learning in GI endoscopic tasks and the inherent challenges of avoiding overfitting.

Mukhtorov et al. [10] examined DenseNet201, MobileNetV2, ResNet18, ResNet152, and VGG16 on 8,000 wireless endoscopic images, identifying overfitting in ResNet152, with a training accuracy of 98.28% versus a validation accuracy of 93.46%. Gunasekaran et al. [2] reported analogous issues using an ensemble of DenseNet201, InceptionV3, and ResNet50 obtaining 95% accuracy but with diminished generalization on validation data. Demirbaş et al. [11] developed a Spatial-Attention ConvMixer (SAC) architecture, surpassing Vanilla ViT, Swin Transformer, and the baseline ConvMixer on the Kvasir dataset, with a final accuracy of 93.37%. This study demonstrated the efficacy of spatial attention mechanisms in enhancing classification performance. In parallel, Ayan [12] investigated the classification of gastrointestinal diseases using thirteen different CNN models and two different ViT architectures. The authors observed that while ViT models reached accuracies of 91.25% and 90.50%, a well-optimized DenseNet201 variant, leveraging optimized transfer learning parameters, recorded an accuracy of 93.13%, a recall of 93.17%, a precision of 93.13%, and an F1 score of 93.11%, thereby outperforming both ViT models. Similarly, Huo et al. [13] proposed Self-Peripheral-Attention (SPA), inspired by human peripheral vision, to improve classification and segmentation on Kvasir and Kvasir-SEG datasets, attaining an accuracy of 92.7%.

If the literature is examined, it can be seen that despite the demonstrated success of transfer learning and deep CNNs in classifying GI endoscopic images, several notable gaps remain unaddressed. First, most existing studies rely on either a single deep network or straightforward ensemble averaging, without systematically exploiting more advanced multi-level ensemble frameworks. As a result, valuable complementary features learned by different architectures may not be fully leveraged, especially for visually subtle classes such as early-stage esophagitis and the Z-line. Second, while overfitting and limited annotated data are frequently acknowledged challenges in GI image analysis, there is comparatively little research into robust strategies—beyond basic augmentation—for mitigating these issues across diverse endoscopic conditions. Finally, few works provide a detailed examination of how stacking ensembles with classical machine-learning meta-classifiers can improve diagnostic consistency and reduce the variance inherent in individual CNNs. Addressing these gaps could lead to more precise classification performance, particularly in clinically challenging contexts where subtle tissue changes are critical for early diagnosis.

This study contributes to the gastrointestinal (GI) endoscopy classification literature in several key ways. First, it proposes a two-level stacking ensemble approach, systematically combining multiple state-of-the-art CNN architectures (ResNet50, DenseNet201, MobileNetV3Large) with classical meta-classifiers (Logistic Regression, Random Forest, SVM, KNN). By moving beyond single-network solutions and basic ensemble averaging, the method fully exploits complementary learned features, which is particularly important for handling the subtle visual distinctions in GI images such as early-stage esophagitis and normal Z-line. Second, the detailed comparison of base CNNs against multiple stacking ensembles offers new insights into how meta-classifiers can improve diagnostic consistency, bridging an existing gap in the literature on advanced ensemble frameworks for GI endoscopic image analysis. By presenting robust evidence that such multi-level ensembles outperform individual CNNs, the study sets a foundation for future research aimed at achieving more accurate and clinically relevant GI disease detection systems.

II. MATERIALS AND METHODS

A. Dataset

This study employs the Kvasir-V2 dataset [14], a collection of gastroenterological endoscopic images gathered by a Norwegian healthcare organization, designed to facilitate research in medical image analysis. The Kvasir-v2 dataset is a comprehensive collection of 8,000 gastrointestinal tract endoscopic images, released in 2017 through the MediaEval Medical Multimedia Challenge. The dataset contains eight balanced classes with 1,000 images each, all annotated and verified by certified endoscopists. These classes are divided into three main categories: anatomical landmarks (pylorus, z-line, and cecum), pathological findings (esophagitis, ulcerative colitis, and polyps), and medical procedures (dyed lifted polyps and dyed resection margins). The classes and image examples used for the application are presented in Fig. 1. The images vary in resolution from 720×576 pixels to 1920×1072 pixels, with some containing annotations in the leftmost quarter and green boxes indicating endoscope location. Each image varies in capture angle, brightness, zoom level, resolution, and centerpoint, making it a challenging dataset for deep learning applications. Despite its relatively small size compared to standard deep learning datasets, Kvasir-v2 has become a crucial benchmark dataset for evaluating machine learning approaches in gastrointestinal image analysis, particularly for testing classification accuracy, developing computer-aided diagnosis tools, and assessing model generalization capabilities. The dataset's standardized format and public availability through Kaggle [15] make it particularly valuable for research in automated gastrointestinal disease detection and medical image classification systems. By applying a random sampling strategy, each category was partitioned into training and test sets at a 70:30 ratio, yielding 5,600 images for training and 2,400 images for testing.

To enhance the robustness and generalization capability of the model, a data augmentation pipeline was implemented, applying a series of random transformations to the input images during training. Each image undergoes random rotations within the range $[-15°, 15°]$ and discrete 90-degree increments, along with adjustments to brightness, contrast, saturation, and hue to introduce variability in illumination and color. Small translations are simulated by padding the image by 20 pixels and cropping it back to its original dimensions, while additive Gaussian noise with a mean of 0.0 and a standard deviation of 0.1 is introduced to mimic real-world imperfections. These augmentations are applied dynamically during training, ensuring that the model is exposed to a diverse range of variations, thereby improving its ability to generalize and reducing the risk of overfitting. Originally, there were 5,600 training instances; after applying these augmentations at a rate of 5×, the total number of augmented samples increases to 28,000. The test dataset remains unchanged at 2,400 samples. All images were resized at 224×224x3 pixels..

B. *Transfer Learning*

Transfer learning is highly beneficial in medical imaging tasks primarily because acquiring large, well-annotated datasets in clinical settings can be challenging due to patient privacy concerns, labeling costs, and the specialized expertise needed for annotation [16]. By leveraging models pre-trained on extensive and diverse non-medical image datasets—such as ImageNet—researchers can repurpose learned features (e.g., edges, textures) and adapt them to medical contexts. This process not only saves significant computation time and resources but also mitigates the risk of overfitting when working with relatively small medical datasets [17]. According to Litjens et al. [18], transfer learning facilitates faster convergence and can enhance classification or detection accuracy in a wide range of medical imaging applications, from lesion identification to organ segmentation. Numerous CNN architectures have been introduced in the literature. In this work, three pre-trained CNN models (ResNet50, DenseNet201 and MobileNetV3Large) are employed to categorize endoscopic images into eight distinct classes.

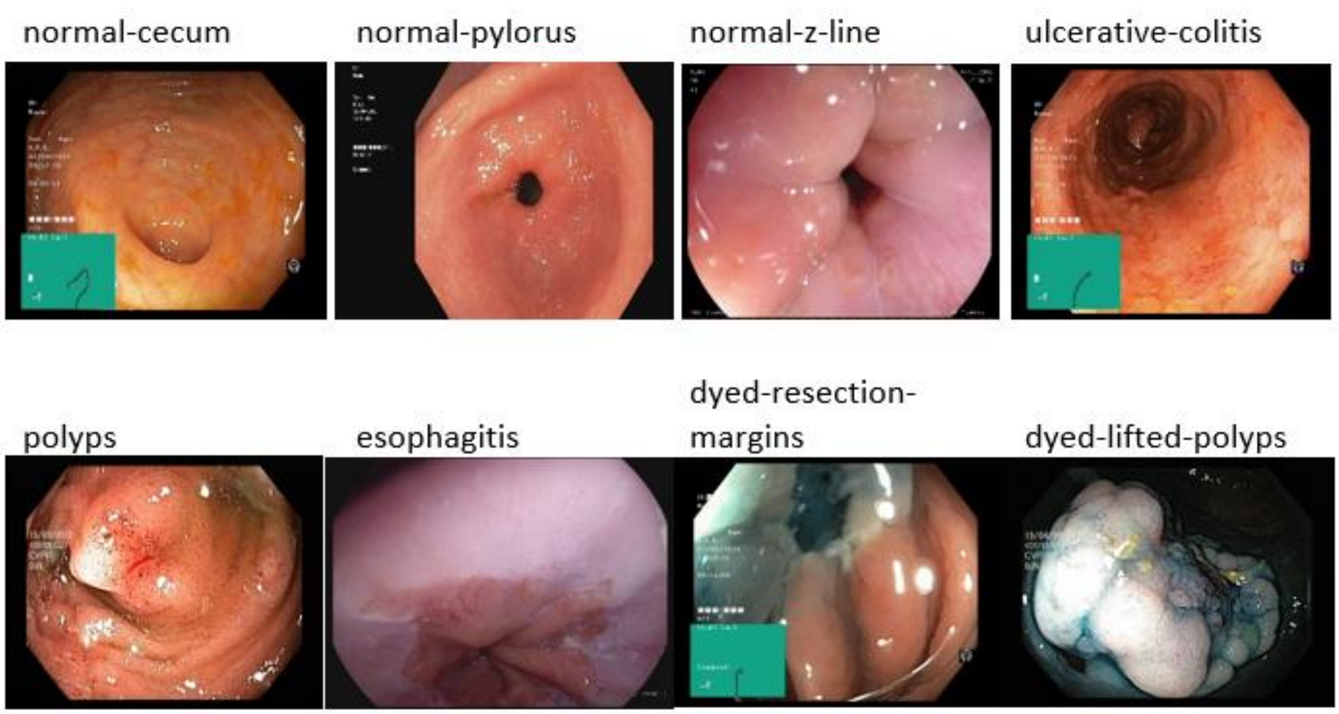


Fig.1. KvasirV2 classes

1. ResNet50

ResNet50 is a 50-layer convolutional neural network introduced by He et al. [19] as part of the ResNet (Residual Network) family, which was primarily designed to overcome the vanishing gradient problem in deeper neural networks. By incorporating residual blocks with identity connections (or skip connections), the architecture allows gradients to flow more effectively, facilitating the training of much deeper networks than earlier CNN models like VGGNet or AlexNet. Owing to its balance between depth and computational efficiency, ResNet50 has become a widely adopted backbone for various computer vision tasks, including image classification, object detection, and segmentation [20]. In medical image analysis, the model's pre-trained weights on large-scale datasets (e.g., ImageNet) have proven highly beneficial when performing transfer learning with limited labeled data, enhancing accuracy and accelerating convergence [18].

2. DenseNet201

DenseNet201 is a 201-layer convolutional neural network that is part of the Densely Connected Convolutional Network (DenseNet) family introduced by Huang et al. [20]. Unlike traditional CNNs, DenseNet layers are directly connected to every other layer in a feed-forward manner, allowing each layer to access the gradients from all preceding layers. This dense connectivity pattern mitigates the vanishing gradient problem and promotes feature reuse, enabling the construction of deeper and more efficient networks without a substantial increase in computational complexity [20]. DenseNet201 has demonstrated strong performance on large-scale image classification tasks such as ImageNet. In medical imaging, transferring pre-trained

DenseNet201 weights to specialized domains has shown to accelerate model convergence and enhance diagnostic accuracy, especially when the available datasets are relatively small [18]. Its depth and ability to capture complex hierarchical features make DenseNet201 particularly suitable for tasks like disease classification, segmentation, and detection, where subtle variations in medical images are critical for accurate predictions.

3. MobileNetV3Large

MobileNetV3Large is a lightweight convolutional neural network introduced as part of the MobileNetV3 family by Howard et al. [21]. It was designed through a combination of neural architecture search (NAS) and platform-aware model optimization (NetAdapt), balancing high accuracy with reduced computational complexity. Key features include the use of squeeze-and-excitation (SE) modules, novel activation functions such as the h-swish, and inverted residual blocks that improve both efficiency and representational power [21]. While originally optimized for resource-constrained devices (e.g., smartphones), MobileNetV3Large has also shown promise in medical imaging contexts, particularly when transferring pre-trained weights to smaller medical datasets for tasks like classification and segmentation [18]. Its efficient design enables faster inference and lower resource usage, which are critical factors for real-time, point-of-care diagnostics.

C. Proposed Stacking Ensemble Model

The proposed framework leverages transfer learning to construct a robust and scalable deep learning model for multi-class classification. Transfer learning is employed to utilize the feature extraction capabilities of pre-trained convolutional neural networks (CNNs), which have been trained on the large-scale ImageNet dataset. This approach not only reduces the computational cost of training from scratch but also enhances the model's ability to generalize to new datasets, particularly when labeled data is limited.

Three state-of-the-art CNN architectures are explored as backbone models: ResNet50, DenseNet201, and MobileNetV3Large. These architectures are chosen for their proven performance in various computer vision tasks, with each offering unique advantages:

1. *ResNet50:* Known for its residual learning framework, which mitigates the vanishing gradient problem and enables the training of very deep networks.
2. *DenseNet201:* Utilizes dense connections between layers, promoting feature reuse and improving parameter efficiency.
3. *MobileNetV3Large:* Designed for efficiency, this architecture is optimized for mobile and edge devices, offering a balance between accuracy and computational cost.

Each backbone model is initialized with pre-trained weights from ImageNet and configured to exclude the fully connected classification head. This allows the model to retain only the feature extraction layers, which are then adapted to the specific task at hand. Global average pooling is applied to the output feature maps to reduce spatial dimensions and produce a fixed-size feature vector. This is followed by a Flatten layer to convert the pooled features into a one-dimensional vector, which is then passed to a Dense layer with eight output units and a softmax activation function. This final layer enables multi-class classification into eight distinct categories.

All layers of the base models are set to be trainable, allowing for fine-tuning of the pre-trained weights during training. This ensures that the model can adapt to the specific characteristics of the target dataset while retaining the generalizable features learned from ImageNet. The models are optimized using Stochastic Gradient Descent (SGD) with a learning rate of 0.01, momentum of 0.9, and Nesterov acceleration. The loss function is defined as categorical cross-entropy, which is well-suited for multi-class classification tasks. Model performance is evaluated using accuracy as the primary metric.

To further enhance classification performance and robustness, a stacking ensemble approach is employed. Stacking combines the predictions of multiple base models (in this case, the three deep learning models) using a meta-classifier, which learns to optimally weigh and combine the predictions. This approach leverages the strengths of diverse models, reducing the risk of overfitting and improving generalization.

The predictions from the three deep learning models are concatenated to form a combined feature representation, which serves as input to the meta-classifier. Four distinct meta-

 ISSN: 2147-284X

classifiers are implemented and evaluated, each chosen for its unique characteristics and suitability for the task:

1 *Logistic Regression (LR):* A linear classifier with L2 regularization (C=10) and the newton-cg solver, configured for one-vs-rest multi-class classification. Logistic regression is chosen for its interpretability and efficiency in handling linearly separable data.
2 *Random Forest (RF):* An ensemble of 300 decision trees with Gini impurity as the splitting criterion, a maximum depth of 20, and balanced class weights to handle class imbalance. Random forests are robust to overfitting and capable of capturing complex, non-linear relationships in the data.
3 *Support Vector Machine (SVM):* A linear SVM with a regularization parameter (C=0.1) and a maximum of 100 iterations, configured to output probability estimates. SVMs are known for their ability to find optimal decision boundaries in high-dimensional spaces.
4 *k-Nearest Neighbors (k-NN):* A non-parametric classifier with three neighbors, Manhattan distance as the metric, and uniform weighting. k-NN is simple yet effective, particularly for datasets with well-defined clusters.

Each meta-classifier is preceded by a StandardScaler to normalize the input features, ensuring consistent scaling across the concatenated predictions. This preprocessing step is critical for algorithms like SVM and k-NN, which are sensitive to the scale of input features.

TABLE I
HYPERPARAMETERS FOR DEEP LEARNING MODELS

Hyperparameter	Value/Configuration
Backbone Architectures	ResNet50, DenseNet201, MobileNetV3Large
Pre-trained Weights	ImageNet
Include Top	False (exclude fully connected layers)
Pooling	Global Average Pooling
Classifier Activation	Softmax
Trainable Layers	All layers trainable
Optimizer	Stochastic Gradient Descent (SGD)
Learning Rate	0.01
Momentum	0.9
Nesterov Acceleration	Enabled
Loss Function	Categorical Crossentropy
Metrics	Accuracy
Epochs	30
Batch Size	32
Early Stopping	Patience = 10 (monitor validation loss)
Learning Rate Scheduler	ReduceLROnPlateau (factor = 0.2, patience = 5, min_lr = 1e-5)
Model Checkpoint	Save best weights based on validation accuracy

The meta-classifiers are trained on the combined predictions from the deep learning models and evaluated using standard classification metrics, including accuracy, precision, recall and F1 score. These metrics provide a comprehensive assessment of model performance. The hyperparameters for the deep learning models are provided in Table 1, while those for the meta-classifiers are detailed in Table 2.

TABLE II
HYPERPARAMETERS FOR META-CLASSIFIERS

Meta-Classifiers	Hyperparameter	Value/Configuration
Logistic Regression	Regularization (C)	10
	Solver	newton-cg
	Max Iterations	100
	Multi-Class Strategy	One-vs-Rest (OvR)
Random Forest	Number of Estimators	300
	Criterion	Gini Impurity
	Max Depth	20
	Max Features	sqrt
	Min Samples Split	2
	Min Samples Leaf	1
	Bootstrap	True
	Class Weight	Balanced
Support Vector Machine (SVM)	Regularization (C)	0.1
	Kernel	Linear
	Max Iterations	100
	Probability Estimates	Enabled
k-Nearest Neighbors (k-NN)	Number of Neighbors (k)	3
	Distance Metric	Manhattan (p=1)
	Weights	Uniform

III. RESULTS

A. Evaluation Metrics

The proposed ensemble model is evaluated using key performance metrics, including *accuracy*, *precision*, *recall*, and *F1 score*. In multiclass classification, where the number of classes exceeds two, the predictions generated by the model can be either correct or incorrect for each class. To evaluate the model's performance, the predictions are analyzed based on the following classification states for each class:

- *True Positive (TP):* The model correctly predicts the class of interest.
- *True Negative (TN):* The model correctly identifies instances that do not belong to the class of interest.
- *False Positive (FP):* The model incorrectly predicts an instance as belonging to the class of interest.
- *False Negative (FN):* The model fails to identify an instance that belongs to the class of interest.

For multiclass classification, these metrics are typically computed using a one-vs-rest approach, where each class is evaluated against the rest of the classes. The formulas for the evaluation metrics are defined as follows:

$$Accuracy = \frac{\sum_{i=1}^{C}(\mathrm{TP}_i + \mathrm{TN}_i)}{\sum_{i=1}^{C}(\mathrm{TP}_i + \mathrm{TN}_i + \mathrm{FP}_i + \mathrm{FN}_i)} \tag{1}$$

$$Precision = \frac{\sum_{i=1}^{C} \mathrm{TP}_i}{\sum_{i=1}^{C} (\mathrm{TP}_i + \mathrm{FP}_i)} \quad (2)$$

$$Recall = \frac{\sum_{i=1}^{C} \mathrm{TP}_i}{\sum_{i=1}^{C} (\mathrm{TP}_i + \mathrm{FN}_i)} \quad (3)$$

$$F1\ Score = \frac{2 \times \mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \quad (4)$$

Here, C represents the total number of classes, and the metrics are aggregated across all classes. *Accuracy* (1) is a metric that measures the overall correctness of the model's predictions across all classes. It is calculated as the ratio of the sum of true positives (TP) and true negatives (TN) for all classes to the total number of instances, including true positives, true negatives, false positives (FP), and false negatives (FN). *Precision* (2) quantifies the proportion of predicted instances for a class that are actually correct. *Recall* (3) measures the proportion of actual instances of a class that the model correctly identifies. The *F1 score* (4) provides a balanced measure by computing the harmonic mean of precision and recall, ensuring that both metrics are equally weighted in the evaluation.

B. Experimental Setup

The experiment was conducted in Google Colab Pro with Python 3 Google Compute Engine backend (GPU-A100) with 40 GB GPU RAM. The deep learning models are trained on the KvasirV2 dataset for a maximum of 30 epochs, with early stopping implemented to prevent overfitting. Training is monitored using validation loss, and the learning rate is dynamically adjusted using the *ReduceLROnPlateau* callback, which reduces the learning rate by a factor of 0.2 if the validation loss does not improve for five consecutive epochs. The best model weights are saved based on validation accuracy using the *ModelCheckpoint* callback.

The meta-classifiers are trained on the concatenated predictions from the deep learning models, ensuring that they learn to effectively combine the strengths of each base model. The performance of the meta-classifiers is evaluated on a held-out test set derived from the KvasirV2 dataset, with results visualized using confusion matrices and summarized using classification metrics.

C. Test Results

Fig. 2, 3, 4, 5, 6 and 7 present the accuracy and loss curves of the base models. The curves demonstrate that each architecture (ResNet, DenseNet, MobileNet) learns the dataset effectively, reaching high overall performance. The differences among models primarily manifest in how smoothly the validation accuracy evolves and how tightly the validation loss tracks the training loss. While the current results already achieve strong classification performance, the remaining gap between training and validation highlights a potential avenue for fine-tuning regularization or data augmentation strategies to bolster robustness further.

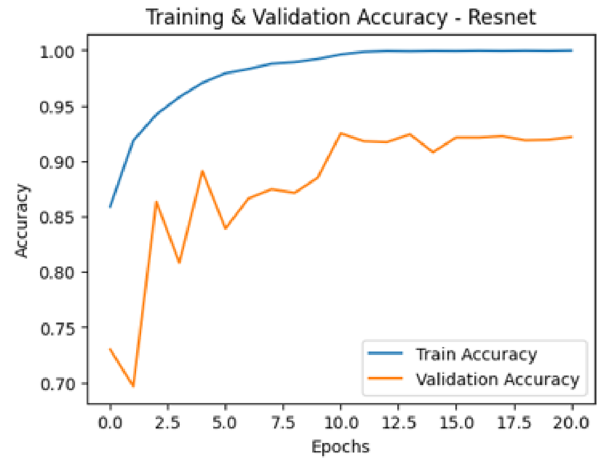


Fig.2. ResNet50 accuracy curve.

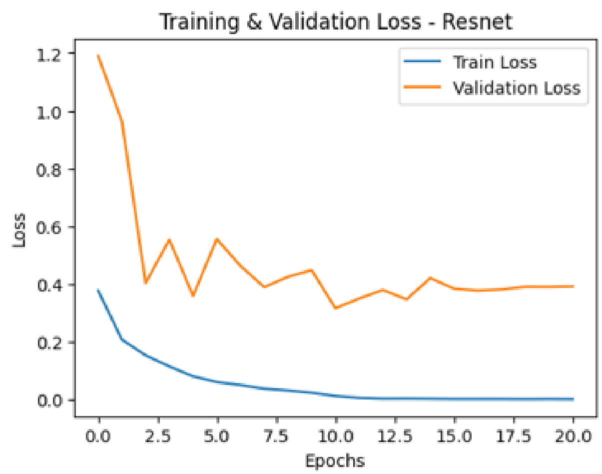


Fig.3. ResNet50 loss curve.

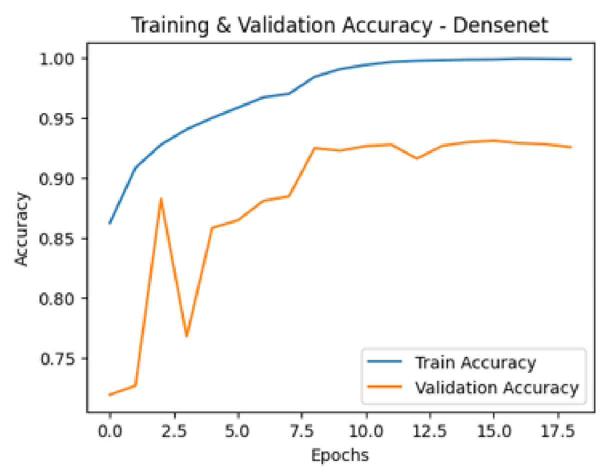


Fig.4 DenseNet201 accuracy curve.

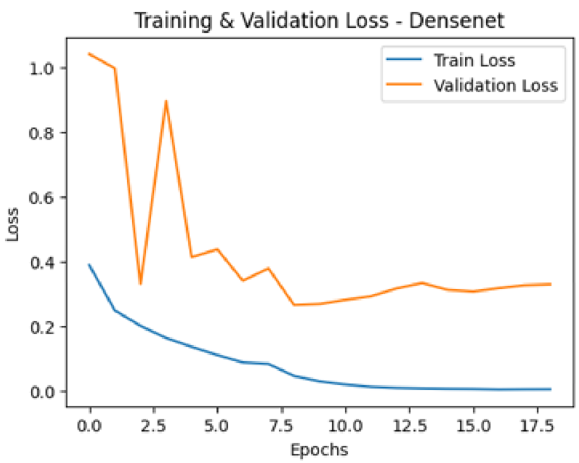


Fig.5 DenseNet201 loss curve.

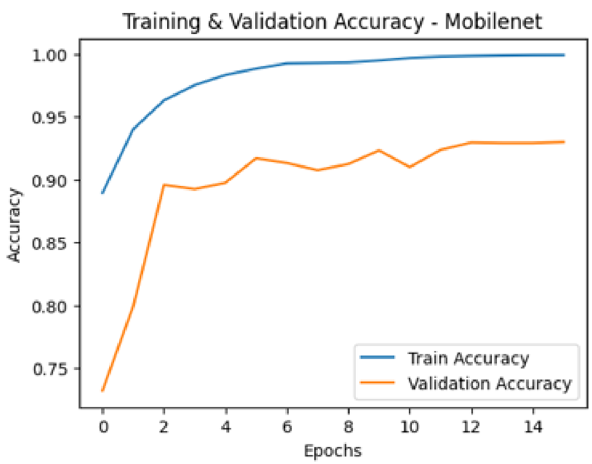


Fig.6 MobileNetV3Large accuracy curve.

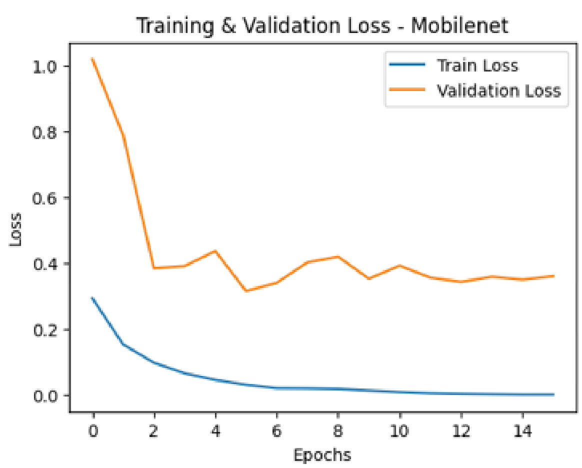


Fig.7 MobileNetV3Large loss curve.

Table III provides a comparison of the models' performances. Among the three base CNNs, MobileNetV3Large achieves the highest accuracy (93.00%), slightly outperforming ResNet50 (92.17%) and DenseNet201 (92.54%). This finding suggests that MobileNetV3Large, which balances depth and efficiency through its inverted residual blocks and attention mechanisms, adapts effectively to the Kvasir dataset. Logistic Regression, Random Forest, and SVM ensembles each hover around 94.3% accuracy, while KNN trails only slightly at 94.17%. Precision, recall, and F1 scores follow a similarly tight range. These minimal differences may reflect the relatively uniform effectiveness of combining three high-performing CNN backbones; once robust feature representations are available, multiple classical classifiers can exploit them effectively.

All stacking ensembles outperform their single CNN counterparts, with the best results reaching 94.33% accuracy (Stacking Ensemble with Logistic Regression or Random Forest). In other words, ensembling the probability outputs from ResNet50, DenseNet201, and MobileNetV3Large typically yields a 1–2% improvement in accuracy, precision, recall, and F1 score. From a healthcare perspective, slight performance gains can be critical, as more reliable diagnoses translate to fewer missed pathologies and better patient outcomes. This is particularly true in endoscopic procedures, where subtle changes can be indicative of early disease progression.

TABLE III
PERFORMANCE EVALUATION OF DEEP LEARNING MODELS ON KVASIR-V2 DATASET.

Models	ACC(%)	Precision(%)	Recall(%)	F_1 Score(%)
ResNet50	92.17	92.06	92.07	92.03
DenseNet201	92.54	92.46	92.46	92.44
MobileNetV3Large	93.00	92.99	92.94	92.95
Stacking Ensemble with SVM	94.29	94.29	94.23	94.22
Stacking Ensemble with LR	94.33	94.33	94.27	94.25
Stacking Ensemble with RF	94.33	94.36	94.27	94.27
Stacking Ensemble with KNN	94.17	94.17	94.10	94.10

Fig. 8, 9, 10 and 11 show the confusion matrices of the stacking ensemble models. These matrices reveal that the stacking ensemble models can identify all classes with high performance, except for the Z-line and esophagitis classes. In many endoscopic images, the visual distinctions between a normal Z-line and mild esophagitis are very subtle, often manifesting as slight color changes or faint lesions. As a result, even advanced CNNs may struggle to reliably differentiate these two classes. The similarity between a normal Z-line and early-stage esophagitis—characterized by minor discoloration or subtle shifts in tissue texture—makes these distinctions less prominent compared to other classes (e.g., polyps or ulcers). Occasional misclassifications persist, indicating the presence of a few particularly challenging cases. Nevertheless, nearly all classes achieve high recall and precision, underscoring the overall robustness of the ensemble approach.

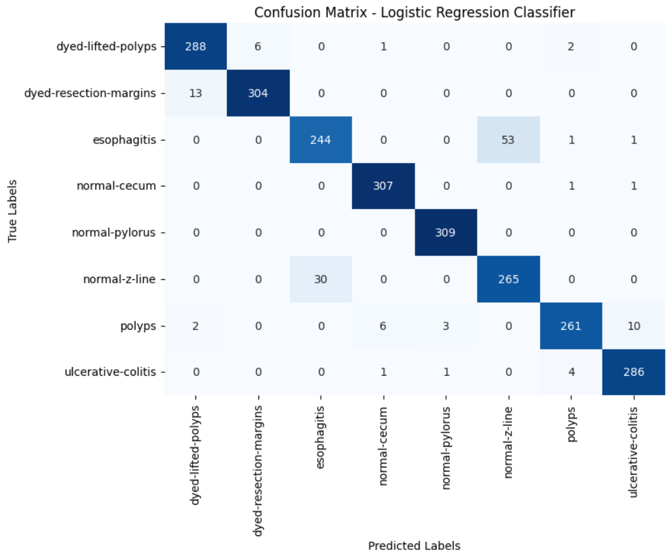


Fig.8 Confusion matrix of LR meta-classifier.

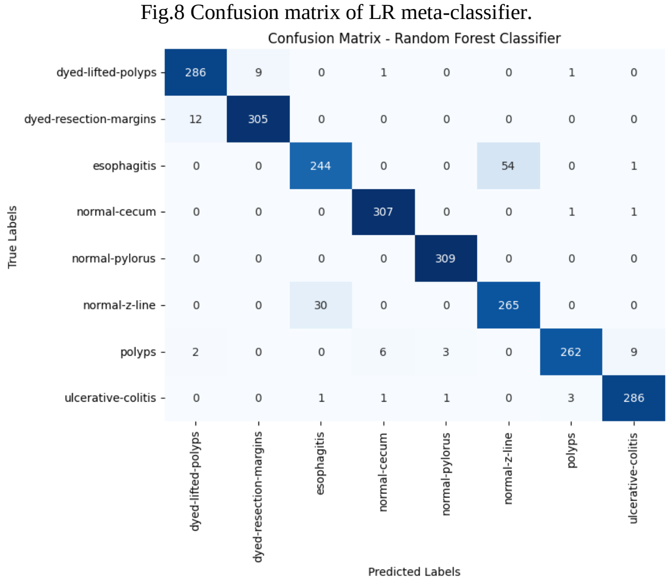


Fig.9 Confusion matrix of RF meta-classifier.

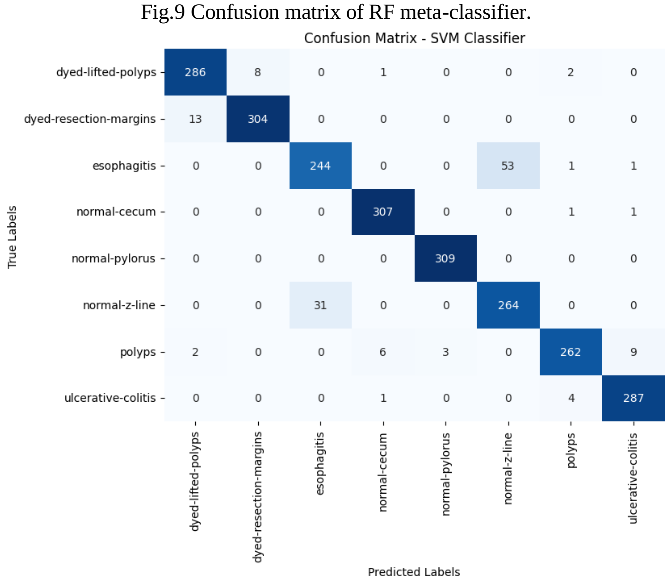


Fig.10 Confusion matrix of SVM meta-classifier.

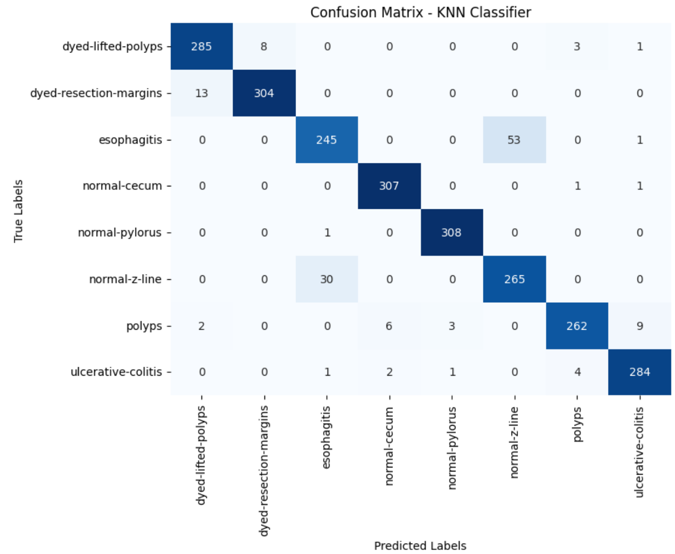


Fig.11 Confusion matrix of KNN meta-classifier.

Table IV compares performance metrics (Accuracy, Precision, Recall, F1-Score) across studies using the Kvasir dataset. Yogapriya et al. [22] achieved the highest accuracy (96.33%) and F1-score (96.50%), setting a strong benchmark. This study ranks second with competitive metrics (Accuracy: 94.33%, F1-Score: 94.27%), showing balanced performance across precision and recall. Other studies like Losenko et al. [6], Demirbaş et al. [11], and Huo et al. [13] also report strong results, while Gupta et al. [9] lags with metrics below 90%. Some studies, such as Mukhtorov et al. [10] and Gunasekaran et al. [2], only report accuracy, limiting comprehensive comparison. Overall, this study demonstrates robust performance, though Yogapriya et al. [22] remains the top performer. The findings highlight the promise of ensemble-based deep learning strategies and underscore the field-wide progress toward robust, clinically relevant models for gastrointestinal disease detection and classification.

TABLE IV
COMPARISON OF PROPOSED MODEL WITH OTHER RECENT MODELS

Study	Accuracy	Precision	Recall	F1-Score
Yogapriya et al [22]	96.33	96.50	96.37	96.50
Losenko et al. [6]	93.19	92.8	92.7	92.8
Gupta et al. [9]	89.3	89	89.3	88.6
Mukhtorov et al. [10]	93.46	-	-	-
Gunasekaran et al. [2]	95.00	-	-	-
Huo et al. [13]	92.87	93.01	92.87	92.88
Demirbaş et al. [11]	93.37	93.66	93.37	93.42
Ayan [12]	93.13	93.17	93.13	93.11
This study	94.33	94.36	94.27	94.27

IV. Discussions

The findings of this study reinforce the value of leveraging transfer learning and ensemble techniques for robust endoscopic image classification. As evidenced by the strong performance of individual CNN backbones (ResNet50, DenseNet201, MobileNetV3Large), pre-trained models offer a reliable starting point when working with relatively small yet challenging medical datasets such as Kvasir v2. The slight variations in baseline performance among these networks likely stem from differences in architecture design—ranging from ResNet's residual connections to DenseNet's dense connectivity and MobileNetV3's efficient inverted residual blocks—each of which provides unique advantages for feature extraction in endoscopic images.

Despite these variations, the proposed stacking framework demonstrates consistent improvements across accuracy, precision, recall, and F1 score. Such gains underscore the ensemble's ability to reconcile the complementary strengths of different CNNs. By uniting multiple feature representations at the meta-classifier level, the method mitigates the variance inherent in individual models and achieves a more robust overall performance. From a clinical perspective, even marginally higher metrics (1–2% above single-model baselines) can be particularly valuable in reducing missed pathologies, given the high-stakes nature of GI disease diagnosis.

The confusion matrices, however, highlight a recurrent challenge in distinguishing subtle classes like early-stage esophagitis versus a normal Z-line. This difficulty points to the inherent complexity of GI endoscopy images, where slight color shifts or minor morphological differences can be easily overlooked. Addressing this gap may require additional strategies, such as more targeted data augmentation, higher-resolution inputs, or region-of-interest (ROI) detection methods that emphasize the gastroesophageal junction. Likewise, incorporating advanced attention mechanisms or domain adaptation techniques may further refine the model's ability to capture faint textural changes indicative of mild esophagitis.

In a broader sense, the results align with existing literature that showcases the benefits of transfer learning in medical imaging, particularly when annotated data are scarce [16], [18]. Pre-trained weights allow the network to capitalize on foundational visual features, reducing the risk of overfitting and expediting convergence. Ensemble approaches, in turn, harness these strengths in a synergistic manner, as documented by related research that reports similar performance boosts when combining models [2].

Overall, this study's findings emphasize two key takeaways: first, that combining multiple CNN architectures through a stacking ensemble is effective in boosting classification metrics on the Kvasir dataset; and second, that additional focus on nuanced, easily confounded classes remains a priority for future work. By refining the proposed framework with enhanced data handling, attention modules, and specialized augmentation, researchers and clinicians can continue to push the boundaries of AI-driven GI diagnostics, ultimately contributing to earlier and more accurate detection of critical gastrointestinal conditions.

V. Conclusion

This work demonstrates that ensemble-based transfer learning can significantly improve the classification of gastrointestinal endoscopic images, addressing both the scarcity of annotated data and the inherent complexity of subtle GI conditions. By combining ResNet50, DenseNet201, and MobileNetV3Large as base architectures and employing a second-level meta-classifier, we achieved higher accuracy, precision, recall, and F1 scores compared to single-network models. These findings underscore the synergy that arises when leveraging diverse CNN features in a stacking framework.

Moreover, the results highlight the practical benefits of enhanced diagnostic accuracy for conditions such as esophagitis and normal Z-line, where visual differences are often minimal. Although occasional misclassifications occur in these subtle classes, the overall performance points to the promise of refined augmentations, region-of-interest approaches, and advanced attention mechanisms in bridging the remaining performance gap.

From a clinical standpoint, the observed improvements in detection rates and classification reliability translate into potentially earlier interventions and reduced workload for gastroenterologists. Future research directions may focus on integrating larger, multi-center datasets, exploring novel attention modules, and automating the identification of key anatomical landmarks. By continuing to refine ensemble strategies and transfer learning pipelines, the field can move closer to real-time, AI-driven diagnostic support that is both efficient and clinically robust.

References

[1] WHO, "The top 10 causes of death." Accessed: Jan. 30, 2025. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

[2] H. Gunasekaran, K. Ramalakshmi, D. K. Swaminathan, A. J, and M. Mazzara, "GIT-Net: An Ensemble Deep Learning-Based GI Tract Classification of Endoscopic Images," *Bioengineering*, vol. 10, no. 7, p. 809, Jul. 2023, doi: 10.3390/bioengineering10070809.

[3] S. Mohapatra, J. Nayak, M. Mishra, G. K. Pati, B. Naik, and T. Swarnkar, "Wavelet Transform and Deep Convolutional Neural Network-Based Smart Healthcare System for Gastrointestinal Disease Detection," *Interdiscip. Sci. Comput. Life Sci.*, vol. 13, no. 2, pp. 212–228, Jun. 2021, doi: 10.1007/s12539-021-00417-8.

[4] A. KahsayGebreslassie, YaecobGirmayGezahegn, M. T. Hagos, AchimIbenthal, and Pooja, "Automated Gastrointestinal Disease Recognition for Endoscopic Images," in *2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, Greater Noida, India: IEEE, Oct. 2019, pp. 312–316. doi: 10.1109/ICCCIS48478.2019.8974458.

[5] S. Poudel, Y. J. Kim, D. M. Vo, and S.-W. Lee, "Colorectal Disease Classification Using Efficiently Scaled Dilation in Convolutional Neural Network," *IEEE Access*, vol. 8, pp. 99227–99238, 2020, doi: 10.1109/ACCESS.2020.2996770.

[6] Z. M. Lonseko *et al.*, "Gastrointestinal Disease Classification in Endoscopic Images Using Attention-Guided Convolutional Neural Networks," *Appl. Sci.*, vol. 11, no. 23, p. 11136, Nov. 2021, doi: 10.3390/app112311136.

[7] A. Musha, R. Hasnat, A. A. Mamun, and T. Ghosh, "Deep Learning-Based Comparative Study to Detect Polyp Removal in Endoscopic Images," in *2022 International Conference on Emerging Smart Computing and Informatics (ESCI)*, Pune, India: IEEE, Mar. 2022, pp. 1–5. doi: 10.1109/ESCI53509.2022.9758254.

[8] M. M. Auzine, P. Bissoonauth-Daiboo, M. H.-M. Khan, S. Baichoo, X. Gao, and N. G. Sahib, "Classification of artefacts in endoscopic images using deep neural network," in *2022 3rd International Conference on*

Next Generation Computing Applications (NextComp), Flic-en-Flac, Mauritius: IEEE, Oct. 2022, pp. 1–5. doi: 10.1109/NextComp55567.2022.9932202.

[9] D. Gupta, G. Anand, P. Kirar, and P. Meel, "Classification of Endoscopic Images and Identification of Gastrointestinal diseases," in *2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON)*, Faridabad, India: IEEE, May 2022, pp. 231–235. doi: 10.1109/COM-IT-CON54601.2022.9850571.

[10] D. Mukhtorov, M. Rakhmonova, S. Muksimova, and Y.-I. Cho, "Endoscopic Image Classification Based on Explainable Deep Learning," *Sensors*, vol. 23, no. 6, p. 3176, Mar. 2023, doi: 10.3390/s23063176.

[11] A. A. Demirbaş, H. Üzen, and H. Fırat, "Spatial-attention ConvMixer architecture for classification and detection of gastrointestinal diseases using the Kvasir dataset," *Health Inf. Sci. Syst.*, vol. 12, no. 1, p. 32, Apr. 2024, doi: 10.1007/s13755-024-00290-x.

[12] E. Ayan, "Classification of Gastrointestinal Diseases in Endoscopic Images: Comparative Analysis of Convolutional Neural Networks and Vision Transformers," *Iğdır Üniversitesi Fen Bilim. Enstitüsü Derg.*, vol. 14, no. 3, pp. 988–999, Sep. 2024, doi: 10.21597/jist.1501787.

[13] X. Huo, S. Tian, Y. Yang, L. Yu, W. Zhang, and A. Li, "SPA: Self-Peripheral-Attention for central–peripheral interactions in endoscopic image classification and segmentation," *Expert Syst. Appl.*, vol. 245, p. 123053, Jul. 2024, doi: 10.1016/j.eswa.2023.123053.

[14] K. Pogorelov *et al.*, "KVASIR: A Multi-Class Image Dataset for Computer Aided Gastrointestinal Disease Detection." Association for Computing Machinery, Haziran 2017. doi: 10.1145/3193289.

[15] Kaggle, "Kvasir v2." Accessed: Jan. 30, 2025. [Online]. Available: https://www.kaggle.com/datasets/plhalvorsen/kvasir-v2-a-gastrointestinal-tract-dataset

[16] N. Tajbakhsh *et al.*, "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1299–1312, May 2016, doi: 10.1109/TMI.2016.2535302.

[17] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010, doi: 10.1109/TKDE.2009.191.

[18] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.

[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.

[20] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI: IEEE, Jul. 2017, pp. 2261–2269. doi: 10.1109/CVPR.2017.243.

[21] A. Howard *et al.*, "Searching for MobileNetV3," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, Korea (South): IEEE, Oct. 2019, pp. 1314–1324. doi: 10.1109/ICCV.2019.00140.

[22] J. Yogapriya, V. Chandran, M. G. Sumithra, P. Anitha, P. Jenopaul, and C. Suresh Gnana Dhas, "Gastrointestinal Tract Disease Classification from Wireless Endoscopy Images Using Pretrained Deep Learning Model," *Comput. Math. Methods Med.*, vol. 2021, pp. 1–12, Sep. 2021, doi: 10.1155/2021/5940433.

BIOGRAPHIES

Şehmus Aslan received his undergraduate degree from İhsan Doğramacı Bilkent University, Department of Industrial Engineering in 2008. He completed his master's degree at Dicle University, Institute of Social Sciences, Department of Business Administration in 2015. In 2018, he began his doctoral studies at Hasan Kalyoncu University, Institute of Social Sciences, Department of Business Administration, and earned his Ph.D. in 2020. He is currently working as an Assistant Professor at Mardin Artuklu University, Faculty of Economics and Administrative Sciences, Department of Business Administration, Production Management Division. His research areas include operations research, combinatorial optimization, evolutionary algorithms and machine learning.

Research Article

Comprehensive Design Criteria and Analysis of Laser Communication System for Underwater Vehicles

Ibrahim Emirhan Delibas and Alper Nabi Akpolat

Abstract— **Underwater vehicles are utilized in various fields, such as exploration, submarine research, and industrial applications. However, underwater communication is often a formidable challenge because conventional communication technologies are ineffective for underwater vehicles. Therefore, laser communication systems for underwater vehicles are receiving more and more attention. This paper presents a novel approach by investigating comprehensive design criteria and analysis of laser communication systems for underwater vehicles. Firstly, the basic principles and working mechanisms of underwater laser communication systems are explained. Then, the main factors affecting system performance and design criteria are discussed in detail. These criteria include communication distance, data rate, power consumption, optical properties of the underwater environment, and system stability. Different laser modulation techniques and communication protocols are also evaluated. The paper also focuses on simulation, channel diversity, and test methods that can be used to assess the performance of laser communication systems in different underwater environments.**

Index Terms— **Laser communication, underwater vehicles, data flow, bit error rate (BER), quality (Q) factor.**

I. INTRODUCTION

IN TODAY'S technology, wired communication has been replaced by wireless communication. Laser communication is preferred in remote distances and places where wiring is not feasible due to environmental conditions. Additionally, the reasons for the preference for laser communication are its faster speed than wired communication and the absence of wiring costs. Underwater communication systems are widely used in military, research, and exploration applications. Various data transmission methods transmit data underwater, such as acoustic, radio frequency, and optical communication. Communication methods using acoustic and radio frequencies are more vulnerable due to high error rates, inability to achieve high data transmission speeds, and susceptibility to cyber-attacks [1]. Radiofrequency and acoustic communication systems underwater communicate with high error rates, low data rates, and limited bandwidth. In such scenarios, optical communication emerges as a viable alternative. Optical communication is preferable for long-distance communication underwater due to features such as light's ability to transmit data at high bandwidth, low latency, high reliability, and low error rates. With this motivation, we design an underwater communication system. This system aims to achieve low error rates and high-quality (Q) factor results in various underwater environments at different distances. This study constructs a fundamental model for defense applications, underwater research, and marine sciences.

Research [2] has presented that amplitude modulation (AM) and frequency modulation (FM) methods transmit signals with distortion. In contrast, phase modulation (PM) and frequency-shift keying (FSK) are more susceptible to phase variations. They used a modulated laser diode (MLD) to generate a laser beam with disruption, and the incoming signals were analyzed using NI myDAQ. In [3], the selection of the transmitter for Free Space Optics (FSO) communication is explored during weather conditions. It compares wavelengths from 1550 nm to 10000 nm.

Similarly, [4] delved into modulation methods employed in communication and evaluated their highest achievable data rates over a distance of 45,000 km. The findings suggest that coherent optical (CO) frequency-shift keying (PSK) performs with the 850 nm wavelength, outperforming the 1064 nm and 1550 nm wavelengths. Furthermore, a simulative analysis of 10 Gbps bandwidth using optical communication channels has been conducted, comparing three optical channels: Optical Wireless Communication (OWC), FSO, Line-of-Sight Free Space Optics (LOS-FSO), using quadrature amplitude modulation (QAM) and PSK modulation formats [5]. A different study has been proposed for underwater visible light communication (UVLC) systems operation, having established that the performance of UVLC is influenced by factors such as receiver transmitter, water type, etc. [6]. Also, this article achieves a high optical density using a UVLC circuit over long distances with the help of Monte Carlo simulations [7]. A gradual attenuation channel is modeled to address the complex nature of vertical underwater links, where water temperature and salinity gradients in various water layers are used to

Ibrahim Emirhan DELİBAŞ, is with the Department of Electrical-Electronics Engineering, Marmara University, Istanbul, Türkiye, (e-mail: idelibas@marun.edu.tr).

https://orcid.org/0009-0007-9561-7538

Alper Nabi AKPOLAT, is with the Department of Electrical-Electronics Engineering, Marmara University, Istanbul, Türkiye, (e-mail: alper.nabi@marmara.edu.tr).

https://orcid.org/0000-0002-6972-2509

Manuscript received Jul 13, 2024; accepted Feb 12, 2025.
DOI: 10.17694/bajece.1515423

consider different attenuation coefficients, with real-time data transmission underwater [8].

The primary goal of this study is to provide a basic methodology that meets the comprehensive design criteria and analysis of laser communication systems for underwater vehicles. The study elaborates on the basic principles, operating mechanisms, and main factors affecting the performance and design criteria of UVLC. One crucial contribution is providing an essential resource for researchers, engineers, and students interested in designing and implementing laser communication systems. It also discusses an example of laser communication on system design and analysis aspects of current technology, future research areas, and development opportunities. This provides essential information for developing more efficient and reliable systems for underwater communications.

The rest of this study is organized as follows: Section II describes a general overview of laser communication systems for FSO and underwater communication. Section III gives the design steps of the proposed system. Section IV explains the findings and discussions that were obtained. Section V provides the conclusion and future remarks.

II. General Overview of Laser Communication Systems

With capabilities such as high-speed data transfer, increased bandwidth, reduced bit error rates, and overall cabling costs, all laser communication systems are at the forefront of communication systems. They are the most preferred option when the distance is too long or cabling could be more convenient. A typical system used in laser communications is FSO.

A. Free Space Optics (FSO)

FSO is a communication method that utilizes light transmission in free space for data transmission. It involves transmitting data through a void, air, or space-like environment. FSO is preferred in environments unsuitable for using fiber optic cables or systems requiring high costs [9]. The attenuation coefficient varies depending on the conditions of the transmission medium. Table I shows the attenuation coefficients in the FSO communication environment.

TABLE I

Weather Attenuation for FSO [10].

Weather Conditions	α (dB/km)
Heavy fog	125
Moderate fog	42.2
Light fog	20
Heavy rain	9.2
Moderate rain	5.8
Haze	4.2
Clear air	0.43

The Beer-Lambert law gives the following formula for the power of light transmitted through free space.

$$I = I_0 \times e^{-kd} \tag{1}$$

where, k is the attenuation coefficient of the medium, d is the length of the medium, I is light power, and I_0 is incoming laser power.

B. Underwater Optic Communication

Underwater Optical Communication (UOC) is a form of communication in which light propagates through the underwater environment for data transmission. This communication type is used in underwater research and exploration applications, where environmental conditions are not conducive to cabling, and high data rates and wide bandwidth are needed. Similar to FSO, light in UOC is subject to the attenuation coefficient of the medium extinction, scattering, and absorption coefficients for different types of water are provided in Table II.

TABLE II

Extinction, Scattering and Absorption Coefficient For Different Types of Water [11-13].

Type of Water	Extinction coefficient (c) (m^{-1})	Scattering coefficient (b) (m^{-1})	Absorption coefficient (α) (m^{-1})
Harbor	1.1	0.913	0.187
Pure Sea	0.043	0.0025	0.0405
Clear Ocean	0.151	0.037	0.114
Coastal Ocean	0.398	0.219	0.179

$$c(\lambda) = \alpha(\lambda) + b(\lambda) \tag{2}$$

Three different modulation techniques, Mach-Zehnder (MZ) modulation, electro-absorption (EA) modulation, and AM, have been investigated to validate the comprehension design for laser communication systems for underwater vehicles.

III. Design Steps, Circuit Description, and Formulation of Studied System

A. System Modeling and Simulation

The general system design is a 1 Watt-480 nm pure sea environment, as given in Fig. 1. It was created using the OptiSystem software environment. The system sends data in bit form with a test pattern to convert to an electrical square wave using the NRZ pulse generator component. The NRZ Pulse Generator produces square wave signals according to the following formulations.

$$\text{Exponential} \quad E(t)\begin{cases} 1-e^{-(t/c_r)}, 0 \le t < t_1 \\ 1, t_1 \le t < t_2 \\ e^{-(t/c_f)}, t_2 \le t < T \end{cases} \tag{3}$$

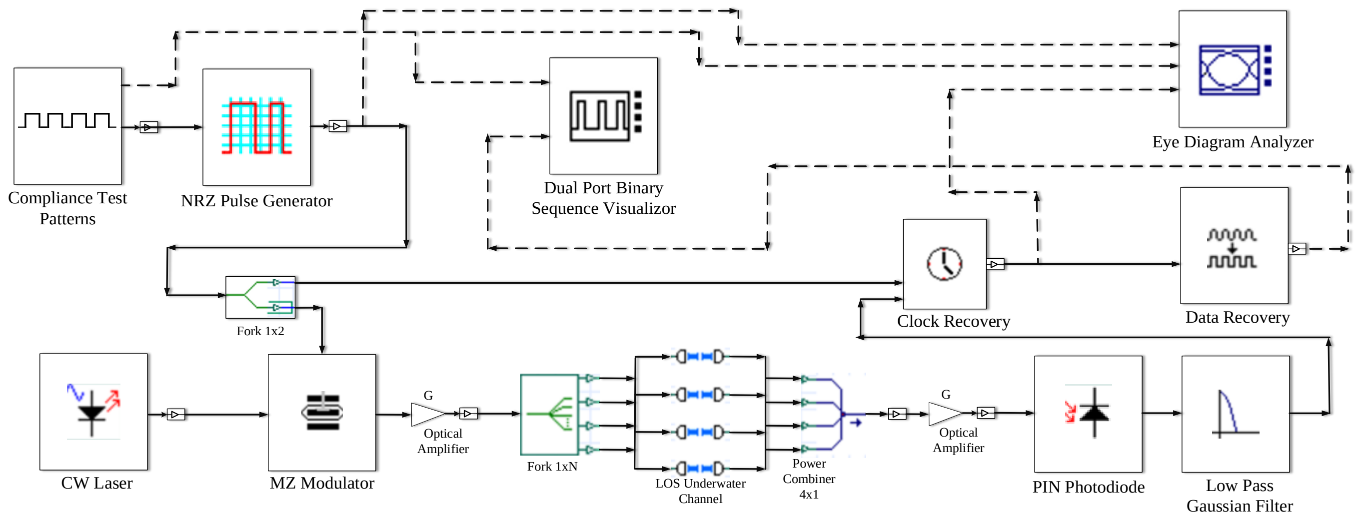


Fig. 1. System design of underwater laser communication in pure sea.

Gaussian
$$E(t)\begin{cases}1-e^{-(t/c_r)^2}, 0\le t<t_1\\ 1, t_1\le t<t_2\\ e^{-(t/c_f)^2}, (t_2\le t<T)\end{cases}\qquad(4)$$

Linear
$$E(t)\begin{cases}t/c_r, 0\le t<t_1\\ 1, t_1\le t<t_2\\ t/c_f, t_2\le t<T\end{cases}\qquad(5)$$

Sine.
$$E(t)\begin{cases}\sin(\pi.t/c_r), 0\le t<t_1\\ 1, t_1\le t<t_2\\ \sin(\pi.t/c_f), t_2\le t<T\end{cases}\qquad(6)$$

where, the rising time coefficient is c_r, the fall time coefficient is c_f, and the bit period is T. The parameters rising time and fall time values are used to numerically set the time numerically points t_1 and t_2, as well as c_r and c_f, to produce pulses. The generated square wave signal is modulated with an MZ modulator using a CW optical laser signal. Data is converted to an optical signal using three different types of modulators. Modulators regarding MZ, EA, and AMAX Fig 2 shows the general scheme of the MZ Modulator.

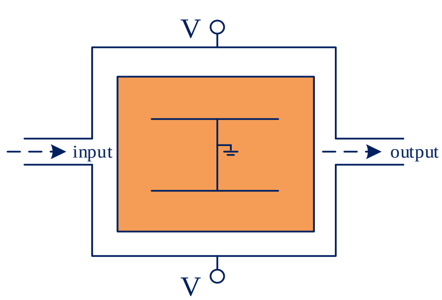


Fig. 2. General overview of MZ modulation.

$$E_{out}(t)=E_{in}(t).\cos(\Delta\theta(t)).\exp(j.\Delta\phi(t))\qquad(7)$$

$$\Delta\theta(t)=\frac{\pi}{2}.(0.5-\alpha.(Modulation(t)-0.5))\qquad(8)$$

$$\alpha=1-\frac{4}{\pi}.\arctan\left(\frac{1}{\sqrt{ER}}\right)\qquad(9)$$

$$\Delta\phi(t)=SC.\Delta\theta(t).(1+SF)/(1-SF)\qquad(10)$$

where, SF is the symmetry factor, SC is equal to minus one (-1) if the negative signal chirp is true or one of the negative signal chirp is false. ER is the signal extinction ratio, and *Modulation(t)* is the modulating input signal. The electrical input modulating signal is normalized internally between 1 and 0. The general scheme of the EA modulator is seen in Fig. 3.

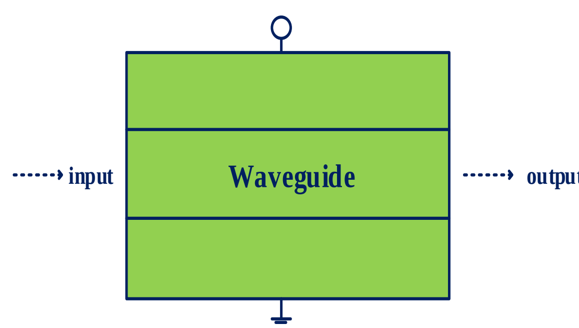


Fig. 3. General overview of EA modulation.

$$E_{out}(t)=E_{in}(t).\sqrt{Mod(t)}.\exp\left(j\frac{\alpha}{2}.\ln(Mod(t))\right)\qquad(11)$$

where, $E_{in}(t)$ is the input signal, $E_{out}(t)$ is the output optical signal, and α is the chirp factor.

$$Mod(t)=(1-MI)+MI.\mathrm{mod}\,ulation(t)\qquad(12)$$

where, *modulation(t)* is the input signal and MI is the modulation index.

- AM Modulator,

$$E_{out}(t)=E_{in}(t).\sqrt{Mod(t)}\qquad(13)$$

where, $E_{in}(t)$ is the input signal, $E_{out}(t)$ is the output optical signal,

$$Mod(t) = (1 - MI) + MI.modulation(t) \tag{14}$$

where, *modulation(t)* is the input signal and *MI* is the modulation index.

The modulated optical signal is amplified using an Optical Amplifier (OA). The reason for using OA is that our transmitted laser signal will be subjected to various noise and attenuation in the underwater channel, so it is necessary to introduce a high-power signal into this channel. The goal is to limit the power of our laser source to a certain level and then amplify it using OA. The modulated and amplified optical signal is split into multiple channels using a Power Splitter and transmitted through four lines of sight (LOS) underwater channels.

$$G = \frac{(P_{OUT} - P_{ASE})}{P_{\sin}} \tag{15}$$

where, G is gain, P_{out} is the total output power, P_{in} is the total input power, and P_{ASE} includes (or does not include) the generated Amplified Spontaneous Emission (ASE).

$$P_{R\_los} = P_t \eta_\tau \eta_R \exp\left[-c(\lambda)\frac{d}{\cos(\theta)}\right]\frac{A_{Rec}\cos(\theta)}{2\pi d^2\left[1-\cos(\theta_0)\right]} \tag{16}$$

where, the laser beam divergence angle is θ_0, the receiver aperture area is A_{Rec}, the perpendicular distance between the transmitter and the receiver plane is d, the optical efficiency of the transmitter is ηT, and the angle between the transmitter-receiver trajectory and the receiver plane is θ.

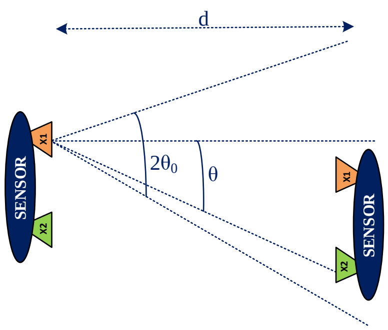


Fig. 4. LOS underwater channel model [14].

Fig. 4 depicts the optical signal transmitted through the LOS Underwater channel, which undergoes attenuation. After being amplified again with an OA, it is converted to an electrical signal using the Pin Photodiode component, as depicted in Fig.5.

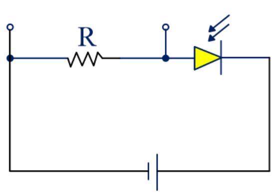


Fig. 5. Pin photodiode circuit.

$$i(t) = rP(t) + i_d \tag{17}$$

$$\sigma^2_{total} = \sigma^2_{th} + \sigma^2_{shot-S} + \sigma^2_{shot-ASE} + \sigma^2_{S-ASE} + \sigma^2_{ASE-ASE} \tag{18}$$

$$\sigma^2_{shot-S} = 2B_e.(rP_s + i_d)\ (Shot - Signal\ Noise) \tag{19}$$

$$\sigma^2_{shot-ASE} = 2B_e.qrP_{ASE}\ (Shot - ASE\ Noise) \tag{20}$$

$$\sigma^2_{S-ASE} = 4B_e.r^2.PSD_{ASE}P_S\ (Signal - ASE\ Noise) \tag{21}$$

$$\sigma^2_{ASE-ASE} = r^2.PSD^2_{ASE}\ (ASE - ASE\ Beat\ Noise) \tag{22}$$

where, B_e is the PIN's equivalent noise bandwidth, r is its responsivity, q is its electron charge value, i_d is its dark current, P_s is its signal power, P_{ASE} is its optical noise power, and PSD_{ASE} is its power spectrum density of the optical noise field (spontaneous emission). The received electrical signal is filtered with a Gaussian low-pass frequency function.

$$H(f) = \alpha e^{-\ln(\sqrt{2})\left(\frac{f}{f_c}\right)^{2N}} \tag{23}$$

where, N is the parameter for Order, α is the parameter for Insertion loss, f is the frequency, $H(f)$ is the filter transfer function, and f_c is the filter cutoff frequency. The filtered electrical signal is equalized in time with the input signal using the Clock Recovery component. Then, the electrical signal is converted to binary data using the Data Recovery component, allowing for comparison between the transmitted and received binary data. The Input and Output signals are compared using the Eye Diagram Analyzer component to obtain parameters such as Q Factor, Min BER, and Max Eye Height and to generate an eye diagram.

IV. Results and Discussions

Before reaching the optimum system design, scenarios were conducted to find the optimal configuration, and each scenario's results are as follows.

In the first scenario, measurements were taken in the system using MZ modulation with a 1-Watt 532 nm wavelength laser; the PIN photodiode's parameters are responsivity 0.5 and dark current 2 nA, and the low-pass Gaussian filter's cutoff frequency=0,75*symbol rate in a Pure Sea environment without using an OA. The other scenarios use the same parameters as the first scenario. Data transmission was performed and compared. The system design of the first scenario can be seen in Fig. 6. The transmitted signal did not adequately reach the receiver in the constructed circuit. The eye diagram output is shown in Fig. 7.

In the second scenario, in addition to the setup in the first scenario, two OAs were added, one at the input and one at the output of the transmission channel. The OA parameters are gain 40 dB and noise Fig. 4 dB. The system design of the second scenario is shown in Fig. 8. As a result of the scenario, the transmitted signal reached the receiver properly, and the Q factor value was measured as 282.337. The eye diagram result is given in Fig. 9.

In the third scenario, when the parameters used in the second scenario system design were simulated in a clear ocean environment, the data was not adequately transmitted to the receiver. The eye diagram result is given in Fig. 10.

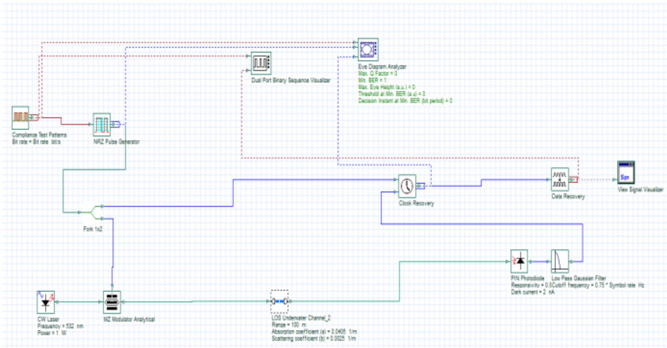


Fig. 6. System design for the first scenario.

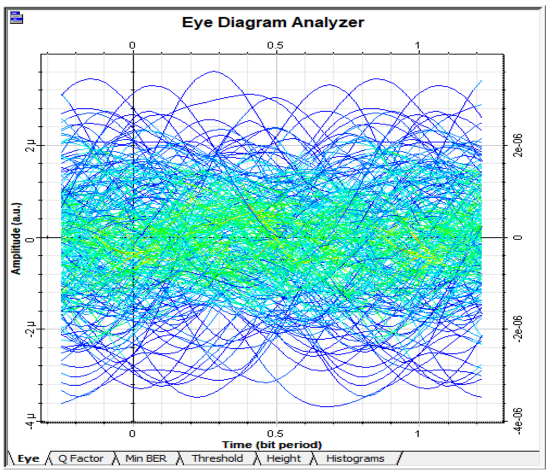


Fig. 7. Eye diagram for the first scenario.

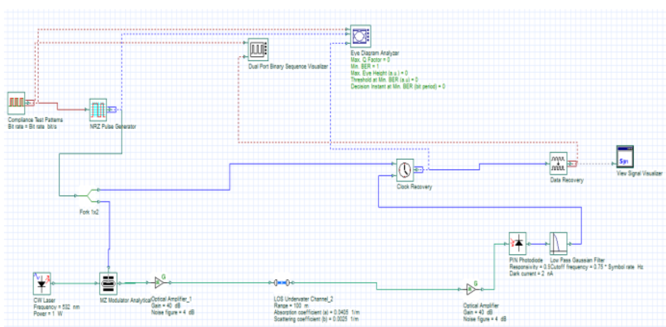


Fig. 8. System design for the second scenario.

In the fourth scenario, the signal was correctly transmitted when the number of transmission channels was increased to 4. The system design of the fourth scenario is shown in Fig. 11. Considering the result for the fourth scenario, Fig. 12 expresses the eye diagram result. The scenarios determined that the optimal circuit design was created with four channels and OA. In the optimal circuit design simulation, two different channels were simulated in the underwater transmission channel: Pure Sea and Clear Ocean. Simulations were conducted with three different modulations at a fixed distance of 100 meters and other wavelengths. The Min BER and Max (M) Q factor values of the transmitted data in these channels with two different attenuation coefficients were measured, and eye diagram graphs were generated.

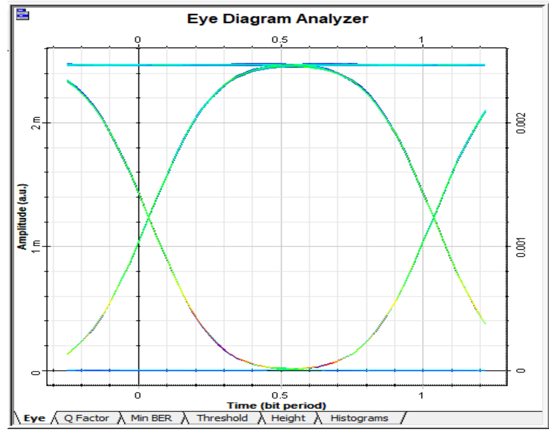


Fig. 9. Eye diagram for the second scenario.

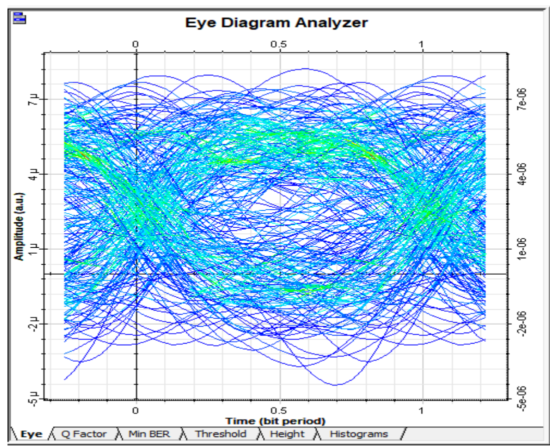


Fig. 10. Eye diagram for the third scenario.

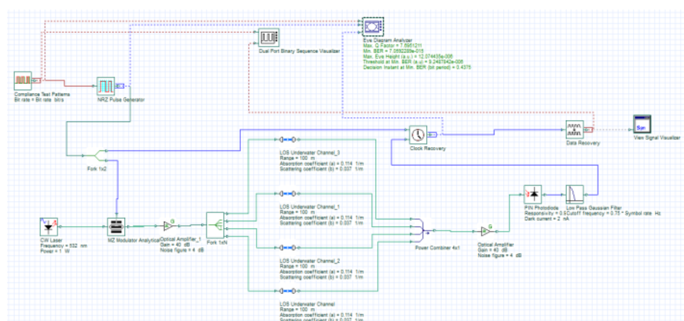


Fig. 11. System design for the fourth scenario.

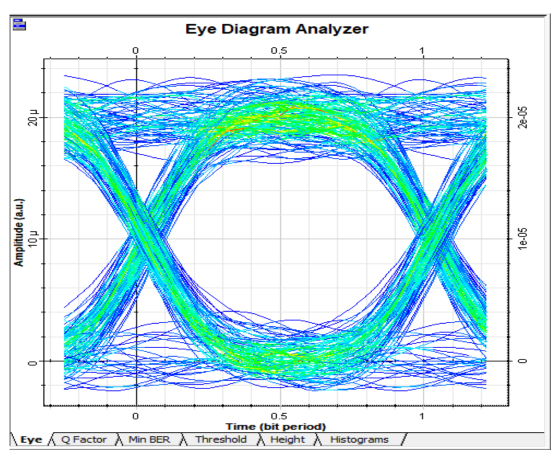


Fig. 12. Eye diagram for the fourth scenario.

TABLE III

PURE SEA, 100 METERS, 1 WATT LASER POWER.

		WAVELENGTH (nm)					
		405	532	635	980	1064	1550
MODULATION TYPE	MZ	Max Q Factor: 354.082	Max Q Factor: 355.331	Max Q Factor: 353.365	Max Q Factor: 353.583	Max Q Factor: 352.925	Max Q Factor: 123.632
		Eye Height: 0.97191	Eye Height: 0.971923	Eye Height: 0.971884	Eye Height: 0.971895	Eye Height: 0.971871	Eye Height: 0.956203
		Min BER : 0	Min BER : 0	Min BER : 0	Min BER : 0	Min BER : 0	Min BER : 0
	EA	Max Q Factor: 344.753	Max Q Factor: 344.87	Max Q Factor: 345.883	Max Q Factor: 345.232	Max Q Factor: 346.688	Max Q Factor: 99.1558
		Eye Height: 0.9248	Eye Height: 0.924788	Eye Height: 0.924822	Eye Height: 0.924807	Eye Height: 0.924838	Eye Height: 0.904289
		Min BER : 0	Min BER : 0	Min BER : 0	Min BER : 0	Min BER : 0	Min BER : 0
	AM	Max Q Factor: 343.996	Max Q Factor: 345.439	Max Q Factor: 344.573	Max Q Factor: 346.414	Max Q Factor: 345.723	Max Q Factor: 115.545
		Eye Height: 0.973449	Eye Height: 0.973485	Eye Height: 0.973471	Eye Height: 0.973515	Eye Height: 0.973482	Eye Height: 0.956898
		Min BER : 0	Min BER : 0	Min BER : 0	Min BER : 0	Min BER : 0	Min BER : 0

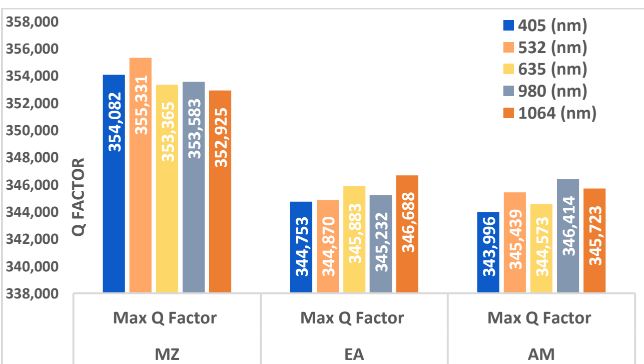


Fig. 13. Q factor table for Pure Sea environment.

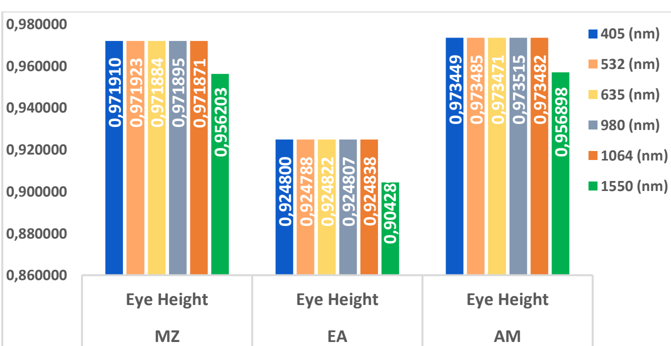


Fig. 14. Eye height table for Pure Sea environment.

As shown in Fig. 13, using MZ modulation and a 532 nm wavelength laser resulted in a higher Q factor in the pure sea environment compared to other scenarios.

Fig. 14 displays that using AM modulation and a 1064 nm wavelength laser resulted in a higher eye height in the pure sea environment than in other scenarios.

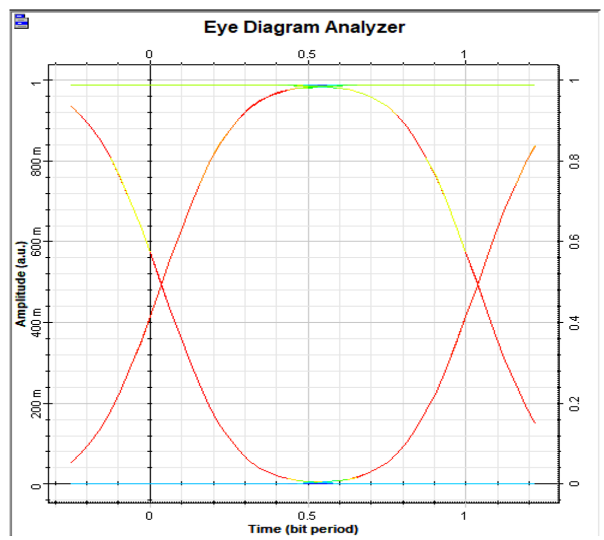


Fig. 15. Eye diagram for Pure Sea environment with MZ modulation (532 nm wavelength).

The scenario conducted in the Pure Sea environment achieved the best result using MZ modulation and a wavelength of 532 nm. The eye diagram of the best scenario can be observed in Fig. 15. As can be seen in Table III, 18 different parameters for the pure sea environment are obtained. Fig. 16 shows that using AM modulation and a 1064 nm wavelength laser resulted in a higher Q factor in the clear ocean environment compared to other scenarios. As seen in Fig. 17, using AM modulation and a 1064 nm wavelength laser resulted in a lower Min BER in the clear ocean environment compared to the other scenarios. Fig. 18 shows that using EA modulation and a 980 nm wavelength laser resulted in a higher eye height in the clear ocean environment compared to different scenarios.

TABLE IV

CLEAR OCEAN, 100 METERS, 1 WATT LASER POWER.

		WAVELENGTH (nm)					
		405	532	635	980	1064	1550
MODULATION TYPE	MZ	Max Q Factor: 7.71839	Max Q Factor: 7.26355	Max Q Factor: 7.55246	Max Q Factor: 8.22645	Max Q Factor: 8.01527	Max Q Factor: 0
		Eye Height : 1.1279e-05	Eye Height : 1.16268e-05	Eye Height : 1.2028e-05	Eye Height : 1.29086e-05	Eye Height: 1.24512e-05	Eye Height : 0
		Min BER: 5.86909e-15	Min BER : 1.88525e-13	Min BER : 2.12939e-14	Min BER : 9.52291e-17	Min BER : 5.46291e-16	Min BER : 1
	EA	Max Q Factor: 7.23672	Max Q Factor: 7.50535	Max Q Factor: 7.48618	Max Q Factor: 7.11237	Max Q Factor: 8.26744	Max Q Factor: 0
		Eye Height: 1.10836e-05	Eye Height: 1.12577e-05	Eye Height: 1.14245e-05	Eye Height: 1.08784e-05	Eye Height: 1.2367e-05	Eye Height : 0
		Min BER : 2.2752e-13	Min BER : 3.05014e-14	Min BER : 3.51692e-14	Min BER : 5.599e-14	Min BER : 6.83795e-17	Min BER : 1
	AM	Max Q Factor: 7.31086	Max Q Factor: 7.97961	Max Q Factor: 8.05461	Max Q Factor: 8.23974	Max Q Factor: 8.61914	Max Q Factor: 0
		Eye Height: 1.27444e-05	Eye Height: 1.06522e-05	Eye Height: 1.24187e-05	Eye Height: 1.26896e-05	Eye Height: 1.29787e-05	Eye Height : 0
		Min BER: 4.66896e-17	Min BER : 7.20103e-16	Min BER : 3.95456e-16	Min BER : 8.52336e-17	Min BER : 3.29406e-18	Min BER : 1

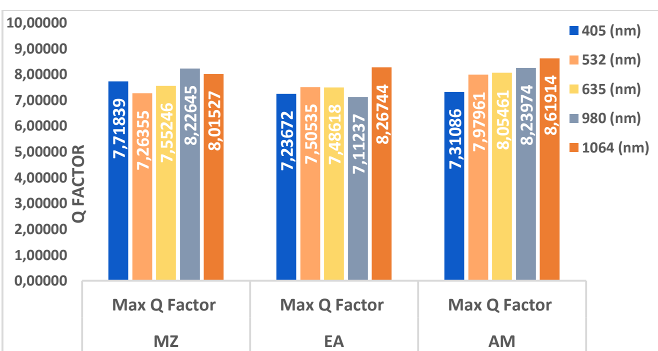


Fig. 16. Q factor table for Clear Ocean environment.

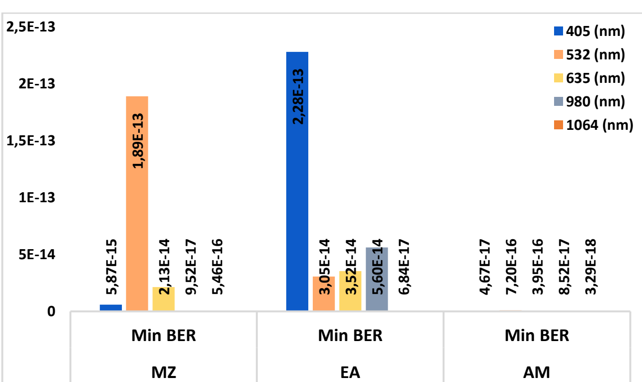


Fig. 17. Min BER table for Clear Ocean environment.

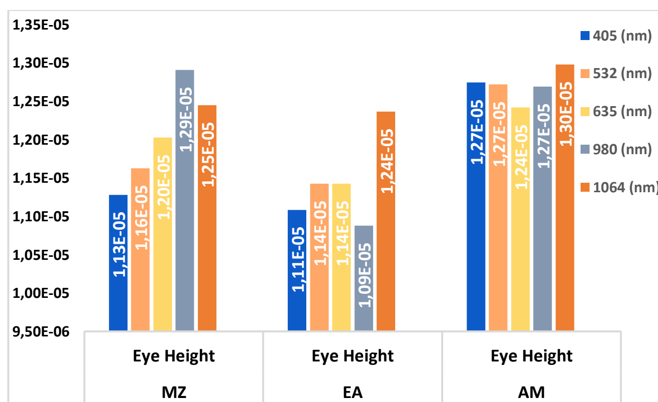


Fig. 18. Eye height table for Clear Ocean environment.

The scenario conducted in the Clear Ocean environment achieved the best result using AM modulation and a wavelength of 1064 nm. The eye diagram of the best scenario is shown in Fig. 19.

Similarly, Table IV presents the max Q factor, eye height, and min BER parameters for a clear ocean environment. Upon reviewing the findings presented in the article, it becomes evident that using optical amplifiers while transmitting data over long distances via laser, both in pure sea and clear ocean environments, ensures a more stable delivery of data to the receiving end. It has been observed that transmitting data over multiple channels rather than a single channel enhances data quality. Additionally, it has been determined that high-wavelength lasers transfer data more effectively in environments with high coefficient values.

V. CONCLUSION AND FUTURE REMARKS

Considering all the experimental results conducted in both pure sea and clear ocean environments, the results of 1550 nm laser were worse than the results of lasers with other wavelengths.

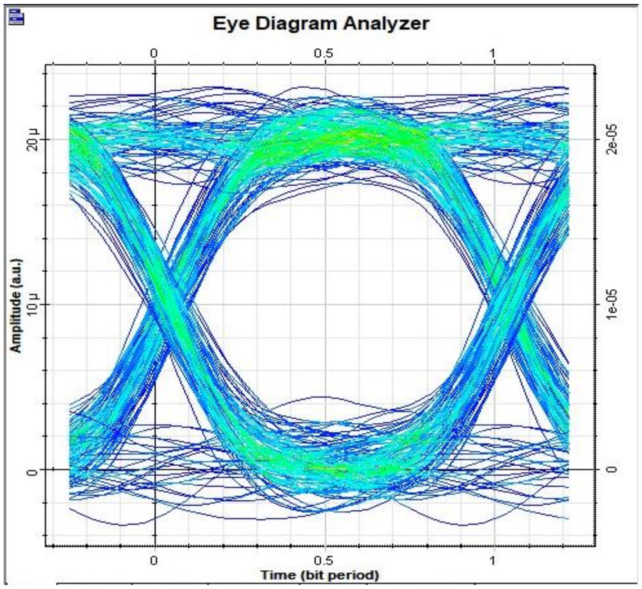


Fig. 19. Eye diagram for Clear Ocean environment with AM modulation (1064 nm wavelength).

Therefore, it has been understood that utilizing lasers with a wavelength of 1550 nm for underwater optical communication is not preferable to lasers with other wavelengths. Briefly, this paper presents an example of laser communication system design and analysis based on current technology and discusses future research areas and development opportunities. It will be a valuable resource for researchers and engineers interested in designing and implementing laser communication systems to meet the communication needs of underwater vehicles.

VI. Acknowledgement

The authors would like to extend their sincere thanks to HAVELSAN for their supportive approaches.

References

[1] P. K. Pandey, "MIMO based high-speed underwater optical wireless communication using WDM," *2022 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, BHOPAL, India, 2022, pp. 1-5.

[2] MAX R. Alam and S. Faruque, "Comparison of different modulation techniques for free space laser communication," *2015 IEEE International Conference on Electro/Information Technology (EIT)*, Dekalb, IL, USA, 2015, pp. 637-640, doi: 10.1109/EIT.2015.7293409.

[3] P. Singal, S. Rai, R. Punia, and D. Kashyap, "Comparison of Different Transmitters Using 1550nm and 10000nm in FSO Communication Systems," *International Journal of Computer Science & Information Technology (IJCSIT)*, vol 7, no 3, June 2015.

[4] J. Kaur, B. Kaur, and K. Singh, "Design and Performance Investigation of Intersatellite Optical Wireless Communication System Employing Modulation Techniques," *Wireless Personal Communication*, vol. 94, pp. 793-807, 2017.

[5] G. Karpagarajesh, R. Santhana Krishnan, Y. Harold Robinson, S. Vimal, S. Kadry, and Y. Nam, "Investigation of digital video broadcasting application employing the modulation formats like QAM and PSK using OWC, FSO, and LOS-FSO channels," *Alexandria Eng. J*. vol. 61, no. 1, pp. 647–657, 2022.

[6] MAX Elamassie, F. Miramirkhani, and MAX Uysal, "Performance Characterization of Underwater Visible Light Communication," *IEEE Transactions on Communications*, vol. 67, no. 1, pp. 543-552, Jan. 2019.

[7] C. Wang, H -Y. Yu, and Y. -J. Zhu, "A Long Distance Underwater Visible Light Communication System With Single Photon Avalanche Diode," *IEEE Photonics Journal*, vol. 8, no. 5, pp. 1-11, Oct. 2016.

[8] MAX Elamassie and MAX Uysal, "Vertical Underwater Visible Light Communication Links: Channel Modeling and Performance Analysis," *IEEE Transactions on Wireless Communications*, vol. 19, no. 10, pp. 6948-6959, Oct. 2020.

[9] S. A. Al-Gailani et al., "A Survey of Free Space Optics (FSO) Communication Systems, Links, and Networks," *IEEE Access*, vol. 9, pp. 7353-7373, 2021.

[10] A. Mansour, R. Mesleh, and MAX Abaza. "New Challenges in Wireless & Free Space Optical Communications," *Optics and Lasers in Engineering*, vol. 89 pp. 95-108, 2017.

[11] R. Boluda-Ruiz, P. Rico-Pinazo, B. Castillo-Vázquez, A. García-Zambrana, and K. Qaraqe, "Impulse Response Modeling of Underwater Optical Scattering Channels for Wireless Communication," *IEEE Photonics Journal*, vol. 12, no. 4, pp. 1-14, Aug. 2020.

[12] C. D. Mobley Light and Water: Radiative Transfer in Natural Waters. New York, NY, USA: Academic, 1994.

[13] V. I. Haltrin, "Chlorophyll-based model of seawater optical properties," *Appl. Opt*., vol. 38, No. 33, pp. 6826–6832, 1999.

[14] S. Arnon, "Optical wireless communication," *Encyclopedia of Optical Engineering*, vol. 2, pp. 1866–1886, 2003.

Biographies

Ibrahim Emirhan DELIBAS received the B.Sc. and MSc. degrees in the Department of Electrical-Electronics Engineering from Marmara University, Istanbul, Turkiye, in 2019 and 2024, respectively. He currently works as a software engineer in the defense industry.

Alper Nabi Akpolat received the B.Sc. and the MSc. degrees from Firat University, Elazig, Turkiye, in 2012 and 2015, respectively. He also received his Ph.D. degree from Marmara University in 2021. He works as an Assistant Professor in the Department of Electrical-Electronics Engineering at the same University, Istanbul, Turkiye. Since March 2019, he has been a Guest Ph.D. student for one year with the Department of Energy Technology, Aalborg University, Denmark. His current research interests include renewable energy systems, PV systems, DC microgrids, control of distributed generation systems, green hydrogen production, neural networks, and applied artificial intelligence in power electronics and power systems.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 1, March 2025
Research Article

Deep Learning based Disease Detection from Potato Leaf Images

Abdulkerim Oztekin and Kenan Almas

Abstract—**This study aims to detect diseases from potato leaf images using deep learning methods. In the proposed work, a large and comprehensive image dataset of healthy and various potato diseases was used to detect common diseases (late blight, early blight) seen in potato plants. Models were developed to detect potato diseases using different Convolutional Neural Network (CNN) architectures and hybrid models. The developed models were trained under various hyperparameters and evaluated using performance metrics such as accuracy, precision, recall, and F1-score. The study shows that deep learning methods can be effectively used in the detection of potato diseases and can contribute to previous studies in this field. The dataset was tested with four different ResNet models and evaluated with various performance metrics. It is thought that the obtained test results can provide a significant information for disease management and productivity increase in potato cultivation. And artificial intelligence (AI)-based disease detection can lead to innovations in the field of agriculture, and can also contribute to machine-human interaction. Our work also highlights the success and importance of ResNet deep learning models in the field of image extraction.**

Index Terms—**Convolutional Neural Network (CNN), Deep Learning, Potato Disease Detection, ResNet.**

I. INTRODUCTION

AGRICULTURE IS one of the oldest and most fundamental activities of humanity, and has made significant contributions to the emergence and development of civilizations. Today, the increasing world population, climate change and limited resources make agricultural production even more important and necessitate the need to increase efficiency. In order to cope with these challenges and ensure sustainable food production, technology has increasingly begun to be used in the agricultural sector. Especially in recent years, rapid developments in the field of artificial intelligence have the potential to revolutionize the agricultural sector. Today, the rapidly increasing world population and changing consumption habits have made food supply and security a global priority. In this context, the efficiency and sustainability of the agricultural sector is of critical importance for the future of humanity. Plant production, which forms the basis of the agricultural sector, is currently faced with human-induced factors such as incorrect planting, incorrect irrigation and incorrect fertilization, as well as environmental challenges such as climate change, water scarcity and soil erosion. These challenges negatively affect the productivity and quality of agricultural products, threatening food security.

Technological developments offer new opportunities to overcome these challenges and increase efficiency in the agricultural sector. Image processing and deep learning techniques, which have developed rapidly in recent years, have the potential to revolutionize agricultural applications. Many different machine learning methods have been successfully used in the literature for early diagnosis, quality control and classification of plant diseases [1-6]. These techniques provide fast, accurate and objective quality assessment thanks to the ability to objectively determine the visual features of food items such as color intensity, color distribution, visual defects, size and shape [7]. One of the biggest disadvantages of machine learning algorithms is the need to pre-process the data and manually determine the features. Deep learning eliminates this laborious process and automatically extracts meaningful features from raw data, thus making it possible to create more powerful and flexible models [8].

Potatoes are very important food sources and one of the most consumed agricultural products in the world. However, there are various difficulties in the potato production and processing process. Post-harvest classification and quality control processes are usually carried out manually, which leads to many problems such as substandard products, human errors, increased labor costs and production waste [9]. The activity of diagnosing and characterizing multiple aspects of potato diseases based on their characteristics, symptoms and appearances is called potato disease classification. Recent studies have proven the success of machine learning-based potato classification, disease detection and quality control systems [10-13].

Deep learning is a more advanced form of artificial neural networks and has the ability to recognize and classify complex patterns by training on large data sets. In this way, more

Abdulkerim Öztekin, is with Department of Electrical&Electronics Engineering, Batman University, Batman, Turkiye, (e-mail: abdulkerimoztekin@gmail.com).
https://orcid.org/0000-0002-0698-3525

Kenan Almas, is with Department of Information Technologies, Batman University, Batman, Turkiye, (e-mail: bajece@hotmail.com).
https://orcid.org/0000-0002-2282-2641

Manuscript received Feb 28, 2025; accepted March 24, 2025.
DOI: 10.17694/bajece.1649068

effective results can be obtained in agricultural applications such as plant disease detection, product classification and quality control [14]. Deep learning models used in potato quality control are generally based on the Convolutional Neural Networks (CNN) architecture, and there are various studies in the literature that uses CNN models in classification, disease and defect detection, and quality control [15-21]. CNNs have the ability to automatically learn and hierarchically represent features in images. In this way, various defects, diseases and quality traits in potatoes can be detected with high accuracy. These studies show that deep learning-based image processing systems can surpass human performance in potato quality control and increase productivity in the agricultural sector. In particular, the use of new imaging technologies such as hyperspectral imaging, thermal imaging, and 3D imaging can provide a more comprehensive analysis of the internal and external quality characteristics of potatoes [22-25]. In addition, studies on the explainability and interpretability of deep learning models will provide a better understanding of the decision-making processes of these models and increase their reliability.

The main purpose of this study is to detect diseases from potato leaf images using deep learning methods and to contribute to previous studies in this field. In this research, an approximation model from deep learning models will be applied to determine healthy and various known potato diseases. For this purpose, models will be developed to detect potato diseases using CNN architectures with the help of a large and comprehensive image dataset of potato diseases and the developed models will be optimized and trained with different parameters. Thus, the effectiveness and potential of deep learning methods in the detection of potato diseases will be evaluated.

This paper is organized as follows: Section 2 explains the dataset and the conducted Deep Learning algorithms for the proposed study. The results are presented in Section 3 and finally, concluding remarks and a comprehensive discussion are made in Section 4.

II. Material And Methods

2.1. Properties of the Dataset

In this study, the analyses on plant diseases were based on the PlantVillage dataset (www.kaggle.com/code/redpen12/cnn-disease-detection/input). This dataset was obtained from the Kaggle platform and consists of three channel 256x256 pixel color images which are rescaled and resized to 224x224 pixels to be used in our models. This dataset comprises high-resolution images of potato plants affected by various diseases, including early blight, late blight, and healthy leaves. It is designed to support the development and evaluation of image recognition models for precise disease detection and classification, contributing to advancements in agricultural diagnostics. Some augmentation techniques such as flipping, rotating, shifting, shearing and brightening has been applied the training data to improve the model's generalization capability and prevent overfitting.

The dataset is a benchmark dataset used to recognize disease and health status from potato leaves. The dataset consists of 3 different image groups as early blight, late blight and healthy. The dataset consists of 2152 images that are visually close to each other and not homogeneous. This situation has a complicating effect on classification tasks. Of these images, there are 1000 images with early blight disease, 1000 images with late blight disease and 152 healthy images. The dataset has been split into training, validation, and test sets with specified ratios (70%-15%-15%). Randomly selected images from the dataset are given in Fig. 1.

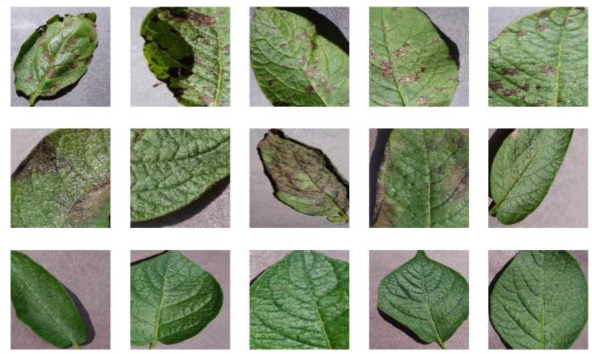


Fig.1. Some images from healthy and diseased potato leaves

2.2. Methods

2.2.1. Transfer learning methods

Transfer learning is a powerful technique that transfers previously learned information to new tasks in machine learning, accelerating the learning process and improving performance [26]. It is based on the principle of using knowledge acquired in a field in a similar field. For example, a model trained for image classification can be used for a related task such as object detection by fine-tuning. This is a great advantage, especially in cases where data sets are limited or training time is limited. The reasons why transfer learning is effective include data efficiency, fast learning, and improved performance. Since a pre-trained model has already learned general features, it needs less data specific to the new task and is trained faster. In addition, it can achieve better results in the target task by generalizing the information it has learned from the source task. Fig. 2 shows the differences between transfer learning and machine learning.

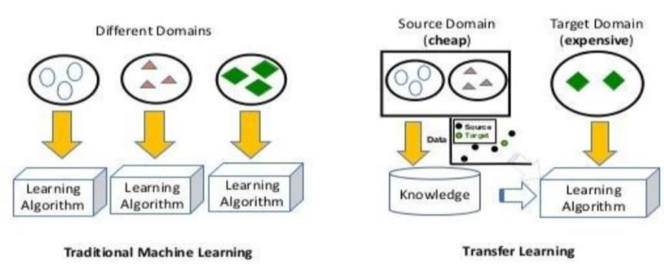


Fig.2. Illustration of process differences between transfer learning and machine learning [27]

Deep learning models benefit greatly from transfer learning, especially in areas such as image processing, natural language processing, and speech recognition. For instance, a CNN trained on a large dataset such as ImageNet can be fine-tuned and used for a different task such as medical image classification. Factors such as the similarity between the source and target tasks, the size and quality of the source dataset, and the model architecture are important for the success of transfer learning. In the transfer learning process, it is important to choose a source dataset with similar characteristics to the target dataset, select a suitable pre-trained model, and fine-tune the model to adapt it to the target task.

As a result, transfer learning leads to a significant paradigm shift in machine learning, enabling faster, more efficient, and more effective learning. With rapid developments in deep learning and increasing data availability, the importance of transfer learning is increasing.

2.2.2. Convolutional Neural Networks

Convolutional neural networks (CNN) are powerful deep learning models that have shown great success in image processing and data classification [28]. Although their foundations were laid in the 1980s, they began to be processed and trained after the 1990s and began to be used in image processing applications in 1995. CNNs provide superior performance by effectively identifying features in visual data thanks to their unique structures [29-30]. Convolutional Neural Networks (CNNs) are deep learning models widely used in processing visual data. They are designed to learn patterns and features in complex images. CNNs perform a shifting operation on the input image using small filters called kernels and thus perform the convolution operation. CNNs, which consist of successive layers such as convolution, pooling, full connection and output layers, extract important features in the image and achieve high success in classification, object recognition and similar tasks [28]. CNNs are a deep learning model that has revolutionized the field of image processing. These models analyze the relationships between pixels, the basic units that make up images, and extract higher-level meanings. CNNs consist of many layers built on top of each other. Each layer represents a different aspect of the image. For example, the first layers detect simple edges, while later layers identify more complex shapes and objects. The learning process of CNNs is optimized through the backpropagation algorithm. This algorithm allows the network to learn from its mistakes and make more accurate predictions. The general structure of the CNN network is shown in Fig. 3.

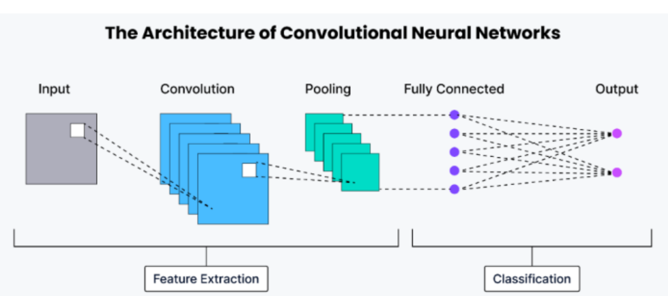


Fig.3. General Structure of CNN Network

CNNs are a powerful tool that allows computers to understand images like humans. These networks identify important features in images, just as a human focuses on specific features (for example, the eyes, nose, and mouth of a face) when examining a picture. Through a process called convolution, CNNs learn patterns, textures, and shapes in images. In this way, they are used in tasks such as recognizing, classifying, and locating objects. CNNs have achieved great success in many areas, such as diagnosing diseases in medical imaging and perceiving the environment in autonomous vehicles. CNNs generally consist of four main layers: convolution layer, pooling layer, correction layer, and fully connected layer [28]. Activation functions are critical components that organize the input signals of neurons, which are the basic building blocks of neural networks, and determine the learning ability of the network. These functions compress the outputs of neurons to certain ranges, allowing the network to work more efficiently. Various activation functions, such as Sigmoid, ReLU, and Hyperbolic Tangent, are selected according to the nature of the problem and control whether neurons are activated or not, directing the network's information processing process. It is important to use a sensitive activation function, especially in areas where matrix operations are intensive, such as image processing. Otherwise, the learning speed may decrease.

The pooling layer is an important component that reduces the number of weights and computational costs by reducing the data size in CNNs. This layer can speed up the learning process while also increasing the generalization ability of the model. No learning takes place in the pooling layer; the purpose of this layer is to summarize the feature maps coming from the convolution layer and reduce their size while preserving important information. In this way, the training time is shortened, thus saving time and resources [31]. These methods are widely used in many areas such as data processing and signal analysis. For example, a maximum pooling layer of 2x2 size takes the maximum value of each 2x2 block in the image matrices, reducing the size and reducing the processing cost. During the pooling process, the "shift" parameter determines how many units the filter will shift on the input data. The size depending on the pooling layer is also very important in controlling the process here. In summary, the pooling layer increases the processing speed of deep learning networks, emphasizes the important features of the data by not performing learning at the same time and increases the generalization ability of the model.

Fully connected layers come after convolution and pooling layers in CNNs. Convolution layers extract important features in the input image. Pooling layers reduce the size of these features and allow the network to do less computation. Fully connected layers take flattened feature maps and use them for classification. Each neuron is connected to each feature in the input. These connections have weights that are adjusted during the learning process of the network.

Dropout layer is a regularization technique used to prevent overfitting in deep learning models. Overfitting is when the model memorizes patterns in the training data and cannot generalize to new, unseen data. The most important requirement of the dropout layer is to prevent overfitting. Dropout works by temporarily removing randomly selected neurons from the network at each iteration during training. This prevents neurons from becoming too dependent on each other and memorizing the noise in the training data. Since a different subset of neurons will be active in each iteration, the model becomes more robust and has a higher ability to generalize.

In CNNs, visual data is processed by convolution and pooling operations. The resulting feature maps are combined and flattened and given as input to the fully connected layer. In this way, the high-dimensional information in the image is represented in a lower-dimensional vector and can be used for tasks such as classification.

ResNet is a deep learning model developed by Microsoft Research in 2015 and is based on the Convolutional Neural Network (CNN) architecture. It has achieved significant success in image classification and object detection tasks, especially on the ImageNet dataset. ResNet is based on the principle of "residual learning". In traditional deep neural networks, each layer receives and processes the output of the previous layer. This structure can lead to the "vanishing gradient" problem in deep networks. Gradient vanishing makes it difficult to update the weights in the lower layers of the network and slows down the learning process. ResNet uses "skip connections" or "residual connections" to solve this problem. These connections directly transmit the output of one layer to several layers later. Thus, gradients can be propagated more effectively to all layers of the network, making it easier to train deep networks.

ResNet-18, as its name suggests, consists of 18 layers. These layers consist of convolutional layers, batch normalization layers, and ReLU activation functions. Each residual block contains two or three convolutional layers. ResNet-18 can be used in various tasks such as image classification, object detection, and image segmentation. It is also often preferred for transfer learning. ResNet-18's fast trainability and good performance have made it a popular model for many applications. It has also formed the basis for the development of larger and deeper ResNet variants such as ResNet-50, ResNet-101, and ResNet-152.

ResNeXt is a convolutional neural network (CNN) model developed by Facebook AI Research in 2016 that has had a significant impact in the field of deep learning, especially in image recognition tasks. An extension of the ResNet architecture, ResNeXt improves the learning capacity and performance of the network by using a technique called "grouped convolutions." The basic idea of ResNeXt is to divide a convolution layer into multiple "groups" of smaller filters. Each group works on a subset of the input channels and learns its own filters. The outputs of the groups are then combined to obtain the result.

2.3. Performance Measures

In the context of deep learning, different metrics are used to measure and compare the performance of models used especially in classification problems. These metrics measure the agreement between the predicted values of the model and the real values and play an important role in evaluating the success of the model. Each metric has its own advantages and disadvantages, and the performance of a model may vary according to different metrics. For example, a model may achieve a high value in the accuracy metric, but may perform lower in the precision or recall metrics. Therefore, it is important to consider more than one metric when evaluating the performance of the model. In addition, the success of the model depends not only on the metrics used, but also on how the training and test data are divided and the distribution of classes in the classification problem. In imbalanced datasets, some classes are more represented than others, and this may affect the performance of the model and the evaluation metrics. In this study, a number of metrics were used to evaluate the performance of deep learning models.

2.3.1. Confusion Matrix

Confusion matrix is an important tool used to evaluate the performance of machine learning and deep learning models [33]. It is used to understand the accuracy of the model's predictions and the types of errors, especially in classification problems. The confusion matrix is created by comparing the true class labels with the labels predicted by the model. These values are used to calculate different metrics to evaluate the performance of the model. For example, metrics such as accuracy, precision, recall, and F1-score can be calculated from the values in the confusion matrix. The confusion matrix can also be used to understand the strengths and weaknesses of the model. For example, a high false positive rate indicates that the model is incorrectly classifying too many examples as positive. This information can help determine which areas to focus on for model improvement. The confusion matrix is a standard method used to measure the performance of a classification model on a test dataset where the true values are known [34]. This matrix allows you to analyze the overall accuracy of the model as well as different types of errors. Table 3.1 shows the confusion matrix. In summary, the confusion matrix is a powerful tool used to evaluate and understand the performance of deep learning and machine learning models. By analyzing the accuracy of the model's predictions and the types of errors, it guides the model to improve and make more accurate predictions. The metrics are calculated using the formulas below:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Specificity = \frac{TN}{TN + FP} \tag{4}$$

$$F1 - score = \frac{2 * recall * precison}{recall + precison} \tag{5}$$

2.3.2. ROC Curve

The ROC curve (Receiver Operating Characteristic) is a graphical tool used to evaluate the performance of a classification model in binary classification problems [35]. This curve shows the relationship between the true positive rate (sensitivity) and the false positive rate (1 - specificity) at different threshold values. The true positive rate is the rate at which samples that are actually positive are correctly classified as positive. The false positive rate is the rate at which samples that are actually negative are incorrectly classified as positive. In the ROC curve, the x-axis shows the false positive rate, and the y-axis shows the true positive rate. The curve consists of the union of points connecting the sensitivity and specificity values obtained at different threshold values. The area under the ROC curve (AUC) is a value that summarizes the overall performance of the model. If the AUC value is greater than 0.5, the model is better than random guessing. The performance of the model increases as the AUC value approaches 1. The advantages of the ROC curve include being independent of the threshold value, being robust to class imbalances, and being able to be used to compare the performances of different classification models [36]. The ROC curve is used in many fields, such as medicine, finance, and machine learning. In summary, the ROC curve is a powerful tool that visually shows how well a classification model does, i.e. how well it distinguishes positive examples from negative examples. Although the ROC curve is a visual representation, it has mathematical calculations underlying it. These calculations are based on the concepts of true positive rate (TPR) and false positive rate (FPR). While TPR shows how accurately positive class examples are predicted, FPR shows the rate at which negative class examples are falsely classified as positive.

III. Results and Discussion

ResNet models trained with the same hyperparameter values are given in Table 1. These values are the ones that provide the best performance obtained as a result of various experiments. The parameters shown in Table 2 belong to these models.

TABLE I
GENERAL RESNET MODEL ARCHITECTURES

```
x=base_model.output (ResNet18, ResNet50,
   Resnet152, ResNeXt)
x =Global Average Pooling 2D()(x)
x =Dropout (0.3) (x)
x =Dense (128, activation='relu')(x)
    predictions = Dense(7, activation='softmax')(x)
```

TABLE II
RESNET MODELS HYPERPARAMETER VALUES

Hyperparameters	Value
batch_size	64
epoch	100
Metrics	accuracy', 'loss', 'precision', 'recall', 'AUC', 'f1_score'
optimizer	Adam
learning_rate	0.001
loss function	SparceCategoricalCrossentropy

In this study, four different ResNet models, namely, ResNet18, ResNet50, ResNet152 and ResNeXt, were used for the detection of potato diseases, and their performances were compared. The images were modeled for 3 classifications of potato leaves: early blight, late blight and healthy. The accuracy and performance metrics collected for the dataset are given in Table 3 and Table 4, respectively.

TABLE 3
ACCURACY RATES OF THE RESNET MODELS

Model	Resnet18	Resnet50	Resnet152	ResNeXt
Accuracy	0.96	0.97	0.98	0.98

When Table 3 is examined, it is seen that nearly all models generally showed high accuracy performance. ResNet152 and ResNeXt models achieved the highest success with an accuracy rate of 98%, and Resnet18 and Resnet50 with 96% and 97% accuracy, respectively.

TABLE 4
PERFORMANCE METRICS OF THE RESNET MODELS

Model	Early Blight		
	Precision	Recall	F1-score
Resnet18	0.96	0.99	0.98
Resnet50	0.97	0.99	0.98
Resnet152	0.99	0.99	0.99
ResNeXt	0.99	0.99	0.99
Model	Late Blight		
	Precision	Recall	F1-score
Resnet18	0.99	0.92	0.95
Resnet50	0.97	0.97	0.97
Resnet152	0.98	0.97	0.97
ResNeXt	0.99	0.97	0.98
Model	Healthy		
	Precision	Recall	F1-score
Resnet18	0.84	0.97	0.90
Resnet50	0.97	0.89	0.93
Resnet152	0.88	0.97	0.92
ResNeXt	0.85	0.92	0.88
Model	Model Metric Averages		
	Precision	Recall	F1-score
Resnet18	0.93	0.96	0.94
Resnet50	0.97	0.95	0.96
Resnet152	0.95	0.97	0.96
ResNeXt	0.94	0.96	0.95
Model	Model Weighted Averages		
	Precision	Recall	F1-score
Resnet18	0.96	0.96	0.96
Resnet50	0.97	0.97	0.97
Resnet152	0.98	0.98	0.98
ResNeXt	0.98	0.98	0.98

As can be seen from Table 4, it is noteworthy that ResNet152 and ResNeXt models have higher precision and F1-score values among all the models, especially in the "Late Blight" and "Healthy" classes. This shows that ResNet152 and ResNeXt can learn complex features better and classify potato diseases more accurately thanks to their deeper architectures. ResNet18 and ResNet50 models reached 96% and 97% accuracy rates, respectively. Since these models have fewer layers than ResNet152 and ResNeXt, they can be trained faster. However, their performance may be lower in some cases due to their less depth. When evaluated in terms of metrics, all models achieved very high precision values (96% and above) in the "Early Blight" class. This shows that the models are quite successful in detecting early blight, ResNet18 and ResNet50 exhibited excellent performance with a sensitivity rate of 99% in the "Early Blight" class. Sensitivity values for other classes are also generally high. All models achieved high F1-score values for different classes.

As a result, the conducted tests show that deep learning models can be used effectively for the detection of potato diseases. ResNet152 and ResNeXt models achieved the highest accuracy rates and also performed successfully in terms of other metrics. It can be seen that the Resnet18 model lags behind other models. These models can be valuable tools for disease detection and control in potato cultivation.

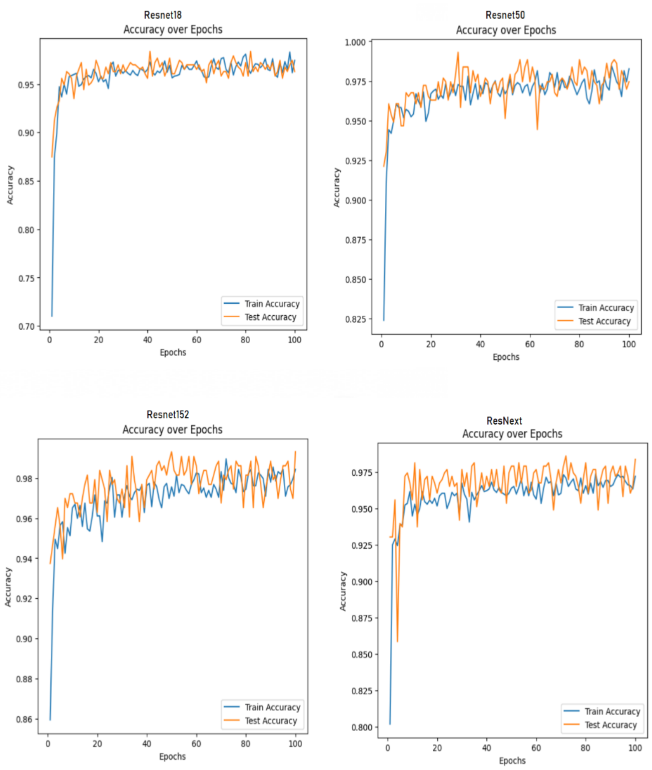


Fig.4. Accuracy of the ResNet models

The results obtained by the models during the training process are examined in detail with the accuracy graphs demonstrated in Fig. 4, and the loss graphs in Fig. 5. These graphs depict how well the models learned and how many errors is made in each iteration. The models were trained for 100 epochs, and early stopping was implemented to monitor validation loss with a patience of 5 epochs, preventing overfitting. The best model weights are saved based on validation loss improvements.

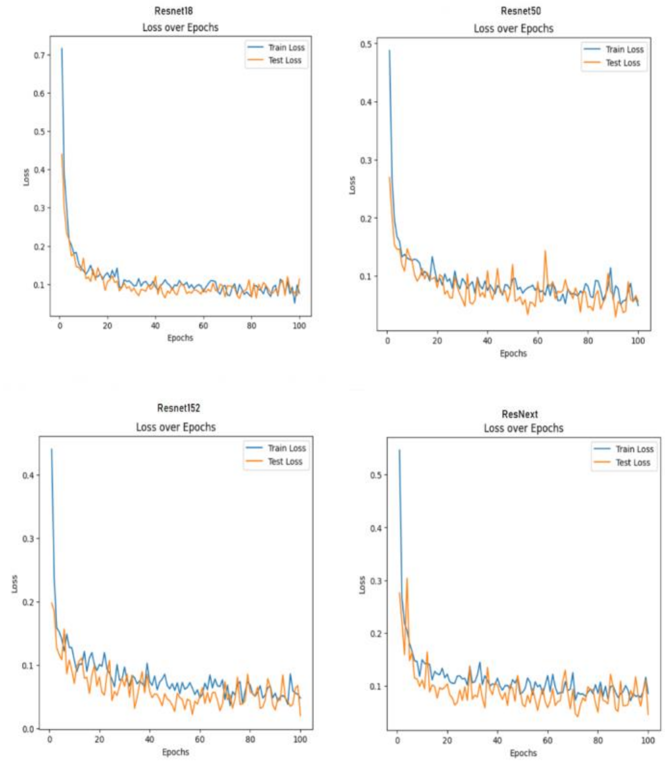


Fig.5. Loss graphs of the ResNet models

The ROC graphs of the models are demonstrated in Figure 6. According to the given graphs, the ROC curves being close to the upper left corner and the AUC values being 1.00 show that the models can distinguish diseases almost perfectly. This shows that the models have both high sensitivity (true positive rate) and high specificity (true negative rate). In other words, they can both correctly detect diseased potatoes and do not misclassify healthy potatoes as diseased. The ROC curves of the ResNet18, ResNet50 and ResNet152 models almost completely overlap and the AUC values are 1.00, indicating that the models are very close to each other in terms of performance. The ROC curve of the ResNeXt model is also very close to the upper left corner and the AUC value is 1.00. However, the curve looks a little "flatter" compared to the other three models. This may indicate that ResNeXt may make slightly more false positive predictions than the other models in some cases.

The confusion matrix can also be used to understand the strengths and weaknesses of the model. For instance, a high false positive rate indicates that the model is incorrectly classifying too many examples as positive. This information can help determine which areas to focus on for model improvement. Figure 7 shows the confusion matrix plots of the evaluated models.

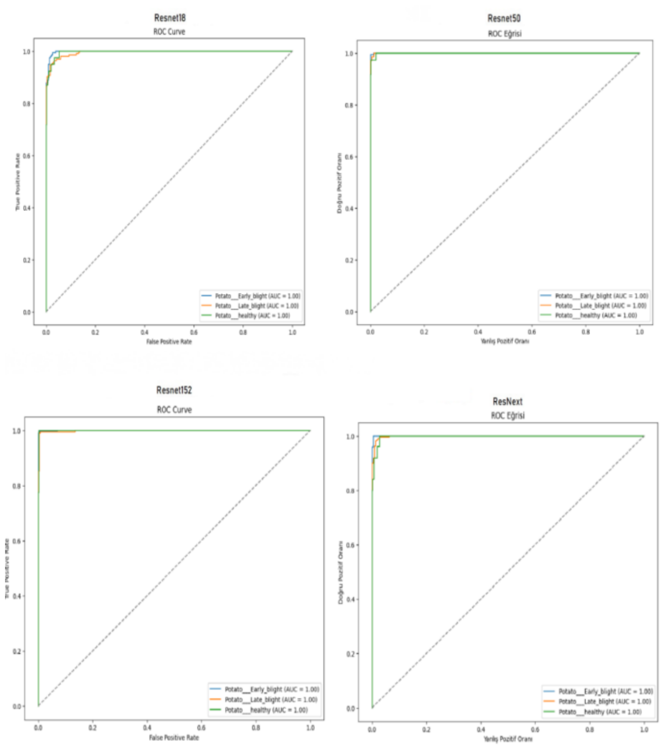


Fig.6. ROC curves of the ResNet Models

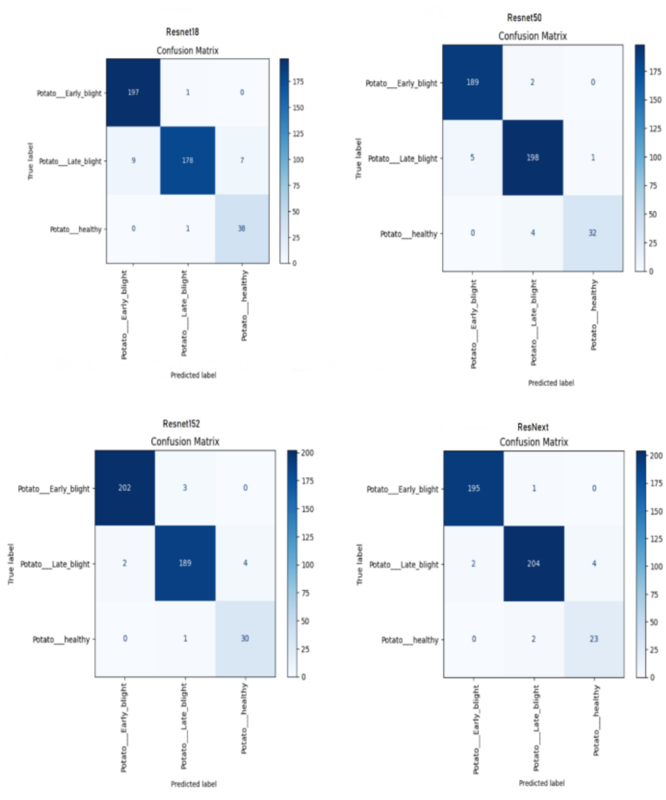


Fig.7. Confusion Matrices

ResNet18 shows high performance in the "Early Blight" class, while making some mistakes in the "Late Blight" and "Healthy" classes. ResNet50 shows a similar performance to ResNet18, but it is observed that it makes fewer mistakes in the "Healthy" class. ResNet152 achieves high accuracy rates for all classes, and it is especially successful in the "Late Blight" class. ResNeXt shows a similar performance to ResNet152, but it makes slightly more mistakes in the "Early Blight" class. In general, it can be said that all models are successful in detecting the "Early Blight" class, but have some difficulties in distinguishing the "Late Blight" and "Healthy" classes. ResNet152 and ResNeXt show higher performance than other models thanks to their deeper architectures.

IV. Conclusion

The aim of this study is to automatically diagnose and classify diseases in potato plants using visual data using deep learning methods. For this purpose, four different models based on ResNet architecture (ResNet18, ResNet50, ResNet152 and ResNeXt) were used and the performances of the models were evaluated with various metrics. The tests performed showed that deep learning models can be used effectively in the detection of potato diseases and that deep models such as ResNet152 and ResNeXt in particular can provide high accuracy rates. It is seen that the highest success rate for the training set belongs to ResNet152 and ResNeXt models. The success rate of both models achieved the highest success with an accuracy rate of 98%. Although the ResNet18 model has lower performance compared to other models, it can be preferred for some applications because it can be trained faster and requires less computational resources. These models show that deep learning methods can achieve high accuracy rates in the detection of potato diseases. In particular, it has been revealed that the deep structure of the ResNet architecture is an effective strategy in learning complex patterns and features. These results show that deep learning methods have high potential for the detection of potato diseases and can be valuable tools to increase efficiency and quality in agricultural production. Besides, running these models on a blended dataset which is prepared from many different datasets, is considered to be as a useful future study.

References

[1] Saleem, M. H., Potgieter, J., & Arif, K. M. (2019). Plant disease detection and classification by deep learning. *Plants*, *8*(11), 468.

[2] Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and electronics in agriculture*, *145*, 311-318.

[3] Shoaib, M., Shah, B., Ei-Sappagh, S., Ali, A., Ullah, A., Alenezi, F., ... & Ali, F. (2023). An advanced deep learning models-based plant disease detection: A review of recent research. *Frontiers in Plant Science*, *14*, 1158933.

[4] Ramesh, S., Hebbar, R., Niveditha, M., Pooja, R., Shashank, N., & Vinod, P. V. (2018, April). Plant disease detection using machine learning. In *2018 International conference on design innovations for 3Cs compute communicate control (ICDI3C)* (pp. 41-45). IEEE.

[5] Shruthi, U., Nagaveni, V., & Raghavendra, B. K. (2019, March). A review on machine learning classification techniques for plant disease detection. In *2019 5th International conference on advanced computing & communication systems (ICACCS)* (pp. 281-284). IEEE.

[6] Ahmed, I., & Yadav, P. K. (2023). Plant disease detection using machine learning approaches. *Expert Systems*, *40*(5), e13136.

[7] Omid, M., Khojastehnazhand, M., & Tabatabaeefar, A. (2010). Estimating volume and mass of citrus fruits by image processing technique. *Journal of food Engineering*, *100*(2), 315-321.

[8] Chauhan, K., Jani, S., Thakkar, D., Dave, R., Bhatia, J., Tanwar, S., & Obaidat, M. S. (2020, March). Automated machine learning: The new wave of machine learning. In *2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)* (pp. 205-212). IEEE.
[9] Blasco, J., Aleixos, N., & Moltó, E. (2003). Machine vision system for automatic quality grading of fruit. *Biosystems engineering*, *85*(4), 415-423.
[10] Gómez, D., Salvador, P., Sanz, J., & Casanova, J. L. (2019). Potato yield prediction using machine learning techniques and sentinel 2 data. *Remote Sensing*, *11*(15), 1745.
[11] Kurek, J., Niedbała, G., Wojciechowski, T., Świderski, B., Antoniuk, I., Piekutowska, M., ... & Bobran, K. (2023). Prediction of potato (solanum tuberosum l.) yield based on machine learning methods. *Agriculture*, *13*(12), 2259.
[12] Gold, K. M., Townsend, P. A., Herrmann, I., & Gevens, A. J. (2020). Investigating potato late blight physiological differences across potato cultivars with spectroscopy and machine learning. *Plant Science*, *295*, 110316.
[13] Kadam, S. U., Dhede, V. M., Khan, V. N., Raj, A., & Takale, D. G. (2022). Machine learning methode for automatic potato disease detection. *NeuroQuantology*, *20*(16), 2102-2106.
[14] Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and electronics in agriculture*, *147*, 70-90.
[15] Lee, H. S., & Shin, B. S. (2020). Potato detection and segmentation based on mask R-CNN. *Journal of Biosystems Engineering*, *45*, 233-238.
[16] Bangari, S., Rachana, P., Gupta, N., Sudi, P. S., & Baniya, K. K. (2022, February). A survey on disease detection of a potato leaf using cnn. In *2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS)* (pp. 144-149). IEEE.
[17] Rozaqi, A. J., & Sunyoto, A. (2020, November). Identification of disease in potato leaves using Convolutional Neural Network (CNN) algorithm. In *2020 3rd International Conference on Information and Communications Technology (ICOIACT)* (pp. 72-76). IEEE.
[18] Xi, R., Hou, J., & Lou, W. (2020). Potato bud detection with improved faster R-CNN. *Transactions of the ASABE*, *63*(3), 557-569.
[19] Gao, W., Xiao, Z., & Bao, T. (2023). Detection and identification of potato-typical diseases based on multidimensional fusion Atrous-CNN and hyperspectral data. *Applied Sciences*, *13*(8), 5023.
[20] Singh, M. K., Kumar, J., Vishwakarma, S., Gupta, V. K., Raghuwanshi, G., Joshi, K., ... & Mishra, K. (2024, March). Identification of Potato Leaves Diseases Using CNN with Transfer Learning. In *2024 2nd International Conference on Disruptive Technologies (ICDT)* (pp. 6-10). IEEE.
[21] Khobragade, P., Shriwas, A., Shinde, S., Mane, A., & Padole, A. (2022, December). Potato leaf disease detection using cnn. In *2022 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)* (pp. 1-5). IEEE.
[22] Griffel, L. M., Delparte, D., & Edwards, J. (2018). Using Support Vector Machines classification to differentiate spectral signatures of potato plants infected with Potato Virus Y. *Computers and electronics in agriculture*, *153*, 318-324.
[23] Islam, M., Dinh, A., Wahid, K., & Bhowmik, P. (2017, April). Detection of potato diseases using image segmentation and multiclass support vector machine. In *2017 IEEE 30th canadian conference on electrical and computer engineering (CCECE)* (pp. 1-4). IEEE.
[24] Ji, Y., Sun, L., Li, Y., Li, J., Liu, S., Xie, X., & Xu, Y. (2019). Non-destructive classification of defective potatoes based on hyperspectral imaging and support vector machine. *Infrared Physics & Technology*, *99*, 71-79.
[25] Shen, D., Zhang, S., Ming, W., He, W., Zhang, G., & Xie, Z. (2022). Development of a new machine vision algorithm to estimate potato's shape and size based on support vector machine. *Journal of Food Process Engineering*, *45*(3), e13974.
[26] Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big data*, *3*, 1-40.
[27] Kulkarni, U., Meena, S. M., Gurlahosur, S. V., & Mudengudi, U. (2019, September). Classification of cultural heritage sites using transfer learning. In *2019 IEEE fifth international conference on multimedia big data (BigMM)* (pp. 391-397). IEEE.
[28] Venkatesan, R., & Li, B. (2017). Convolutional neural networks in visual computing: a concise guide. CRC Press.
[29] Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. (2023). *Dive into deep learning*. Cambridge University Press.
[30] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, *521*(7553), 436-444.
[31] Scherer, D., Müller, A., & Behnke, S. (2010, September). Evaluation of pooling operations in convolutional architectures for object recognition. In *International conference on artificial neural networks* (pp. 92-101). Berlin, Heidelberg: Springer Berlin Heidelberg.
[32] 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* (pp. 770-778).
[33] Narkhede, S. (2018). Understanding confusion matrix. *Towards Data Science*, *180*(1), 1-12.
[34] Powers, D. M. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*.
[35] Bento, C. (2022). ROC analysis and the AUC—Area under the curve. Towards Data Science.
[36] Cook, N. R. (2007). Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation, 115(7), 928-935.

BIOGRAPHIES

Abdulkerim Öztekin was born in Berlin, Germany in 1978. He received the B.S. degree in Electrical and Electronics Engineering from Hacettepe University, Ankara, Turkiye, in 2001, and the M.S. and the Ph.D. degrees in Electrical and Electronics Engineering from Gaziantep University, Gaziantep, Turkiye, in 2018.

He is currently working as an Assistant Professor at Batman University and also serves as the General Manager of Batman University Technology Development Zone. He has several publications in the field and his research interests include signal processing, artificial intelligence, optimization, video coding, and communications.

Kenan Almas received the B.S. degree in Department of Mathematics from Dicle University, Diyarbakır, Turkiye, in 2005 and the M.S. degree in Department of Information Technologies from Batman University, Batman, Turkiye, in 2025. He is currently working as a primary school mathematics teacher at Batman Science and Art Center affiliated with the Ministry of National Education.

Blockchain-Integrated Framework for Data Security: An Application Based on IoT Data and Deep Learning

Doygun Demirol, Resul Daş, Mehmet Özdem, Ceren Nur Cansel, Davut Hanbay

***Abstract*—The rapid development of the IoT (Internet of Things) ecosystem leads to the creation of big data environments that require real-time analysis. In this comprehensive data ecosystem, anomaly detection and data security emerge as critical requirements. This paper presents a comprehensive approach that integrates a deep learning model developed for anomaly detection in IoT network traffic and a blockchain-based data storage structure designed to ensure data integrity. In the research, network traffic data of a sample device from the N-BaIoT dataset is used. The developed deep learning model was able to classify attack and normal traffic patterns with high accuracy. The proposed model achieved an accuracy rate of 99.79% in anomaly detection, demonstrating its effectiveness in classifying IoT network traffic. Data security is ensured with the Fernet encryption algorithm, while data integrity is protected using blockchain technology. Experimental results show that the proposed system achieves significant performance metrics in terms of both anomaly detection accuracy and data security verification. The proposed framework contributes to the development of more secure and reliable IoT systems by providing an innovative solution to anomaly detection and data security challenges in IoT environments.**

***Index Terms*—Internet of Things (IoT), Deep Learning, Network Security, Blockchain, Data Integrity, Anomaly Detection.**

I. Introduction

The Internet of Things (IoT) has become a widely used technology in every aspect of daily life. IoT devices have revolutionised many different sectors, from home automation to healthcare, from energy management to smart city applications. These devices have become an important component of the big data ecosystem by continuously generating data. However, this large data ecosystem has also brought security and privacy risks. The vulnerability of IoT devices to cyber attacks has made data security and data integrity issues critical.

In this study, an innovative approach is proposed to detect anomalies in IoT network traffic and ensure data security.

Doygun Demirol is with the Department of Computer Technologies, Bingol University, Bingol, 12000 Türkiye. Email: ddemirol@bingol.edu.tr

Resul Daş is with the Department of Software Engineering, Technology Faculty, Firat University, Elazig, 23119 Türkiye. Email: rdas@firat.edu.tr

Mehmet Özdem Türk Telekom, Ankara, Türkiye. Email: mehmet.ozdem@turktelekom.com.tr

Ceren Nur Cansel is with the Industrial Engineering, College of Engineering, Koç University, Istanbul, 34450 Türkiye. Email: ccansel19@ku.edu.tr

Davut Hanbay is with the Department of Computer Engineering, Faculty of Engineering, Inonu University, Malatya, 44000 Türkiye. Email: davut.hanbay@inonu.edu.tr

Manuscript received Jan 27, 2025; accepted March 5, 2025.
DOI: 10.17694/bajece.1625208

Network traffic obtained from IoT devices is analysed with a deep learning based model and normal and attack traffic are successfully classified. To ensure confidentiality in data transmission, encryption operations are performed with a confidentiality algorithm and data integrity is guaranteed by using a blockchain-based structure. Blockchain provides a structure that prevents unauthorised access to data while ensuring secure storage of data.

In this study, a dataset containing the network traffic of a device within the N-BaIoT dataset is used. The data is cleaned, scaled and labelled as attack/normal traffic in preprocessing steps. Then, a deep learning based model is trained to detect anomalies. In the simulation environment, the test data is encrypted at random time intervals and transmitted to a central server and data integrity is ensured by using a blockchain-based structure.

The results obtained show that the proposed system exhibits a high performance in terms of both anomaly detection accuracy and data security. This study provides an effective solution for eliminating security vulnerabilities in IoT networks and increasing data security. Future studies can investigate the applicability of the proposed system on different IoT devices and the scalability of the blockchain structure.

II. Background

A. Key Stages of Big Data Analysis

Nowadays, it has become important to understand the meaning and importance of data and to produce and collect data. Understanding data correctly plays an important role in decision-making. Big data analysis is the process by which algorithms are applied to data to extract useful and unknown patterns, relationships, and information [1]. Extracting and analyzing useful information from big data sets requires smart and scalable analysis services, programming tools, and algorithms [2]. Also, big data analytics is used to extract previously unknown, useful, valid, and hidden patterns and data from large data sets and identify important relationships between stored variables. This section examines four main phases of big data analysis processes, whose infographic are examined. The key stages of big data analytics in Figure 1 are presented.

1) Data Collection: For big data analysis, data can be collected from semi-structured sources such as data warehouses, business transactions, sensors, log files, unstructured sources such as social media platforms, and structured sources such

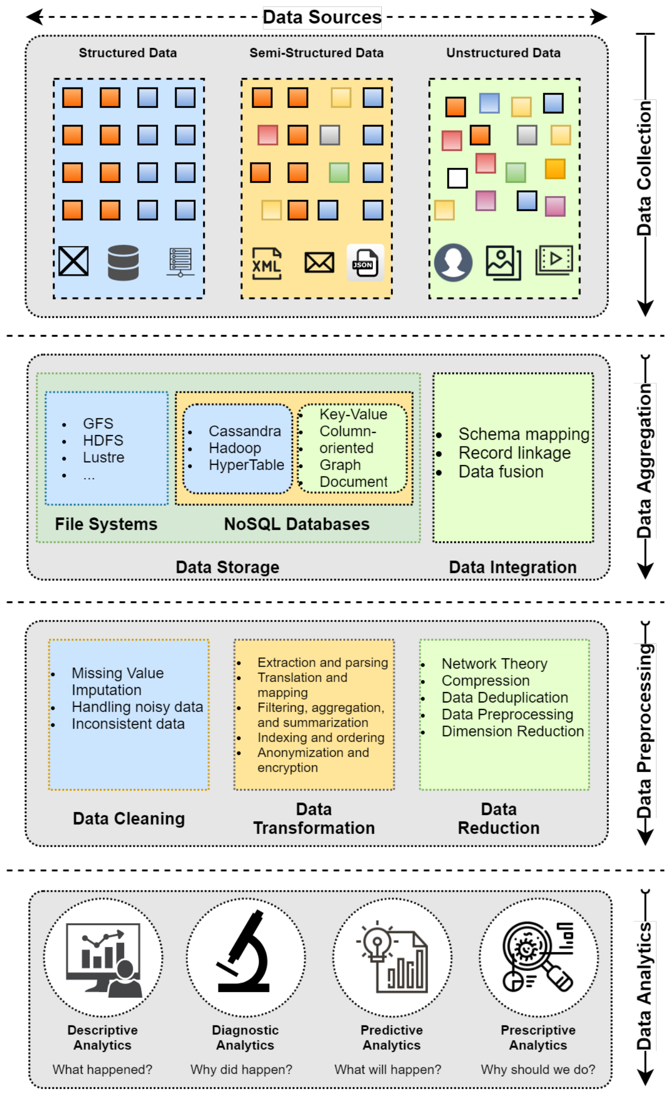


Fig. 1. Key stages of the big data analysis.

as multimedia resources. In this step, obtaining reliable data is critical for the analysis process.

2) Data Aggregation: The accuracy of useful information obtained from data analysis depends on the sufficient amount of data used in the analysis and the accuracy of high-quality data. Data aggregation is a crucial step in big data analytics, as the transform of different data collected from multiple sources into a more readable and analyzable format in the data collection step directly affects the quality and accuracy of the data to be used in the analysis.

3) Data Preprocessing: Due to the wide variety of data sources, undesirable factors may negatively affect data quality, such as noise, inconsistency, and missing information in the collected datasets. Storing these meaningless data is unnecessary both in terms of storage cost and analysis quality. Therefore, data must be preprocessed to perform effective data analysis. Pre-processing of data not only increases storage costs but also increases analysis accuracy [3].

4) Data Analytics: Analyzing big data sets directly without realizing other phases means time-consuming navigation for users across a very large search area. In addition, to analyzing big data quickly, it is also very important to analyze the data on time and to use the right type of analysis. Four types of data analytics are taxonomied for big data analysis: Descriptive, Diagnostic, Predictive, and Prescriptive.

B. Big Data Storage Technologies

Sensor networks, scientific experiments, websites, and many other applications generate semi-structured or unstructured data in various forms [4]. The transition from structured data to unstructured data renders traditional databases useless for storage [5]. This disadvantage of traditional databases contributes to the development of efficient and robust storage systems, and many new technologies are being developed. Ensuring high scalability, reliability, robust, and efficient storage for rapid growth data is the main goal of technologies designed for Big Data storage [6].

This section briefly describes, categorizes, and compares storage technologies developed for Big Data.

1) Column-Based Databases:

- **BigTable**: BigTable is a flexible, reliable and applicable storage system developed by Google for managing petabyte-scale data on thousands of machines to provide high storage capacity for huge amounts of structural data. It also provides dynamic control over data positioning, viewing, and clustering [7].
- **HBase**: HBase is a column-based data storage technology developed by Apache and designed to address the storage requirements of Big Data of the Apache project [8]. It uses HDFS storage technology to store table data and a clustering approach for data management. Therefore, it is a potential solution for projects that contain many rows of data.
- **Hypertable**: Hypertable is a distributed database that provides support for the consistency of stored data. It is designed to work with many distributed file systems such as GFS, HDFS, and CloudStore. It stores the data in tabular form and manages the tables by splitting them to ensure scalability with distribution.
- **Cassandra**: Cassandra is a decentralized, highly accessible, and structured key-value storage system [9]. It provides scalable storage and improved performance for applications that perform intensive reading and writing operations. Fault tolerance, reduced latency, scalability, clustering, segmentation and replication are the highlights of Cassandra [10]. The social media platform Facebook where scalability depends linearly on the number of users and requires intensive reading and writing operations, uses Cassandra for messaging applications [11]

2) Document-Based Databases:

- **MongoDB**: MongoDB, an open source and free to use under the GNU AGPL license is a document-based NoSQL database that stores each entry in JSON documents. It has a query language that uses JSON structure.

It is also named a grid-based file system (GridFS), which allows storage of large objects [12].

- **Terrastore**: Terrastore is a document-based, open-source, distributed storage system developed by Terracotta Inc. hlthat provides scalability, consistency, and dynamic clustering support while running. It automatically distributes data through server nodes and has the ability to redistribute data automatically when new servers are included or servers are removed. It also has replication and switching as backup system features [13].
- **DynamoDB**: DynamoDB is a distributed, schemaless NoSQL database technology used by Amazon [14]. It is widely used to store unstructured, mutable and scalable data. It provides infinite capacity for data storage and access rate. It offers updating, efficient indexing, adaptability and scalability in distributed systems. Therefore, it is suitable for rapidly growing data and scalable applications [15].
- **RethinkDB**: RethinkDB is the first open-source document-based database that efficiently supports complex queries for real-time web applications [16]. It successfully responds to real-time requests from users and also uses the JSON data structure.
- **OrientDB**: OrientDB is the first multi-model, open source, and highly scalable database technology for document data with an extended and transparently managed graph layer that provides relationships between records. Thanks to the embedded graphic layer, data circulation and data relationship management can be implemented quickly and without increasing costs. OrientDB, supports all operating systems, has features such as clustering and replication on heterogeneous servers to ensure performance and scalability [17].
- **CouchDB**: NoSQL document-based database using JSON, JavaScript for MapReduce queries, and HTTP for its API, known for replication and offline capabilities.

3) Graph-Based Databases:

- **HyperGraphDB**: HyperGraphDB is an open source graphical database system designed for artificial intelligence and web semantics projects [18].The storage infrastructure is provided by Berkeley DB. Also, it has a data mapping system among storage systems and hosts. It uses a key-value mechanism to store graph information [19]. Unlike other master-slave storage technologies, HyperGraphDB uses peer-to-peer data distribution architecture. It has features such as being able to run each peer independently and performing updates asynchronously [18].
- **Neo4j**: Highly popular, open-source, providing ACID(Atomicity, Consistency, Isolation, Durability) compliance, widely used in social networks, fraud detection, and network management.

4) Key-Value Stores:

- **Aerospike**: Aerospike is an open-source data storage technology for real-time data that offers scalability and reliability at low cost. It has a ”share-nothing” architecture that supports petabyte-scale data volumes with reliability and scalability.
- **Redis**: In-memory key-value store known for high performance and support for multiple data structures, often used for caching, real-time analytics, and message brokering.

5) Time-Series Databases:

- **InfluxDB**: Open-source time-series database designed for high write and query loads, used in monitoring and real-time analytics.
- **OpenTSDB**: Scalable, open-source time-series database built on top of HBase, optimized for storing and serving large amounts of time-series data.
- **Prometheus**: Open-source system monitoring and alerting toolkit, designed for reliability and simplicity.
- **TimescaleDB**: Open-source time-series database optimized for fast ingest and complex queries, built as an extension to PostgreSQL.

6) NewSQL Databases:

- **Google Spanner**: Globally distributed, strongly consistent database service by Google, offering the scalability of NoSQL with the transactional consistency of SQL.
- **CockroachDB**: Open-source, distributed SQL database designed for high availability and horizontal scalability.
- **VoltDB**: High-speed, in-memory, NewSQL database designed for real-time analytics and decision-making.
- **NuoDB**: Distributed SQL database providing horizontal scalability and continuous availability.

7) Object-Oriented Databases:

- **db4o**: Open-source object database for Java and .NET developers, allowing the storage and retrieval of complex data types.
- **ObjectDB**: High-performance object database for Java, providing a simple and intuitive API.
- **InterSystems Caché**: High-performance, multi-model database that supports object and relational data models.

8) Multi-Model Databases:

- **ArangoDB**: Native multi-model database supporting graph, document, and key-value data models.
- **OrientDB**: OrientDB is the first multi-model, open source, and highly scalable database technology for document data with an extended and transparently managed graph layer that provides relationships between records. Thanks to the embedded graphic layer, data circulation and data relationship management can be implemented quickly and without increasing costs. OrientDB, supports all operating systems, has features such as clustering and replication on heterogeneous servers to ensure performance and scalability [17].
- **MarkLogic**: Enterprise NoSQL database with multi-model capabilities, providing a single platform for data integration.
- **Couchbase**: Distributed NoSQL cloud database supporting multiple data models including document, key-value, and SQL for JSON.

9) Distributed File Systems:

- **Google File System (GFS)**: Proprietary distributed file system developed by Google to provide efficient, reliable access to data using large clusters of commodity hardware. Google File System (GFS) is a storage system developed by Google Inc. to manage data-intensive applications. Along with its innovative features, it is designed to meet the ever-increasing data storage requirements. For large files, GFS guarantees that writing new data simultaneously, even when reading and writing are intensive [20].

- **Hadoop Distributed File System (HDFS)**: Open-source distributed file system designed to run on commodity hardware, part of the Apache Hadoop project, and used for reliable, scalable, and distributed storage. Hadoop Distributed File System (HDFS) inspired by Google File System and meets the data storage requirements of Hadoop applications for Map-Reduce [21]. It is a suitable solution that works with high efficiency in Big Data for data-intensive applications. Although it contains a large number of components, it has fast error detection and automatic recovery features.

C. Big Data Applications

Today, there are many business and industrial areas where big data analysis is applied. These sectors generate large amounts of Big Data that must be analysed for effective and efficient decision-making processes. Some of these areas of application are shown in Fig. 2 and are explained below.

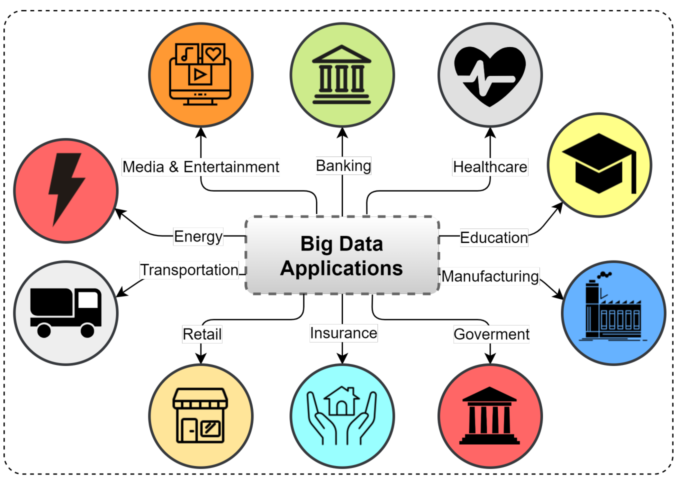


Fig. 2. Important applications fields of the big data

1) Health Applications: Big Data is a part of the healthcare industry, as it is in many other industries, and is actively used in healthcare applications. Big Data are generated from heterogeneous sources such as laboratory records, patient symptoms, devices used, and patient health record data in healthcare applications. These data are crucial to facilitating and improving decision-making for both physicians and other healthcare employees. Advanced analysis of centralized medical data provides great benefits to healthcare. For example, doctors can monitor patients' symptoms and prescribe the medication in real time [22].

Today, many applications are developed with the analysis of big data gathering from the healthcare sector. The analysis of Big Data plays a vital role in predicting the outbreak of diseases such as the Ebola virus, using patient records and sensor data to improve the quality of healthcare, and providing a feedback mechanism [23], analyzing of genome [24]. Applications for the Covid 19 virus epidemic, which is still having its effect, are also developed [25]. Through these applications, Big Data is in many steps from the identification of infected cases to analyzing the risks, storing the travel history of the people and identifying the people who may be in contact with the infected patient, keeping the data of the patients' fever and other symptoms, whether medical intervention is required, and the detection of the virus at early stages, are used [26].

2) Recommendation Systems: Recommendation systems play an important role in our daily life. These systems provide users with recommendations based on their interests. YouTube, for example, analyzes the user's Watch history and predicts in which category to watch videos in the future, and offers video suggestions accordingly. Similarly, e-commerce applications analyze users' product search and review histories and offer product recommendations to users.

3) Smart Cities: Smart Cities is a concept that encompasses the economy, governance, mobility, people, environment and life [27]. In order to increase the quality, performance, and interaction of urban services, it is necessary to use information technology in smart cities. Big Data technologies are used in smart cities in a variety of fields, including traffic statistics, smart agriculture, healthcare, transportation, and some other purposes [28]. In smart cities, data is collected from power poles, water pipes, buses, trains, and traffic lights with sensors on them [29]. With the data collected, faults in the specified systems can be prevented in advance or the faults that occur in the future can be fixed. Also, energy can be managed according to needs, and attacks on smart grids or possible attempted attacks can be prevented in advance.

4) Shipping and Logistics: Public transport companies use GPS devices in vehicles to monitor their vehicles and schedule transportation. With big data analysis, data such as using different routes, identification of optimized routes, passenger travel frequencies, and the number of passengers are collected with GPS devices and processed and analyzed with big data analysis methos. Various real-time systems not only provide advice to passengers but also provide useful information about when to wait for the next vehicle that will take them to the right destination. In India, for example, which has one of the largest railway networks in the world, the total number of reserved seats are allocated every day is around 250,000 and can be booked 60 days in advance. It is quite complicated to make predictions on such data due to factors such as weekends, festivals, night trains, starting or intermediate stations, etc. Considering the factors mentioned above, the calculation of the probability of approval of the ticket to be purchased for any train to solve the passengers' ticket purchase decision problem is performed using Big Data analysis methods [30].

5) Social Networks and Media: Social media platforms, which consist of many Internet applications such as social and business-oriented networks, online mobile photo and video sharing services, are another data source that requires real-time data processing. Analyzing and processing of Big Data from these sources requires a range of methods and algorithms such as text analysis, sentiment analysis, information fusion, and network analytics [31]. Many governments use sentiment analysis on social media data to predict election outcomes by monitoring political trends.Such systems, where a lot of social media data is aggregated and used, also provide to identify national and local problems [32].

Apart from the areas mentioned in Section II-C, there are many business sectors and industries that benefit from the possibilities of big data analysis today. These industries that use Big Data produce huge amounts of data that require Big Data analytics for effective and efficient decision making.

D. Privacy of Big Data

People around the world use various web services and networks such as social media, e-government, e-health, etc. Users save personal information such as name, surname, user name, password, contact information, address, phone number, user behavior, habits, political or religious tendencies, and locations in these systems. The use of such services increases risks such as user privacy and data disclosure. Companies use these personal, sensitive data which can be obtained by using various data mining methods to determine customer needs and customer habits and to get more benefits from customers. As a result, hackers can pose a security threat at any phase of the big data life-cycle to compromise data security or privacy.

1) Protecting Privacy of Big Data: Data analysis activities lead to data privacy issues throughout the life cycle of big data. Although there are privacy protection laws in many countries, malicious actors violate the privacy of personal information. For example, most mobile applications require access permission for contact lists, files, cameras, microphones, etc. Users also often accept all terms and conditions without reading the privacy terms. Such permissions may lead to the leakage of vital information about individuals and breach of data privacy. In this section, preservation methods used in the privacy of personal information categorized and summarized under three main headings. These categories are anonymization, encryption, and perturbation methods.

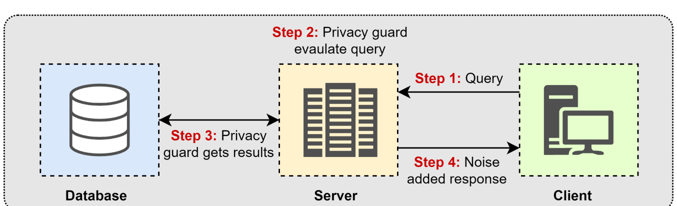


Fig. 3. Differential privacy mechanism.

2) Anonymization Techniques: Anonymization techniques preserve privacy by masking or altering personal, sensitive data. As with most data applications, raw data should first be preprocessed to secure sensitive information. In anonymization approaches, basically, data are represented in four categories. Explicit Identifiers are the set of attributes that directly concern individuals, such as name, surname, and telephone number. The Quasi Identifier (QID) is a set of attributes such as age, gender, postal code, which can identify individuals when combined with other attributes, although not enough by itself to identify individuals. Sensitive attributes (SA) is an attribute set that includes sensitive information about individuals as disease and income information. Non-Sensitive attributes (NSA) include attributes not fall into the other categories. This section discusses K-Anonymity, L-Diversity, T-Closeness, and Differential Privacy privacy-preserving methods [33].

K-Anonymity [34], is developed to hide information specific to individuals that show singular characteristics in a data set. In the K-anonymity algorithm, suppression and generalization methods are used to anonymize the data. In the suppression method, values of categorical variables such as names are completely removed from the data set. With the generalization method, quantitative variables such as age or height are expressed within a range. This makes each record in a data set indistinguishable from other records with at least (k-1). One of the biggest disadvantages of the K-anonymity method is that it is possible to extract identity from the data set if certain properties are already known.

Table I shows non-anonymized dataset of the fictional patient records. In suppression method, ”Name” and ”Religion” columns are replaced with ”*” character as its shown in Table II. In generalization method attributies are replaced a specific value with a more generic one. In Table I, values in ”age” column replaced with a wider range. As a result Table I has become 2-anonymity. Any combination of ”Age”, ”Gender” and ”Country” attributes is any row of the Table I that exists at least two rows with these attributes.

TABLE I
NON-ANONYMIZED OF FICTIONAL PATIENT RECORDS [?].

Name	Age	Gender	Country	Religion	Disease
Ramsha	30	Female	Tamil	Hindu	Cancer
Yadu	24	Female	Kerala	Hindu	Viral infection
Salima	28	Female	Tamil	Muslim	Tuberculosis
Sunny	27	Male	Karnataka	Parsi	No illness
Joan	24	Female	Kerala	Christian	Heart-related
Bahuksana	23	Male	Karnataka	Buddhist	Tuberculosis
Rambha	19	Male	Kerala	Hindu	Cancer
Kishor	29	Male	Karnataka	Hindu	Heart-related
Johnson	17	Male	Kerala	Christian	Heart-related
John	19	Male	Kerala	Christian	Viral infection

TABLE III
L-DIVERSITY ORIGINAL DATASET.

Zip Code	Age	Country	Disease
25044	38	Hindu	Heart-related
25059	39	Japan	Heart-related
25059	31	Chinese	Viral infection
25044	33	Japan	Viral infection
26844	60	English	Cancer
26844	65	Hindu	Heart-related
26841	67	Japan	Viral infection
26841	69	Japan	Viral infection
25044	51	Japan	Cancer
25044	57	English	Cancer
25059	56	Turkish	Cancer
25059	55	Japan	Cancer

TABLE II
ANONYMIZED METHODS APPLIED RECORDS [?].

Name	Age	Gender	Country	Religion	Disease
*	$20 < Age \leq 30$	Female	Tamil	*	Cancer
*	$20 < Age \leq 30$	Female	Kerala	*	Viral infection
*	$20 < Age \leq 30$	Female	Tamil	*	Tuberculosis
*	$20 < Age \leq 30$	Male	Karnataka	*	No illness
*	$20 < Age \leq 30$	Female	Kerala	*	Heart-related
*	$20 < Age \leq 30$	Male	Karnataka	*	Tuberculosis
*	$Age \leq 20$	Male	Kerala	*	Cancer
*	$20 < Age \leq 30$	Male	Karnataka	*	Heart-related
*	$Age \leq 20$	Male	Kerala	*	Heart-related
*	$Age \leq 20$	Male	Kerala	*	Viral infection

L-Diversity [35] is a method developed to overcome the disadvantages of k-anonymity. L-Diversity model is developed to diversify SA attributes of k records in QID groups created in the k-anonymity approach. If SA attributes in QID groups are combined with other information, it is possible to extract sensitive information. To prevent this problem, the L-Diversity model is developed by diversifying the "l" piece of the SA attributes. In Table III, patients' private attributes(name, gender etc.) are hidden to apply anonymization method to avoid leak. But other attributes are given like zip code, age, country, and disease. So, adversaries may be possible to detect patients' private information. In Table IV, k=4 anonymity method applied to avoid this problem. However, if the last 4 records are examined, it is seen that everyone whose zip code starts with 250 and who is in their 50s has cancer. If a malicious person has this information and additional information, it may be possible to guess that any of the patients has cancer. For this reason, the confidentiality of personal information is ensured by the l-diversity method, which is applied by creating 3 variations within each group of 4 in Table V.

T-Closeness [36] model is proposed by Li et al. because k-anonymity and l-Diversity approaches could not provide enough protection. The proposed model guarantees that the semantic closeness of SA attributes in each QID group is t-closeness to each other also in SA attributes as in the QID group. In other words, the T-Closeness approach argues that the distribution of SA attributes in the data set should also be preserved in QID groups [37].

TABLE IV
K=4 ANONYMIZATION APPLIED DATASET.

Zip Code	Age	Country	Disease
250**	< 40	*	Heart-related
250**	< 40	*	Heart-related
250**	< 40	*	Viral infection
250**	< 40	*	Viral infection
268**	≥ 60	*	Cancer
268**	≥ 60	*	Heart-related
268**	≥ 60	*	Viral infection
250**	≥ 60	*	Viral infection
250**	5*	*	Cancer
250**	5*	*	Cancer
250**	5*	*	Cancer
250**	5*	*	Cancer

TABLE V
K=4 ANONYMIZATION AND L=3 DIVERSITY APPLIED DATASET.

Zip Code	Age	Country	Disease
2505*	≤ 60	*	Viral infection
2505*	≤ 60	*	Heart-related
2505*	≤ 60	*	Cancer
250**	≤ 60	*	Cancer
2685*	> 60	*	Cancer
2585*	> 60	*	Heart-related
2685*	> 60	*	Viral infection
250**	> 60	*	Viral infection
2506*	≤ 60	*	Heart-related
2504*	≤ 60	*	Viral infection
2506*	≤ 60	*	Cancer
250**	≤ 60	*	Cancer

Differential Privacy [38] minimizes the chance of personal information being identified. It also maximizes the accuracy of queries from databases. With this technique, researchers can make queries from databases containing personal information and protect the privacy of personal information. As seen in Fig. 3, in the first step, the client sends a query. In the second step, the privacy protection tool between the client and the Database evaluates this query for risk. In the third step, the privacy protection tool retrieves the results from the database, and in the fourth step, the server sends the noise added results to the client. The amount of noise appended to pure data is directly related to the assessed privacy risk. If the risk of data privacy is low, the added noise is too small to influence the quality of the result set. However, in the opposite case, the amount of noise to be added is large enough to protect individual privacy.

3) Data Perturbation Techniques: Data perturbation techniques protect the privacy of personal data by applying systematic changes to the records in the database. Data perturbation can be defined in three categories as input, output, and algorithm perturbation according to the phases to which noise is added. The phase where noise is added before performing any computation is called input and the phase where noise is added to an algorithm result is called algorithm, and the phase where the noise is added directly to the data during the algorithm and computations is called the output. Data perturbation techniques apply certain changes to the original data, such as adding randomization and noise to protect privacy. *Randomization* is a technique for privacy-preserving data generally defined as adding noise to the data to mask the attribute values of records with the probability distribution [39]. This technique is generally used in surveys, emotion analysis, etc. Records in the data set are evaluated independently from each other. The method can be implemented during data collection and preprocessing. It is stated in [40], that the randomization technique is proved to be not suitable for large data sets due to its time complexity and data utility. These techniques are known to have lower computational complexity than cryptographic methods to protect privacy. The dataset to which the technique is implemented is generally not distinguishable from the original dataset [41]. The most major disadvantage of data perturbation techniques is that they cannot process high volumes of data effectively. This technique provides sufficient privacy on the data, but it takes a significant amount of time to get excellent results [42], [43].

4) Cryptographic Techniques: In the context of big data, the most crucial issue to preserve the privacy of data is to ensure communication security. Encryption methods meet communication security to a great extent [44]. However, it is very difficult to encrypt large-scale data using traditional encryption techniques. So, privacy protection methods based on data encryption are mostly used for distributed applications. Ensuring the security of sensitive private data in the cloud and other big data storage platforms is generally possible today with methods such as homomorphic encryption, attribute-based encryption, and image encryption. Although homomorphic encryption allows some operations without decrypting the encrypted data, the computational efficiency and scalability of this encryption method need improvement to handle large data [45], [46]. In order to address the complex access control mechanism and fine-grained data sharing issue over encrypted data, Sahai et al. [47] introduced an attribute-based encryption (ABE) concept. Next, ABE is categorized into Key-Policy ABE (KP-ABE) [47] and the Ciphertext Policy ABE (CP-ABE) [48]. KP-ABE technique uses identifier attributes to encrypt data and can be accessed if the user's key contains an access policy that includes which user is allowed to access the data [49]. In the CP-ABE technique, users' secret keys are associated with their attributes, access control policies are defined by the data owner, and ciphertext is generated under the access structure [50]. An image encryption algorithm transforms users' private images into texture-like or noise-like encrypted images, namely, incomprehensible form. In this way, encrypted images become withstand various attacks by hackers. So, they can not obtain the original image without the correct secret keys.

III. Literature Review

Anomaly detection and data integrity in IoT systems are of critical importance in the field of cyber security. In this context, the integration of machine learning and deep learning-based methods with blockchain technology offers promising approaches to improve the security of IoT networks. There are various studies in this direction in the literature. In this study, the literature is reviewed and categorised under three headings: machine and deep learning based anomaly detection, ensuring data integrity with blockchain technology and integration of machine learning, deep learning and blockchain.

Anomaly detection in IoT networks plays a critical role in terms of security and performance of the systems. For this purpose, various machine learning and deep learning models are used [51]. Satılmış et al. [52] investigated machine learning and deep learning models proposed to detect anomaly-based attacks in IoT networks. In their study, the advantages and disadvantages of the models used are discussed. Gokdemir et al. [53] investigated the effectiveness of deep learning methods for anomaly detection in IoT time series data. In their study, it was shown that LSTM (Long Short-Term Memory) networks can successfully detect anomalies in time series data. In another study [54], the authors propose a deep learning based model to detect and classify anomalies in IoT networks. The proposed model is based on Residual Networks and Bi-directional GRU architectures and can fully utilise the spatial and temporal characteristics of network traffic data. In addition, the attention mechanism is used to extract key features to improve the performance of the model. Experimental results show that the proposed model performs well in anomaly detection and classification. Ferrag and Maglaras [55] developed a blockchain-based collaborative model for anomaly detection in IoT networks. The study provides a low-cost and highly accurate approach for resource-constrained IoT devices while improving data integrity and inter-device security with blockchain technology. Golomb et al. [56] developed a lightweight framework called CIoTA for anomaly detection in IoT devices. This framework provides a reliable collaboration between IoT devices and a continuously updated anomaly detection model using blockchain technology. The study shows that devices with limited resources can perform anomaly detection securely within a distributed structure. Diro et al. [57] extensively investigated machine learning-based solutions for anomaly detection in IoT networks. While addressing the advantages and limitations of existing algorithms, the study emphasises that blockchain-based anomaly detection systems can be effectively used by integrating with machine learning models. Emec and Özcanhan [58] developed a BLSTM-GRU based hybrid deep learning model for intrusion detection in IoT networks. The model was tested on CIC-IDS-2018 and BoT-IoT datasets and showed superior performance with 98.78% and 99.99% accuracy rates, respectively. Dongxing et al. [59] developed a blockchain-based mechanism for authentication and security of IoT devices. In the study, a unique ID is created for each device and stored on the blockchain, eliminating the dependency on centralised authorities. The proposed system has been tested using the Hyperledger Fabric platform and shown to provide a successful solution for ensuring the data integrity of IoT devices. In another study [60] developed a blockchain-based data protection framework to secure IoT data in untrusted storage. The proposed framework aims to reduce storage overhead and increase security through lightweight verification structures and smart contracts on the blockchain. Simulation results show that the system effectively reduces the storage overhead on the blockchain. Türker et all. [61] employs blockchain technology to secure data obtained from IoT devices in smart home systems. Blockchain is implemented as a distributed database for data privacy and integrity, with results indicating preserved data integrity and no data loss. This paper [62] explores the potential of blockchain technology to address critical security and trust challenges within the rapidly growing IoT ecosystem. It examines how blockchain's decentralized, immutable, and transparent features contribute to enhancing security and trust in IoT networks. This study [63] proposes a scalable blockchain-based framework for efficient data management in IoT networks accommodating a large number of devices. By utilizing the Delegated Proof of Stake (DPoS) consensus algorithm, performance and efficiency in resource-constrained IoT networks are enhanced. This paper [64] explores the pivotal role of artificial intelligence in enhancing network security and privacy within blockchain-enabled IoT systems. Blockchain provides a decentralized and immutable ledger for secure management of device identities and transactions in IoT networks. When combined with artificial intelligence, these systems gain the ability to adaptively respond to new and evolving cyber threats, thereby enhancing the resilience of networks against cyber-attacks.

Literature reviews show that the combination of machine learning, deep learning and blockchain technologies can be useful for anomaly detection and data security in IoT networks. In addition to these studies, it is thought that the proposed system can contribute to anomaly detection and data integrity in IoT network traffic. It is also considered to provide a solution for preventing data manipulation and protecting data confidentiality by using data privacy methods in data transmission. The system has been tested with experiments on the N-BaIoT dataset and is considered as an approach that can address security vulnerabilities in IoT networks.

Unlike previous studies that focus solely on anomaly detection, our approach integrates blockchain to ensure data integrity. This combination enhances security by preventing data manipulation, which is a limitation in traditional IDS approaches.

IV. Proposed System

In this study, a system is proposed to detect anomalies in IoT network traffic and improve data security. The proposed system includes a blockchain-based data storage structure for processing data collected from IoT devices, using a deep learning model for anomaly detection, and securely storing the data.

A. Dataset

In this study, the dataset used to detect anomalies in IoT network traffic is obtained from the N-BaIoT dataset, which contains the network traffic of the Philips_B120N10_Baby_Monitor device and is available on the Internet. The dataset includes normal traffic data of the device and various types of attacks (e.g., Mirai and Gafgyt attacks). In the data preprocessing stage, the raw data is cleaned, scaled and labelled as attack/normal traffic. This was done in order to ensure that the model works correctly in the training and testing processes. The data set is divided into 80% training and 20% testing. The training data is further divided into a subset to evaluate the validation performance of the model. This structure is used to objectively evaluate the performance of the model and increase its generalisation ability.

B. System Architecture

The proposed system consists of two main layers to analyse and securely manage network traffic from IoT devices: anomaly detection and blockchain-based data management. In the first layer, IoT network traffic is analysed using a Fully Connected Neural Network (FNN). The model consists of two hidden layers and contains 128 and 64 neurons, respectively. ReLU activation function and Dropout regularisation are used to increase the generalisation capability of the model. Adam algorithm and binary cross-entropy loss function were used to optimise the model. High accuracy was achieved throughout the training process.

In the blockchain layer, the data analysed by the model is stored by encrypting the hash functions of each data block. This SHA-256 based structure guarantees the reliability of the chain and the immutability of the data. Fernet algorithm was used to encrypt the data and privacy was ensured by this method. This approach made it possible to manage data from IoT networks both securely and transparently.

C. Implementation

The implementation of the proposed system involves a series of steps for the detection of anomalies in IoT network traffic and secure management of data. The dataset used in the study is obtained from the N-BaIoT dataset available on the Internet and the network traffic of the Philips_B120N10_Baby_Monitor device is analysed. The dataset contains both normal traffic and various attack types such as Mirai and Gafgyt. The data was preprocessed to make it suitable for training the model. In this process, normal traffic is labelled as '0' and attack traffic is labelled as '1'; unnecessary columns are removed and only features that will facilitate the learning of the model are left. Then, the dataset was split into 80% training and 20% testing, and the training data was further split into a subset for validation. All features were scaled using StandardScaler, which ensured that the model evaluated each input on the same scale.

The model is structured as a fully connected neural network. In the model, there are two hidden layers following the input layer. There are 128 neurons in the first layer and 64 neurons in the second layer and ReLU activation function is used in both layers. These layers are designed to learn the complex relationships between the data and improve classification accuracy. To prevent overfitting, dropout was applied at each fully connected layer. The application of dropout effectively mitigated overfitting, as demonstrated by the high accuracy and stable validation results obtained during training. No significant performance degradation was observed, indicating that the chosen dropout rate was sufficient to maintain model generalization. The output layer is structured with sigmoid activation function for classification of anomalies. The model uses Adam algorithm for optimisation and binary cross-entropy function for loss computation. During the training process, the model was optimised for 10 epochs and the accuracy and loss values were monitored at the end of each epoch. The results show that the training accuracy is 99.11%, the test accuracy is 99.79% and the low test loss (0.0006) values prove the generalisation capacity of the model.

During the simulation process, 10% of the test data was encrypted with the Fernet algorithm at random time intervals and sent to the central server. This method aims to protect data confidentiality. On the server, the data after decryption was analysed by the model. These analysis results integrated into the blockchain structure enabled each data block to be encrypted and stored with SHA-256 hash function. This structure of the blockchain prevented the manipulation of data and made it possible to store it securely. In addition, the transparency and immutability of the data within the chain increased the reliability of the system.

In conclusion, the proposed application effectively uses a deep learning model to detect anomalies in IoT networks, while utilising blockchain technology to ensure data confidentiality and integrity. The high accuracy of the model and the security contributions of blockchain are promising for the usability of the system in IoT networks.

V. Results and Discussion

In this section, the performance of the proposed system is discussed in detail and the accuracy metrics of the model and blockchain-based data management results are evaluated. The integration of the deep learning model that detects anomalies in IoT network traffic and the blockchain-based data storage structure shows that the proposed system offers a powerful solution in terms of both technical and security.

A. Model Performance Metrics

The deep learning model was trained and tested using N-BaIoT dataset. The dataset consisting of normal traffic and various attack types of 'Philips_B120N10_Baby_Monitor' device was used in the training of the model. During the training process, the dataset was labelled as '0' for normal traffic and '1' for attack traffic, thus increasing the classification capacity of the model. The training of the model was optimised for 10 epochs and the accuracy and loss values were evaluated at the end of each epoch.

The results show that the model performs well on both training and test data. The training accuracy increased with each epoch and reached 99.11%. In the evaluation results on the test data, the accuracy of the model was recorded as 99.79% and the test loss value was recorded as 0.0006. These results show that the model learnt the data effectively and the classification performance is quite high. Figure 4 shows that the training and validation accuracy of the model shows a parallel increase throughout each epoch, reflecting a stable learning process.

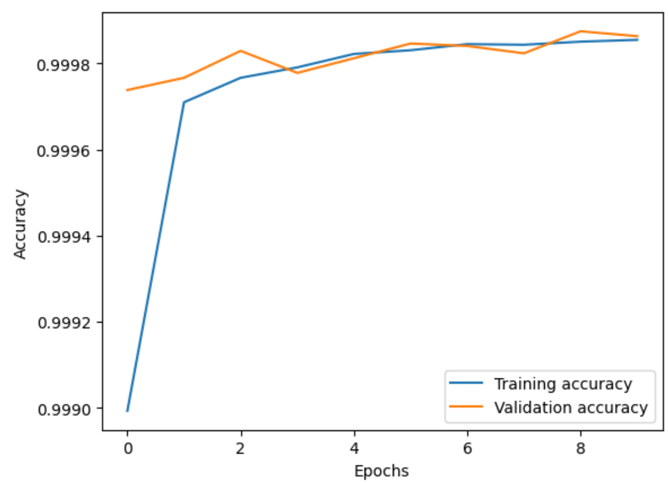


Fig. 4. Train and Validation Accuracy

The loss values decreased steadily during training and validation. As can be seen in Figure 2, the training loss of the model decreased at the end of each epoch and the validation loss also decreased steadily. This shows that the model is optimised and does not overfit the data. This high accuracy and low loss of the model is due to the significant differences between normal traffic and attack traffic in the dataset. In particular, the clear distinction between the characteristics of Mirai and Gafgyt attacks led to the successful classification of the model.

B. Blockchain Data Results

The blockchain-based data storage structure ensured secure storage of data from IoT devices and protection of data integrity. In this system, the data generated by the deep learning model that detects anomalies are added to the blockchain structure in blocks. Each block is associated with the hash function of the previous block, thus ensuring the integrity and security of the chain.

During the simulation process, 10% of the test data was encrypted at random time intervals and transmitted to the central server. The encryption process was performed using the Fernet algorithm and the data was protected against unauthorized access. After the decryption process was completed on the central server, this data was integrated into the blockchain structure. The blockchain system ensured that the data was not only stored with accuracy, but also managed transparently.

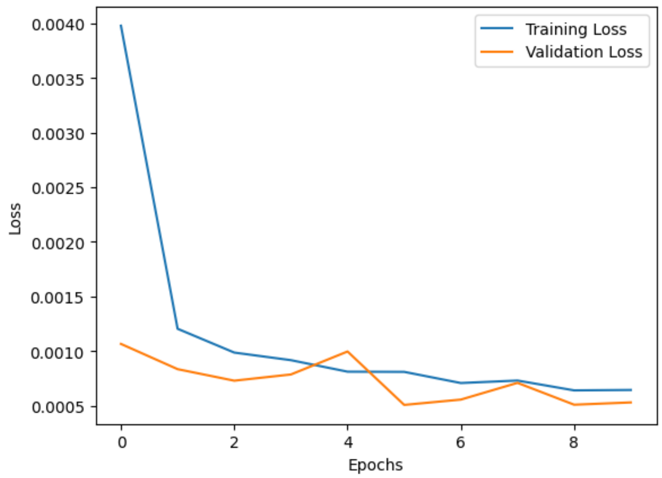


Fig. 5. Train and Validation Loss

As a result, the blockchain-based data storage system guaranteed the accuracy of each block on the chain and made data manipulation impossible. During the simulation, it was observed that data encryption and decryption processes were performed with high performance. Storing data on the chain with hash values proved to be an effective method to secure the large data flow from IoT devices.

The impact of blockchain on processing time, energy consumption, and scalability has been discussed. While blockchain improves data security, it introduces computational overhead, which must be considered in real-time IoT applications. Future studies will explore optimizations for lightweight blockchain implementations.

VI. Conclusion and Future Work

In this study, a system is proposed to detect anomalies in IoT network traffic and ensure data integrity. The proposed system combines a deep learning based anomaly detection model with a blockchain based data storage structure. In this study, network traffic data of Philips_B120N10_Baby_Monitor device is analysed using N-BaIoT dataset. The deep learning model used for anomaly detection provided an effective classification in IoT networks with a high accuracy rate (99.11%) and low loss value. The blockchain structure guarantees data integrity by securely storing and preventing data modification.

The results show that the proposed system provides a powerful solution for the analysis and security of data from IoT devices. Deep learning-based anomaly detection offers the opportunity to prevent potential threats by detecting attacks early in IoT networks. The integration of the blockchain structure not only prevented the manipulation of data, but also provided transparency and traceability. The encryption methods used in the simulation process have increased the effectiveness of the system in terms of data privacy and security.

However, the clear class distinctions in the dataset used in this study played an important role in the high performance of the model. Future work should include testing the proposed system on more complex and noisy IoT datasets to increase its generalisability. Furthermore, analysing data from different IoT devices and the compatibility of the system with these devices is another important issue that needs to be investigated.

Blockchain data structure is an effective solution to ensure data security and integrity in IoT data and critical network infrastructures. With its decentralised and immutable structure, it prevents data manipulation and unauthorized access, but has limitations such as processing speed and energy consumption. In this study, blockchain was used for a relatively limited data flow. However, the performance and energy efficiency of blockchain for large volumes of IoT data should be further investigated in the future. In addition, the integration of more complex encryption algorithms and privacy enhancing technologies can further increase the security level of the system.

The N-BaIoT dataset includes data from nine different IoT devices; however, this study used data from only one device. Future work will expand the evaluation to multiple IoT devices to assess the model's robustness in diverse environments.

In conclusion, this paper presents an innovative approach to provide anomaly detection and data security in IoT network traffic. The integrated use of deep learning and blockchain has made significant progress in IoT data management and security. Future work can enhance the scalability and generalisation capabilities of the proposed system, allowing it to find a wider application in the IoT ecosystem.

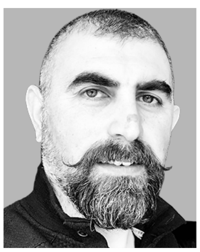
References

[1] N. M. Adams, "Perspectives on data mining," *International Journal of Market Research*, vol. 52, no. 1, 2010.
[2] D. Talia, "Clouds for scalable big data analytics," *Computer*, vol. 46, no. 5, pp. 98–101, 2013.
[3] D. Demirol, R. Das, and D. Hanbay, "Büyük veri üzerine perspektif bir bakış," in *2019 International Artificial Intelligence and Data Processing Symposium (IDAP)*, 2019, pp. 1–9.
[4] A. Abouzeid, K. Bajda-Pawlikowski, D. Abadi, A. Silberschatz, and A. Rasin, "HadoopDB: An architectural hybrid of mapreduce and DBMS technologies for analytical workloads," *Proceedings of the VLDB Endowment*, vol. 2, no. 1, pp. 922–933, aug 2009.
[5] V. Subramaniyaswamy, V. Vijayakumar, R. Logesh, and V. Indragandhi, "Unstructured data analysis on big data using map reduce," *Procedia Computer Science*, vol. 50, pp. 456–465, 2015.
[6] S. F. Oliveira, K. Fürlinger, and D. Kranzlmüller, "Trends in computation, communication and storage and the consequences for data-intensive science," in *Proceedings of the 14th IEEE International Conference on High Performance Computing and Communications, HPCC-2012 - 9th IEEE International Conference on Embedded Software and Systems, ICESS-2012*. IEEE, jun 2012, pp. 572–579.
[7] F. Chang, J. Dean, S. Ghemawat, W. C. Hsieh, D. A. Wallach, M. Burrows, T. Chandra, A. Fikes, and R. E. Gruber, "Bigtable: A distributed storage system for structured data," *ACM Transactions on Computer Systems*, vol. 26, no. 2, pp. 1–26, 2008.
[8] L. George, *HBase: The Definitive Guide*. O'Reilly Media, Inc., 2016.
[9] A. Lakshman and P. Malik, "Cassandra - a decentralized structured storage system," *Operating Systems Review (ACM)*, vol. 44, pp. 35–40, 2010.
[10] V. Abramova and J. Bernardino, "Nosql databases: Mongodb vs cassandra," in *C3S2E '13: Proceedings of the International C* Conference on Computer Science and Software Engineering*, 2013, pp. 14–22.
[11] M. Armbrust, A. Fox, D. Patterson, N. Lanham, B. Trushkowsky, J. Trutna, and H. Oh, "Scads: Scale-independent storage for social computing applications," in *CIDR 2009 - 4th Biennial Conference on Innovative Data Systems Research*, January 2009.
[12] K. Pattnaik and B. Mishra, "Introduction to big data analysis," in *Techniques and Environments for Big Data Analysis*, 2016, pp. 1–20.

[13] R. Cattell, "Scalable sql and nosql data stores," *SIGMOD Record*, vol. 39, pp. 12–27, May 2010.
[14] S. Sivasubramanian, "Amazon dynamodb: A seamlessly scalable non-relational database service," in *Proceedings of the 2012 International Conference on Management of Data - SIGMOD '12*, 2012, pp. 729–730.
[15] U. Vyas and P. Kuppusamy, *DynamoDB Applied Design Patterns*. Packt Publishing Ltd., 2014.
[16] R. Paul, "An introduction to building realtime apps with rethinkdb," March 2018, accessed: 2025-01-15. [Online]. Available: https://jaxenter.com/building-realtime-apps-rethinkdb-115254.html
[17] OrientDB, "Orientdb nosql models," 2021. [Online]. Available: http://orientdb.com/docs/3.0.x/gettingstarted/
[18] B. Iordanov, "Hypergraphdb: A generalized graph database," in *Web-Age Information Management. WAIM 2010 Workshops*, ser. 6185 LNCS, 2010, pp. 25–36.
[19] D. Dominguez-Sal, P. Urbón-Bayes, A. Giménez-Vañó, S. Gómez-Villamor, N. Martínez-Bazán, and J. Larriba-Pey, "Survey of graph database performance on the hpc scalable graph analysis benchmark," in *Web-Age Information Management. WAIM 2010 Workshops*, ser. 6185 LNCS, 2010, pp. 37–48.
[20] S. Ghemawat, H. Gobioff, and S. Leung, "The google file system," *SIGOPS Oper. Syst. Rev.*, vol. 37, pp. 29–43, 2003.
[21] N. Gemayel, "Analyzing google file system and hadoop distributed file system," *Research Journal of Information Technology*, vol. 8, pp. 66–74, 2016.
[22] O. Kisi, J. Shiri, S. Karimi, and R. M. Adnan, *Big Data in Engineering Applications*. Springer Singapore, May 2018, vol. 44.
[23] E. J. Khatib, R. Barco, P. Munoz, I. D. La Bandera, and I. Serrano, "Self-healing in mobile networks with big data," *IEEE Communications Magazine*, vol. 54, no. 1, pp. 114–120, jan 2016.
[24] B. Das, "A deep learning model for identification of diabetes type 2 based on nucleotide signals," *Neural Computing and Applications*, vol. 22, no. 1, pp. 1–5, 2022.
[25] Q. V. Pham, D. C. Nguyen, T. Huynh-The, W. J. Hwang, and P. N. Pathirana, "Artificial Intelligence (AI) and Big Data for Coronavirus (COVID-19) Pandemic: A Survey on the State-of-the-Arts," *IEEE Access*, vol. 8, pp. 130 820–130 839, 2020.
[26] A. Haleem, M. Javaid, I. H. Khan, and R. Vaishya, "Significant Applications of Big Data in COVID-19 Pandemic," *Indian Journal of Orthopaedics*, vol. 54, no. 4, pp. 526–528, jul 2020.
[27] J. H. Lee, R. Phaal, and S. H. Lee, "An integrated service-device-technology roadmap for smart city development," *Technological Forecasting and Social Change*, vol. 80, no. 2, pp. 286–306, feb 2013.
[28] C. L. Stimmel, *Building smart cities: Analytics, ICT, and design thinking*. Auerbach Publications, aug 2015.
[29] C. T. Yin, Z. Xiong, H. Chen, J. Y. Wang, D. Cooper, and B. David, "A literature survey on smart cities," *Science China Information Sciences*, vol. 58, no. 10, pp. 1–18, oct 2015.
[30] S. R., P. (2015, aug) 8 innovative examples of big data usage in india. [Online]. Available: https://www.dqindia.com/8-innovative-examples-of-big-data-usage-in-india/
[31] M. Assefi, E. Behravesh, G. Liu, and A. P. Tafti, "Big data machine learning using apache spark MLlib," in *Proceedings - 2017 IEEE International Conference on Big Data, Big Data 2017*, vol. 2018-January. IEEE, dec 2017, pp. 3492–3498.
[32] A. Oussous, F. Z. Benjelloun, A. Ait Lahcen, and S. Belfkih, "Big Data technologies: A survey," *Journal of King Saud University - Computer and Information Sciences*, vol. 30, no. 4, pp. 431–448, oct 2018.
[33] W. Fang, X. Z. Wen, Y. Zheng, and M. Zhou, "A Survey of Big Data Security and Privacy Preserving," *IETE Technical Review (Institution of Electronics and Telecommunication Engineers, India)*, vol. 34, no. 5, pp. 544–560, sep 2017.
[34] P. Samarati and L. Sweeney, "Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression," *Technical Report, SRI International Computer Science Laboratory*, 1 1998.
[35] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam, "l-diversity: Privacy beyond k-anonymity," *ACM Transactions on Knowledge Discovery from Data*, vol. 1, no. 1, p. 3, mar 2007.
[36] L. Ninghui, L. Tiancheng, and S. Venkatasubramanian, "t-Closeness: Privacy beyond k-anonymity and l-diversity," in *Proceedings - International Conference on Data Engineering*. IEEE, apr 2007, pp. 106–115.
[37] H. Y. Tran and J. Hu, "Privacy-preserving big data analytics a comprehensive survey," *Journal of Parallel and Distributed Computing*, vol. 134, pp. 207–218, dec 2019.
[38] C. Dwork, "Differential privacy," in *Lecture Notes in Computer Science*, 2006, vol. 4052 LNCS, pp. 1–12.
[39] C. C. Aggarwal and P. S. Yu, *A General Survey of Privacy-Preserving Data Mining Models and Algorithms*. Boston, MA: Springer US, 2008, pp. 11–52.
[40] P. Ram Mohan Rao, S. Murali Krishna, and A. P. Siva Kumar, "Privacy preservation techniques in big data analytics: a survey," *Journal of Big Data*, vol. 5, no. 1, 2018.
[41] M. A. Chamikara, P. Bertok, D. Liu, S. Camtepe, and I. Khalil, "Efficient data perturbation for privacy preserving and accurate data stream mining," *Pervasive and Mobile Computing*, vol. 48, pp. 1–19, aug 2018.
[42] K. Chen and L. Liu, "A random rotation perturbation approach to privacy preserving data classification," *International Conference on Data Mining*, 2005.
[43] Chen, Keke and Liu, Ling, "Geometric data perturbation for privacy preserving outsourced data mining," *Knowledge and Information Systems*, vol. 29, no. 3, pp. 657–695, dec 2011.
[44] M. Z. Gündüz, D. Demirol, R. Daş, and K. Hanbay, "Frameworks for smart grid cyber security analysis," in *Cyber Security Solutions for Protecting and Building the Future Smart Grid*, D. Asija, R. Viral, R. Daş, and G. Tuna, Eds. Elsevier, 2025, pp. 191–214.
[45] S. Venkatraman and R. Venkatraman, "Big data security challenges and strategies," *AIMS Mathematics*, vol. 4, no. 3, pp. 860–879, 2019.
[46] M. Zhao E and Y. Geng, "Homomorphic Encryption Technology for Cloud Computing," *Procedia Computer Science*, vol. 154, pp. 73–83, 2019.
[47] A. Sahai and B. Waters, "Fuzzy identity-based encryption," in *Advances in Cryptology – EUROCRYPT 2005*, R. Cramer, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 457–473.
[48] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-Policy Attribute-Based Encryption," in *2007 IEEE Symposium on Security and Privacy (SP '07)*. IEEE, may 2007, pp. 321–334.
[49] T. Bouabana-Tebibel and A. Kaci, "Parallel search over encrypted data under attribute based encryption on the Cloud Computing," *Computers & Security*, vol. 54, pp. 77–91, oct 2015.
[50] M. Ali, J. Mohajeri, M.-R. Sadeghi, and X. Liu, "A fully distributed hierarchical attribute-based encryption scheme," *Theoretical Computer Science*, vol. 815, pp. 25–46, may 2020.
[51] M. Z. Gunduz and R. Das, "Cyber-security on smart grid: Threats and potential solutions," *Computer Networks*, vol. 169, p. 107094, 2020.
[52] H. Satılmış and S. Akleylek, "Iot güvenliği İçin kullanılan makine Öğrenimi ve derin Öğrenme modelleri Üzerine bir derleme," *Bilişim Teknolojileri Dergisi*, vol. 14, no. 4, p. 457–481, 2021.
[53] A. Gökdemr and A. Çalhan, "Nesnelerin interneti ortamlarında derin öğrenme ve makine öğrenmesi tabanlı anomali tespiti," *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, vol. 37, no. 4, p. 1945–1956, 2022.
[54] L. E. Dai, X. Wang, and S. B. Xu, "A deep learning based anomaly detection model for iot networks," in *Proceedings of the 2nd International Conference on Internet of Things, Communication and Intelligent Technology*, J. Dong, L. Zhang, and D. Cheng, Eds. Singapore: Springer Nature Singapore, 2024, pp. 187–196.
[55] Y. Mirsky, T. Golomb, and Y. Elovici, "Lightweight collaborative anomaly detection for the iot using blockchain," *Journal of Parallel and Distributed Computing*, vol. 145, pp. 75–97, 2020.
[56] T. Golomb, Y. Mirsky, and Y. Elovici, "Ciota: Collaborative iot anomaly detection via blockchain," *CoRR*, vol. abs/1803.03807, 2018. [Online]. Available: http://arxiv.org/abs/1803.03807
[57] A. Diro, N. Chilamkurti, V.-D. Nguyen, and W. Heyne, "A comprehensive study of anomaly detection schemes in iot networks using machine learning algorithms," *Sensors*, vol. 21, no. 24, 2021.
[58] M. Emeç and M. H. Özcanhan, "A hybrid deep learning approach for intrusion detection in iot networks," *Advances in Electrical and Computer Engineering*, vol. 22, no. 1, pp. 3–12, 2022.
[59] D. Li, W. Peng, W. Deng, and F. Gai, "A blockchain-based authentication and security mechanism for iot," in *2018 27th International Conference on Computer Communication and Networks (ICCCN)*, 2018, pp. 1–6.
[60] T. Xu, Z. Fu, M. Yu, J. Wang, H. Liu, and T. Qiu, "Blockchain based data protection framework for iot in untrusted storage," in *2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, 2021, pp. 813–818.
[61] G. F. Türker and K. Tanyeri, "Blokzincir teknolojisi ile nesnelerin İnterneti tabanlı (iot) sistemlerin veri güvenliğinin sağlanması," *Gazi Üniv. Fen Bilim. Derg. C Tasar. ve Teknol.*, May 2023.

[62] S. Almarri and A. Aljughaiman, “Blockchain technology for iot security and trust: A comprehensive slr,” *Sustainability*, vol. 16, no. 23, 2024. [Online]. Available: https://www.mdpi.com/2071-1050/16/23/10177
[63] E. U. Haque, A. Shah, J. Iqbal, S. S. Ullah, R. Alroobaea, and S. Hussain, “A scalable blockchain based framework for efficient IoT data management using lightweight consensus,” *Sci. Rep.*, vol. 14, no. 1, p. 7841, Apr. 2024.
[64] A. M. Ruzbahani, “Ai-protected blockchain-based iot environments: Harnessing the future of network security and privacy,” 2024. [Online]. Available: https://arxiv.org/abs/2405.13847

Doygun Demirol received his Bachelor’s and Master’s degrees in Electronics and Computer Education from Firat University in 2010 and 2012, respectively. He is currently pursuing his Ph.D. in Computer Engineering at Inonu University. He works as a lecturer at Bingol University. His research interests include cybersecurity, deep learning, natural language processing, and graph data science.



Resul Das is a full professor in the Department of Software Engineering at the Faculty of Technology, Fırat University. He earned his B.Sc. and M.Sc. degrees in Computer Science from Fırat 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 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, teaching CCNA and CCNP courses. From September 2017 to June 2018, he researched as a visiting professor at the University of Alberta, Edmonton, Canada, under the TÜBİTAK-BİDEB 2219 Postdoctoral Research Fellowship program. He served as Head of the Department of Software Engineering from March 2020 to April 2023.

Prof. Das has held editorial roles in several prestigious academic journals. He was an Associate Editor for IEEE Access and the Turkish Journal of Electrical Engineering and Computer Science. Currently, he serves as an Associate Editor for Elsevier journals (Internet of Things, Alexandria Engineering Journal, and Telematics and Informatics Reports), IEEE Open Journal of the Communications Society (OJ-COMS), and the International Journal of Grid and Utility Computing (Inderscience). Globally recognized for his contributions, Prof. Daş has been included in the top 2% of the ”World’s Most Influential Scientists” list, compiled by Stanford University researchers, for five consecutive years (2019–2023). His research interests encompass computer networks, cybersecurity, IoT and systems engineering, data science and visualization, software quality assurance, and testing.

Ceren Nur Cansel graduated from Malatya Doğa College Science High School as valedictorian in 2019. She graduated from Koç University Industrial Engineering and Economics departments with Vehbi Koç Honour Award in 2024. She has successfully completed three different specialisation programmes: Finance, Financial Engineering and Macroeconomic Policies and Financial Markets. She is fluent in English. After her university education focused on data analysis, modelling and forecasting, she continues his studies in the field of financial analysis and pricing.

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.

Davut Hanbay received his Bachelor’s and Master’s degrees in the Department of Electronics Education and Department of Electric–Electronic Engineering from Firat University, Elazig, Turkey, in 1999, and 2003 respectively. He completed his Ph.D. at Fırat University, Department of Electrical and Electronics Engineering in 2007. He is currently Professor in the computer engineering of Inonu University. His research interests include computeraided systems, operating systems, deep learning, machine learning, image processing, and signal processing.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 1, March 2025
Research Article

Classification of Images in Bad Weather Conditions with Convolutional Neural Networks

Yasin Demir, Nagihan Severoglu, Nur Huseyin Kaplan, Sefa Kucuk

***Abstract*—Weather conditions are one of the major factors significantly influencing the daily lives of individuals. Unfavorable weather conditions adversely affect their lives and directly impede the progress of the subsequent image-processing steps necessary for real-world vision tasks such as object detection and autonomous driving. For this reason, the correct classification of the weather conditions is of great importance. Although traditional classification methods achieve high accuracy in various tasks, they cannot achieve the same success in classifying weather conditions. In this paper, we propose a novel convolutional neural network (CNN) framework for the classification of weather conditions with high accuracy. The proposed network outperforms the existing methods with 95.50% accuracy for a classification problem with five different scenarios.**

***Index Terms*—Multi-class classification, deep learning, convolutional neural networks, weather classification**

I. Introduction

WEATHER conditions influence outdoor imaging systems, leading to low-contrast and reduced image visibility. It can also directly affect the operation of many real-world visual systems, such as autonomous vehicles, intelligent driver assistance systems, and outdoor video analysis. Most research in computer vision is based on the assumption that the weather is clear in the processed images. However, one of the most critical issues in developing these systems is their poor performance in adverse weather conditions such as rain, snow, fog, and haze [1], [2]. Therefore, weather classification applications have great importance in providing more reliable and better visibility of imagery.

Over the past decades, the authors have generally focused on the single-class weather recognition problem in which they try to determine whether an image belongs to a particular category or not [3], [4], [5], [6]. It is a fact that single-class classification tasks are unable to provide a comprehensive description of weather conditions. On the other hand, some of them deal with two-class recognition problems, e.g., sunny and cloudy weather [7], [8]. To the best of our knowledge, despite its numerous application areas, a limited number of multi-class weather classification studies [9], [10] have been carried out.

With the rapid development of machine learning, learning-based models have been widely used in classification problems. Furthermore, collecting large data sets has become more accessible in recent years, owing to the progress in image acquisition systems and their increased accessibility. These developments have facilitated the training process of learning-based networks. Weather classification is a multi-class classification problem for which numerous learning-based methods have been proposed for such problems. Generally speaking, learning-based approaches are roughly divided into four main groups [11]. The first group, supervised learning methods, utilize information from labeled training data to predict output classes [12]. The second category, unsupervised learning methods, is applied when the training data lacks labeling. In such cases, these methods classify based on certain features and inferences from the available data [13]. The third group, semi-supervised learning methods, are between supervised and unsupervised techniques as they use a combination of labeled and unlabelled data for training, and the amount of unlabelled data is higher in this techniques [14]. The last category, reinforcement learning methods, estimates the consequences of system actions in environments that lead from one situation to another through rewards and punishments scheme [15]. These learning-based methods provide significant advantages in terms of flexibility in a wide range of applications.

Weather conditions classification holds crucial importance for various computer vision applications in outdoor surveillance systems, robotic vision, and driver assistance systems, to name a few. It can play a vital role in deciding which pre-processing steps to execute for an application. For example, computer vision processes suffer from hazy environments, and dehazing methods deal with these circumstances to improve the visual quality of images. Detecting a hazy environment during autonomous driving will allow the proper functioning of required pre-processing steps.

In this paper, we propose an effective convolutional neural network (CNN) framework to classify images captured in adverse weather scenarios. By doing this, we attempt to address multi-class weather classification problems for the benefit of further image processing algorithms such as dehazing, defogging, and low-light image enhancement methods. To summarize, the main contributions of this paper can be summarized as follows:

- As part of this work, we have compiled a dataset from several publicly available datasets for those interested in further research on this task.
- We propose a novel multi-label CNN-based weather classification network that can accurately categorize poor weather images, including haze, snow, and rain.
- Low-light illuminations degrade the performance of vision applications. Considering this fact, we aim to extend

Yasin DEMİR, **Nagihan SEVEROGLU**, **Nur Hüseyin KAPLAN**, and **Sefa KÜÇÜK** are with the Department of Electrical and Electronics Engineering, Erzurum Technical University, Erzurum, 25050, TURKEY, e-mails: yasin.demir@erzurum.edu.tr, nagihan.severoglu@erzurum.edu.tr, huseyin.kaplan@erzurum.edu.tr, sefa.kucuk@erzurum.edu.tr.

*Corresponding author.
Manuscript received Jan 04,2024; revised Feb 17, 2025.

DOI: 10.17694/bajece.1415025

the applicability of the proposed model by adding a new category to the dataset called *low light class* for bad weather conditions.

- We have created a normal class devoid of any adverse weather conditions. By doing this, we have tried completely separating bad weather conditions from clear ones.

The rest of the paper is organized as follows. Section II gives the most common deep learning methods used specifically in image classification tasks. Section III details the proposed architecture for classifying different weather scenarios. Section IV presents the experimental results, and the conclusion is given in Section V.

II. Related work

Many deep learning (DL) models have been proposed for image classification including convolutional neural networks, dynamic Bayesian networks, autoencoders, and restricted Boltzmann machine models [16]. In this section, we first provide a brief overview of commonly used deep learning models for image classification. Then, we review the image-based weather detection and classification works in the literature.

A. Convolutional Neural Networks

Convolutional neural networks (CNNs) are one of the most widely used DL models, generally consisting of convolution layers, pooling layers, and fully connected layers. LeNet-5, a leading model, is composed of two convolutional layers, two fully connected layers, multiple pooling layers, and a Gaussian connection layer. With large-scale datasets and significant advances in computational capabilities, more advanced networks have been proposed, such as AlexNet [17], which leverages the ImageNet dataset [18]. AlexNet is structured with five convolutional layers and three subsequent fully connected layers. VGGNet [19], another inspiring model, has been proposed to achieve better performance by increasing the depth of the network while reducing the number of model parameters. It has also introduced innovations such as modular networks, smaller convolution, and multi-scale training. In contrast to previous approaches, the Network in Network (NIN) [20] model adopts a combination of multi-layer perceptron and convolution, resulting in a more complex micro-neural network structure than the traditional convolutional layers.

B. Dynamic Bayesian Networks

Bayesian Networks (BNs) play a significant role in various applications, including anomaly detection, classification, and clustering. BNs provide a better efficient representation of the joint probability distribution over a group of random variables [21]. Dynamic Bayesian Networks (DBNs), a specialized variant of BNs, recursively capture the dynamics of the system in a time-dependent fashion [22]. In a DBN, the first layer is referred to as the input layer, the middle layer is the hidden layer, and the final layer is designated as the output layer [23].

C. Autoencoders

Autoencoders serve as a technique for extracting principal components within large data distributions [24]. Due to its adaptable network structure that can be customized for various domains, it has the ability to create deeper networks. The autoencoder stands out as one of the most effective preprocessing techniques for image classification. Sparse autoencoder is a commonly used deep learning approach for automatically extracting features from unlabeled data [25]. Since deep learning applications are not robust against noisy data, pre-training with noisy data is necessary. To cope with these circumstances, denoising autoencoder structures [26] have been proposed. In the denoising autoencoder, the input is distorted by adding random noise. The model is then trained to generate predictions for the original, uncorrupted data. Deep Wavelet Autoencoder is an autoencoder architecture that has gained interest in recent years [27]. It integrates concepts from wavelet transforms into its design, employing these transforms within the network's operations to enhance its capacity for acquiring hierarchical and multi-scale data representations.

D. Restricted Boltzman Machine

Restricted Boltzmann Machine (RBM) is frequently used as a feature extractor in image classification. RBM shares parametrization with the layers of the deep belief network and is therefore considered the building block of the deep belief network. This model was first introduced under the name Harmonium [28]. Recently, many deep learning algorithms have been proposed using the RBM model [29], [30]. Model structures generally consist of two layers: the visible layer and the hidden layer. While RBMs may not effectively represent certain distributions, they demonstrate the capability to represent any discrete distribution when an adequate number of hidden layers are employed [31]. This characteristic renders RBMs one of the most suitable types of deep networks for feature extraction in unsupervised learning applications.

E. Image-based Weather Detection and Classification

Weather classification involves the identification of different weather conditions from a single image, which is a challenging task due to the diversity of weather phenomena and the lack of distinctive features. Early attempts are based on handcrafted features and traditional machine-learning techniques. For example, Kurihata $et\ al.$ defined raindrop characteristics using image features from principal component analysis (PCA) that represent the essential characters of raindrops [4]. Roser and Moosmann proposed a classification method on single color images based on Support Vector Machines (SVM) using some features of the image such as contrast, minimum brightness, sharpness, and color [7]; The authors of [5] proposed an algorithm based on Real Adaboost that combines three features: histogram of gradient amplitude (HGA), histogram of HSV color space and road information.

These works are pioneers in the task of weather classification, but can only recognize rainy weather, while applications are limited due to fixed target scenes. Over the last decades,

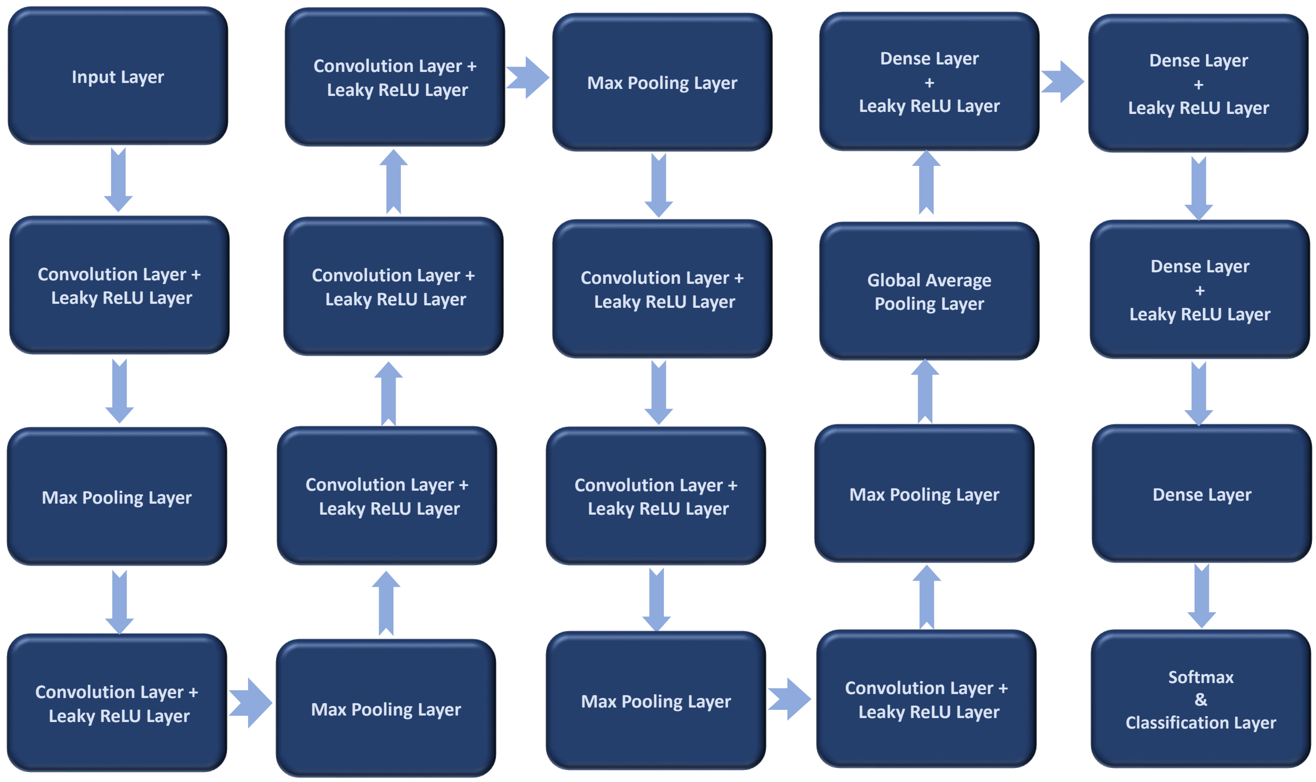


Fig. 1. Illustration of the proposed method.

researchers have focused on two-class weather classification to estimate weather conditions from images. For example; the authors of [32] use a correlation between weather and scene illumination to categorize classes; the authors of [8] proposed a collaborative learning framework that uses some specific image features such as sky, shadow, and haze to identify images as sunny or cloudy; the authors of [6] used the bag of words method for feature extraction from images and multiple kernel learning (MKL) to classify images into three class; Elhoseiny *et al.* introduced a CNN architecture for two-class weather classification [33]; Kang *et al.* [34] utilized a deep learning-based method to classify weather images into one of four classes: hazy, rainy, snowy, and other. They have achieved better performance than traditional methods by using the GoogleNet deep CNN model. Zhao *and* Wu [35] proposed a weather forecast classifier for four classes (rainy, snowy, sunny, and foggy) using the CNN method to extract high-dimensional features from images. The images were preprocessed with Mask R-CNN to improve classification performance.

III. The Proposed Framework

Here, we give the details of the proposed network. CNNs are widely used in applications such as object recognition [36] and classification [37], predominantly owing to their superior classification accuracy. This paper proposes a CNN-based structure as shown in Fig. 1.

A. Convolution layer

The main purpose of the layer is to obtain the filtered image by moving the filters of certain sizes over the entire image. The dimensions of these filters are generally chosen as 3×3, 5×5, and 7×7. This process results in an output image with higher-level features in a hierarchical manner. We set the filter sizes to 3×3 and 5×5 in convolution layers.

B. Pooling layer

Reducing the number of hyper-parameters is crucial to prevent the model from memorizing the training data and to alleviate the computational burden. To do this, CNNs utilize pooling layers. Similar to the convolution operation, the pooling process is also carried out via specific filters. These filters execute maximum, minimum, and averaging operations with certain window sizes. In the proposed model, we set the size of the pooling layers as 2×2.

C. Leaky ReLU

Activation functions play a crucial role in introducing non-linearities to CNNs. Different activation functions, including Linear, Tanh, ReLU, Swish, and Leaky ReLU (LReLU), have been employed in numerous studies. To investigate the influence of activation functions on the proposed model, we employ different activation functions and present the corresponding training accuracy in Table I. As seen from Table, we have achieved the highest training accuracy with LReLU. Thus, we choose LReLU as the activation function. LReLU, a variant of

TABLE I
ABLATION STUDY ON ACTIVATION FUNCTIONS

Activation function	Training accuracy(%)
Linear	93.12
Tanh	93.44
ReLU	96.18
Swish	96.88
LReLU	**97.69**

the classical ReLU, also has a small slope for negative inputs. LReLU addresses the issue of neurons struggling to learn after entering the negative range, thanks to its small slope. It is also known that although it is slower than classical ReLU, it demonstrates better performance [38]. Mathematically, LReLU is defined as

$$f(x) = \begin{cases} 0.01x & x < 0 \\ x & x \geq 0 \end{cases} \tag{1}$$

and visually it looks as shown in Fig. 2.

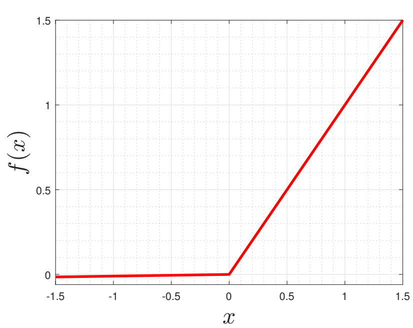


Fig. 2. Graph of the LReLU activation function.

D. Fully connected layer

The fully connected layer utilizes a weight matrix for each neuron to perform a linear transformation to the input. Through this process, all possible connections between layers ensure that the entries of the input influence every entry of the output. Typically, the fully connected layer is used at the end of CNN models to optimize classification scores.

E. Classification layer

As its name implies, the classification layer is used in DL networks for classification tasks. The output dimension of the classification layer is equivalent to the number of classified objects. Among the classifiers used in CNNs, *softmax* is the most commonly used and highly effective one. Softmax generates a probability output value within the range of 0 to 1 for each object to be classified. Softmax assigns the output as the class of the item for a probability value close to 1.

TABLE II
DETAILS OF THE CONSTRUCTED DATASET

Image class	# of image	Dataset
Hazy	340	[39]-[40]
Rainy	381	[41]-[42]
Snowy	430	[43]
Low-light	444	[44]
Normal	500	[45]
Test	200	40 of each class

IV. EXPERIMENTS

In this section, we first provide the procedure for creating the dataset and then give the implementation details. Finally, we present the evaluation metrics and experimental results in turn.

A. Dataset

For the multi-class weather classification experiment, we identified five different weather-related attributes: hazy, rainy, snowy, low-light, and normal. We choose images for individual classes from various datasets that are publicly accessible. To enhance the effectiveness of the proposed model, we introduce challenging scene images to diversify the dataset. For example, the model should recognize an image taken in snowy conditions as snowy weather and classify an image containing snow but not snowfall as representing normal weather. In doing so, we allow for the examination of weather conditions that affect the visibility of objects in the scene. Considering these conditions, we have constructed the dataset as follows:

- *Hazy images:* We select the images for this class from the O-Haze [39] dataset and the hazy weather [40] dataset of road images captured in foggy weather.
- *Rainy images:* Images are chosen from the datasets given in [41] and [42].
- *Snowy images:* We pick snowy images from the Snow100k [43].
- *Low-light images:* We randomly select low-light images from the ExDark [44], which contains images taken in various low-light conditions.
- *Normal images:* Finally, normal images without any adverse weather conditions have been chosen from the Part2 Subset [45].

When determining images within each class from the datasets, we consider the presence of only one feature in a single image. For example, we disregard hazy information present in a snowy image. Detailed information on the dataset is presented in Table II. The last row of Table II corresponds to the number of test images, which includes 40 randomly selected images for each class.

B. Implementation results

It is common knowledge that fine-tuning the hyperparameters is crucial for achieving high performance in DL methods. We have conducted several experiments to demonstrate the

TABLE III
ABLATION STUDY ON MODEL TRAINING WITH K-FOLD

# of fold	Average accuracy (%)
5	89
10	91.5

validity of the proposed CNN model. The test set has been created from unseen images not used in the training phase. We set the batch size to 32, and the number of epochs to 200. With these parameters, we use the k-fold cross-validation technique to evaluate the performance of the proposed model. We set k to 5 and 10, and have obtained the results given in Table III. We have achieved a training accuracy of 97.88% and a test accuracy of 95.50% using the model parameters that provide the highest classification accuracy among the k-folding results. Moreover, we present the confusion matrix in Fig. 3 to evaluate the performance of the proposed model on the test set. As can be seen from Fig. 3, the proposed model demonstrates a classification performance exceeding 97% for all classes, except for the Hazy class. Fig. 4 illustrates a few examples where we give some failure cases from different k-fold validation results.

It is apparent from Fig. 4 that hazy weather tends to be misclassified as normal. We believe that this tendency arises from the insufficient learning of the airlight in hazy images. We note from Fig. 3 and Fig. 4 that the hazy class is the most challenging weather condition to classify among the five weather conditions.

The proposed CNN architecture was modeled using the open-source TensorFlow library in version 3.9.13 of the Python programming language [46]. Experimental operations such as training and testing were carried out on a personal computer with 11^{th} Generation Intel® Core$^{\text{TM}}$ i7-11800H CPU and NVIDIA Geforce RTX 3060 GPU.

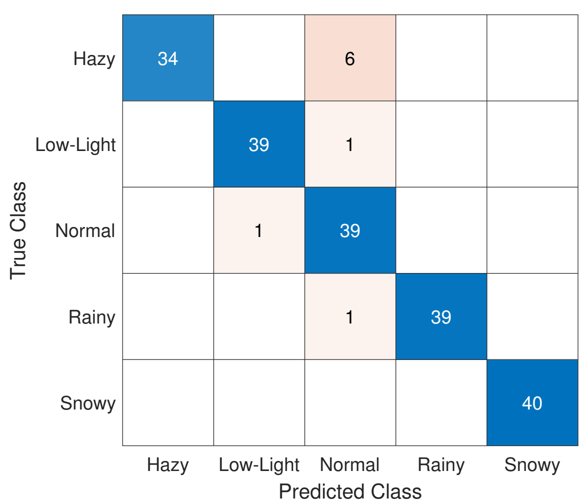


Fig. 3. Confusion matrix of five-class classification results.

C. Comparison with related methods

We have chosen four assessment criteria to quantitatively measure the performance of the compared methods. The first evaluation criterion is accuracy, which is popular in multi-class classification. Accuracy is defined as the ratio of correctly predicted data to the total amount of data and is calculated as follows.

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{2}$$

Here, $\text{TP}, \text{TN}, \text{FP},$ and FN represent true positive, true negative, false positive, and false negative, respectively. Using Eq. (2), we calculate the accuracy of the proposed model as 0.955. Although accuracy is a useful metric, it may not be sufficient when evaluating datasets characterized by uneven distributions or unbiased data. Therefore, precision has been used for a more comprehensive evaluation. It is particularly employed in scenarios where the cost of making a false positive prediction is high. Mathematically, the precision metric is computed as follows.

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{3}$$

We get the precision as 0.9646 by Eq. (3). Another assessment metric is recall, which becomes particularly crucial when the cost of predicting false negatives is substantial. The recall is defined by the following expression.

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{4}$$

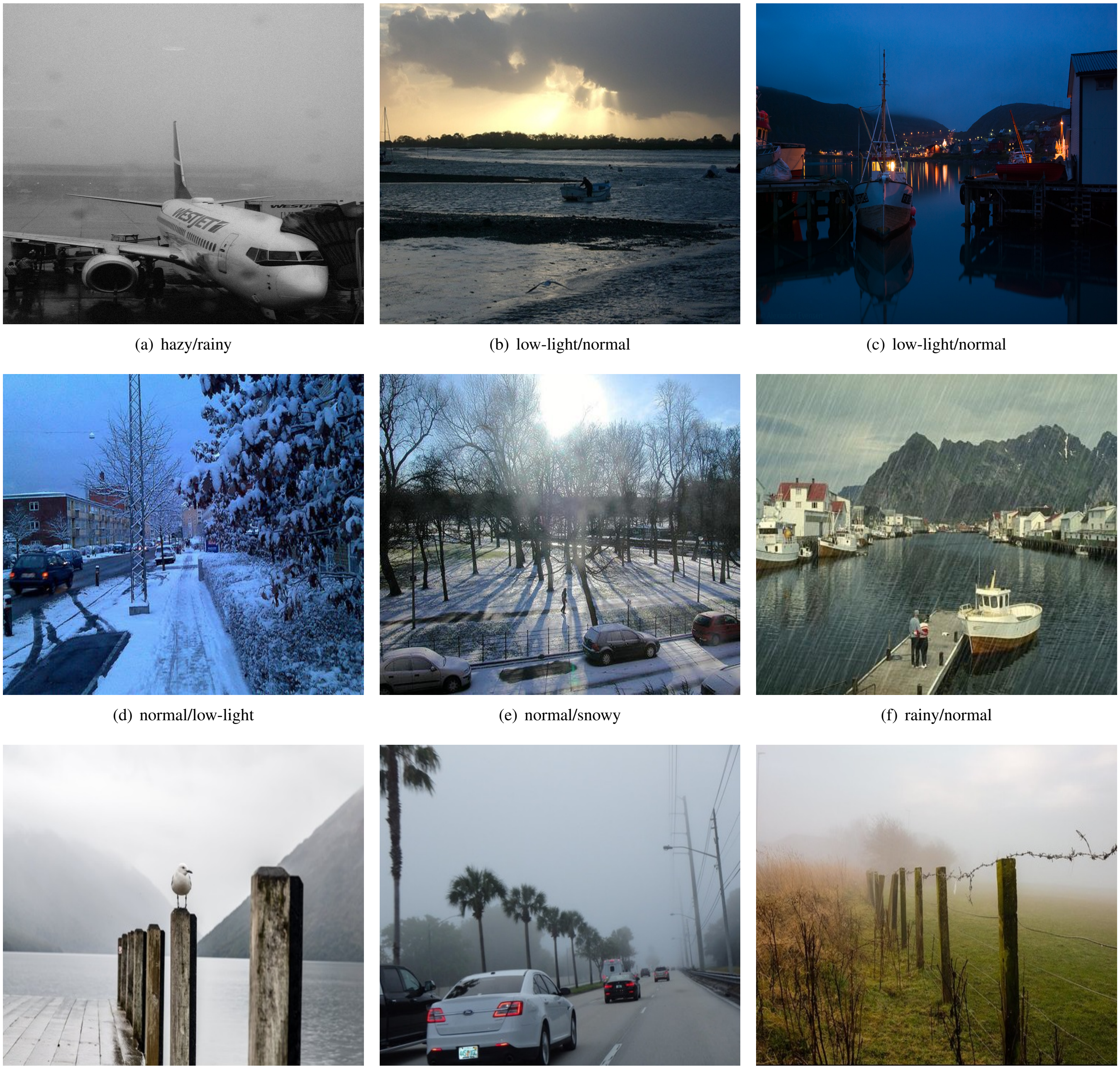
The recall value is evaluated as 0.9845. The last performance discriminator is the F1 score, which comprehensively evaluates every aspect of prediction success on a dataset, as it considers all error costs. The F_1 score [49] is calculated as in Eq. (5). The F_1 score for the proposed model is computed as 0.9744.

$$\text{F}_1 = 2 \times \left(\frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \right) \tag{5}$$

We compare the performance of the proposed method with several competing methods, including AlexNet [17], VGGNet [19], ResNet101 [33], ML-KNN [47], SRN [48] and CNN-RNN [9]. We adopt overall precision (OP), overall recall (OR), and overall F1 (OF1) as evaluation metrics, and tabulate the results in Table IV. As highlighted in Table IV, the proposed model outperforms the compared methods in all metrics, except for the rainy class. Moreover, we achieve superior performance in the newly introduced classes, namely low-light and normal.

V. CONCLUSION

In this paper, we have proposed a CNN framework for multi-label weather classification tasks. We have identified five different classes for weather conditions. To train the proposed network model, we have created a new training and test set by selecting images suitable for five classes from public datasets. On the constructed dataset, we have demonstrated the effectiveness of the proposed method and proved that it yields



(a) hazy/rainy (b) low-light/normal (c) low-light/normal

(d) normal/low-light (e) normal/snowy (f) rainy/normal

(g) hazy/normal (h) hazy/normal (i) hazy/normal

Fig. 4. Examples of some failure cases. The captions of the sub-figures represent the true class/predicted class.

TABLE IV
PRECISION/RECALL COMPARISONS WITH THE STATE-OF-THE-ART METHODS. WE HIGHLIGHT THE BEST AND THE SECOND-BEST RESULTS IN BOLD AND ITALIC, RESPECTIVELY

Method	Hazy	Snowy	Rainy	Low-light	Normal	OP	OR	OF1
AlexNet [17]	0.735/0.890	0.784/0.685	0.876/0.905	-	-	0.9007	0.8668	0.8834
VGGNet [19]	0.867/0.728	0.814/0.701	0.887/0.931	-	-	0.9087	0.8494	0.8780
ResNet101 [33]	0.841/0.855	0.776/0.882	0.947/0.938	-	-	0.8876	0.8861	0.8868
ML-KNN [47]	0.819/0.834	0.794/0.736	0.918/0.934	-	-	0.9138	0.8766	0.8948
SRN [48]	-	-	-	-	-	0.8988	0.865	0.8816
CNN-RNN [9]	0.856/0.861	0.856/0.758	0.894/0.938	-	-	*0.9263*	*0.8946*	*0.9135*
Proposed method	1/0.850	1/1	1/0.975	0.975/0.975	0.829/0.975	**0.9646**	**0.9845**	**0.9744**

quantitatively superior results compared to other competing algorithms. The proposed model, designed to predict weather conditions for specified classes, is expected to improve the utilization of further image enhancement algorithms. In this way, it will be possible to prevent applications such as object detection and target tracking against the disruptive effect of adverse weather conditions.

We are aware that the proposed framework has limitations in classifying multiple weather types within a single image. In further studies, we focus on labeling multi-weather types in a single image to provide a more comprehensive description of weather conditions.

ACKNOWLEDGMENT

The source code and pre-trained model, along with the constructed dataset, are available at https://spars.erzurum.edu.tr.

REFERENCES

[1] S. Zang, M. Ding, D. Smith, P. Tyler, T. Rakotoarivelo, and M. A. Kaafar, "The impact of adverse weather conditions on autonomous vehicles: How rain, snow, fog, and hail affect the performance of a self-driving car," *IEEE vehicular technology magazine*, vol. 14, no. 2, pp. 103–111, 2019.

[2] X. Zhao, P. Liu, J. Liu, and X. Tang, "A time, space and color-based classification of different weather conditions," in *2011 Visual Communications and Image Processing (VCIP)*. IEEE, 2011, pp. 1–4.

[3] Q. Li, Y. Kong, and S.-m. Xia, "A method of weather recognition based on outdoor images," in *2014 International Conference on Computer Vision Theory and Applications (VISAPP)*, vol. 2. IEEE, 2014, pp. 510–516.

[4] H. Kurihata, T. Takahashi, I. Ide, Y. Mekada, H. Murase, Y. Tamatsu, and T. Miyahara, "Rainy weather recognition from in-vehicle camera images for driver assistance," in *IEEE Proceedings. Intelligent Vehicles Symposium, 2005.* IEEE, 2005, pp. 205–210.

[5] X. Yan, Y. Luo, and X. Zheng, "Weather recognition based on images captured by vision system in vehicle," in *Advances in Neural Networks–ISNN 2009: 6th International Symposium on Neural Networks, ISNN 2009 Wuhan, China, May 26-29, 2009 Proceedings, Part III 6*. Springer, 2009, pp. 390–398.

[6] Z. Chen, F. Yang, A. Lindner, G. Barrenetxea, and M. Vetterli, "How is the weather: Automatic inference from images," in *2012 19th IEEE International conference on image processing*. IEEE, 2012, pp. 1853–1856.

[7] M. Roser and F. Moosmann, "Classification of weather situations on single color images," in *2008 IEEE intelligent vehicles symposium*. IEEE, 2008, pp. 798–803.

[8] C. Lu, D. Lin, J. Jia, and C.-K. Tang, "Two-class weather classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 3718–3725.

[9] B. Zhao, X. Li, X. Lu, and Z. Wang, "A cnn–rnn architecture for multi-label weather recognition," *Neurocomputing*, vol. 322, pp. 47–57, 2018.

[10] S. Chen, T. Shu, H. Zhao, and Y. Y. Tang, "Mask-cnn-transformer for real-time multi-label weather recognition," *Knowledge-Based Systems*, vol. 278, p. 110881, 2023.

[11] R. Saravanan and P. Sujatha, "A state of art techniques on machine learning algorithms: a perspective of supervised learning approaches in data classification," in *2018 Second international conference on intelligent computing and control systems (ICICCS)*. IEEE, 2018, pp. 945–949.

[12] M. Cord and P. Cunningham, Eds., *Machine Learning Techniques for Multimedia - Case Studies on Organization and Retrieval*. Springer, 2008.

[13] M. A. T. Figueiredo and A. K. Jain, "Unsupervised learning of finite mixture models," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 3, pp. 381–396, 2002.

[14] P. Thomas, "Review of" semi-supervised learning" by o. chapelle, b. schölkopf, and a. zien, eds. london, uk, mit press, 2006," *IEEE Transactions on Neural Networks*, vol. 20, no. 3, pp. 542–542, 2009.

[15] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *Journal of artificial intelligence research*, vol. 4, pp. 237–285, 1996.

[16] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural networks*, vol. 61, pp. 85–117, 2015.

[17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.

[18] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009, pp. 248–255.

[19] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[20] M. Lin, Q. Chen, and S. Yan, "Network in network. arxiv 2013," *arXiv preprint arXiv:1312.4400*, 2013.

[21] B. Mihaljević, C. Bielza, and P. Larrañaga, "Bayesian networks for interpretable machine learning and optimization," *Neurocomputing*, vol. 456, pp. 648–665, 2021.

[22] P. Sarshar, J. Radianti, and J. J. Gonzalez, "Modeling panic in ship fire evacuation using dynamic bayesian network," in *Third International Conference on Innovative Computing Technology (INTECH 2013)*. IEEE, 2013, pp. 301–307.

[23] P. Wang, E. Fan, and P. Wang, "Comparative analysis of image classification algorithms based on traditional machine learning and deep learning," *Pattern Recognition Letters*, vol. 141, pp. 61–67, 2021.

[24] C.-Y. Liou, W.-C. Cheng, J.-W. Liou, and D.-R. Liou, "Autoencoder for words," *Neurocomputing*, vol. 139, pp. 84–96, 2014.

[25] Y. Xiao, J. Wu, Z. Lin, and X. Zhao, "A semi-supervised deep learning method based on stacked sparse auto-encoder for cancer prediction using rna-seq data," *Computer methods and programs in biomedicine*, vol. 166, pp. 99–105, 2018.

[26] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P.-A. Manzagol, and L. Bottou, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion." *Journal of machine learning research*, vol. 11, no. 12, 2010.

[27] P. Kumar Mallick, S. H. Ryu, S. K. Satapathy, S. Mishra, G. N. Nguyen, and P. Tiwari, "Brain mri image classification for cancer detection using deep wavelet autoencoder-based deep neural network," *IEEE Access*, vol. 7, pp. 46 278–46 287, 2019.

[28] P. Smolensky *et al.*, "Information processing in dynamical systems: Foundations of harmony theory," 1986.

[29] L.-W. Kim, "Deepx: Deep learning accelerator for restricted boltzmann machine artificial neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 5, pp. 1441–1453, 2018.

[30] C. L. P. Chen, C.-Y. Zhang, L. Chen, and M. Gan, "Fuzzy restricted boltzmann machine for the enhancement of deep learning," *IEEE Transactions on Fuzzy Systems*, vol. 23, no. 6, pp. 2163–2173, 2015.

[31] N. Le Roux and Y. Bengio, "Representational power of restricted boltzmann machines and deep belief networks," *Neural computation*, vol. 20, no. 6, pp. 1631–1649, 2008.

[32] L. Shen and P. Tan, "Photometric stereo and weather estimation using internet images," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2009, pp. 1850–1857.

[33] M. Elhoseiny, S. Huang, and A. Elgammal, "Weather classification with deep convolutional neural networks," in *2015 IEEE international conference on image processing (ICIP)*. IEEE, 2015, pp. 3349–3353.

[34] L.-W. Kang, K.-L. Chou, and R.-H. Fu, "Deep learning-based weather image recognition," in *2018 International Symposium on Computer, Consumer and Control (IS3C)*. IEEE, 2018, pp. 384–387.

[35] X. Zhao and C. Wu, "Weather classification based on convolutional neural networks," in *2021 International Conference on Wireless Communications and Smart Grid (ICWCSG)*. IEEE, 2021, pp. 293–296.

[36] J. M. Gandarias, A. J. Garcia-Cerezo, and J. M. Gómez-de Gabriel, "Cnn-based methods for object recognition with high-resolution tactile sensors," *IEEE Sensors Journal*, vol. 19, no. 16, pp. 6872–6882, 2019.

[37] R. Vaddi and P. Manoharan, "Hyperspectral image classification using cnn with spectral and spatial features integration," *Infrared Physics & Technology*, vol. 107, p. 103296, 2020.

[38] J. Xu, Z. Li, B. Du, M. Zhang, and J. Liu, "Reluplex made more practical: Leaky relu," in *2020 IEEE Symposium on Computers and communications (ISCC)*. IEEE, 2020, pp. 1–7.

[39] C. O. Ancuti, C. Ancuti, R. Timofte, and C. De Vleeschouwer, "O-haze: a dehazing benchmark with real hazy and haze-free outdoor images," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2018, pp. 754–762.

[40] H. Halmaoui, A. Cord, and N. Hautière, "Contrast restoration of road images taken in foggy weather," in *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*. IEEE, 2011, pp. 2057–2063.

[41] G. Viyaj, "Keras-multiclass-image-classification," https://github.com/vijayg15/Keras-MultiClass-Image-Classification/tree/master, 2021.

[42] X. Fu, J. Huang, X. Ding, Y. Liao, and J. Paisley, "Clearing the skies: A deep network architecture for single-image rain removal," *IEEE Transactions on Image Processing*, vol. 26, no. 6, pp. 2944–2956, 2017.

[43] Y.-F. Liu, D.-W. Jaw, S.-C. Huang, and J.-N. Hwang, "Desnownet: Context-aware deep network for snow removal," *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 3064–3073, 2018.

[44] Y. P. Loh and C. S. Chan, "Getting to know low-light images with the exclusively dark dataset," *Computer Vision and Image Understanding*, vol. 178, pp. 30–42, 2019.

[45] J. Cai, S. Gu, and L. Zhang, "Learning a deep single image contrast enhancer from multi-exposure images," *IEEE Transactions on Image Processing*, vol. 27, no. 4, pp. 2049–2062, 2018.

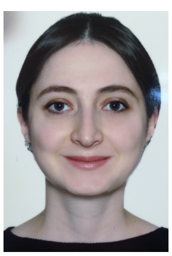
[46] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin *et al.*, "Tensorflow: Large-scale machine learning on heterogeneous distributed systems," *arXiv preprint arXiv:1603.04467*, 2016.

[47] M.-L. Zhang and Z.-H. Zhou, "Ml-knn: A lazy learning approach to multi-label learning," *Pattern recognition*, vol. 40, no. 7, pp. 2038–2048, 2007.

[48] F. Zhu, H. Li, W. Ouyang, N. Yu, and X. Wang, "Learning spatial regularization with image-level supervisions for multi-label image classification," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 5513–5522.

[49] P. Flach and M. Kull, "Precision-recall-gain curves: Pr analysis done right," *Advances in neural information processing systems*, vol. 28, 2015.

Yasin DEMİR received his B. Sc., degree in Electrical and Electronics Engineering from Karadeniz Technical University, Turkey. He received his M.Sc., degree in Electrical and Electronics Engineering from Erzurum Technical University, Turkey. He is a Research Assistant at the Electrical and Electronics Engineering Department at Erzurum Technical University, Erzurum, Turkey. His primary research interests include digital signal and image processing.



Nagihan SEVEROGLU received the B.Sc. and M.Sc. degrees in Electrical and Electronics Engineering from Atatürk University, Erzurum, Turkey, in 2012 and 2014, respectively, and the Ph.D. degree in Electrical and Electronics Engineering from Gazi University, Ankara, Turkey, in 2020. Her main research interest has been biomedical engineering during her B.Sc. studies, whereas she has focused on signal processing for power system applications during her Ph.D. research. She is currently an Assistant Professor at the Department of Electrical and Electronics Engineering at Erzurum Technical University, Erzurum, Turkey. Her research interests include signal processing, deep learning techniques, and image processing.

Nur Huseyin KAPLAN received his B. Sc., M. Sc, and Ph.D degrees in Electronics and Telecommunication Engineering from Istanbul Technical University in 2005, 2008 and in 2015, respectively. He worked for TURKSAT A.Ş. as a specialist between 2005-2015. Since 2024, he is a Professor at Electrical and Electronic Engineering Department, Erzurum Technical University, Erzurum, Turkiye, where he taughts undergraduate/graduate level courses about signal and image processing. His primary research interests include digital signal and image processing.

Sefa KUCUK received the B.Sc degree from the Department of Electrical and Electronics Engineering, Atatürk University, Erzurum, Turkey, in 2012, the M.Sc and Ph.D. degrees from the Department of Electrical and Electronics Engineering, Hacettepe University, Ankara, Turkey in 2015 and 2021, respectively. He is currently an Assistant Professor with the Electrical and Electronic Engineering Department, Erzurum Technical University, Erzurum, Turkey. His research interests include computer vision, hyperspectral imaging, image processing, inverse problems, sparse theory, and machine learning.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 1, March 2025
Research Article

Improve MPPT in Organic Photovoltaics with Chaos-Based Nonlinear MPC

Mohammad Mahdi Borhan Elmi, Osman Yildirim

Abstract—**Environmental pollution, climate changes such as melting of natural glaciers and rising sea levels are only instances of the challenges of using fossil fuels. Therefore, increasing use of clean renewable energy sources such as photovoltaic systems has a great importance. In this paper, the chaotic-based nonlinear model predictive control approach is used for extracting the maximum power of organic photovoltaic cells, which has not only a suitable tracking speed but also in fault conditions, can be useful to improve the operation level of the distribution network. This approach is a feedback-based recursive control strategy which capable of predict the proper operating state that minimizes its cost function. The proposed approach consists of two stages of estimating the reference point and regulating the operatimg point according to it. In this regard, the Lagrange function is used for managing the performance of the estimator and chaotic neural network model predictive controller to control the operation of boost converter. By using the chaos-based nonlinear model predictive controller, the amount of overvoltage is reduced by more than 1.3%. In fact, without using of control methods, the voltage range exceeds its allowable values with increasing of the OPV panels penetration. According to the obtained results, with the reduction of network losses, the capacity of distribution feeders is increased and the level of system efficiency is also improved.**

Index Terms—**Chaos Theory, Model predictive controller, MPP tracking, Nonlinear MPC, Organic Photovoltaic.**

I. Introduction

IN THE past years, the capacity of photovoltaic (PV) systems used in the power system has grown significantly. According to the report of the European Photovoltaic Industry Association (EPIA), the amount of installed PV system capacity has increased to about 345 GW by 2020 [1]. With the increase in the penetration of clean renewable sources, there are concerns about their effect on power quality indices, reliability, stability against possible faults and network losses [2].

Mohammad Mahdi Borhan Elmi, is with Department of Electrical and Electronic Engineering Faculty of Istanbul Aydin University, Istanbul, Turkey,(e-mail: mmahdielmi@stu.aydin.edu.tr).
https://orcid.org/0009-0007-1705-0214

Osman Yildirim, is with Department of Electrical and Electronic Engineering Faculty of of Istanbul Aydin University, Istanbul, Turkey, (e-mail: osmanyildirim@stu.aydin.edu.tr).
https://orcid.org/0000-0002-8900-3050
Manuscript received Jan 12, 2024; accepted Jan 09, 2025.
DOI: 10.17694/bajece.1418574

For instance, a sudden drop in the voltage level can activate the islanding mode and lead to the isolation of PV systems from the main grid. In fact, an adaptive control model not only can reduce power quality problems but also provide a reliable performance for transient stability [3]. Although the traditional methods for controlling the performance of solar arrays have a simple implementation, they have disadvantages such as slow tracking, continuous fluctuations even when reaching the maximum power point (MPP), low efficiency and lack of proper control in the transient states [4]. In [5-6] the angle of incremental conductance with type-2 fuzzy was combined to transfer the maximum power under long-term weather conditions such as broken clouds. Also, in [7] the adaptive type-2 fuzzy neural network (AT2FNN) was applied to EN 50530 test procedure system to extract maximum power under dynamic performance of the PV generation systems. According to the literature review, model predictive controller (MPC) based MPPT methods present a unique performance that increases operational efficiency and improves the speed of convergence for tracking the reference point [8] (Table I).

In [9], the MPC as a model-based control method has achieved remarkable success in industrial applications, especially in variable working conditions and the presence of limitations. In addition, these approaches have a fast dynamic response. This controller has been improved in recent decades. A brief overview of MPC applications can be found in [10]. However, at the presence of severe variations in the environmental conditions and undesired disturbances, using a linear model for approximation can not provide the proper efficiency. So, the non-linear model can be used to provide a better mapping between the current conditions and the desired outputs. Also, using of chaos theory is increasing the speed of convergence with improving the exploring ability in search space. Among the most important contributions of the proposed paper, the following can be mentioned:

- Designing an innovative chaotic-based nonlinear MPC (CNMPC) approach to control the switching process of the converter from organic solar arrays for operating at MPP with considering variable environmental conditions, network load changes and unwanted disturbances.
- Prediction of converter performance during fault occurrence in order to improve stability
- Using the chaos-based neural network to improve the quality of the reference point tracking
- Implementation of the proposed approach on the organic PVs as a new generation of solar panels with 100% recycling capability.

In the rest of the article, first, details about organic PVs, nonlinear model preictive controller, chaos-based systems are presented. Then the characteristics of the proposed approach include how to generate reference values, the designed process for regulating the operating point of the converter, The mathematical model of organic PV and Boost converter is stated and the objective function is expressed. Finally, the obtained results are presented in the form of figures and after evaluating the efficiency of the proposed approach, the conclusion is stated.

II. Basic concepts

A. Organic Photovoltaic

Organic Photovoltaic (OPV) technology has several advantages compared to silicon-based PV, including light weight, semi-transparency, high-flexibility, cheaper and simpler production process with low energy consumption, the possibility of OPV production through 3D printing on flexible substrates at low temperature, the ability to mount on flexible surfaces [11], using in indoor and outdoor environments and full recyclability (100%) [12].

An OPV cell consists of an absorber-receiver layer connected to transmission layers. Electrical contacts are used on both sides of the absorbent layers. This layer absorbs the incoming photons and if the energy absorbed from the electrons is greater than their band gap, excitation occurs. This excitability causes the occurrence of current between the layers [13]. The specifications of the different models are compared with each other in Table II.

TABLE II
COMPARISON OF ORGANIC PV CHARACTERISTICS

Model	No. of Parameters	Complexity	Ref.
One Diode	5	Low	[14]
Two Diodes	7	Medium	[15]
Two Inverted Diodes	8	Medium	[16]

B. Nonlinear Model Predictive Controller

Although the traditional methods have a simple implementation, they have disadvantages such as slow tracking, constant fluctuations even when reaching the maximum power point, and low efficiency. In this regard, several solutions have been proposed to overcome the above challenges. But these methods, unlike linear systems modeling methods, did not provide a specific method to identify the system. Recently, researchers, inspired by the human brain, presented neural networks to model nonlinear, uncertain and complex systems [17]. Among the nonlinear models, neural networks are the required model for control due to the feature of high non-linear adaptation can provide the forecast appropriately. In this project, the NMPC approach will be used, which has several advantages, some of which are as follows:

- The ability to explore precisely among various operating states of the system
- Providing the necessary flexibility to control complex systems
- Improving the level of system stability during disturbances

C. Chaotic systems

Chaos is a fundamental concept in modern science, which is observed in many phenomena of the real world, including systems with apparently random and disorderly behavior, as well as systems with definite behavior. The main essence of chaos theory is that there is order in every disorder. In the sense that order should not be sought on a small scale. Excessive sensitivity to initial conditions and having a continuous frequency spectrum are special characteristics of the dynamic behavior of chaotic systems [17]. In recent years, random value generation has been combined with chaotic mappings to use their better dynamic and statistical properties. Indeed, chaotic mapping is applied in the main structure of these proposed algorithms in order to improve the process of generating random numbers for improving the exploring capability.

TABLE I
COMPARISON OF DIFFERENT MPPT METHODS

Mppt Method	Speed	Accuracy	Sensor Type	Cost	Status	Complexity	Characteristic
VOC	Medium	Low	Voltage	Low	Offline	Simple	Linear
ISC	Medium	Low	Current	Low	Offline	Simple	Linear
P&O	Medium	Medium	Voltage Current	High	Online	Simple	Linear
INC	High	Good	Voltage Current	High	Online	Medium	Linear
ANN	High	Medium	Voltage Current	High	Online	Complex	Non-linear
Fuzzy	Medium	Medium	Voltage Current	Low	Online	Complex	Non-linear
MPC	High	Medium	Voltage Current	Low	Online	Simple	Linear

III. PROBLEM MODELLING

Due to the non-linearity of the I-V characteristic of the PV curve resulting from variation of irridiation, temperature and other effective factors, tracking the maximum power point would be a challenging task. In this regard, perturb and observe (P&O) or incremental conductivity (IC) methods have been used to generate reference values. However, these two methods can not capable to track severe changes in the environmental conditions properly. Also, the generated power of these methods fluctuates at its steady state point, which increase the amount of power loss. The strategy of the proposed approach is separated into two stages of estimating the reference point for each time step and adjusting the operating point of the converter

A. Generation of reference values

In MPC techniques, discrete equations of the system are used to predict the future state of the controlled variable. In the case of the solar array, designing an accurate model of the system at each time step is a very challenging task due to the presence of various effective factors that change continuously. Among these factors, variable solar irradiations, sudden temperature changes, and erosion of PV modules can be mentioned. In order to improve the performance efficiency, a predictor model with Lagrange polynomial transform function is used to model the P-V curve whose characteristics are stated in Equation 1. Also, the Vandermonde matrix also is used to determine the values of the coefficients, which are defined according to Eq. (1) [18]. These coefficients are updated at each sample step to obtain an accurate estimation for the P-V curve in any weather condition. In fact, it is necessary to model the variations of environmental conditions continuously in predictor model. Then the obtained points are applied to the control system for tracking. In this paper, the logistic map is used to improve the search process of extracting the global optimal point. In this mapping, a second-order polynomial function is used according to Eq. (3). In this relation, r is the control parameter ($0 \le r \le 4$) and x_0 is the initial value ($0 \le x_0 \le 1$).

$$v_{pv}(k) = a_2 . i(k)_{pv}^2 + a_1 .\ i_{pv}(k) + a_0 \tag{1}$$

$$\begin{bmatrix} i(k-2)_{pv}^2 & i_{pv}(k-2) & 1 \\ i(k-1)_{pv}^2 & i_{pv}(k-1) & 1 \\ i(k)_{pv}^2 & i_{pv}(k) & 1 \end{bmatrix} \begin{bmatrix} a_2 \\ a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} v_{pv}(k-2) \\ v_{pv}(k-1) \\ v_{pv}(k) \end{bmatrix} \tag{2}$$

$$x_{k+1} = r . x_k (1 - x_k) \tag{3}$$

B. Regulating the operating point

In this project, the combination of chaotic neural network with non-linear MPC system is used to track the reference values according to the conditions created in each time step and received values from predictor model. Neural networks has the ability to produce dynamic responses in a suitable time frame, which has made it possible to use them in systems with severe changes. In this project, the Boost converter is used to achieve high voltage gain, and the dynamic model of the converter's behavior is described by Eq. (4) to Eq. (5). In these equations, s indicates the switching mode, $v_c(t)$ is the capacitor voltage and $v_{pv}(t)$ is the PV array voltage.

In other words, the control process in the proposed CNMPC controller as shown in Fig. 1 can be described as follows:

- Determining the reference for active/reactive powers by the upper hand control level
- Calculation of the reference current of the controller
- Determination of predicted current
- Apply two currents to the cost function
- Testing the values in the cost function and choosing the desired voltage value as the controller voltage reference
- Calculation of switching times

$$L \frac{di_L}{dt}(t) = -(1-s) v_c(t) + v_{pv}(t) \tag{4}$$

$$C \frac{dv_c}{dt}(t) = (1-s) i_L(t) + \frac{1}{R} v_c(t) \tag{5}$$

C. Organic PV model

In recent years, several models such as single diode, two and three diode models have been introduced for organic PV cells [19]. Although single-diode and double-diode models are commonly used to model silicon PV cells, they can not provide proper results for describing the behavior of organic cells. So, the model of two inverted diodes is used for this purpose (Fig. 2). The related equations are stated in Eq. (6) to Eq. (10). The current-voltage characteristic of the organic PV is described by Lambert W function. In these equations, V_{D1} and V_{D2} indicate the voltage of diode D_1 and D_2, respectively.

$$V_{D1} = V - IR_s - V_{D2} \tag{6}$$

$$I = \frac{1}{R_s + R_{sH1} + R_{sH2}} [-(I_{ir} + I_{D1}) R_s + (I_{D2} . R_{sH2}) - V] \tag{7}$$

$$V_{D2} = (I - I_{D2}) . R_{sH2} \tag{8}$$

$$I_{D1} = I_{o1} \left[e^{\frac{V - IR_S - V_{D2}}{n_1 V_T}} - 1 \right] \tag{9}$$

$$I_{D2} = I_{o2} \left[e^{\frac{V_{D2}}{n_2 V_T}} - 1 \right] \tag{10}$$

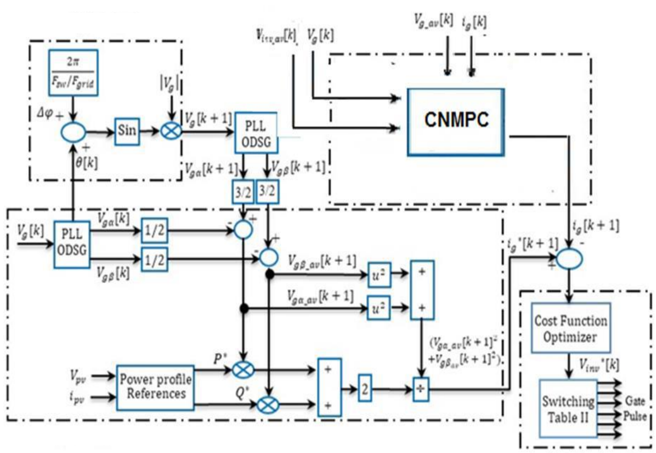


Fig.1. The structure of proposed approach

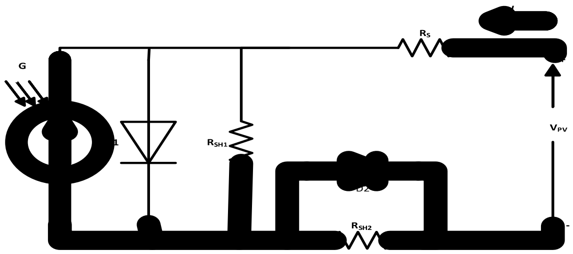


Fig.2. Two diode model of organic PV [20]

D. Boost Converter Model

A boost converter is a step-up converter that is usually used to convert voltage from low levels to higher levels. The schematic of this converter is shown in Fig. 3. In order to extract the maximum power at each time step, the designed signal commands are applied to this converter. In this project, a discrete-time model is used for modelling of the converter. In this regard, the inductor current is estimated for the next time step and then the obtained results are evaluated. According to the designed model, Eq. (11) and Eq. (12) are presented the ON state and Eq. (13) and Eq. (14) are presented the OFF state of the converter.

$$i_{pv}(k+1) = i_{pv}(k) + \frac{T_s}{L}\left(v_{pv}(k) - v_c(k)\right) \tag{11}$$

$$v_c(k+1) = \left[1 - \frac{T}{RC}\right] v_c(k) + \frac{T_s}{C}\left(i_{pv}(k)\right) \tag{12}$$

$$i_{pv}(k+1) = i_{pv}(k) + \frac{T_s}{L}\left(v_{pv}(k)\right) \tag{13}$$

$$v_c(k+1) = \left[1 - \frac{T}{RC}\right] v_c(k) \tag{14}$$

IV. Objective Function

In the proposed approach, at the first stage, the voltage and current values of the reference operating point are estimated by model predictor according to the intensity level, ambient temperature, dynamic variations of load and other effecting factors. Then, the switching pattern of the converter is determined by using the chaos-based nonlinear predictive controller.

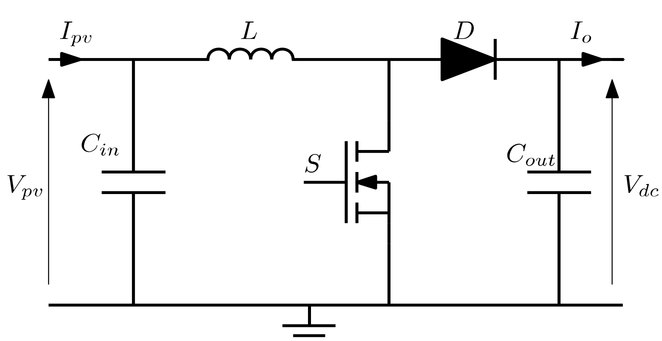


Fig.3. Schematic of boost converter [21]

The steps of the proposed approach are as follows:

- Measuring the current/voltage of the panel and sending it to the control unit
- Calculation of current and voltage variations according to Eq. (15) and Eq. (16).
- Calculation of the reference voltage by using predictor model (Eq. (17))
- Regulating the duty cycle of converter (Eq. (18))

Eq. (18) shows the objective function, which tries to minimize the difference between the current/voltage of the operating point with reference ones in each time step.

$$\frac{di_L}{dt} = \frac{v_{pv}}{L} - \frac{v_o}{L} \tag{15}$$

$$\frac{dv_{pv}}{dt} = -\frac{i_L}{C_1} + \frac{i_{pv}}{C_1} \tag{16}$$

$$\frac{dv_o}{dt} = \frac{i_L}{C_2} - \frac{v_o}{RC_2} \tag{17}$$

$$\min O.F = W_A.|V_o(k+1) - V_{MPP}|^2 + W_B.|i_o(k+1) - i_{MPP}|^2 \tag{18}$$

V. Simulation Results

In order to evaluate the proposed controller approach, the simulation was done in MATLAB/Simulink environment (Fig. 4.)

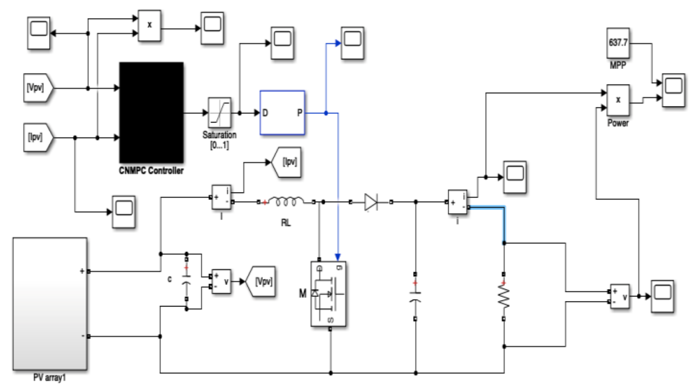


Fig.4. Simulink model of under study system

V.1 Performance evaluation during disturbances

When one of the modules receives less light due to several reasons such as dust and shading, the voltage at that point drops. In this situation, the mentioned point acts as a generator, which is known as a hotspot. Normally, in this situation, a bypass diode is used in parallel with each of the OPV modules to protect the module. In addition, a blocking diode is used at the end of each string (combination of series modules in a current path) to prevent reverse current caused by voltage mismatch between parallel strings. The voltage profile in these conditions is shown in Fig. 5.

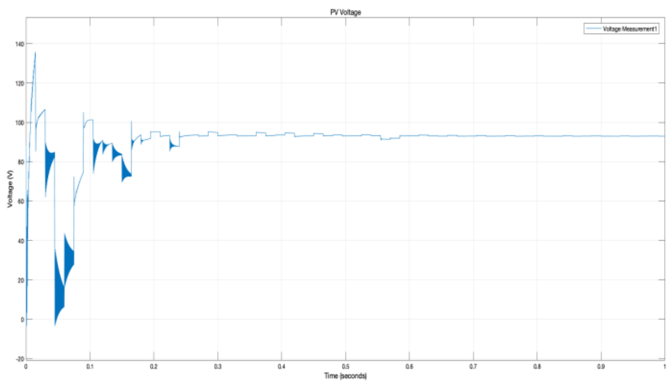


Fig.5. Output voltage of OPV panel (With considering disturbance)

The output signal of PWM modulator and the output power of the controller as well as reference current are shown in Fig. 6 and Fig. 7, respectively.

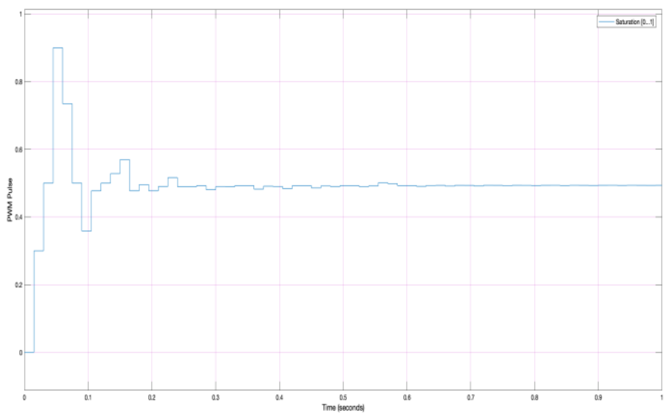


Fig.6. Output signal of PWM converter

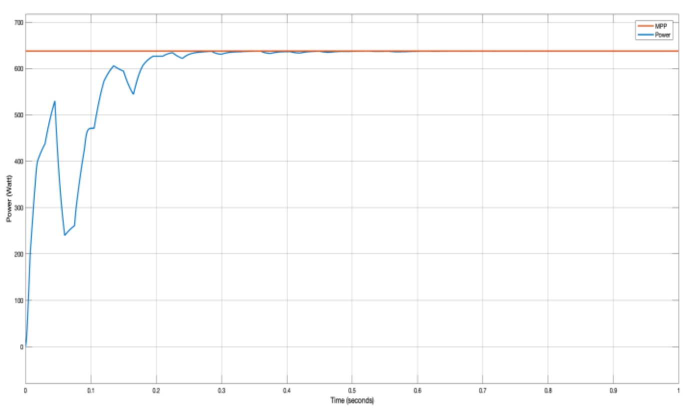


Fig.7. Tracking the desired output power

In order to tracking the desired values, a 100s time interval is considered. In the first 25s, the reference signal has a value of 1 p.u. and then in the period of 25s to 50s, the value of the amplitude is increased to 2 p.u. then in the period between 50s and 75s is decreased to -1 p.u. Finally, is converged to 0 p.u. The results of using the linear MPC shows that when the desired value changes, the tracking signal consists of significant undershoots and overshoots, which are reduced when the nonlinear MPC control method is used (Fig. 8).

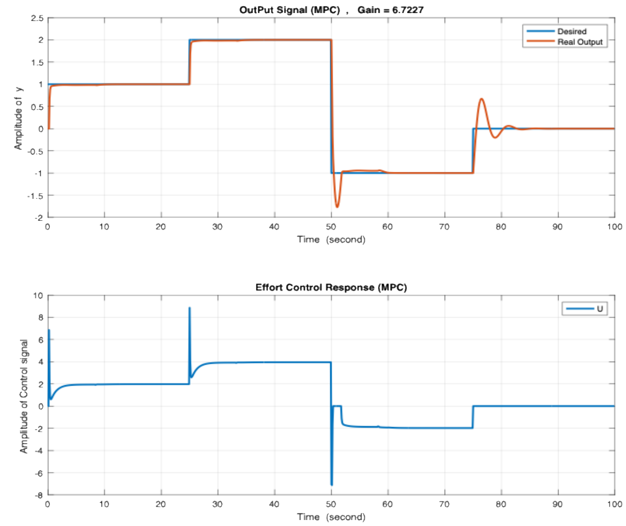


Fig.8. Linear MPC control method

(Top: Tracking of desired output – Bottom: Variation of control signal)

The output signal obtained from the implementation of the nonlinear MPC approach and the chaos-based nonlinear MPC approach are also shown in Fig. 9 and Fig. 10, respectively. As it is shown from the results, in the linear MPC, the rate of convergence of the output signal insists of severe variations, but by using the nonlinear MPC approach and chaos-based non-linear MPC, this is greatly reduced and lead to faster convergence. The output power of the controller in two states (real output and reference mode) is shown in Fig. 8. In this situation, it can be seen that tracking has been done properly using the proposed approach.

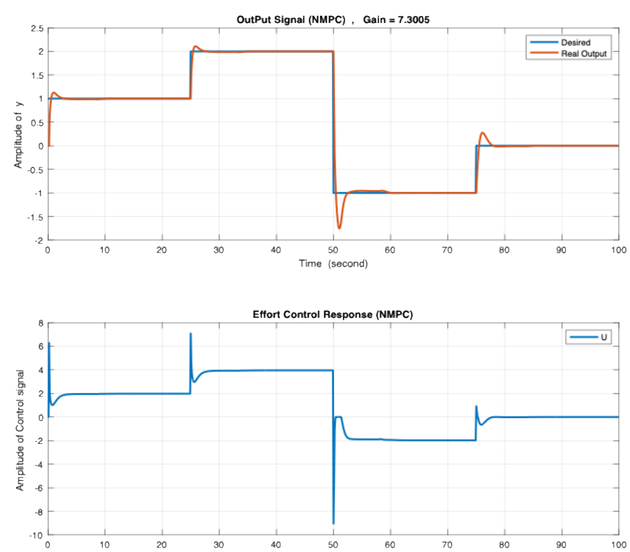


Fig.9. Non-linear MPC control method

(Top: Tracking of desired output – Bottom: Variation of control signal)

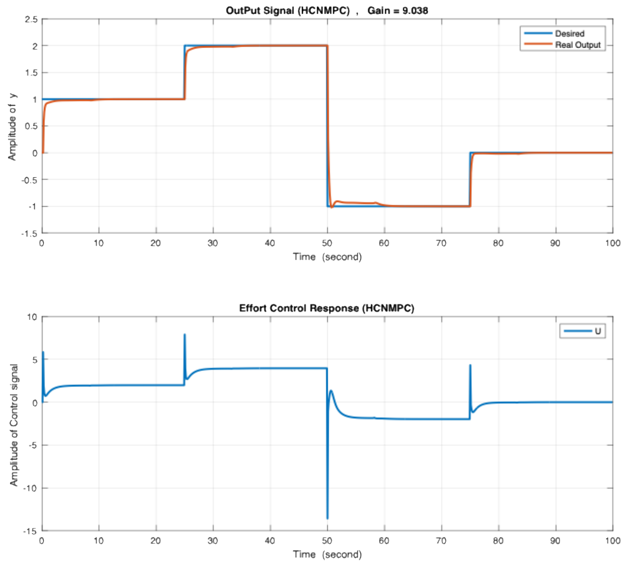


Fig.10. Chaos-based Non-linear MPC control method

(Top: Tracking of desired output – Bottom: Variation of control signal)

V.2 Results comparison

In this section, the numerical results obtained by using different methods are compared with each other. As can be seen in Table. 3, by using the CNMPC, the amount of overvoltage is reduced by more than 1.3%. In fact, without using of control methods, the voltage range exceeds its allowable values with increasing of the OPV panels penetration.

TABLE III
COMPARISON OF VOLTAGE PROFILES (P.U.)

Status	Phase 1	Phase 2	Phase 3
Without using OPV arrays	0.985	0.984	0.983
OPV arrays + P&O controller	1.024	1.035	1.017
OPV arrays + CNMPC controller	1.010	1.002	1.005

VI. CONCLUSION

Reactive power injection can be used to restore voltage drop in distribution networks. In this paper, the constant peak current adjustment method is used to protect the converter system against overcurrent and if the voltage reduction is less than 50%, all injected power will be converted into reactive power. In other words, although the function of the MPPT controller is to provide the maximum power that can be extracted from the solar panels, but in the fault duration, the reference values for active and reactive power is estimated by predictor model and the operation of the converter is controlled by chaos-based NMPC controller. Indeed, the strategy of the proposed approach is divided into two stages, the estimation of the reference point in each time step and the adjustment of the operating point of the converter in order to extract the maximum power. In this regard, the Lagrange polynomial tarnsform function was used to model the power-voltage curve predictor, whose coefficients are updated at each sampling step. Also, the combination of chaotic neural network with NMPC system has been used to track the reference values according to the dynamic conditions of environment. In addition, the modeling of the control approach has been done on organic solar panels, which are not only fully recyclable, but also have a higher efficiency than the silicon-based PV cells.

REFERENCES

[1] Menna, P., R. Gambi, T. Howes, W. Gillett, G. Tondi, F. Belloni, P. De Bonis, and M. Getsiou. "European Photovoltaic Actions and Programmes-2011”." *EU PVSEEC* (2011).

[2] Razavi, Ali Bannae, and Mohammad Mahdi Borhan Elmi. "Improvement of Voltage Profiles in Mashhad Distribution Systemwith Presence of Rooftop PV." In *2020 10th Smart Grid Conference (SGC)*, pp. 1-6. IEEE, 2020.

[3] Borhan, Elmi Mohammad Mahdi, Hamed Lotfi, and Amir Hossein Lotfi. "Economic-environmental evaluation of using green energy resources with considering pollution limitations." (2023).

[4] Bouaouaou, Hemza, Djaafer Lalili, and Nasserdine Boudjerda. "Model predictive control and ANN-based MPPT for a multi-level grid-connected photovoltaic inverter." *Electrical Engineering* 104, no. 3 (2022): 1229-1246.

[5] Gani, Ahmet, and Mustafa Sekkeli. "Experimental evaluation of type-2 fuzzy logic controller adapted to real environmental conditions for maximum power point tracking of solar energy systems." *International Journal of Circuit Theory and Applications* 50, no. 11 (2022): 4131-4145.

[6] Kececioglu, O. Fatih, Ahmet Gani, and Mustafa Sekkeli. "Design and hardware implementation based on hybrid structure for MPPT of PV system using an interval type-2 TSK fuzzy logic controller." *Energies* 13, no. 7 (2020): 1842.

[7] Gani, Ahmet. "Improving dynamic efficiency of photovoltaic generation systems using adaptive type 2 fuzzy-neural network via EN 50530 test procedure." *International Journal of Circuit Theory and Applications* 49, no. 11 (2021): 3922-3940.

[8] Derbeli, Mohamed, Cristian Napole, Oscar Barambones, Jesus Sanchez, Isidro Calvo, and Pablo Fernández-Bustamante. "Maximum power point tracking techniques for photovoltaic panel: A review and experimental applications." *Energies* 14, no. 22 (2021): 7806.

[9] Abdel-Rahim, Omar, and Haoyu Wang. "A new high gain DC-DC converter with model-predictive-control based MPPT technique for photovoltaic systems." *CPSS Transactions on Power Electronics and Applications* 5, no. 2 (2020): 191-200.

[10] Rampradesh, T., and C. Christober Asir Rajan. "Performance assessment of NMPC based MPPT controller and extended Kalman filter for a Wind/PV hybrid system." In *2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, pp. 1-4. IEEE, 2021.

[11] Välimäki, Marja K., Laura I. Sokka, Heidi B. Peltola, Sami S. Ihme, Teijo MJ Rokkonen, Timo J. Kurkela, Jyrki T. Ollila, Arttu T. Korhonen, and Jukka T. Hast. "Printed and hybrid integrated electronics using bio-based and recycled materials—increasing sustainability with greener materials and technologies." *The International Journal of Advanced Manufacturing Technology* 111 (2020): 325-339.

[12] Piccoli, Emanuele, Alessandro Dama, Alberto Dolara, and Sonia Leva. "Experimental validation of a model for PV systems under partial shading for building integrated applications." *Solar Energy* 183 (2019): 356-370.

[13] Mei, Hongkun, Chunjuan Jia, Jun Fu, and Xiaoming Luan. "Low voltage ride through control strategy for MMC photovoltaic system based on model predictive control." *International Journal of Electrical Power & Energy Systems*125 (2021): 106530.

[14] Mehta, Hitesh K., Himanshu Warke, Kaushik Kukadiya, and Ashish K. Panchal. "Accurate expressions for single-diode-model solar cell parameterization." *IEEE Journal of Photovoltaics* 9, no. 3 (2019): 803-810.

[15] Hejri, Mohammad, Hossein Mokhtari, Mohammad Reza Azizian, Mehrdad Ghandhari, and Lennart Söder. "On the parameter extraction of a five-parameter double-diode model of photovoltaic cells and modules." *IEEE Journal of Photovoltaics* 4, no. 3 (2014): 915-923.

[16] Moliton, André, and Jean-Michel Nunzi. "How to model the behaviour of organic photovoltaic cells." *Polymer International*55, no. 6 (2006): 583-600.

[17] Tahir, Furqan, and Imad M. Jaimoukha. "Robust feedback model predictive control of constrained uncertain systems." *Journal of Process Control* 23, no. 2 (2013): 189-200.

[18] Pecora, Louis M., and Thomas L. Carroll. "Synchronization in chaotic systems." *Physical review letters* 64, no. 8 (1990): 821.

[19] Liu, Zhijian, Zhao Jin, Guiqiang Li, Xudong Zhao, and Ali Badiei. "Study on the performance of a novel photovoltaic/thermal system combining photocatalytic and organic photovoltaic cells." *Energy Conversion and Management* 251 (2022): 114967.

[20] Uvais, Mohd, Asif Jamil Ansari, and Mohammed Asim. "Modeling and Analysis of Organic Solar Cells Using Multiple Diodes." In *2022 2nd International Conference on Emerging Frontiers in Electrical and Electronic Technologies (ICEFEET)*, pp. 1-5. IEEE, 2022.

[21] Meddour, Sami, Djamel Rahem, Ali Yahia Cherif, Walid Hachelfi, and Laib Hichem. "A novel approach for PV system based on metaheuristic algorithm connected to the grid using FS-MPC controller." *Energy Procedia* 162 (2019): 57-66.

BIOGRAPHIES

Mohammad M. Borhan Elmi Dubai, UAE, in 1984. He received the B.S. degree in power electrical engineering from the Islamic Azad University of Mashhad, Iran in 2007, first M.S. degree in electrical engineering from the Islamic Azad University of Saveh, Iran in 2010 and second M.S. degree in electronic and elecric emgineering from Istanbul Aydin University of Istanbul, Turkey, in 2024. From 2009 to 2014, he was an Educational Assistant with the Islamic Azad University. From 2014 to 2022, he had been teaching as the head of electrical department in Sobhan Institute of Higher Education. He is the author of more than 20 articles. His research interests include renewable resources, smart grids, home energy management, maximum power point tracking, and optimization algorithms.

Osman Yildirim has received two PhD, one is in the field of Engineering, one is in the field of Business Management, specifically in Human Resources Management. He has international journal papers in both disciplines. He has been dealing with OHSAS for a long time by teaching the OHSAS training courses and giving consultancy in OHSAS to companies. He has both trainer certificate and consultant certificate which are accredited by Turkish Employee Ministry. He is currently working with Istanbul Aydin University, Istanbul, Turkey.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 1, March 2025
Research Article

Optimization of Credit Card Fraud Based on Roulette and Tournament Selection by Genetic Algorithm and Artificial Bee Colony

Ridvan Cubukcuogullari, Hasan Temurtas and Cigdem Bakir

***Abstract*— Credit card fraud has become a major problem, especially with the spread of e-commerce and the increase in online shopping during the COVID-19 period. Theft of credit card information and transactions made with cards belonging to others lead to legal and economic problems. The aim of this study is to develop Genetic Algorithm (GA) and Artificial Bee Colony (ABC) methods based on roulette and tournament selection instead of random selection in the determination and classification of these illegal credit card fraud transactions. A data set consisting of 28 attributes and 284,807 credit card transactions was used in the study. Fraud in credit card transactions was estimated using genetic algorithm and artificial bee colony and the obtained results were compared. Optimization methods such as GA and ABC developed based on roulette and tournament selection were analyzed separately according to linear, quadratic and exponential functions. Using the linear, quadratic and exponential functions for the test data, the GA based on roulette selection showed success rates of 98.6%, 98.46% and 98.6% in identifying credit card fraudulent transactions, respectively; With the GA based on tournament selection, it shows 98.53%, 98.33% and 98.66% success. In addition, for the test data, using linear, quadratic and exponential functions, ABC method achieved 98.6%, 97.86% and 97.93% success in determining credit card fraudulent transactions, respectively. Success results were calculated with different evaluation criteria such as accuracy, recall and F1-Score and performance evaluation was presented for each proposed method.**

Index Terms— **Artificial Bee Colony, Genetic Algorithm, linear, roulette, tournament.**

Rıdvan Çubukcuğulları, is with Department of Computer Engineering University of Dumlupinar University, Kutahya, Turkey, (e-mail: ridvan.cubukcuogullari0@ogr.dpu.edu.tr).
https://orcid.org/0000-0001-7889-9114

Hasan Temurtaş, is with Department of Computer Engineering University of Dumlupinar University, Kutahya, Turkey, (e-mail: hasan.temurtas@dpu.edu.tr).
https://orcid.org/0000-0001-6738-3024

Çiğdem Bakır, is with Department of Software Engineering University of Dumlupinar University, Kutahya, Turkey, (e-mail: cigdem.bakir@dpu.edu.tr).
https://orcid.org/0000-0001-8482-2412

Manuscript received Mar 06, 2024; accepted Mar 04, 2025.
DOI: 10.17694/bajece.1447975

I. INTRODUCTION

IN THE past, people used to meet their needs with money or other means of transformation, but now, with the development of technology, they have slowly begun to change their habits. Although it has not fully spread to the base, people now fulfill their daily needs by using credit cards or debit cards without buying paper money or coins. Nowadays, using credit cards has become very popular. Some people even make purchases with credit cards by using the NFC feature of their phones or by scanning the barcode on their phones [1].

As people's use of credit cards increased, they automatically became the target of fraudsters who made their living by stealing people's savings. That's why detecting credit card fraud has become an issue that requires precautions for both people and the organizations that issue these cards to people [2]. While preventing credit card fraud or fraud has become a matter of prestige for institutions, proving that the cards provided by banks or institutions are reliable and cannot be the target of fraudsters has become important both financially and legally [3].

Some studies have been conducted in the literature on the detection of credit card fraud. Yeşilyurt et al. focused on the increase in credit card usage with the development of digital technologies and the threat of fraud becoming an important problem, and carried out a study on the detection of credit card fraud [4]. Artificial Neural Networks (ANN) have been used to detect and prevent credit card fraud. 284,807 credit card transaction data obtained from the database obtained via Kaggle were studied. The data set was normalized via MATLAB and training, validation and test sets were created. Modeling was done using different ANN structures and success rates were evaluated. It was stated that the model proposed for the best results had [28 20 15 10 2] neuron numbers and 98.44% success was achieved, and it was also emphasized that the use of different training and performance functions also affected the success. As a result, it has been stated that Artificial Neural Networks are an effective tool in detecting credit card fraud and constitute an important area for future studies.

In the study conducted by Çilburunoğlu, various machine learning and deep learning techniques are examined to detect credit card fraud and the effectiveness of methods such as Logistic Regression, Support Vector Machines (SVM), Decision Trees, Random Forest and Artificial Neural Networks (ANN) is evaluated [5]. The dataset was collected in collaboration between Worldline and the Universite Libre de Bruxelles (ULB) and contains a total of 284,807 transactions. Fraudulent transactions account for 0.172% of total transactions. Logistic Regression and SVM achieved 95% accuracy rates, while Decision Trees and Random Forest showed 98% and 97% accuracy rates, respectively. Artificial Neural Networks have achieved a perfect 100% accuracy rate. The results show that different algorithms exhibit different performances and that the choice of model should be appropriate to the characteristics of the algorithm and the dataset. This research can contribute to cost savings and secure shopping experiences for financial institutions.

In the study conducted by Hüssein, machine learning and deep learning methods were used to detect credit card fraud [6]. The dataset is the Worldline-ULB Machine Learning Group dataset containing 284,807 transactions. Of these transactions, 492 are fraudulent and only 0.172% are actually fraudulent, creating a serious imbalance. Most transactions are favored by the prediction algorithm when fed with highly skewed class distribution data, resulting in fraudulent transactions often being hidden. To address this issue, a data level strategy was implemented using multiple sampling methods. The proposed system was compared with logistic regression, support vector machine, naive Bayes classifier, decision tree, random forest, XGBoost, modified XGBoost, KNN and multilayer perceptron. These studies used many educational characteristics and class distributions. Bagging and boosting algorithms were used to solve class imbalance. Several experiments were conducted to determine the optimal classifier. Random forest emerged as the strongest model, achieving 99.957% accuracy, 0.875 F1 score, 0.857 recall, 0.893 precision, and 0.98 ROC AUC values. Measurements in other models varied.

In the study by Jovanovic et.al, they used sophisticated machine learning techniques to prevent increases caused by credit card fraud and detect fraud [7]. Additionally, by applying developed meta-heuristic methods and swarm intelligence approaches, they compared the success of these methods. The Firefly algorithm was applied to synthetic and real credit card datasets and its classification success was evaluated. However, the study needs to be evaluated with larger data sets and more parameters need to be determined to evaluate the performance of the methods used.

In the study conducted by Akyüz, it was aimed to create a prediction model that accurately detects theft transactions by analyzing the transactions of credit card customers [8]. Various steps were followed, from data preparation to model building and selecting performance parameters using machine learning algorithms. First, an initial prediction model was produced and then the effects of different observation sizes and theft rates on the test data were examined. Due to the observation imbalance problem, the F1 score was determined as the primary performance measure. Then, features such as customer number, expenditure amount, transaction date and auxiliary features covering various periods were produced. Classification predictions were made with these produced features. After examining the results, three different grouping methods were proposed. The selected single prediction models were grouped by various methods and it was observed that the success level of the grouped models increased compared to the single models.

The study conducted by Shah was conducted in the Python programming language and used the data set on credit card transactions of European credit card holders in 2013 obtained via Kaggle [9]. Since the data set was unbalanced, SMOTE (Synthesized Minority Oversampling Technique) technique was used. Sampling techniques such as Oversampling and Undersampling were also used to compare results with SMOTE and determine which sampling technique performed better. It has been concluded that many Machine Learning algorithms can be used to detect fraudulent transactions. These algorithms are classification algorithms such as Logistic Regression, Naive Bayes, Random Forest, K-Nearest Neighbor and Artificial Neural Network. A comparative analysis is performed to determine which model performs best and provides the most appropriate solution among these algorithms.

The study conducted by Soylu aims to detect fraudulent credit card transactions using automation systems [10]. These systems must have a high rate of correctly identifying fraudulent transactions. For this purpose, credit card transactions were classified using machine learning methods. In a study conducted in September 2013 using deep learning, random forest and classifier stack methods on a dataset containing credit card transactions of European cardholders, the AUC was 96.3% with the deep learning method, 95.6% with the random forest method, and 97.9% with the classifier stack method. It has achieved successful results with its (Area Under the Curve) value. Considering these results, it is concluded that it is important considering that different sampling methods were used.

Save et.al, conducted a study using a mixed method to detect credit card fraud [11]. Researchers obtained results using decision tree, Luhn and Hunt algorithms on previous data. Using a decision tree, customer and transaction details were matched. If the details match the previous transaction details, the transaction is considered real. Based on these details, previous transactions were clustered. The new transaction data is matched with the previous clusters; If the details match, the transaction is considered genuine, but if there is a mismatch, it is marked as fraud. Luhn and Hunt algorithms were used to verify card details and customer's address with transaction details. Luhn and Hunt algorithms calculate the sum of card numbers and compare it to the sum of odd and even numbers, preventing fraudsters from using false information. The researchers claimed that frauds can be minimized with this technique and false alarms will be minimal as each transaction data is matched with the set based on customer details and previous data.

In the study conducted by Shakya using machine learning techniques, ensemble models are used with resampling methods [12]. A real dataset containing 284807 transactions obtained from Kaggle was combined with different machine learning algorithms and resampling techniques to create a model for detecting fraudulent transactions. Only 492 of these transactions were detected as fraudulent transactions. The random forest algorithm performs better than other models, especially when used with SMOTE and Tomek Connections.

Taha et al. proposed an optimized light gradient boosting machine (OLightGBM) method to detect fraudulent transactions in credit card usage [13]. This method used Bayes and lightweight gradient boosting machine (LightGBM) for hyperparameter optimization. This study was carried out on two different datasets and approximately 98.40% accuracy was obtained. In addition, it was observed that more successful results were obtained compared to other methods (regression, kNN, SVM) with different evaluation criteria such as 92.88% sensitivity and 92.88% AUC. In this study, it is emphasized that more successful results will be achieved with more effective parameter optimization in the proposed method. However, fine-tuning parameter optimization methods may be needed to produce more accurate results in larger data sets.

Prabhakaran and Nedunchelian used machine learning and deep learning methods to detect credit card fraud accurately and reliably [14]. This proposed method is called deep learning-based cat herd optimization-based feature selection (OCSODL - CCFD). The purpose of this model is to effectively classify and detect fraudulent and genuine transactions in credit card use. It proposes an OCSO-based algorithm for optimal feature selection. Parameter adjustment of the proposed model was made with the Chaotic Krill Herd algorithm (CKHA). Numerous analyzes were carried out with different evaluation criteria and the proposed model was compared with other methods. 99.97% success was achieved in detecting credit card fraudulent transactions with the OCSODL-CCFD technique. Additionally, clustering and outlier analysis methods can be used to increase classification success in the future.

Nowadays, with the development of technology in recent years, e-commerce applications have also increased significantly. Accordingly, credit card usage is increasing and credit card fraud is increasing day by day. Especially companies and customers who frequently use credit cards lose large amounts of money due to credit card fraud and face financial difficulties. In addition, credit card fraudsters are constantly adding new fraud transactions every day by using different technologies. For this reason, it is becoming more and more difficult to detect fraudulent credit card transactions. Prusti and his colleagues performed credit card fraud detection with genetic algorithm and potential solution-based particle swarm optimization (PSO) [15]. The PSO algorithm has been tested with different population and herd numbers and the most appropriate parameters have been tried to be determined. Additionally, PSO and genetic algorithm results are presented in comparison with each other. This proposed model achieved a prediction success of 91.58%. By increasing the number of iterations, the mean square error (MSE) was reduced to 0.67% in both models.

Optimization methods are used to solve problems that are impossible or very difficult to solve. Optimization algorithms are also referred to as metaheuristic methods. The purpose of these methods is to estimate the best solution to the problem under certain conditions. In our study, credit card fraud was tried to be detected by using the artificial bee colony algorithm, which is one of the optimization algorithms, and the genetic algorithm and artificial bee colony developed based on roulette and tournament selection. The performance results of these optimization algorithms according to different function types were compared and various analyzes were carried out to reveal the best result. In addition, box plot analyzes according to the linear, quadratic and exponential functions of the optimization methods used, graphs showing comparative cost functions and error rates for 100 iterations are presented, and each optimization method is given in detail.

This study presents a new approach to credit card fraud detection using a combination of genetic algorithms and artificial bee colony algorithms. In particular, the speed and accuracy of genetic algorithms were increased by tournament selection, while the artificial bee colony algorithm provided continuous improvement of the solutions. The study compared the performance of optimization algorithms with three different types of functions and determined the most suitable parameters of each. This method clearly demonstrated how the algorithms performed in different scenarios and allowed the development of more efficient models. In addition, the success of the algorithms was made more understandable by using advanced visualization techniques and practical evaluations were made. In conclusion, this study makes a significant contribution to the literature by proposing an effective and innovative model for credit card fraud detection.

This study provides significant scientific contributions in the field of credit card fraud detection. The difference of the study from other studies is that it detects fraudulent transactions using an innovative combination of optimization algorithms. In particular, the combination of ABC with roulette and tournament selection-based GA, which is a rarely used approach in the literature, provides more flexible and high-performance results. In addition, a more comprehensive optimization analysis was performed using three different types of functions (linear, quadratic and exponential) in the performance evaluation of the algorithms. This reveals how the algorithms perform under different conditions and guides the selection of the most appropriate algorithm. The study also examines the strengths and weaknesses of the algorithms in detail by evaluating their performance separately for each class with metrics such as Precision, Recall and F1 Score. The analysis conducted with real-world data consisting of approximately 284,807 credit card transactions obtained from Kaggle increased the validity of the study and the training and test data ratio was determined as 70-30% to test the generalization ability of the model. The boxplot analyses used in the study visually present the performance of the algorithms, making the results more understandable. These elements distinguish the study from other credit card fraud detection studies in the existing literature and provide an important scientific contribution to obtaining more efficient and accurate results.

II. Material and Method

In this study, genetic algorithm based on roulette and tournament selection and artificial bee colony were used.

A. Genetic Algorithm

Genetic algorithm is a heuristic search algorithm inspired by natural selection and genetics, aiming to solve the problem using mutation and crossover methods [16]. Just as individuals who live organically in nature and adapt to the environment continue to live longer and sustain themselves and their species, the same logic underlies the genetic algorithm. It is a very logical and reasonable option for stronger, more durable individuals to continue their generation. An equal number of individuals who are considered to have a higher chance in the population are selected and matched. A new individual is formed by crossing over between the chromosomes of these matched individuals. So what actually happens is a child. Mutation is performed on the genes of the new individual that occurs under certain probability conditions. It is observed that strong and suitable individuals continue their lives because they have a high chance of mating, while weak and unsuitable individuals are eliminated from the population. The natural selection process is operated in the algorithm. The genetic algorithm consists of the following steps [17].

1. The number of the initial population is determined depending on the desired problem.

2. The fitness value is calculated for each sequence or chromosome.

3. The following steps are repeated until a new population is formed.

a. Two populations are randomly selected among the populations, taking into account the fitness value.

b. To create new individuals, a crossover process is performed for n individuals. If the resulting individuals and their parents are desired to be different, a crossover must be made.

c. Mutation is done for diversity.

d. The resulting individual is transferred to the population pool.

4. The old population is replaced with the newly formed population.

5. Testing is carried out and if the termination criteria are met, the study is terminated.

6. If the termination criterion is not met, return to item number 2.

The genetic algorithm is performed with the following steps:

Coding: The continuation of the generation continues thanks to the coded genetic information on DNA. If we are going to model with a genetic algorithm, of course the parameters of the planned algorithm must be encoded with specific genetic functions. Before the parameters are optimized, they are converted to the desired shape.

Creating the Initial Population: A solution group is created. In this solution group, possible solutions are coded [18]. This group of solutions created is called population, and the codes of the solutions are called chromosomes. Binary coding is used to represent chromosomes. Random number generators can be used to create the initial population. The random number generator is called and the generated value is looked at. If the value is less than 0.5, it is set to 0, and if it is greater than 0.5, it is set to 1. As a result of this process, the population and the sequences/chromosomes within the population are created.

Calculation of Fitness Value: After the initial population is created, the next step is to calculate the fitness value [19]. For example, for a maximization problem i. The individual's fitness value f(i) is generally the value of his intended function at that point.

Application of the Crossover Process: In order to determine the quality and potential of the new generation to be obtained from the existing population, the crossover process is carried out by selecting individuals with high fitness values.

Application of the Mutation Process: If we cannot get a satisfactory solution from crossover, or if the coded chromosome structure in the population does not reach a sufficient level of information, or if the results we want and the results obtained are far from the solution, a new operator is needed to obtain a new chromosome, and this operator is called the mutation operator. .

Formation of the New Generation and Stopping the Cycle: Starting from the initial population until the stopping criterion is terminated for a certain number of iterations, a new generation of individuals is created as a result of reproduction, crossover and mutation steps, and this process continues in this way.

Termination Operator: Termination operator is used to decide what to do with the genetic algorithm, whether to continue the research or stop the research. The genetic algorithm checks the termination criterion after each generation, and depending on the value here, the research either continues or ends.

Selection: The selection process is actually the process of determining which individuals will be selected from the population while producing new individuals. Fitness value is also important here. Good, strong individuals should continue their lives, new individuals should emerge from these strong individuals, and this cycle should continue in this way. Selection operators are used when creating new individuals. We can list some of these as roulette selection, boltzman selection, tournament selection and sequential selection.

Roulette Wheel/Circle Method: In this selection, chromosomes are gathered around the roulette according to their fitness value. By summing the fitness values of all individuals in the population and dividing the fitness value by this total value, the probability of selection of each individual emerges. The reason why this method is called the roulette wheel is because it simulates the results obtained when the circle is rotated by slicing the individuals in the population in proportion to their fitness values.

$$TUD = \sum_{i=1}^{n} UD(i) \quad (1)$$

TUD = Sum of total fitness values of individuals in the population
UD(i)= i. fitness value of the individual

Tournament Method: It is similar to the tournament system in football matches. Two randomly selected individuals are entered into the tournament. Whoever is better in the tournament wins. Those who are not good lose. Just like in the tournament system in football, the team that won the matches first on its own field and then on the opponent's field was placed in the next round. In this system, each individual, that is, the series, has the opportunity to participate in the tournament twice. The individual who is successful in both tournaments moves to the next population and is thrown into the matching pool, creating a new population here.

B. Artificial Bee Colony

The Artificial Bee Colony Algorithm (ABC) is a recently introduced optimization algorithm that simulates the foraging behavior of a bee colony [20]. ABC divided artificial foraging bees into three groups. These; worker bees, onlooker bees and scout bees. In fact, it has been observed that half of the colony consists of worker bees and the other half consists of onlooker bees. Worker bees search for nutrients around the food source in their memory and transfer the nutritional information they find to the onlooker bees. Onlooker bees tend to choose good food sources found by employed bees and search for nutrients around these selected food sources. A worker bee assigned to a food source turns into a scout bee after the food source is consumed [21].

1. Population initiation: The initial population of solutions is populated with the SN number of randomly generated n-dimensional real-valued vectors (i.e., food sources) [22]. Xi = {xi,1,xi,2, . . . ,xi,n} i in the population. represent the food source, and then each food source is generated as follows:

$$xi,j = xminj + rand(0,1)(xmaxj - xminj) \quad (2)$$

Here i = 1,2,…,SN, j=1,2,…,n. and xminj and xmaxj are the lower and upper bounds, respectively, for dimension j. These food sources are randomly assigned to the SN numbers of worker bees and their suitability is evaluated.

2. Initiation of bee phase: In this phase, each employed bee Xi produces a new food source Vi near its current location using the following solution search equation [23]:

$$vi,j = xi,j + \phi i,j(xi,j - xk,j) \quad (3)$$

Here $k \in \{1,2,...,SN\}$ and $j \in \{1,2,...,D\}$ are randomly selected index values. k must be different from i and Φij is a random number in the range [1, 1].

Once Vi is obtained, it will be evaluated and compared with Xi. If the fitness of Vi is equal to or better than that of Xi, Vi will replace Xi and become a new member of the population; otherwise Xi will be preserved. In other words, a greedy selection mechanism is used between old and candidate solutions.

3. Calculation of probability values involved in probabilistic selection: After all employed bees finish their searches, they share their information about nectar amounts and the location of their sources with the onlooker bees. The onlooker bee evaluates the nectar information it receives from all employed bees and then determines the food source location with a probability depending on the amount of nectar. This choice depends on the fitness values of the solutions in the population. It could also be a fitness-based selection scheme, a roulette wheel, ranking-based, stochastic universal sampling, tournament selection, or any other selection scheme.

4. Onlooker bee stage: The onlooker bee evaluates the nectar information it receives from all employed bees and selects a food source Xi based on the probability value pi. After the viewer chooses the food source Xi, he produces a change on Xi using the equation. If the modified food source has a better or equal amount of nectar than Xi, as in employed bees, the modified food source will replace Xi and become a new member of the population.

5. Scout bee stage: If a food source Xi cannot be improved further by a predetermined number of trials, it is concluded that the food source has been abandoned and the corresponding worker bee becomes a scout. The watcher generates a random food source as shown below:

$$xi,j = xminj + rand(0,1)(xmaxj - xminj) \quad (4)$$

III. EXPERIMENTAL STUDY

In this study, genetic algorithm based on roulette and tournament selection and artificial bee colony were used. In this study, a dataset consisting of 284,807 credit card transactions in 2013 was used (https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud). There are two classes in this dataset: fraudulent transaction and normal transaction. The dataset used is divided into 70% training and 30% testing.

A. Genetic Algorithm Results Based on Roulette Selection

The parameter values of the proposed genetic algorithm based on roulette selection are given in Table 1. These parameter values gave the most optimum results for the dataset used in the study.

TABLE I
GENETIC ALGORITHM PARAMETER VALUES ON ROULETTE SELECTON

Genetic Algorithm Based on Roulette Selection	
Number of Runes	50
Number of Iterations	100
Population Number	40
Number of Variables	29
Number of Inputs	28
Number of Outputs	2
Number of Bits	10
Crossover Rate	0,9
Number of Mutations	8
Mutation Rate	0,15
Number of Crossovers	0 (Random)
Number of Data	5000
Number of Class 1 Samples	4508
Number of Class 2 Samples	492
Number of Training Samples	%70 – 3500
Number of Test Samples	%30 - 1500

It is aimed to obtain the results by training and testing the data on the credit card data set with a linear function using the genetic algorithm (Roulette Method). A credit card fraud data set with a total of 284,807 transactions was used, and 492 of these transactions were fake (fraudulent) transactions. The Roulette method, a type of Genetic Algorithm, was used and 50 runs were studied, each with 100 iterations and a population of 40 individuals. At the end of each study, performance measures such as Mean Squared Error and classification accuracy were reported. The best result was found in the 26th run and the worst result was found in the 44th run. In Table 2, the performance of the classification model in training, testing and all data in terms of Accuracy, Recall, Specificity, Precision and F1 Score, according to the fraud and non-fraud cases. has been evaluated. It has been observed that the samples without credit card fraud give more successful results when the results are compared with the samples with credit card fraud according to different evaluation criteria for training, testing and all data.

TABLE II
GENETIC ALGORITHM RESULTS BASED ON ROULETTE SELECTION ACCORDING TO LINEAR FUNCTION

Class / Results	Lineer (Rulet)			
		Train	Test	All
Class 1 (No Fraud)	Accuracy	0.9828	0.9860	0.9838
	Recall	0.9987	0.9970	0.9982
	Specificity	0.8372	0.8851	0.8516
	Precision	0.9825	0.9875	0.9840
	F1Score	0.9905	0.9922	0.9910
Class 2 (Fraud)	Accuracy	0.9828	0.9860	0.9838
	Recall	0.8372	0.8851	0.8516
	Specificity	0.9987	0.9970	0.9982
	Precision	0.9863	0.9703	0.9812
	F1Score	0.9056	0.9257	0.9118

The results were obtained by training and testing the data on the credit card data set with the Exponential Function using the genetic algorithm (Roulette Method) in Table 3. The best result was found in the 22nd run and the worst result was found in the 9th run. According to the exponential function, initially, as the number of non-forged samples increases, the model's accuracy increases; however, as the number of forged samples increases, the model's accuracy decreases. This shows a trend where the model initially exhibits higher accuracy but then fails after a point due to the difficulty of detecting forgery.

TABLE III
GENETIC ALGORITHM RESULTS BASED ON ROULETTE SELECTION ACCORDING TO EXPONENTIAL FUNCTION

Class / Results	Exponential (Rulet)			
		Train	Test	All
Class 1 (No Fraud)	Accuracy	0.9828	0.9860	0.9838
	Recall	0.9984	0.9948	0.9973
	Specificity	0.8401	0.9054	0.8597
	Precision	0.9828	0.9896	0.9848
	F1Score	0.9905	0.9922	0.9910
Class 2 (Fraud)	Accuracy	0.9828	0.9860	0.9838
	Recall	0.8401	0.9054	0.8597
	Specificity	0.9984	0.9948	0.9973
	Precision	0.9829	0.9503	0.9724
	F1Score	0.9059	0.9273	0.9126

The results were obtained by training and testing the data on the credit card data set with the quadratic function using the genetic algorithm (Roulette Method) in Table 4. The best result was found in the 22nd run and the worst result was found in the 40th run.

TABLE IV
GENETIC ALGORITHM RESULTS BASED ON ROULETTE SELECTION ACCORDING TO QUADRATIC FUNCTION

Class / Results	Quadratic (Rulet)			
		Train	Test	All
Class 1 (No Fraud)	Accuracy	0.9842	0.9846	0.9844
	Recall	0.9987	0.9963	0.9980
	Specificity	0.8517	0.8783	0.8597
	Precision	0.9840	0.9868	0.9848
	F1Score	0.9913	0.9915	0.9914
Class 2 (Fraud)	Accuracy	0.9842	0.9846	0.9844
	Recall	0.8517	0.8783	0.8597
	Specificity	0.9987	0.9963	0.9980
	Precision	0.9865	0.9629	0.9791
	F1Score	0.9141	0.9187	0.9155

From the perspective of the quadratic function, initially, increasing the number of non-forged samples increases the accuracy of the model. However, as the number of forged samples increases, the accuracy of the model begins to decrease. This shows that while the model initially achieves high accuracy rates, it eventually fails due to difficulty in detecting forgery. In other words, the performance of the model exhibits a decreasing trend as the forgery data increases.

B. *Genetic Algorithm Results Based on Tournament Selection*

The parameter values of the genetic algorithm based on the proposed tournament selection are given in Table 5.

TABLE V
GENETIC ALGORITHM PARAMETER VALUES ON TOURNAMENT SELECTON

Genetic Algorithm Based on Tournament Selection	
Number of Runes	50
Number of Iterations	100
Population Number	40
Number of Variables	29
Number of Inputs	10
Number of Outputs	0,9
Number of Bits	8
Crossover Rate	0,15
Number of Mutations	0 (Random)
Mutation Rate	3
Number of Crossovers	5000
Number of Data	4508
Number of Class 1 Samples	492
Number of Class 2 Samples	%70 – 3500
Number of Training Samples	%30 - 1500
Number of Test Samples	50

It is aimed to obtain the results by training and testing the data on the credit card data set with a linear function using the genetic algorithm (Tournament Method). A credit card fraud data set with a total of 284,807 transactions was used, and 492 of these transactions were fake (fraudulent) transactions. Tournament method, which is a type of Genetic Algorithm, was used and 50 runs were studied, each with 100 iterations and a population of 40 individuals. At the end of each study, performance measures such as Mean Squared Error and classification accuracy were reported. The best result was found in the 20th run and the worst result was found in the 6th run. In Table 6, the performance of the classification model in terms of training, testing and all was evaluated in terms of Accuracy, Recall, Specificity, Precision and F1 Score, according to the fraud and non-fraud cases. In the linear function framework, the accuracy of the model increases as the number of non-forged samples increases. However, when the number of forged samples increases, the accuracy starts to decrease. While high accuracy rates are obtained in the beginning, the model starts to fail after a while due to the difficulty in detecting forgery. This shows that the accuracy of the model decreases with the increase in forged data.

TABLE VI
GENETIC ALGORITHM RESULTS BASED ON TOURNAMENT SELECTION ACCORDING TO LINEAR FUNCTION

Class / Results	**Tournament (Lineer)**			
		Train	**Test**	**All**
Class 1 (No Fraud)	Accuracy	0.9822	0.9853	0.9832
	Recall	0.9984	0.9940	0.9971
	Specificity	0.8343	0.9054	0.8556
	Precision	0.9822	0.9896	0.9844
	F1Score	0.9902	0.9918	0.9907
Class 2 (Fraud)	Accuracy	0.9822	0.9853	0.9832
	Recall	0.8343	0.9054	0.8556
	Specificity	0.9984	0.9940	0.9971
	Precision	0.9828	0.9436	0.9700
	F1Score	0.9025	0.9241	0.9092

The results were obtained by training and testing the data on the credit card data set with the Exponential Function using the genetic algorithm (Tournament Method) in Table 7. The best result was found in the 16th run and the worst result was found in the 1st run. In the exponential function framework, the model's accuracy increases rapidly as the number of non-forged samples increases, but the accuracy decreases rapidly as the number of forged samples increases. This suggests that although the model is initially successful, it fails because of its difficulty in detecting forgery, and its accuracy decreases proportionally to the number of forged data.

TABLE VII
GENETIC ALGORITHM RESULTS BASED ON TOURNAMENT SELECTION ACCORDING TO EXPONENTIAL FUNCTION

Class / Results	**Tournament (Exponential)**			
		Train	**Test**	**All**
Class 1 (No Fraud)	Accuracy	0.9825	0.9866	0.9838
	Recall	0.9974	0.9955	0.9968
	Specificity	0.8459	0.9054	0.8638
	Precision	0.9834	0.9897	0.9853
	F1Score	0.9904	0.9926	0.9910
Class 2 (Fraud)	Accuracy	0.9825	0.9866	0.9838
	Recall	0.8459	0.9054	0.8638
	Specificity	0.9974	0.9955	0.9968
	Precision	0.9732	0.9571	0.9681
	F1Score	0.9051	0.9305	0.9129

The results were obtained by training and testing the data on the credit card data set with the quadratic function using the genetic algorithm (Tournament Method) in Table 8. The best result was found in the 11th run and the worst result was found in the 20th run. In the quadratic function framework, as the number of non-forged examples increases, the model's accuracy increases rapidly, because the model learns better on these examples. However, as the number of forged examples increases, the model's accuracy decreases rapidly, because forged examples are more complex and difficult to learn data points.

TABLE VIII
GENETIC ALGORITHM RESULTS BASED ON TOURNAMENT SELECTION ACCORDING TO QUADRATIC FUNCTION

Class / Results	Tournament (Quadratic)			
		Train	**Test**	**All**
Class 1 (No Fraud)	Accuracy	0.9837	0.9833	0.9836
	Recall	0.9984	0.9940	0.9971
	Specificity	0.8488	0.8851	0.8597
	Precision	0.9837	0.9875	0.9848
	F1Score	0.9910	0.9907	0.9909
Class 2 (Fraud)	Accuracy	0.9837	0.9833	0.9836
	Recall	0.8488	0.8851	0.8597
	Specificity	0.9984	0.9940	0.9971
	Precision	0.9831	0.9424	0.9701
	F1Score	0.9110	0.9128	0.9116

C. *Artificial Bee Colony*

Parameter values according to the proposed Artificial Bee Colony are given in Table 9.

TABLE IX
PARAMETER VALUES ACCORDING TO ABC

Artificial Bee Colony	
Number of Runes	50
Number of Iterations	100
Population Number	40
Number of Variables	29 (Number of columns in the data set)
Number of Entries	28
Number of Outputs	2
Number of Data	5000
Number of Class 1 Samples	4508
Number of Class 2 Samples	492
Number of Training Samples	%70 – 3500
Number of Test Samples	%30 - 1500

The results were obtained by training and testing the data on the credit card data set using the artificial bee colony algorithm (ABC). A credit card fraud data set with a total of 284,807 transactions was used, and 492 of these transactions were fake (fraudulent) transactions. ABC algorithm was used and 50 runs were studied, each with 100 iterations and a population of 40 individuals. At the end of each study, performance measures such as Mean Squared Error and classification accuracy were reported. The best result was found in the 44th run and the worst result was found in the 21st run. In Table 10, the performance of the classification model in terms of training, testing and all was evaluated in terms of Accuracy, Recall, Specificity, Precision and F1 Score, according to the fraud and non-fraud cases. In the ABC algorithm framework, the accuracy of the model increases rapidly as the number of non-forged examples increases, because the model learns more information and patterns in these examples. However, when the number of forged examples increases, the accuracy decreases rapidly, because these examples, with their more complex structures and less obvious features, make the learning process of the model difficult. The exploration and exploitation mechanisms of the ABC algorithm are difficult in forged data, so the accuracy rate decreases.

TABLE X
RESULTS ACCORDING TO ABC ALGORITHM ACCORDING TO LINEAR FUNCTION

Class / Results	ABC (Lineer)			
		Train	**Test**	**All**
Class 1 (No Fraud)	Accuracy	0.9782	0.9806	0.9790
	Recall	0.9974	0.9963	0.9971
	Specificity	0.8023	0.8378	0.8130
	Precision	0.9788	0.9824	0.9799
	F1Score	0.9880	0.9893	0.9884
Class 2 (Fraud)	Accuracy	0.9782	0.9806	0.9790
	Recall	0.8023	0.8378	0.8130
	Specificity	0.9974	0.9963	0.9971
	Precision	0.9718	0.9612	0.9685
	F1Score	0.8789	0.8953	0.8839

The results were obtained by training and testing the data on the credit card data set with the Exponential Function using the artificial bee colony algorithm (ABC) in Table 11. The best result was found in the 30th run and the worst result was found in the 9th run.

TABLE XI
RESULTS ACCORDING TO ABC ALGORITHM ACCORDING TO EXPONENTIAL FUNCTION

Class / Results	ABC (Exponential)			
		Train	**Test**	**All**
Class 1 (No Fraud)	Accuracy	0.9780	0.9793	0.9784
	Recall	0.9990	0.9977	0.9986
	Specificity	0.7848	0.8108	0.7926
	Precision	0.9770	0.9796	0.9778
	F1Score	0.9879	0.9886	0.9881
Class 2 (Fraud)	Accuracy	0.9780	0.9793	0.9784
	Recall	0.7848	0.8108	0.7926
	Specificity	0.9990	0.9977	0.9986
	Precision	0.9890	0.9756	0.9848
	F1Score	0.8752	0.8856	0.8783

In terms of the exponential function, the accuracy of the model increases rapidly as the number of non-forged examples increases, because the model learns more effectively on such examples. However, when the number of forged examples increases, the accuracy drops rapidly because forged data are more complex and harder for the model to learn accurately (Table 11).

The results were obtained by training and testing the data on the credit card data set with the quadratic function using the artificial bee colony algorithm (ABC) in Table 12. The best result was found in the 7th run and the worst result was found in the 33rd run. In terms of quadratic function, the accuracy of the model increases rapidly as the number of non-forged examples increases, because the model performs more effective learning on these examples. However, the accuracy decreases rapidly when the number of forged examples increases, because forged examples have more complex structures and require more effort for the model to classify correctly.

TABLE XII
RESULTS ACCORDING TO ABC ALGORITHM ACCORDING TO QUADRATIC FUNCTION

Class / Results	ABC (Quadratic)			
		Train	Test	All
Class 1 (No Fraud)	Accuracy	0.9780	0.9786	0.9782
	Recall	1.0000	0.9977	0.9993
	Specificity	0.7761	0.8040	0.7845
	Precision	0.9761	0.9789	0.9770
	F1Score	0.9879	0.9882	0.9880
Class 2 (Fraud)	Accuracy	0.9780	0.9786	0.9782
	Recall	0.7761	0.8040	0.7845
	Specificity	1.0000	0.9977	0.9993
	Precision	1.0000	0.9754	0.9922
	F1Score	0.8739	0.8814	0.8762

D. *Performance Evaluation*

Performance evaluation and box plot analyzes according to the mean square error of the training data in 100 iterations are given for the linear, exponential and quadratic functions of the optimization algorithms developed in this study. Looking at the performance of the optimization algorithms according to the linear function in Figure 1, the genetic algorithm based on roulette and tournament selection is more successful than the ABC algorithm and has a lower error rate. In addition, the performance in roulette and tournament selection is close to each other. This may indicate that GA provides more efficient optimization on linear functions and explores the solution space better. It is also observed that the performance difference between roulette and tournament selection is very small, and both selection methods are similarly effective. This may suggest that the selection mechanism does not have a great effect on linear problems, and that both approaches are powerful. In general, it can be concluded that GA is more successful than ABC on linear functions, but the selection mechanisms exhibit similar performance.

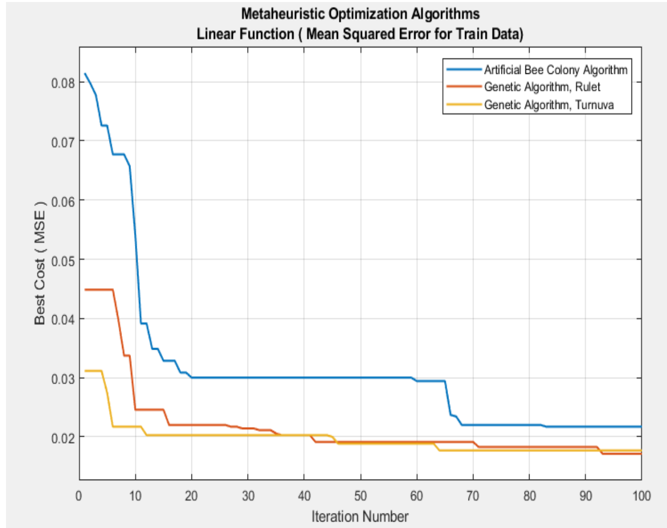


Fig.1. Performances of the proposed optimization algorithms according to the linear function

Figure 2 shows the box plot analysis according to the linear function of the proposed optimization algorithms. Looking at the graph, ABC is spread over a wider area than other methods and shows a more heterogeneous distribution than other methods. This means that the ABC algorithm has more variation on linear functions and its performance may be more inconsistent. The other methods are more concentrated, i.e., show a more homogeneous distribution of performance. This may indicate that the ABC algorithm tends to explore larger solution spaces on linear functions but may give higher error rates in some cases.

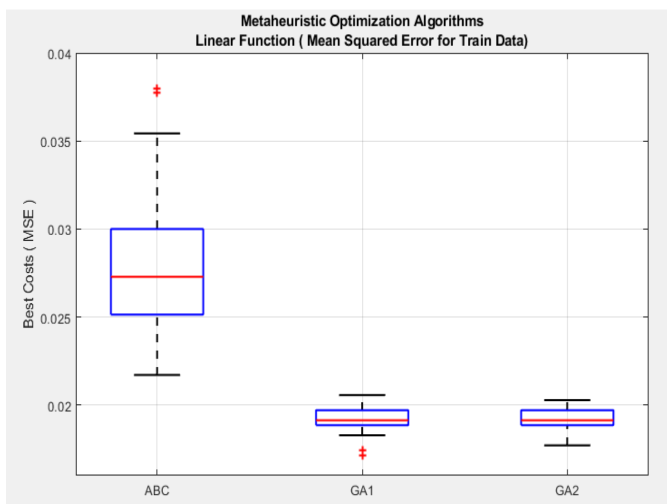


Fig.2. Box plot analyzes according to the linear function of the proposed optimization algorithms

Looking at the performance of the optimization algorithms according to the exponential function in Figure 3, the genetic algorithm based on roulette and tournament selection is more successful than the ABC algorithm and has a lower error rate. In addition, the performance in roulette and tournament selection is close to each other. This shows that GA can optimize more effectively on exponential functions and obtain more accurate results in the search for solutions. Furthermore, the minimal difference in performance between roulette and tournament selection indicates that both selection methods are similarly efficient for exponential functions and do not make a significant difference in terms of optimization success.

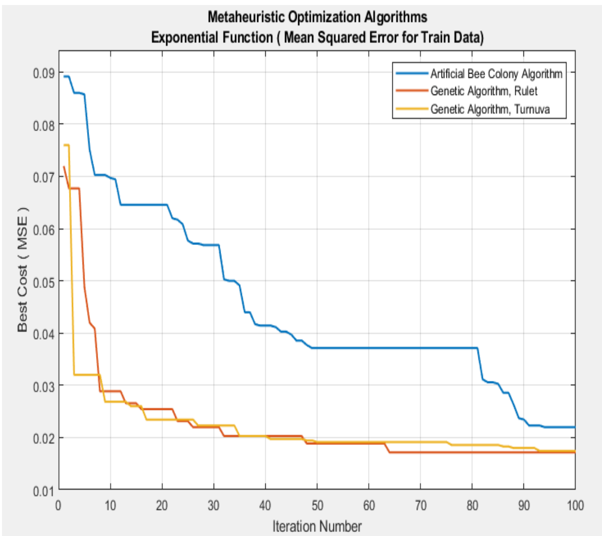


Fig.3. Performances of the proposed optimization algorithms according to the exponential function

Figure 4 shows the box plot analysis of the proposed optimization algorithms according to the exponential function. Looking at the graph, ABC is spread over a wider area than other methods and shows a more heterogeneous distribution than other methods. This means that the ABC algorithm exhibits greater variation in exponential functions and its results can be more inconsistent. The other methods are more concentrated, i.e., have more consistent and uniform performance and generally have lower error rates. This may indicate that the ABC algorithm may give better results in some cases, but it also tends to have higher error values.

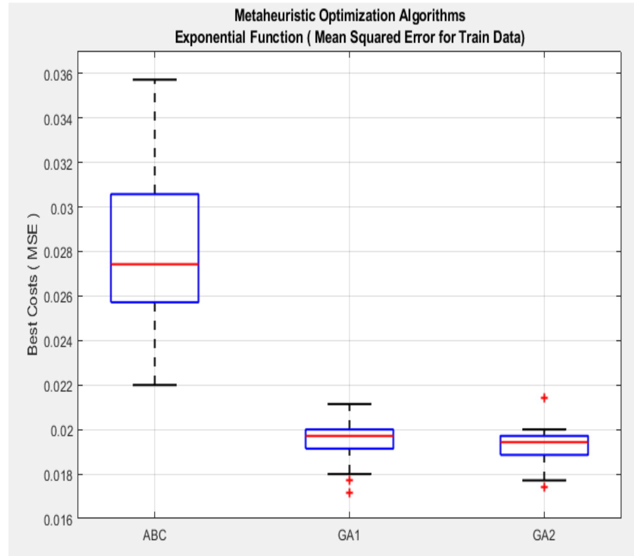


Fig.4. Box plot analyzes according to the exponential function of the proposed optimization algorithms

Looking at the performance of the optimization algorithms according to the quadratic function in Figure 5, the genetic algorithm based on roulette and tournament selection is more successful than the ABC algorithm and has a lower error rate. In addition, the performance in roulette and tournament selection is close to each other. This shows that GA optimizes more effectively on quadratic functions and explores the solution space more accurately. Furthermore, the minimal performance difference between roulette and tournament selection indicates that both selection mechanisms are similarly efficient and GA achieves high accuracy with both selection methods.

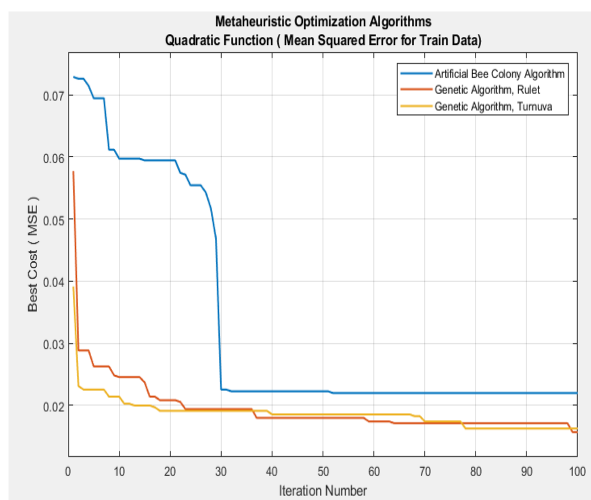


Fig.5. Performances of the proposed optimization algorithms according to the quadratic function

Figure 6 shows the box plot analysis of the proposed optimization algorithms according to the quadratic function. Looking at the graph, ABC is spread over a wider area than other methods and shows a more heterogeneous distribution than other methods. This means that the ABC algorithm has more variation on quadratic functions and its results may be more inconsistent. Other methods concentrate on a narrower error range and obtain more consistent and homogeneous results. This may indicate that the performance of the ABC algorithm is more variable, but may give better results in some cases.

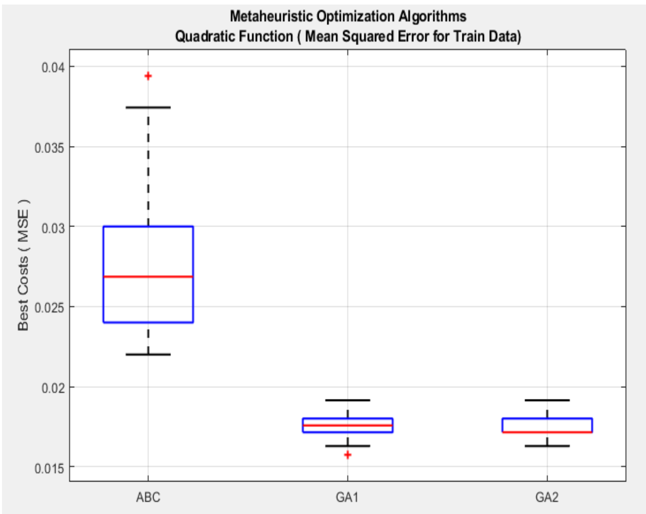


Fig.6. Box plot analyzes of the proposed optimization algorithms according to the quadratic function

IV. Conclusion and Discussion

With technological developments in the world and the widespread increase in Internet use, people have made their lives easier by carrying out their transactions through virtual online means. However, due to all these positive developments brought by technology, fraud transactions, especially credit card fraud, have increased rapidly. Many illegal activities are carried out, such as stealing someone else's credit card information, copying it, accessing password information, performing fraudulent transactions, and these illegal activities lead to large financial losses for producers and consumers every year. For this reason, various necessary precautions and precautions are taken to prevent credit card fraud. However, the measures taken are insufficient as attackers constantly develop different strategies. Due to these problems, preventing financial and financial losses by combating credit card fraud is very important in all sectors. In this study, this problem was discussed and the studies in this field in the literature were examined. In addition, using the proposed genetic algorithm and artificial bee colony developed based on roulette and tournament selection, abnormal behavior on the credit card was detected. In order to increase the success and performance in fraud detection, the hyperparameters used in the genetic algorithm have been improved. The aim of the study is to protect the financial rights of individuals and institutions by minimizing fraud in credit card cases as much as possible. In addition, the success of the optimization algorithms developed with different evaluation criteria was measured and their advantages and disadvantages were presented comparatively.

The results obtained in credit card fraud detection for the test data within the scope of our study are shown in Table XIII. When the functions used in the optimization algorithms are examined, the linear function is more successful than the other functions (quadratic and exponential). In addition, the genetic algorithm method based on roulette and tournament selection gave more successful results than the artificial bee colony algorithm method. The results of all optimization algorithms used in this study were calculated according to different function types (linear, exponential and quadratic) and it was analyzed which method was more successful according to the function type used. The test success of the genetic algorithm and artificial bee colony optimization methods based on roulette and tournament selection using the quadratic function are 0.9846, 0.9833 and 0.9786, respectively.

TABLE XIII
SUCCESS RESULTS OF ALL PROPOSED METHODS ACCORDING TO TEST DATA

	Linear	**Quadratic**	**Exponential**
GA (roulette)	0.9860	0.9846	0.9860
GA (tournament)	0.9853	0.9833	0.9866
ABC	0.9806	0.9786	0.9793

In this study, the comparative performance of optimization algorithms used to classify the credit card dataset was examined in detail. In Table 14, the test accuracy results of optimization methods such as Slime Mold Optimization Algorithm (SMA), Crow Search Algorithm (CSA) and Harmony Search Algorithm (HS) applied to this dataset are presented. These methods have become an important reference when compared to other popular optimization techniques in the literature. The findings of the study reveal that the proposed Genetic Algorithm (GA) and Artificial Bee Colony (ABC) methods exhibit better performance than classical optimization algorithms such as SMA, CSA and HS. This shows that GA and ABC algorithms achieve higher accuracy rates on the credit card dataset and increase the overall success of the model. In particular, it was observed that GA and ABC algorithms are more effective in avoiding local minima and reaching global optimum by showing faster convergence. In addition, although methods such as SMA, CSA and HS are frequently used in the literature, the advantages provided by GA and ABC-based approaches in optimization processes presented in this study show that these algorithms provide more flexible and robust results in the face of evolving data sets and problem types. Therefore, GA and ABC algorithms stand out as potentially more efficient alternatives in classification problems such as credit card fraud detection.

In conclusion, the advantages of GA and ABC algorithms reveal that these optimization methods not only provide increased performance but also have more general applicability in various data sets. These findings reinforce the preference of such optimization techniques in critical application areas such as credit card fraud.

TABLE XIV
SUCCESS RESULTS OF ALL PROPOSED METHODS ACCORDING TO TEST DATA

	Linear	Quadratic	Exponential
Slime Mould Optimization Algorithm (SMA)	0.9228	0.9263	0.9223
Crow Search Algorithm (CSA)	0.9387	0.9391	0.9385
Harmony Search Algorithm (HS)	0.9315	0.9307	0.9367

Table XV shows the results of the study conducted within the scope of the thesis and the studies conducted in the literature in recent years. The accuracy of the genetic algorithm based on roulette selection using a linear function is 0.9860.

TABLE XV
SUCCESS RESULTS OF STUDIES CONDUCTED IN RECENT YEARS IN LITERATURE

Study	Dataset	Method	Accuracy
Huang et al. [24]	Credit card transactions received from Chinese bank	Hybrid Neural Network Based on Clustering Method (HNN-CUHIT)	0.9121
Duan et al. [25]	YelpChi, Amazon,	Causal Temporal Graph Neural Network (CaTGNN)	0.9035 0.9706 0.8281
Raphael et al. [26]	FFSD	Artificial Neural Networks (ANN)	0.9184
Karthika and Senthilselvi [27]	Synthetic data	Improved Convolutional Neural Network	0.9727 0.9739 0.9739
Khidmat et al. [28]	CCF, UCSDFICO, Small Card	XGBoost	0.9600
Prusti et al. [29]	CSMAR	Particle Swarm Optimization (PSO) Based Genetic Algorithm	0.9158
Karthikeyan et al. [30]	The data we used in our study	PSO-Recurrent Neural Network (RNN)	0.9703
Our study	Kaggle data	GA (Linear function) based on Roulette selection	0.9860

It is planned to perform more analyzes with different optimization algorithms in the future. In addition, accuracy will be increased by fine-tuning the parameters and it is aimed to work on larger data sets.

This study makes significant scientific contributions by presenting innovative approaches that increase the effectiveness of optimization algorithms in credit card fraud detection. In particular, a new perspective has been brought to the process of detecting fraudulent transactions by combining genetic algorithms and artificial bee colony algorithms. The use of tournament selection method in genetic algorithms accelerates the solution process while also providing more efficient results. In tournament selection, selecting the best individual among randomly selected individuals stands out as a strategy that increases the accuracy of the algorithm. On the other hand, artificial bee colony algorithm provides better solutions in each iteration by continuously improving individuals according to their fitness values. With the adaptation of this algorithm, the accuracy rate in detecting fraudulent transactions has been significantly increased. The performance of the optimization algorithms used in the study was compared over linear, quadratic and exponential function types, and the most appropriate parameters were determined for each function type. This reveals in detail how the algorithms perform in different scenarios and enables the development of more efficient algorithms. In addition, supporting these optimization methods with advanced visualization techniques (e.g. box plot analysis) makes the success of the algorithms more understandable and provides a more effective evaluation process for users. In conclusion, this study has made significant contributions to the existing literature with the proposed algorithms aimed at solving an important security problem such as credit card fraud, and a useful model has been developed at the theoretical and practical level with the results obtained on real-world data.

References

[1] Aytekin, A., & Yücel, Y. B. (2017). Yeni Ödeme Teknolojilerinin Iş Hayatina Etkileri. Avrasya Sosyal ve Ekonomi Araştırmaları Dergisi, 4(12), 11-33.

[2] Sikkandar, H., Saroja, S., Suseandhiran, N., & Manikandan, B. (2023). An intelligent approach for anomaly detection in credit card data using bat optimization algorithm. Inteligencia Artificial, 26(72), 202-222.

[3] Gadi, M. F. A., Wang, X., & do Lago, A. P. (2008, December). Comparison with parametric optimization in credit card fraud detection. In 2008 Seventh International Conference on Machine Learning and Applications (pp. 279-285). IEEE.

[4] Bakır, Ç., Temurtaş, H., & Yeşilyurt, F. (2023, September). Detection of Credit Card Fraud with Artificial Neural Networks. In Proceedings of the International Conference on Advanced Technologies (Vol. 11, pp. 38-43.

[5] Çılburunoğlu, K. (2023). Kredi Kartı Dolandırıcılık Tespitinde Makine Öğrenme Algoritmalarının Karşılaştırmalı Analizi. İstanbul Gedik Üniversitesi. Yüksek Lisans Tezi.

[6] Hussein A. (2023). Developing an Artificial Inteligence Based Aystem to Detect Fraud in Credit Card Transactions. Altınbaş Üniversitesi. Yüksek Lisans Tezi.

[7] Jovanovic, D., Antonijevic, M., Stankovic, M., Zivkovic, M., Tanaskovic, M., & Bacanin, N. (2022). Tuning machine learning models using a group

search firefly algorithm for credit card fraud detection. *Mathematics*, *10*(13), 2272.
[8] Akyüz, B. (2022). Topluluk Öğrenimi ile Kredi Kartı Dolandırıcılık Tespiti: Karşılaştırmalı Bir Analiz. Yıldız Teknik Üniversitesi. Yüksek Lisans Tezi.
[9] Shah, H. B. (2020). Comparing Machine Learning Algorithms For Credit Card Fraud Detection.
[10] Soylu, K. (2018). Kredi kartı sahte işlem tespiti.
[11] Save, P., Tiwarekar, P., Jain, K. N., & Mahyavanshi, N. (2017). A novel idea for credit card fraud detection using decision tree. International Journal of Computer Applications, 161(13).
[12] Shakya, R. (2018). Application of machine learning techniques in credit card fraud detection (Doctoral dissertation, University of Nevada, Las Vegas).
[13] Taha, A. A., & Malebary, S. J. (2020). An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine. IEEE Access, 8, 25579-25587.
[14] Prabhakaran, N., & Nedunchelian, R. (2023). Oppositional Cat Swarm Optimization-Based Feature Selection Approach for Credit Card Fraud Detection. Computational Intelligence and Neuroscience, 2023.
[15] Prusti, D., Rout, J. K., & Rath, S. K. (2023, January). Detection of credit card fraud by applying genetic algorithm and particle swarm optimization. In Machine Learning, Image Processing, Network Security and Data Sciences: Select Proceedings of 3rd International Conference on MIND 2021 (pp. 357-369). Singapore: Springer Nature Singapore.
[16] Tayebi, M., & El Kafhali, S. (2021, May). Hyperparameter optimization using genetic algorithms to detect frauds transactions. In *The International Conference on Artificial Intelligence and Computer Vision* (pp. 288-297). Cham: Springer International Publishing.
[17] Prusti, D., Rout, J. K., & Rath, S. K. (2023, January). Detection of credit card fraud by applying genetic algorithm and particle swarm optimization. In *Machine Learning, Image Processing, Network Security and Data Sciences: Select Proceedings of 3rd International Conference on MIND 2021* (pp. 357-369). Singapore: Springer Nature Singapore.
[18] Patidar, R., & Sharma, L. (2011). Credit card fraud detection using neural network. International Journal of Soft Computing and Engineering (IJSCE), 1(32-38).
[19] Vats, S. A. T. V. I. K., Dubey, S. K., & Pandey, N. K. (2013, July). Genetic algorithms for credit card fraud detection. In International Conference on Education and Educational Technologies.
[20] Geetha, N., & Dheepa, G. (2022). Transaction fraud detection using Artificial Bee Colony (ABC) based feature selection and Enhanced Neural Network (ENN) classifier. International Journal of Mechanical Engineering, 7(3).
[21] Karthikeyan, T., Govindarajan, M., & Vijayakumar, V. (2023). Intelligent Financial Fraud Detection Using Artificial Bee Colony Optimization Based Recurrent Neural Network. Intelligent Automation & Soft Computing, 37(2).
[22] Furlanetto, G. C., Gomes, V. Z., & Breve, F. A. (2023, June). Artificial Bee Colony Algorithm for Feature Selection in Fraud Detection Process. In International Conference on Computational Science and Its Applications (pp. 535-549). Cham: Springer Nature Switzerland.
[23] Küçüksille, E., & Tokmak, M. (2011). Yapay arı kolonisi algoritması kullanarak otomatik ders çizelgeleme. Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 15(3), 203-210.
[24] Huang, H., vd. (2024). Imbalanced credit card fraud detection data: A solution based on hybrid neural network and clustering-based undersampling technique. Applied Soft Computing, 154 (2024), 1-11.
[25] Duan, Y., vd. (2024). CaT-GNN: Enhancing Credit Card Fraud Detection via Causal Temporal Graph Neural Networks. arXiv preprint arXiv:2402.14708, 1-10.
[26] Raphael, B. A., vd. (2023). Card fraud detection using artificial neural network and multilayer perception algorithm. International Journal of Algorithms Design and Analysis Review, 1(1), 21-30.
[27] Karthika, J., & Senthilselvi, A. (2023). Smart credit card fraud detection system based on dilated convolutional neural network with sampling technique. Multimedia Tools and Applications, 82(20), 31691-31708.
[28] Khidmat, W. B., vd. (2023). Machine learning in the boardroom: gender diversity prediction using boosting and undersampling methods. Research in International Business and Finance, 66(2023), 1-17.
[29] Prusti, D., vd. (2023, January). Detection of credit card fraud by applying genetic algorithm and particle swarm optimization. In Machine Learning, Image Processing, Network Security and Data Sciences: Select Proceedings of 3rd International Conference on MIND 2021, Singapore, 357-369.
[30] Karthikeyan, vd. (2023). Intelligent financial fraud detection using artificial bee colony optimization based recurrent neural network. Intelligent Automation & Soft Computing, 37(2), 1485, 1498.

BIOGRAPHIES

RIDVAN ÇUBUKCUOĞULLARI completed his primary and secondary education at Tokat Artova Regional Boarding Primary School (YİBO), his high school education at Tokat Erbaa Industrial Vocational High School, and his associate degree at Gaziosmanpaşa University Tokat Vocational School Computer Technology and Programming department. He completed his undergraduate education at Kütahya Dumlupınar University, Faculty of Engineering, Department of Computer Engineering. He has been working as a civil servant in a public institution since 2008. His research areas are optimization algorithms, machine learning, data mining.

HASAN TEMURTAŞ is Associate Professor of Computer Engineering at the Engineering Faculty- Kütahya University, Turkey. His research interests include artificial intelligence, deep learning and optimization algorithm.

ÇİĞDEM BAKIR is a Assistant Professor of Software Engineering at the Engineering Faculty – Kutahya University, Turkey. She received the B.S. degree in computer engineering from the University of Sakarya, in 2010, and the M.S. and Ph.D. degrees in computer engineering from Yildiz Technical University, Istanbul. She is currently pursuing the doctorate degree in computer science with the University of Yildiz Technical, Istanbul. She was a Research Assistant at Yildiz Technical University and Igdir University. She was an Instructor at Erzincan Binali Yildirim, from 2020 to 2021. She has been an Assistant Professor with the Software Engineering Department, Dumlupinar University, since 2021. Her research interests include information security, distributed database, big data, blockchain technology, cloud computing, and computer networks.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 1, March 2025
Research Article

A Comparative Analysis of Chest X-ray Examination with AI Enhancement Using XAI Techniques

Cem Ozkurt

Abstract— **Chest X-ray analysis plays a vital role in diagnosing pneumonia, and recent advancements in Deep Learning (DL) methods have significantly improved the accuracy of automated diagnosis. This study explores the intersection of DL and explainable artificial intelligence (XAI) in the context of pneumonia diagnosis through chest X-rays. The dataset used in this study consists of 1,341 training images of healthy individuals and 3,875 images of pneumonia cases, with the test set comprising 234 healthy and 390 pneumonia cases. Additionally, the validation set includes 8 images for both categories. This diversity aims to enhance the model's ability to generalize across different scenarios. The Convolutional Neural Network (CNN) and Transfer Learning (TL) methods utilizing the ResNet50 model achieved accuracies of 95.23 and 96.67, respectively. Subsequently, the models were explained using XAI methods such as SHAP and Grad-CAM. The study concludes by highlighting the potential of DL and XAI to enhance the interpretability and reliability of pneumonia diagnoses through chest X-ray analysis, aiming to contribute to future research in this field.**

Index Terms— **Chest X-ray analysis, Convolutional Neural Networks (CNNs), Deep Learning (DL) methods, Explainable AI (XAI), Grad-CAM (Gradient-weighted Class Activation Mapping).**

I. INTRODUCTION

TODAY, ARTIFICIAL intelligence (AI) technologies play a crucial role in the diagnosis and treatment of critical health issues such as pneumonia [1]. However, concerns regarding the safety, transparency, and comprehensibility of these technologies in clinical applications are significant. The primary objective of this study is to evaluate the role of AI models in pneumonia diagnosis and treatment, with a particular emphasis on the potential utilization of Explainable Artificial Intelligence (XAI) methods [2].

With the widespread use of AI technologies in the healthcare sector, the understandability and traceability of decision-making processes are becoming increasingly important. In this context, understanding the role of explainable AI models in pneumonia diagnosis and treatment is of critical importance. This study aims to highlight the potential of XAI in ensuring the effective use of AI technologies in pneumonia diagnosis.

Cem Ozkurt, is with Department of Computer Engineering at Sakarya University of Applied Sciences is located in Sakarya, Turkey (cemozkurt@subu.edu.tr).
https://orcid.org/0000-0002-1251-7715

Manuscript received Marc 7, 2024; accepted Feb 25, 2025.
DOI: 10.17694/bajece.1448546

Deep learning (DL) methods, especially Convolutional Neural Networks (CNNs), have become the cornerstone of pneumonia diagnosis due to their remarkable ability to process complex visual data such as chest X-rays (CXR). CNNs are inspired by the human brain's visual processing mechanisms and are known for their proficiency in extracting hierarchical features from images. This adaptability allows CNNs to excel in medical image analysis, identifying subtle visual patterns critical to diagnosing pneumonia accurately [3]. Various layers within CNN architectures, including convolutional, pooling, and fully connected layers, contribute to this capability by analyzing features ranging from basic edges to more intricate structures indicative of disease. The use of CNNs in the healthcare field has extended far beyond pneumonia diagnosis, as they have proven their effectiveness in diverse domains, from medical imaging to wireless resource allocation [4,5].

Therefore, this article presents an analysis to evaluate the current status and future potential of AI models in pneumonia diagnosis and treatment, with particular emphasis on the role of CNNs and XAI techniques. A review of significant studies in the literature will be conducted to compile existing knowledge on how explainable AI methods can be utilized in pneumonia diagnosis and treatment to guide future research efforts. The findings of this study could contribute significantly to enhancing the effectiveness of AI technologies in pneumonia diagnosis and ensuring reliability in clinical applications.

Pneumonia is an inflammation of the lung parenchyma caused by infectious microorganisms and non-infective agents. It can affect all age groups but is particularly severe in fragile populations such as children and the elderly. Early and accurate detection of pneumonia is crucial to prevent fatal outcomes.Recent advancements in deep learning (DL) methods have significantly improved the accuracy of automated pneumonia diagnosis through chest X-rays (CXR).

Yang et al. proposed a deep learning approach that considers the background factors of lung X-ray images to improve pneumonia identification accuracy. Using VGG16, they achieved an accuracy of 95.6 and emphasized the importance of considering background factors in the diagnostic process while using Grad-CAM to highlight model explainability [6].

De Moura et al. utilized SHAP and Grad-CAM to differentiate chest X-ray images of COVID-19-based pneumonia from other lung patterns. Their approach achieved an accuracy of 82 with the XGBoost model, underscoring the importance of explainable AI in distinguishing COVID-19 pneumonia from other types [7].

Ren et al. explored an interpretable approach combining deep learning with Bayesian Networks, achieving high performance in pneumonia detection using a dataset of 35,389 cases. This study emphasized the necessity of interpretability in AI models for clinical applications [8].

Zou et al. presented an ensemble AI explainability method combining SHAP and Grad-CAM++ to provide visual explanations for a deep learning prognostic model predicting the mortality risk of pneumonia patients. Their method showed high trust and localization effectiveness among radiologists, demonstrating the value of explainability in clinical decision-making [9].

Stephen et al. developed a convolutional neural network from scratch for pneumonia classification, achieving significant validation accuracy through data augmentation techniques. Their work addressed the challenges of reliability and interpretability in medical imagery by using a large, well-augmented dataset [10].

Alsharif et al. introduced PneumoniaNet, a novel CNN-based framework for automated detection and classification of pediatric pneumonia, achieving an accuracy of 99.7. Their model distinguished between viral, bacterial, and normal cases, demonstrating the potential of deep learning in improving diagnostic accuracy, especially in remote areas lacking expert radiologists [11].

Varshni et al. evaluated the functionality of pre-trained CNN models for pneumonia detection, highlighting the effectiveness of using these models as feature extractors in conjunction with supervised classifiers. Their results indicated that pre-trained CNN models are highly beneficial for analyzing CXR images [12].

These studies collectively demonstrate the advancements and applications of deep learning and explainable AI in pneumonia diagnosis. The integration of SHAP and Grad-CAM provides comprehensive insights into model decision-making, which is critical for clinical acceptance and reliability. Our study builds upon these findings by employing a combination of ResNet50, SHAP, and Grad-CAM to enhance the interpretability and accuracy of pneumonia diagnosis models.

Table 1 includes a literature review table that outlines the dataset, methods (architectures and XAI techniques), and results of various related studies, providing a comprehensive overview of the field:

TABLE I
LITERATURE REVİEW TABLE

	SHAP	Grad-Cam	LIME	DL	TL	Chest Xray
[3]	-	✓	-	-	✓	-
[4]	✓	-	-	✓	-	✓
[5]	✓	-	-	✓	-	✓
[6]	✓	✓	✓	-	✓	-
[7]	-	-	-	✓	-	✓
[8]	-	-	-	✓	-	✓
[9]	✓	-	-	✓	✓	✓
Ours	✓	✓	-	✓	✓	✓

II. Materials And Methodology

A. DataSet

The dataset used in this study consists of chest X-ray images, focusing on the diagnosis of pneumonia. The test set includes 234 images of healthy individuals and 390 images of pneumonia cases. The training set comprises 1,341 images of healthy subjects and 3,875 images of pneumonia cases. Additionally, the validation set includes 8 images for both healthy and pneumonia cases. All images are in X-ray format, capturing various aspects of chest conditions. The diversity in the dataset aims to enhance the model's ability to generalize across different scenarios. Figure 1 provides a visual representation of selected images from the dataset. The composition of the dataset serves as a critical foundation for training, validating, and testing the models in subsequent phases of our methodology. Figure 1 illustrates selected examples from the dataset, providing a glimpse into the diversity of chest X-ray images used in this study. The dataset was obtained from https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia.

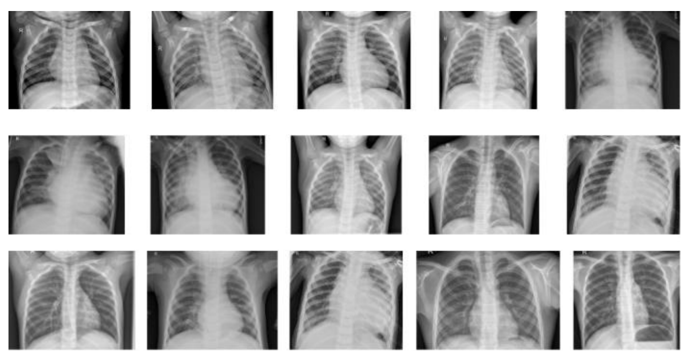


Fig.1. Pneumonia Chest X-ray Database dataset examples

B. DL Methods

Deep Learning (DL) is a subfield of machine learning known for its automatic learning capability, particularly in complex datasets. Emphasizing the inclusion of multi-layered neural networks and the capacity to learn from extensive datasets, deep learning highlights its ability to successfully accomplish complex tasks. Among these methods,CNNs stand out as a significant component extensively used in areas involving visual data, such as image recognition and classification. In Transfer Learning, the ResNet50 model has been used. Transfer Learning is another method within this framework, allowing the adaptation of learned general features for another task.

1) Convolutional Neural Networks (CNN)

CNNs form the core of our pneumonia diagnosis methodology, leveraging their ability to comprehend complex visual information [13,14,15]. Inspired by the visual processing mechanisms of the human brain, CNNs demonstrate remarkable proficiency in image analysis tasks [16,17]. In the context of chest X-ray analysis, CNNs excel at capturing intricate patterns and nuanced features crucial for accurate diagnosis [3]. The architecture of our CNN model, illustrated in Figure 2, consists of multiple layers, each contributing to the network's capacity to understand and interpret hierarchical visual information [16]. The convolutional layers, essential for feature extraction, utilize filters to detect hierarchical features, progressing from basic edges and textures to more intricate structures indicative of pneumonia [15].

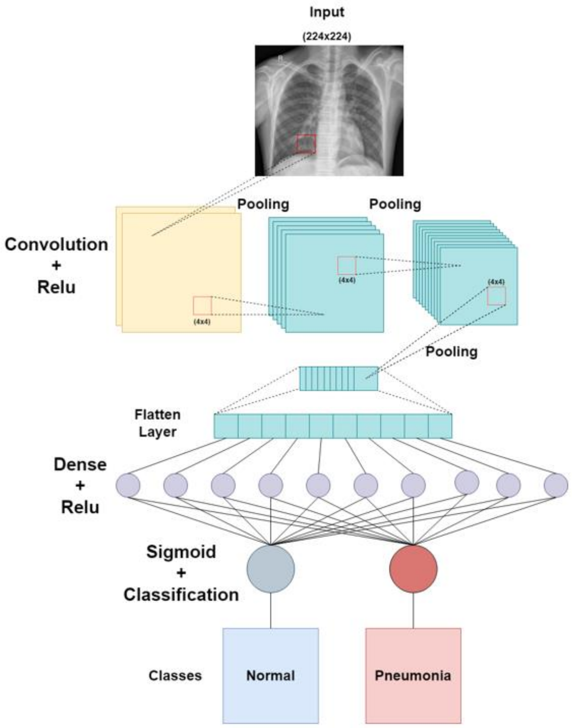


Fig.2. Used Convolutional Neural Network (CNN) Model

In this study, we adopt a CNN architecture with carefully designed layers, including convolutional layers for feature extraction, pooling layers for spatial down-sampling, and fully connected layers for decision-making [16,14]. These layers collectively contribute to the network’s capability to comprehend chest X-ray images and make accurate diagnostic predictions. Our choice of leveraging CNNs is rooted in their proven efficacy in various domains, from medical image analysis [4,5] to wireless resource allocation [3] and beyond. The adaptability and versatility of CNNs make them a valuable tool in addressing the complexities associated with pneumonia diagnosis, where identifying subtle visual cues is paramount [35]. The inherent hierarchical feature learning capabilities of CNNs allow them to discern patterns in medical images, making them particularly well-suited for tasks requiring nuanced understanding, such as the diagnosis of pneumonia from chest X-ray images [4]. This adaptability is further enhanced by fine-tuning the pre-trained models on specific medical datasets, optimizing the network for the intricacies of pneumonia diagnosis [16,4].

This study utilized a Convolutional Neural Network (CNN) model that was carefully structured by adjusting several hyperparameters. In the first two Conv2D layers, 32 and 64 filters were used, respectively, with a filter size of 3×3. Each Conv2D layer was employed to extract specialized features, while MaxPooling2D layers were applied to reduce spatial dimensions. For instance, the first MaxPooling2D layer reduced the output dimensions to 111×111×32. All these layers utilized the ReLU activation function, allowing the model to learn non-linear relationships. Additionally, the final stages of the model included fully connected (Dense) layers, which were adjusted to enable complex decision-making within the CNN. The Flatten layer transformed the output of the convolutional layers into a single vector, providing the necessary structure for classification. This model, which includes dense layers with a large number of parameters (e.g., 22,151,424 parameters in the dense1 layer), exhibits a strong learning capacity and delivers effective results in high-dimensional data processing.

A Convolutional Neural Network (CNN) model was trained on the dataset, and the layers of the CNN model are provided in Table 2.

TABLE II
CONVOLUTIONAL NEURAL NETWORK (CNN) LAYERS.

Layer(type)	Output Shape	Param
conv3d (Conv2D)	None, 222, 222, 32)	896
max_pooling2d (MaxPooling2D	(None, 111, 111, 32)	0
conv2d_1 (Conv2D)	(None, 109, 109, 64)	18496
max_pooling2d_1	(None, 54, 54, 64)	0
conv2d_2 (Conv2D)	(None, 52, 52, 128)	73856
max_pooling2d_2	(None, 26, 26, 128)	0
flatten (Flatten)	(None, 86528)	0
dense (Dense)	(None, 256)	22151 424
dense_1 (Dense)	(None, 1)	257

2) *Transfer Learning in Pneumonia Diagnosis*

Transfer Learning (TL) stands as a pivotal component in our methodology, facilitating the seamless adaptation of knowledge acquired from pre-training on a broader dataset to the specific task of pneumonia diagnosis. In our study, we adopted the ResNet50 model, pre-trained on a vast array of images encompassing diverse categories [18,19,20,21,22]. The weights obtained during the training of ResNet50 were then transferred to our pneumonia diagnosis model, serving as a foundational starting point. Transfer Learning is instrumental in addressing challenges associated with limited datasets specific to a particular medical domain [23,24,25,26,27].

By leveraging the learned features from ResNet50, our model gains a robust understanding of general image patterns and structures, significantly enhancing its ability to recognize relevant features in chest X-ray images. The process involves fine-tuning the pre-trained model on our pneumonia dataset, allowing the model to adapt its learned features to the intricacies of pneumonia diagnosis. This strategic integration not only accelerates the training process but also promotes better convergence and performance on our specific task. Transfer Learning, with its ability to transfer knowledge across domains, proves particularly beneficial in medical image analysis, where labeled datasets are often limited [18,19,20,21,22]. Our approach showcases the effectiveness of transferring pre-learned features, emphasizing the adaptability and enhanced performance that Transfer Learning brings to the realm of pneumonia diagnosis.

C. DataSet

In recent years, Explainable AI (XAI) has emerged as a crucial tool to address the interpretability challenges posed by complex machine learning models [27]. Particularly in critical domains like medical image analysis, where transparency is paramount, XAI aims to demystify the decision-making processes of these models. Our methodology places a strong emphasis on XAI principles, leveraging SHAP (SHapley Additive exPlanations) for enhanced interpretability [28]. Figure 3 provides a visual representation illustrating the working mechanism of explainable Artificial Intelligence (XAI) methods, showcasing the transparency and interpretability aspects integrated into our approach.

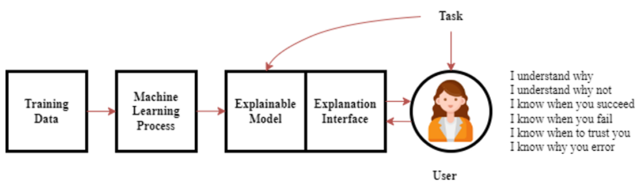


Fig.3. Provides a visual overview of explainable Artificial Intelligence (XAI) methods, illustrating their working mechanisms.

1) SHAP (SHapley Additive exPlanations)

SHAP, rooted in cooperative game theory, offers a principled way to allocate contributions among features for each prediction made by a model. In the context of pneumonia diagnosis, SHAP facilitates the identification and visualization of critical regions within chest X-ray images that heavily influence the model's classification decision. The SHAP value ϕ for feature prediction is mathematically defined as seen in Equation 1:

$$\phi_g^j(f) = \sum s \subseteq \{x^1, \ldots, x^p\}\{x^j\} \frac{|S|!(p-|S|-1)!}{P!}\left(g(f, S \cup \{x^i\}, \Omega) - g(f, S, \Omega)\right) \quad (1)$$

This equation serves as a foundational tool in our methodology, enabling the systematic evaluation of each feature's impact on the model's predictions. Here, $\phi_g^j(f)$ represents the SHAP value for feature j in a given prediction. The summation across subsets S involves the consideration of different combinations of features, and the coefficients $\frac{|S|!(p-|S|-1)!}{P!}$ balance the combinatorial interactions. The terms $g(f, S \cup \{x^i\}, \Omega)$ and $g(f, S, \Omega)$ signify the model's predictions when including and excluding feature i, respectively.

This integration provides valuable insights into the regions of chest X-ray images that contribute most to the final classification decision, ranging from basic edges and textures in the early layers to more complex structures indicative of pneumonia in the deeper layers.

By visualizing SHAP values, we gain insights into regions of X-ray images that are crucial for explaining the model's decision-making process. Our approach utilizes masking techniques, such as the "inpaint telea" method, to identify specific areas of interest, facilitating a comprehensive understanding of the model's interpretability [28,29].

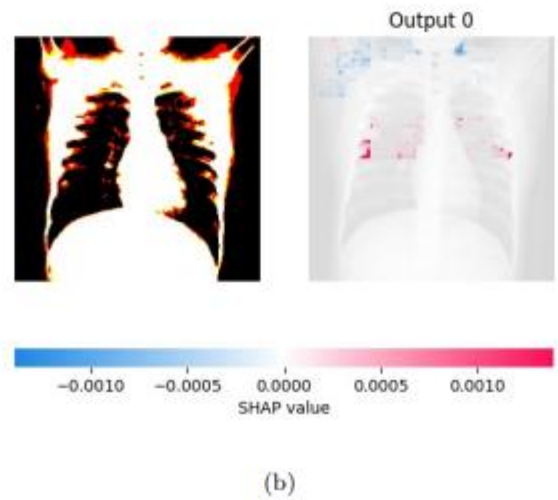


Fig.4. SHAP interpretation of randomly selected image from the dataset.

2) Grad - CAM (Gradient-weighted Class Activation Mapping)

In our exploration of Explainable AI (XAI) methodologies, we incorporated Grad-CAM (Gradient-weighted Class Activation Mapping) [20] alongside SHAP. This technique plays a pivotal role in unveiling the decision-making processes of CNNs, particularly in the realm of medical image analysis.

Grad-CAM provides valuable visual insights into the influential regions of an input image affecting the model's final classification decision. This is achieved through the computation of gradients ∇ of target class scores Y^c with respect to the feature maps A^k of the final convolutional layer.

The resulting gradient-weighted activation maps $L^c_{Grad-CAM}$ are obtained using global average pooling, as expressed in Equation 2:

$$L^c_{Grad-CAM} = \sum_{k=1}^{K} w_k^c \ . \ ReLU\left(\frac{\partial Y^c}{\partial A^k}\right) . \ F^k(x, y) \quad (2)$$

Here, w represents the importance weight associated with the feature maps. This weight is determined by summing the gradients with respect to the corresponding feature maps, normalized by Z as seen in Equation 3:

$$w_k^c = \frac{1}{Z}\sum_i \sum_j \frac{\partial Y^c}{\partial A_{ij}^k} . ReLU\left(\frac{\partial Y^c}{\partial A_{ij}^k}\right) \quad (3)$$

Furthermore, the calculation of the gradient $\frac{\partial Y^c}{\partial A_{ij}^k}$ involves the weighted summation of gradient values across spatial dimensions, denoted by α_{kcij} as seen in Equation 4:

$$\frac{\partial Y^c}{\partial A_{ij}^k} = \frac{1}{Z}\sum_{i'}\sum_{j'} \alpha_{kcij}.ReLU\left(\frac{\partial Y^c}{\partial A_{i'j'k}}\right) \quad (4)$$

These mathematical formulations contribute to a comprehensive understanding of Grad-CAM's mechanism, elucidating the significance of each feature map in influencing the final decision. In our pneumonia diagnosis framework, Grad-CAM proved instrumental in uncovering the specific regions within chest X-ray images that played a critical role in the CNN's classification decisions.

By applying Grad-CAM to various layers of the network, ranging from top to mid layers, we gained insights into the hierarchical features learned by the model. The transparency introduced by Grad-CAM enhances the interpretability of our pneumonia diagnosis model, which is crucial in medical applications for fostering trust among healthcare practitioners.

This integration aligns with the current trend of leveraging XAI techniques to bridge the gap between complex model architectures and interpretability, promoting the responsible and ethical deployment of AI, especially in critical domains such as healthcare. Our work is inspired by prior research successfully applying XAI techniques in medical image analysis [30,31,32,33,34], emphasizing the significance of transparent and interpretable models in AI-based medical diagnosis systems.

III. Materials And Methodology

In the context of this study, the dataset contains chest X-ray images focusing on pneumonia diagnosis. This versatile research employs an advanced approach that combines the Transfer Learning model ResNet50 with SHAP (SHapley Additive exPlanations) and the CNN model with Grad-CAM (Gradient-weighted Class Activation Mapping).

The confusion matrix is a matrix used, particularly, to evaluate the performance of a classification model, focusing on comparing the model's predictions with the actual classes. This matrix includes True Positive (TP) and True Negative (TN) values, representing cases where the model accurately predicts Pneumonia classes, along with False Positive (FP) and False Negative (FN) predictions. Each cell represents a combination of the true class and the predicted class. This visual is used to understand which classes the model predicted correctly and in which cases it made errors. The resulting confusion matrix is presented in Figure 5.

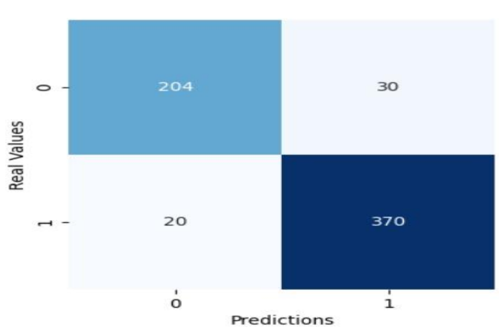


Fig.5. Confusion matrix results

• True Positive (TP): Represents the case where the model accurately predicts the positive class.

TP = Numerical Value

• True Negative (TN): Represents the case where the model accurately predicts the negative class.

TN = Numerical Value

• False Positive (FP): Represents the case where the model incorrectly predicts the positive class.

FP = Numerical Value

• False Negative (FN): Represents the case where the model incorrectly predicts the negative class.

FN = Numerical Value

Performance metrics such as precision, recall, and accuracy are obtained from the confusion matrix values.

Precision measures how many of the samples predicted as positive are actually positive. It expresses the ratio of true positives to the total positive predictions, as shown in Formula 5.

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

Recall measures how many of the true positives are detected. It expresses the ratio of true positives to the total number of positive examples, as shown in Formula 6.

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

Accuracy expresses the ratio of correctly predicted examples to the total number of examples. It is a metric that evaluates the overall model performance, as shown in Formula 7.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

F1-score balances precision and recall. This metric tends to minimize both false positives and false negatives, especially in balanced classification problems. Formula 8 illustrates the F1-score.

$$F! = 2\ x\ \frac{Precision\ x\ Recall}{Precision + \ Recall} \quad (8)$$

When the ResNet50 model, used as a Transfer Learning (TL) method, is applied to the same dataset, the F1 score, precision, recall, and support metrics are provided in Table 3.

TABLE III
CONFUSİON MATRIX RESULTS

	Accuracy	precision	Recall	F1-score	Support
NORMAL	0.92	0.91	0.87	0.89	234
PNEUMANIA	0.92	0.93	0.95	0.94	234

Subsequently, Grad-CAM was applied to visualize activation patterns within the convolutional layers of the CNN. This study clarified each step by emphasizing specific layers such as "conv2d2" for Grad-CAM Top and "conv2d1" for Grad-CAM Mid.

The integration of SHAP's analysis involved applying an image mask to the model's predictions, enabling a deeper understanding of the decision-making process with 15,000 evaluations. An example containing 15,000 evaluations is presented in Figure 6.

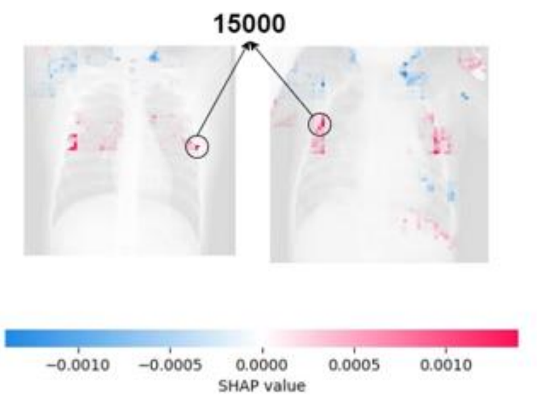


Fig.6. tems contributing to model misinterpretation in Shap analysis with 15,000 evaluations.

A notable issue encountered in this study is the model's susceptibility to misinterpretations. SHAP analysis highlights situations where the model may be misled by seemingly insignificant artifacts or misplaced objects in the image mask. This underscores the importance of meticulous preprocessing and awareness of potential errors in medical image analysis.

To illustrate this situation, we present a visual representation of misclassifications, showing instances where the model made correct predictions and errors in Figure 7.

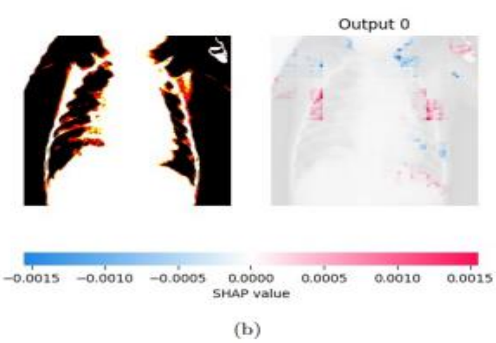


Fig.7. Misleading items causing model misinterpretation in Shap analysis.

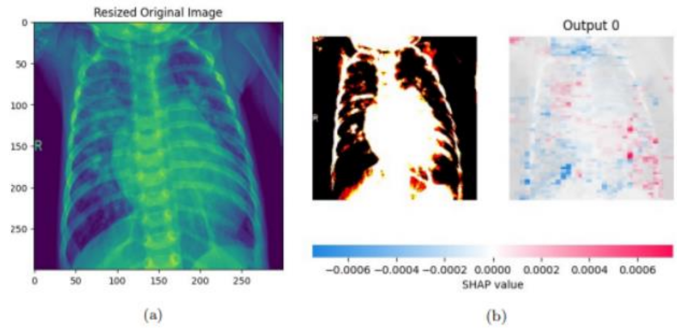


Fig.8. Comparison of (a) Original Image and (b) SHAP Output

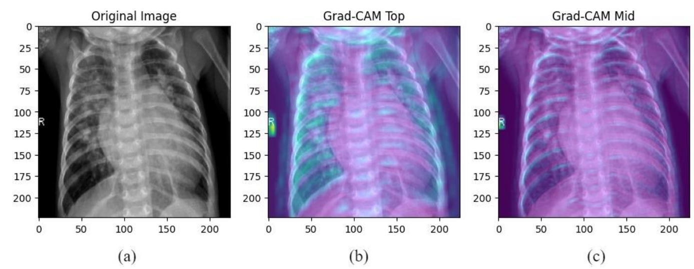


Fig.9. Grad-CAM Output

True Positive (TP): In the case of True Positive, where the model correctly identified instances of the disease, we initiated the analysis with the original image and progressed through subsequent stages. Figure 8 (a) shows our starting point with the original image. SHAP elucidated why the ResNet50 model classified this example as positive, highlighting the features supporting the positive decision. Figure 8 (b) presents the SHAP output, offering a detailed glimpse into the model's decision rationale. Grad-CAM, as illustrated in Figure 9, plays a crucial role in unraveling the rationale behind the Convolutional Neural Network (CNN) recognizing the provided example as positive. Figure 9 (a) serves as the anchor, showcasing the original image for our analysis. Figures 9 (b) and 9 (c) spotlight the two phases of Grad-CAM. Figure 9 (b) shows the completed Grad-CAM highlighting the top contributing regions, while Figure 9 (c) captures the ongoing process, emphasizing the mid regions. The integration of SHAP and Grad-CAM not only enriches our understanding of positive classification but also provides a unique perspective on the model's decision-making process.

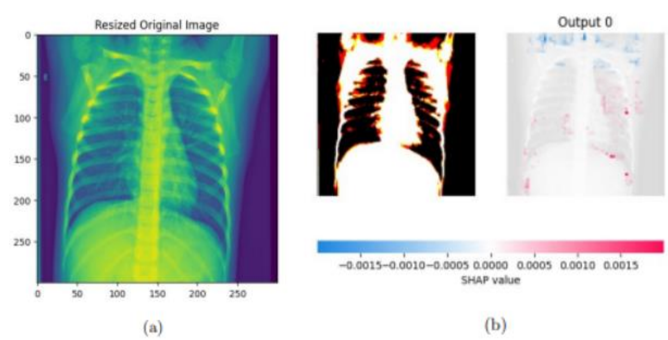


Fig.10. Comparison of (a) Original Image and (b) SHAP Output

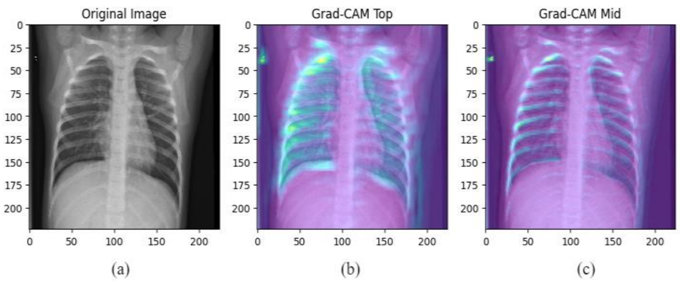


Fig.11. Grad-CAM Output

True Negative (TN): In the case of True Negative, where the model correctly identified a healthy state, our analysis began with the original image, setting the stage for subsequent examinations. Figure 10 (a) depicts the initial step with the original image, followed by Figure 10 (b), presenting the SHAP output that sheds light on the reasons behind the negative classification. Grad-CAM, featured in Figure 11, was employed to unravel the Convolutional Neural Network's (CNN) reasoning behind identifying this instance as negative, visually emphasizing the contributing regions. Figure 11 (a) presents the original image, forming the basis for our analysis. Figures 11 (b) and 11c shed light on the two distinctive phases of Grad-CAM. In Figure 11 (b), the completed Grad-CAM showcases the highlighted top contributing regions, while Figure 11 (c) captures the ongoing process, accentuating the mid regions. The strategic fusion of SHAP and Grad-CAM not only enhances our understanding of negative classifications but also introduces a fresh dimension to the interpretability of the model.

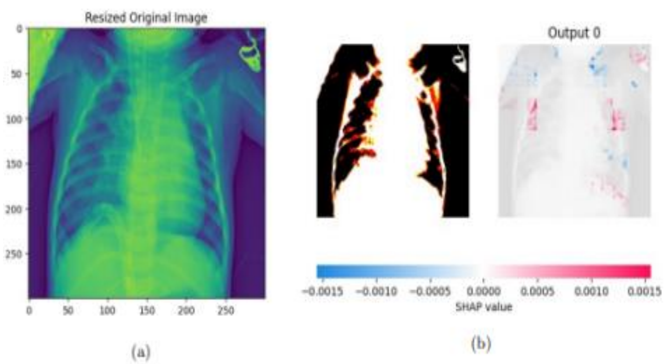


Fig.12. Comparison of (a) Original Image and (b) SHAP Output

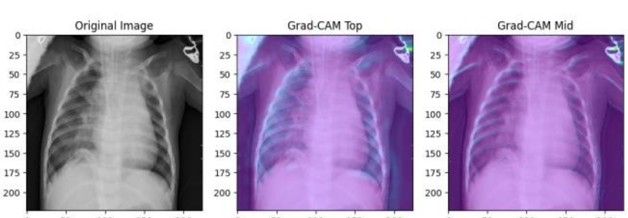


Fig.13. Grad-CAM Output

False Positive (FP): For False Positive instances where the model inaccurately predicted disease presence, our investigation began with the original image, offering insights into the misclassification. Figure 12 (a) captures the starting point with the original image, followed by Figure 12 (b), showcasing the SHAP output that explains the false positive classification by the ResNet50 model. Grad-CAM, showcased in Figure 13, extensively explored the reasons behind the Convolutional Neural Network (CNN) model's erroneous positive identification, visually spotlighting the regions responsible for the misclassification. Figure 13 (a) exhibits the original image, forming the basis for our analysis. Figures 13 (b) and 13 (c) illustrate the two significant phases of Grad-CAM. In Figure 13 (b), the completed Grad-CAM reveals the highlighted regions contributing to the false positive identification, while Figure 13 (c) captures the ongoing process, emphasizing mid regions that played a role in the misclassification. The strategic amalgamation of SHAP and Grad-CAM not only exposes the misclassification patterns of the model but also introduces innovation in comprehending false positives.

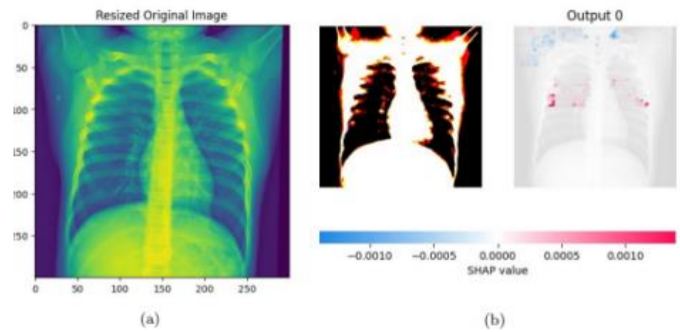


Fig.14. Comparison of (a) Original Image and (b) SHAP Output

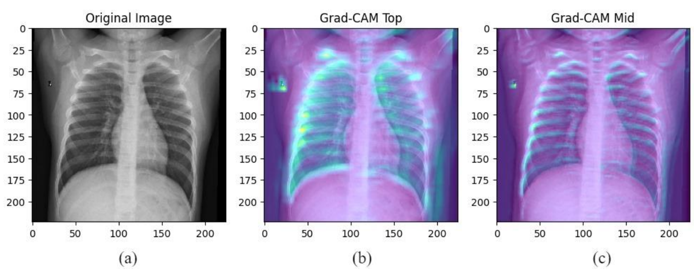


Fig.15. Grad-CAM Output

False Negative (FN): In cases of False Negative, where the model wrongly predicted a healthy state, our scrutiny began with the original image, providing insights into the misclassification. Figure 14 (a) marks our starting point with the original image, succeeded by Figure 14 (b), showcasing the SHAP output that explains the false-negative classification by the ResNet50 model. Grad-CAM, depicted in Figure 15, delved deeper into understanding the reasons behind the Convolutional Neural Network (CNN) model's erroneous negative identification, visually representing the regions accountable for the misjudgment. Figure 15 (a) showcases the original image as the foundation for our analysis. Figures 15 (b) and 15 (c) delineate the two key phases of Grad-CAM. In Figure 15 (b), the completed Grad-CAM unveils the highlighted regions contributing to the false negative identification, while Figure 15 (c) captures the ongoing process, emphasizing mid regions that played a role in the misjudgment. The collaborative synergy of SHAP and Grad-CAM not only brings to light the model's misjudgments but also introduces a fresh perspective on comprehending false negatives.

The results obtained in this study are noteworthy when compared to other studies in the literature, revealing both similarities and differences. For instance, Yang et al. used the VGG16 model, considering background factors in pneumonia diagnosis, and achieved an accuracy of 95.6 [6]. Their results, particularly in terms of model explainability using Grad-CAM, align with our ResNet50-based model. However, our model incorporates an additional layer of explainability through SHAP, allowing for a deeper understanding of the decision-making process. This added explainability supports the model's reliability in clinical applications.

In the study conducted by De Moura et al. on COVID-19 pneumonia, SHAP and Grad-CAM were utilized with the XGBoost model, resulting in an accuracy of 82 [7]. In comparison, the integration of SHAP and Grad-CAM in our model achieved higher accuracy, underscoring the robustness of ResNet50 as a transfer learning model in medical image analysis tasks.

Moreover, Zou et al.'s work, which combined Grad-CAM++ and SHAP to predict the mortality risk in pneumonia patients [9], highlighted the critical role of explainability in clinical decision-making. Similarly, in our study, the integration of SHAP and Grad-CAM significantly enhanced the explainability of the model, thereby strengthening its reliability for clinical use.

Finally, while Alsharif et al. achieved a remarkable accuracy of 99.7 with their CNN-based PneumoniaNet framework for pediatric pneumonia diagnosis [11], our results are comparably strong. Although their model demonstrated high accuracy, the addition of explainability techniques such as SHAP and Grad-CAM in our model offers a clearer understanding of the decision-making process. These explainability methods not only improve the reliability of deep learning models but also enhance their acceptance in clinical settings by providing greater transparency.

IV. Conclusions and the Scope for Future Work

Chest X-ray analysis plays a vital role in pneumonia diagnosis, and recent advancements in Deep Learning (DL) methods have significantly increased the accuracy of automated diagnosis. This article explores the intersection of DL and explainable artificial intelligence (XAI) in the context of pneumonia diagnosis through chest X-rays. Using the ResNet50 model from CNNs and Transfer Learning (TL) methods, an accuracy of 95.23 was achieved.

In this study, we found that the combination of ResNet50, SHAP, and Grad-CAM provides a robust methodology for interpreting and explaining pneumonia diagnosis model decisions. SHAP's ability to individually evaluate the contribution of each input to the model output allows medical professionals to better understand why the model made a particular decision and assess the reliability of that decision. Grad-CAM, on the other hand, visually shows which regions the model considers, but it may sometimes be insufficient for a deep understanding of the decision-making process.

Our results show that when SHAP and Grad-CAM are used together, they can enhance interpretability in medical imaging analyses. Future studies should aim to further explore the synergy between these two methods to develop and improve methodologies for medical imaging analyses. The combination of SHAP's quantitative explanations with Grad-CAM's visual explanations can provide medical professionals with a more holistic and reliable interpretation. Such integration could enhance the reliability of machine learning models in clinical applications and contribute to the development of more effective decision support systems.

Funding Information
No funding

Data Availability

The Kaggle link of dataset is "https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia".

Conflict of Interest

The author declare that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical Considerations

This research adheres to ethical principles and guidelines in conducting a comparative analysis of Explainable Artificial Intelligence (XAI) techniques, specifically SHAP (SHapley Additive exPlanations) and Grad-CAM (Gradient-weighted Class Activation Mapping), on pneumonia X-ray dataset.

Declarations

Ethics Approval Not Applicable

Competing interests The author declare no competing interests.

References

[1] H. Sharma, et al., "Feature Extraction and Classification of Chest X-Ray Images Using CNN to Detect Pneumonia," in *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, 2020, pp. 227–231.

[2] S.-H. Lo and Y. Yin, "A novel interaction-based methodology towards explainable AI with better understanding of Pneumonia Chest X-ray Images," *Discover Artificial Intelligence*, vol. 1, no. 1, p. 16, 2021.

[3] M. Eisen and A. Ribeiro, "Optimal Wireless Resource Allocation With Random Edge Graph Neural Networks," *ARXIV-EESS.SP*, 2019.

[4] M. Rahimzadeh and A. Attar, "A Modified Deep Convolutional Neural Network for Detecting COVID-19 and Pneumonia from Chest X-ray Images Based on The Concatenation of Xception and ResNet50V2," *ARXIV-EESS.IV*, 2020.

[5] G. F. Elsayed, B. Wohlberg, and S. Jastrzębski, "Deep Double Descent: Where Bigger Models and More Data Hurt," *ARXIV-EESS.ST*, 2020.

[6] Y. Yang, G. Mei, and F. Piccialli, "A Deep Learning Approach Considering Image Background for Pneumonia Identification Using Explainable AI (XAI)," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. PP, 2022, doi: 10.1109/TCBB.2022.3190265.

[7] L. V. de Moura, C. Mattjie, C. M. Dartora, R. C. Barros, and A. M. Marques da Silva, "Explainable Machine Learning for COVID-19 Pneumonia Classification With Texture-Based Features Extraction in Chest Radiography," *Frontiers in Digital Health*, vol. 3, 2022, doi: 10.3389/fdgth.2021.662343.

[8] H. Ren, et al., "Interpretable Pneumonia Detection by Combining Deep Learning and Explainable Models With Multisource Data," *IEEE Access*, vol. 9, pp. 95872–95883, 2021, doi: 10.1109/ACCESS.2021.3094025.

[9] L. Zou, et al., "Ensemble image explainable AI (XAI) algorithm for severe community-acquired pneumonia and COVID-19 respiratory

infections," *IEEE Transactions on Artificial Intelligence*, vol. 4, no. 2, pp. 242–254, 2022, doi: 10.1109/TAI.2022.3154871.
[10] O. Stephen, M. Sain, U. J. Maduh, and D. U. Jeong, "An efficient deep learning approach to pneumonia classification in healthcare," *Journal of Healthcare Engineering*, vol. 2019, Article ID 4180949, 2019, doi: 10.1155/2019/4180949.
[11] R. Alsharif, et al., "PneumoniaNet: Automated detection and classification of pediatric pneumonia using chest X-ray images and CNN approach," *Electronics*, vol. 10, no. 23, p. 2949, 2021, doi: 10.3390/electronics10232949.
[12] D. Varshni, K. Thakral, L. Agarwal, R. Nijhawan, and A. Mittal, "Pneumonia detection using CNN based feature extraction," in *2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, 2019, pp. 1–7, doi: 10.1109/ICECCT.2019.8869364.
[13] W. Zhang, et al., "Blind Image Quality Assessment Using A Deep Bilinear Convolutional Neural Network," *ARXIV-EESS.IV*, 2019.
[14] D. Valsesia, et al., "Deep Graph-Convolutional Image Denoising," *ARXIV-EESS.IV*, 2019.
[15] M. Gil-Martín, J. Montero, and R. San-Segundo, "Parkinson's Disease Detection from Drawing Movements Using Convolutional Neural Networks," *ELECTRONICS*, 2019.
[16] H. Gao, et al., "PhyGeoNet: Physics-Informed Geometry-Adaptive Convolutional Neural Networks For Solving Parameterized Steady-State PDEs On Irregular Domain," *ARXIV-EESS.IV*, 2020.
[17] F. Eitel, K. Ritter, and Alzheimer's Disease Neuroimaging Initiative (ADNI), "Testing the robustness of attribution methods for convolutional neural networks in MRI-based Alzheimer's disease classification," in *Interpretability of Machine Intelligence in Medical Image Computing and Multimodal Learning for Clinical Decision Support: Second International Workshop, iMIMIC 2019, and 9th International Workshop, ML-CDS 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 17, 2019, Proceedings*, 2019, pp. 3–11.
[18] I. D. Apostolopoulos and T. Bessiana, "Covid-19: Automatic Detection From X-Ray Images Utilizing Transfer Learning With Convolutional Neural Networks," *ARXIV-EESS.IV*, 2020.
[19] H. S. Maghdid, et al., "Diagnosing COVID-19 Pneumonia From X-Ray And CT Images Using Deep Learning And Transfer Learning Algorithms," *ARXIV-EESS.IV*, 2020.
[20] N. E. M. Khalifa, et al., "Detection of Coronavirus (COVID-19) Associated Pneumonia Based on Generative Adversarial Networks and A Fine-Tuned Deep Transfer Learning Model Using Chest X-ray Dataset," *ARXIV*, 2020.
[21] T. Rahman, et al., "Transfer Learning With Deep Convolutional Neural Network (CNN) For Pneumonia Detection Using Chest X-ray," *ARXIV-EESS.IV*, 2020.
[22] P. R. A. S. Bassi and R. Attux, "A Deep Convolutional Neural Network for COVID-19 Detection Using Chest X-Rays," *ARXIV-EESS.IV*, 2020.
[23] Z. Zhou, et al., "Models Genesis: Generic Autodidactic Models For 3D Medical Image Analysis," *ARXIV-EESS.IV*, 2019.
[24] Y.-A. Chung and J. Glass, "Generative Pre-Training For Speech With Autoregressive Predictive Coding," *ARXIV-EESS.AS*, 2019.
[25] C. L. Srinidhi, O. Ciga, and A. L. Martel, "Deep Neural Network Models For Computational Histopathology: A Survey," *ARXIV-EESS.IV*, 2019.
[26] Z. Zhao, et al., "Applications of Unsupervised Deep Transfer Learning to Intelligent Fault Diagnosis: A Survey," *ARXIV-EESS.SP*, 2019.
[27] M. Goyal, et al., "Artificial Intelligence-Based Image Classification For Diagnosis Of Skin Cancer: Challenges And Opportunities," *ARXIV-EESS.IV*, 2019.
[28] S. M. Lundberg, et al., "A Unified Approach to Interpreting Model Predictions," in *Advances in Neural Information Processing Systems*, 2017.
[29] SHAP Contributors, "SHAP (SHapley Additive exPlanations) Documentation," 2020. [Online]. Available: https://shap.readthedocs.io/en/latest/
[30] S. Basu, S. Mitra, and N. Saha, "Deep Learning For Screening COVID-19 Using Chest X-Ray Images," *ARXIV-EESS.IV*, 2020.
[31] C. Xia, et al., "Vision Based Defects Detection for Keyhole TIG Welding Using Deep Learning with Visual Explanation," *Journal of Manufacturing Processes*, 2020.
[32] M. R. Karim, et al., "DeepCOVIDExplainer: Explainable COVID-19 Diagnosis Based On Chest X-ray Images," *ARXIV-EESS.IV*, 2020.
[33] S. Vijayarangan, et al., "Interpreting Deep Neural Networks For Single-Lead ECG Arrhythmia Classification," *ARXIV-EESS.SP*, 2020.
[34] M. Kim, et al., "Medinoid: Computer-Aided Diagnosis and Localization of Glaucoma Using Deep Learning," *Applied Sciences*, 2019.
[35] I. Elbouknify, et al., "CT-xCOV: A CT-scan Based Explainable Framework for COVID-19 Diagnosis," *ARXIV-EESS.IV*, 2023.

BIOGRAPHIES

Cem Ozkurt, received his B.Sc. in Industrial Engineering from Sakarya University in 2002. He completed two M.Sc. degrees in Industrial Engineering and Curriculum & Instruction in 2016, and earned his Ph.D. in Industrial Engineering in 2020. Since 2018, he has held academic positions at Sakarya University of Applied Sciences, currently serving in the Department of Computer Engineering. His research focuses on digital transformation, artificial intelligence in industry, and Industry 4.0.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 1, March 2025
Research Article

A Novel Fuzzy Logic Based Hand Gesture Recognition System

Harun Sumbul

Abstract—**In this proposed study, a Fuzzy Logic System (FLS) was developed to classify and detect hand movements. The designed FLS system consists of a Fuzzifier, Inference Engine, Knowledge Base, and Defuzzifier. The Mamdani technique was used as the Inference Engine, and the centroid method was used for Defuzzification. Five input variables (Flex1-5) and one output variable (Sign) were used to create a rule base with 94 rules. A sensor array was placed on a glove to generate data, and a data collection circuit was established. Movements were performed through this circuit to create the rule bases. A total of 15.030 data points were analyzed to develop the FLS. According to the results, the movements (97.5%) were detected successfully.**

Index Terms—**Flex sensor, fuzzy logic, hand gesture recognition, sign language**

I. Introduction

RECENTLY, RESEARCHERS have shown increasing interest in gesture recognition due to its significant potential in various fields such as medical, military, and commercial applications. Most studies have adopted traditional learning model approaches. To achieve this, multiple interfaces such as data gloves, motion sensors, and position trackers have been developed to gather hand movement data. In recent years, the development of more robust sign language prediction for effective communication has required detecting joints in the hand, face, head, and entire body [1]. Sign language, used by deaf individuals, has been employed as a means of interaction and exchange of ideas parallel to spoken language since its inception [2]. Over 5% of the global population experiences hearing impairment, and recent projections by the World Health Organization (WHO) suggest that the prevalence of hearing impairment is expected to rise due to various factors [3].

Various communication techniques incorporating human-machine interfaces have been developed, particularly for hearing-impaired individuals. Among these, methods such as image processing [4], computer vision [5], artificial neural networks [6], deep learning [7], and machine learning [8] are frequently used, although fuzzy logic applications are also observed in similar studies. Fuzzy logic is also one of the methods that has recently become popular and is frequently utilized in significant applications [9].

In conclusion, research aimed at developing computational methods for the recognition and analysis of hand movements is a significant area of study today. This study presents a novel FLS that automates the processes of recognizing hand movements with linguistic expressions. This approach will result in time and cost savings.

To substantiate our assertions, we outline the principal characteristics of our fuzzy logic model. Section 2 offers a comprehensive description of the materials and methods utilized in developing the FLS model structure. Section 3 presents the findings and analyses derived from testing the proposed model with real-world data. In the conclusion section, we delineate the unique contributions of our study within the existing literature, provide a summary of the research, and underscore its significant findings.

Harun Sümbül, is with Department of Biomedical Device Technologies, Ondokuz Mayis University, Samsun, Turkey, (e-mail: harun.sumbul@omu.edu.tr).
https://orcid.org/0000-0001-5135-3410

Manuscript received Jun 20, 2024; accepted Oct 02, 2024.
DOI: 10.17694/bajece.1519693

II. MATERIAL AND METHOD

To provide real-time measurement of hand movements, a sensor array was placed on a plastic glove, and data were collected and digitized using a microcontroller (Arduino UNO). The sensor array consists of flexible sensors extending along each finger. This setup allows data collection from each finger during every movement, providing a stable measurement system. An overview of the measurement system architecture is illustrated in Fig. 1.

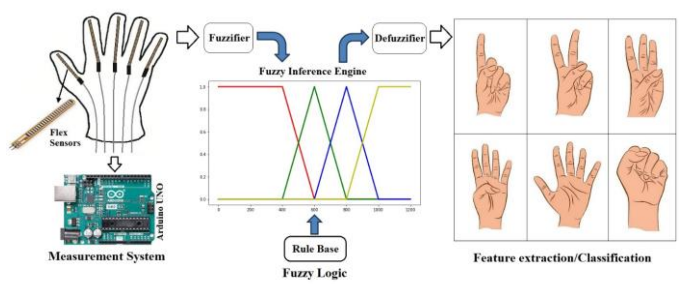


Fig.1. The methodology of the proposed study

A. Generating a hand movements dataset and labeling

Various hand movements were performed using a glove equipped with sensors, and the measured signals were recorded in real-time on a computer. To label the hand movements, six different movements were performed (signs 1-5 and the fist-clenching movement). Each movement was labeled with a numerical digit. Fig. 2 illustrates the hand movements representing the numerical digits used in creating the database.

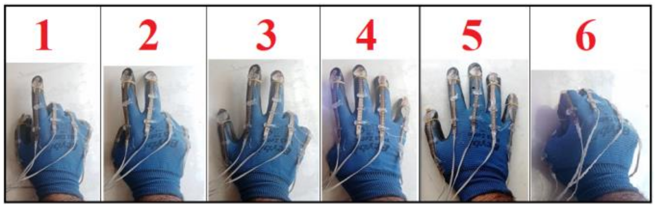


Fig.2. Hand gesture recognition from sensor arrays based on flex sensor

B. *The structure of FLS*

Fuzzy systems leverage fuzzy logic principles to handle imprecise and ambiguous data. These systems find applications across various fields such as healthcare, engineering, and finance. An expert system is composed of two core elements: the Inference Engine, which applies rules and logical operations for decision-making, and the Knowledge Base, which houses validated information used by the system to generate recommendations. Fig. 3 depicts the structure of the fuzzy expert system employed in this research [10].

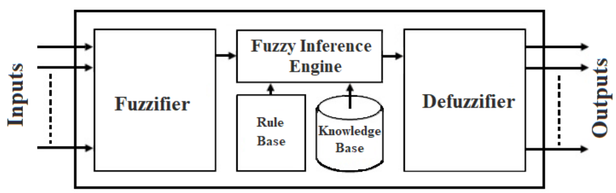


Fig.3. The structure of FLS

The components of the FLS include:

1) *Fuzzifier*: This module derives input data into a format that aligns with fuzzy logic principles, enabling it to be expressed in linguistic terms.

2) *Inference Engine*: This module generates fuzzy outputs by applying the fuzzifier's inputs to the knowledge rule base. The Mamdani method, widely employed in fuzzy logic applications, has been selected for this research.

3) *Knowledge Base*: This element organizes the results utilized for decision-making.

Various fuzzy systems implement an IF-THEN programming framework to apply knowledge. The knowledge base is divided into two sections: the database and the rule base. Rules within the knowledge base generally follow this format [11]:

IF (conditions) THEN (actions)

4) *Defuzzifier*: Commonly known as the defuzzifier, this module converts the fuzzy output from the inference engine into a precise or non-fuzzy value that is applicable for practical use.

C. *Design of FLS*

During finger movements, five key parameters measured with flex sensors (Flex1, Flex2, Flex3, Flex4, Flex5) were chosen as input variables for classifying and predicting the sign performed in this study. The output variable, representing the numerical value of the sign being made, is denoted as (Sign), with digits selected as follows: 1 for the sign of 1; 2 for the sign of 2; 3 for the sign of 3; 4 for the sign of 4; 5 for the sign of 5 and 6 for the fist sign. Fig. 4 illustrates the structure of the FLS regulator, including the input and output variables.

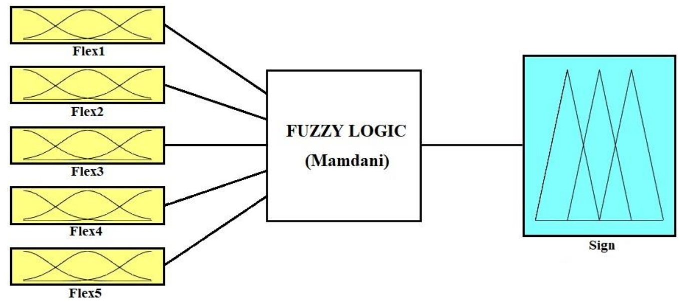


Fig.4. The architecture of FLS editor

1) *Input variables*

In the constructed expert system, five input variables (Flex1-5) were designated to represent the movement information for each finger. The boundary values for each fuzzy expression are detailed below. These parameters are classified symbolically. Membership functions and their degrees have been predefined in the software. Given that the same type of sensor is employed for each finger, the value ranges from 0 to 10 is divided into five fuzzy sets and variable ranges as follows:

VS (Very Slow) – S (Slow) - N (Normal) – F (Fast) - VF (Very Fast)

2) *Output variables*

The output variables have been divided into six fuzzy sets and variable ranges, selected empirically within the range of 0 to 6, which is considered the threshold value.

S1 ("sign 1") – S2 ("sign 2") – S3 ("sign 3") – S4 ("sign 4") – S5 ("sign 5") – S6 ("fist" sign)

Table 1 presents the input parameters, output parameters, and their corresponding linguistic expressions.

TABLE I
FUZZY SETS AND SHORT FORM LINGUISTIC EXPRESSIONS

Input Variables					*Output Variables*
Flex1	*Flex2*	*Flex3*	*Flex4*	*Flex5*	*Sign*
VS [0 2 4]	*VS [0 2 4]*	*VS [0 2 4]*	*VS [0 2 4]*	*VS [0 2 4]*	*S1 [0-2]*
S [2 4 6]	*S [2 4 6]*	*S [2 4 6]*	*S [2 4 6]*	*S [2 4 6]*	*S2 [1-3]*
N [4 6 8]	*N [4 6 8]*	*N [4 6 8]*	*N [4 6 8]*	*N [4 6 8]*	*S3 [2-4]*
F [6 7.5 9]	*F [6 7.5 9]*	*F [6 7.5 9]*	*F [6 7.5 9]*	*F [6 7.5 9]*	*S4 [3-5]*
VF [8 9 10]	*VF [8 9 10]*	*VF [8 9 10]*	*VF [8 9 10]*	*VF [8 9 10]*	*S5 [4-6]*
					S6 [5-7]

3) Rule Base

Classification uncertainties present significant challenges for deep learning models. However, fuzzy rule-based (FRB) approaches are particularly adept at managing these uncertainties. FRB methods are dependable for making inferences and are characterized by internal structures that are both efficient and straightforward to interpret [12]. Fuzzy set theory, an extension of classical crisp set theory, deals with the notion of partial truth using values that span from 0 to 1. A value of 0 denotes an entirely false statement, whereas a value of 1 signifies an entirely true statement [13]. After establishing linguistic variables, defining linguistic terms, and constructing membership functions, the final step in designing a fuzzy system is to develop a rule base. Rules verbally articulate the relationships between input and output linguistic variables according to their linguistic terms. A rule base constitutes the collection of rules for a fuzzy system [14]. In the creation of the conventional algorithm, the Rule Base and Expert System methodologies were employed as inputs for the fuzzy inference system. The initial phase in designing a fuzzy logic system typically involves the creation of an "IF-THEN" rule table. Using the membership functions we developed, a rule base consisting of 94 rules was obtained with the help of an expert. The rules derived from these membership functions are outlined in Table 2.

TABLE II
A SUMMARIZED SET OF RULES WAS DERIVED FROM THE MEMBERSHIP FUNCTIONS

Rule No:	***Rule structure***	***Input variables***					***Rule structure***	***Output variables***
	condition	*Flex 1*	*Flex 2*	*Flex 3*	*Flex 4*	*Flex 5*	*action*	***Sign***
1	***if***	*VS*	*VS*	*S*	***S***	*VS*	***then***	***S1***
2		*VS*	*VS*	*M*	*VS*	*S*		***S1***
3		*VS*	*VS*	*H*	*S*	*S*		***S1***
…		…	…	…				**…**
16		*S*	*N*	*N*	*N*	*N*		***S3***
17		*S*	*N*	*F*	*N*	*S*		***S2***
18		*S*	*N*	*F*	*F*	*F*		***S4***
…		…	…	…				**…**
92		*F*	*VF*	*N*	*VF*	*VF*		***S6***
93		*F*	*VF*	*F*	*VF*	*VF*		***S5***
94		*VF*	*F*	*VF*	*VF*	*VF*		***S6***

According to Table 2, a total of 94 fuzzy rules have been formulated, with examples provided below. The accuracy value for each rule has also been determined.

Rule 1: If (Flex1 is VF) and (Flex2 is N) and (Flex3 is VF) and (Flex4 is VF) and (Flex5 is VF) then (Sign is S1).

Rule 2: If (Flex1 is F) and (Flex2 is N) and (Flex3 is VF) and (Flex4 is VF) and (Flex5 is VF) then (Sign is S1).

…

Rule 93: If (Flex1 is F) and (Flex2 is F) and (Flex3 is VF) and (Flex4 is F) and (Flex5 is VF) then (Sign is S6).

Rule 94: If (Flex1 is F) and (Flex2 is F) and (Flex3 is VF) and (Flex4 is F) and (Flex5 is F) then (Sign is S6).

The Mamdani approach is employed as the inference engine. The degrees of validity (α) for each rule are computed using the Mamdani max-min method, as specified by the following formulas.

$\alpha_1 = min\ (Low\ (x), Slow\ (y))$
$\alpha_2 = min\ (Low\ (x), Medium\ (y))$
...
$\alpha_8 = min\ (High\ (x), Medium\ (y))$
$\alpha_9 = min\ (High\ (x), Fast(y))$

The maximum validity scores for the activated rules are computed using the formulas provided below.

$\alpha_{1,2 \ldots n} = max\ (\alpha_1, \alpha_2, \ldots \alpha_n)$

The centroid method, also known as the center-of-gravity or center-of-area technique, is utilized as the defuzzification approach in this study [15]. This method involves calculating the output of each membership function along with its associated maximum membership value (z*) using the formula specified in Equation (1). The range value relative to the centroid of the respective membership functions is represented by ž [16].

$$z^* = \frac{\sum \mu r(\check{z})\check{z}}{\sum \mu r(\check{z})} \qquad (1)$$

Following the definition of the input-output parameters and their respective boundary ranges, the Triangular model depicted in Fig. 5 was employed as the method for fuzzifying membership functions, transforming the data into fuzzy sets. This model block utilizes a membership function with a triangular shape.

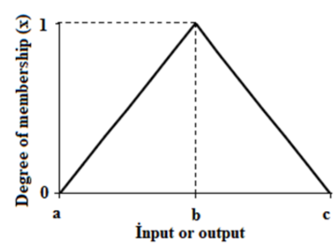


Fig.5. Triangular membership model

The defuzzification function for the triangular membership model is defined by Equation (2).

$$\mu(x) = \mu(x,a,b,c) = f(x) = \begin{cases} \frac{(x-a)}{(x-b)}, & a \le x < b; \\ \frac{(c-x)}{(c-b)}, & b \le x < c; \\ 0, & x > c \text{ or } x < a; \end{cases} \qquad (2)$$

The membership functions for the fuzzy sets, derived using the centroid method, are aggregated using a weighted average approach. Each function is weighted by its peak membership degree. The equations governing the membership functions for both the input and output parameters are provided below. The detailed equations for the membership functions of the Flex1 variable are presented in Equations (3-7).

$$\mu_{VS}(x) = \begin{cases} \frac{(4-X)}{4}, & 0 \le x \le 4; \\ 0, & x \ge 4 \; or \; x < 0; \end{cases} \quad (3)$$

$$\mu_{S}(x) = \begin{cases} \frac{X}{4}, & 2 \le x \le 4; \\ \frac{(6-X)}{4}, & 4 \le x \le 6; \\ 0, & x \ge 6 \; or \; x < 0; \end{cases} \quad (4)$$

$$\mu_{N}(x) = \begin{cases} \frac{(X-4)}{4}, & 4 \le x \le 6; \\ \frac{(8-X)}{6}, & 6 \le x \le; 8 \\ 0, & x \ge 8 \; or \; x \le 4; \end{cases} \quad (5)$$

$$\mu_{F}(x) = \begin{cases} \frac{(X-6)}{6}, & 6 \le x \le 7.5; \\ \frac{(9-X)}{7.5}, & 7.5 \le x < 9; \\ 0, & x \le 7.5 \; or \; x \ge 9; \end{cases} \quad (6)$$

$$\mu_{VF}(x) = \begin{cases} 0, & x \le 8; \\ \frac{(X-8)}{4}, & 8 \le x \le 10; \\ 1, & 10 \le x \end{cases} \quad (7)$$

The membership functions for the Sign output are specified in Equations (8-13).

$$\mu_{S1}(x) = \begin{cases} \frac{(2-X)}{2}, & 0 \le x \le 2; \\ 0, & x \ge 2 \; or \; x < 0; \end{cases} \quad (8)$$

$$\mu_{S2}(x) = \begin{cases} \frac{X}{2}, & 1 \le x \le 2; \\ \frac{(3-X)}{2}, & 2 \le x \le 3; \\ 0, & x \ge 3 \; or \; x < 0; \end{cases} \quad (9)$$

$$\mu_{S3}(x) = \begin{cases} \frac{(X-2)}{2}, & 2 \le x \le 3; \\ \frac{(4-X)}{3}, & 3 \le x < 4; \\ 0, & x \ge 4 \; or \; x \le 2; \end{cases} \quad (10)$$

$$\mu_{S4}(x) = \begin{cases} \frac{(X-3)}{3}, & 3 \le x \le 4; \\ \frac{(5-X)}{4}, & 4 \le x < 5; \\ 0, & x \le 4 \; or \; x \ge 3; \end{cases} \quad (11)$$

$$\mu_{S5}(x) = \begin{cases} \frac{(X-4)}{4}, & 4 \le x \le 5; \\ \frac{(6-X)}{5}, & 5 \le x < 6; \\ 0, & x \le 6 \; or \; x \ge 4; \end{cases} \quad (12)$$

$$\mu_{S6}(x) = \begin{cases} 0, & x \le 7; \\ \frac{(X-7)}{7}, & 7 \le x < 5; \\ 1, & 5 \le x \end{cases} \quad (13)$$

The graphical representation of the fuzzy input membership functions for the variable "Flex1" is shown in Fig. 6.

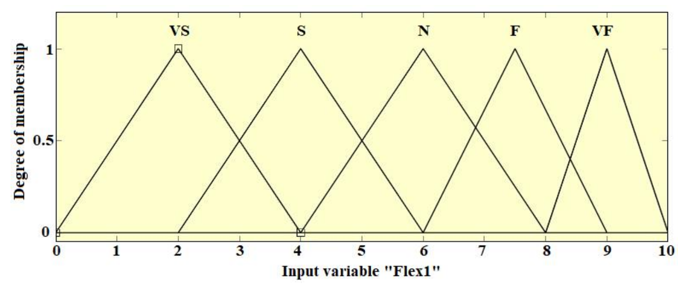


Fig.6. Graphical representation of the membership functions for the fuzzy input variable "Flex1".

Fig. 7, Graphical representation of the membership functions for the fuzzy output variable "Sign".

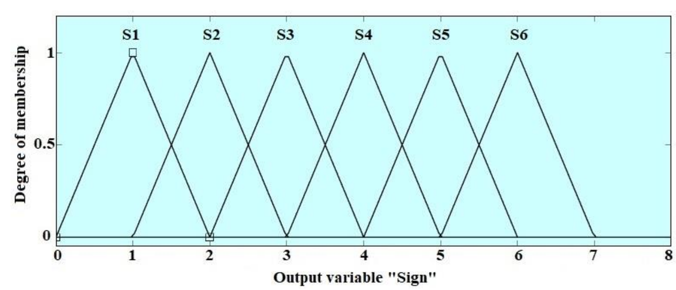


Fig.7. Graphical representation of the membership functions for the fuzzy output variable "Sign"

D. Evaluation Metrics

To assess the performance of our proposed method, we calculated sensitivity, specificity, and accuracy rates using the following formulas. In Equations (13-16), "TP" stands for true positives, "TN" refers to true negatives, "FP" denotes false positives, and "FN" indicates false negatives. These terms are standard in binary classification for evaluating prediction outcomes [17].

$$\text{Accuracy} = \frac{TN + TP}{TN + TP + FP + FN} \; X \; 100 \quad (14)$$

$$\text{Specificity} = \frac{TN}{TN + FP}\ X\ 100 \qquad (15)$$

$$\text{Sensitivity / Recall} = \frac{TP}{TP + FN}\ X\ 100 \qquad (16)$$

E. Statistical Analysis

Table 3 presents the complete list of features extracted from the hand gesture dataset along with their respective equations. Here, *i* ranges from 1 to *K*, representing *K* pieces of information assumed to be represented by the data u_i. Here, *u* denotes the mean, and n_i represents the frequency of occurrence of u_i in the dataset. The results have been statistically interpreted based on the formulas provided here.

TABLE III
DESCRIPTIVE STATISTICAL FEATURES

	Feature	*Equation*
1	*Mean*	$\frac{1}{K}\sum_{i=1}^{K} u_i$
2	*Range*	$u_{max} - u_{min}$
3	*Standard error*	$\frac{Standard\ Deviation}{\sqrt{K}} = \frac{\sqrt{\frac{1}{K-1}\sum_{i=1}^{K}(u_i - \bar{u})^2}}{\sqrt{K}}$
4	*Median*	*if K is odd:* $\left(\frac{K+1}{2}\right)^{th}$ *term* *if K is even:* $\frac{\left(\frac{K}{2}\right)^{th} term + \left(\frac{K+1}{2}\right)^{th} term}{2}$
5	*Mode*	$\{u_i: n_i = max\}, i = 1, 2, \ldots, K$
6	*Standard deviation*	$\sqrt{\frac{1}{K-1}\sum_{i=1}^{K}(u_i - \bar{u})^2}$
7	*Sample Variance*	$\frac{1}{K-1}\sum_{i=1}^{K}(u_i - \bar{u})^2$
8	*Kurtosis*	$\frac{\frac{1}{K}\sum_{i=1}^{K}(u_i - \bar{u})^4}{\left(\sqrt{\frac{1}{K-1}\sum_{i=1}^{K}(u_i - \bar{u})^2}\right)^2}$
9	*Skewness*	$\frac{\frac{1}{K}\sum_{i=1}^{K}(u_i - \bar{u})^3}{\left(\sqrt{\frac{1}{K-1}\sum_{i=1}^{K}(u_i - \bar{u})^2}\right)^3}$
10	*Maximum*	u_{max}
11	*Minimum*	u_{min}
12	*Sum*	$\sum_{i=1}^{K} u_i$

III. EXPERIMENTAL RESULTS

The developed glove model initially performed numerical signs from 1 to 5, followed by the fist clenching movement, during which the flexible sensor data were recorded in real-time on a computer for subsequent analysis. The resulting signal patterns can be seen in Fig. 8.

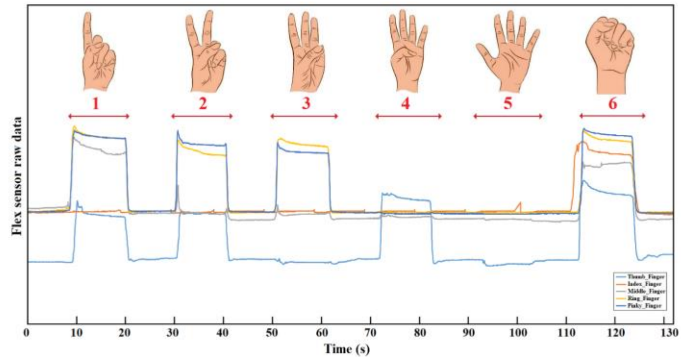
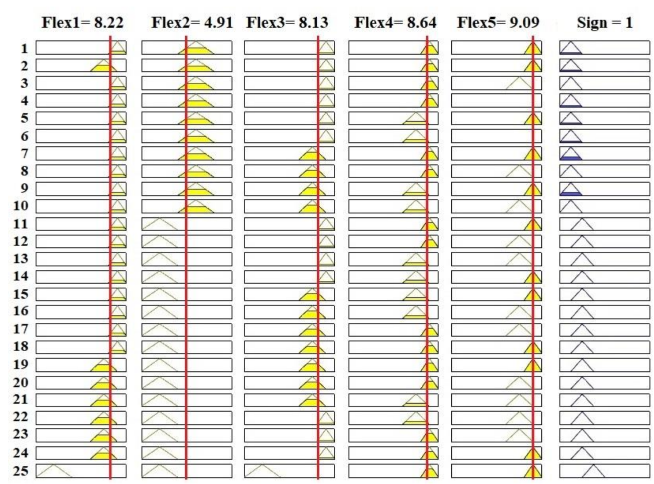
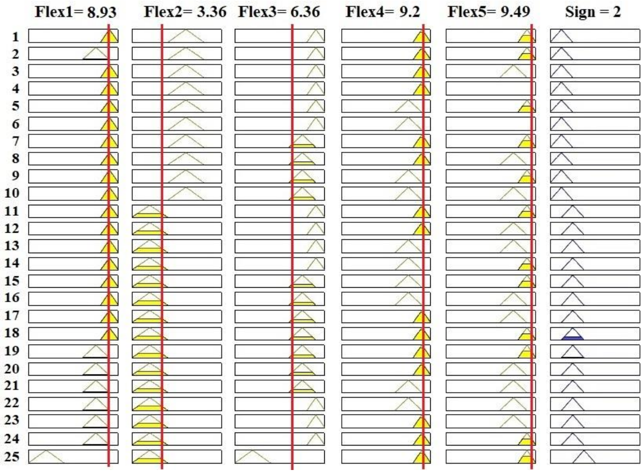


Fig.8. Flex sensor response for signs

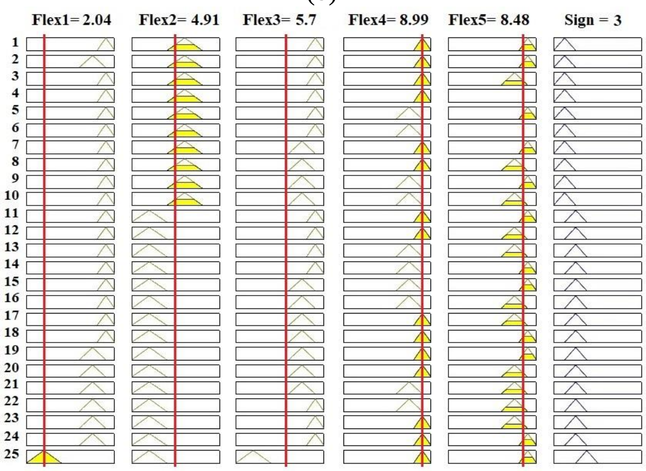
In this study, a fuzzy logic system was designed to classify and detect hand movements. The developed FLS system includes a Fuzzifier, Inference Engine, Knowledge Base, and Defuzzifier. The Mamdani technique was employed for the Inference Engine, while the centroid method was utilized for Defuzzification. A rule base comprising 94 rules was established using five input variables (Flex1-5) and one output variable (Sign). A sensor array was placed on a glove to generate data, and a data collection circuit was established. Movements were performed through this circuit to create the rule bases. A total of 15.030 data points were analyzed to develop the FLS. According to the results obtained from this study, for example, Flex1: 8.22; Flex2: 4.91; Flex3: 8.13; Flex4: 8.64; and Flex5: 9.09, the linguistic output value of Sign1 was found to be 1. This means that the sign performed according to these inputs is "1." The results obtained from fuzzy inputs for all output states (1-6) are collectively as shown in Fig. 9.



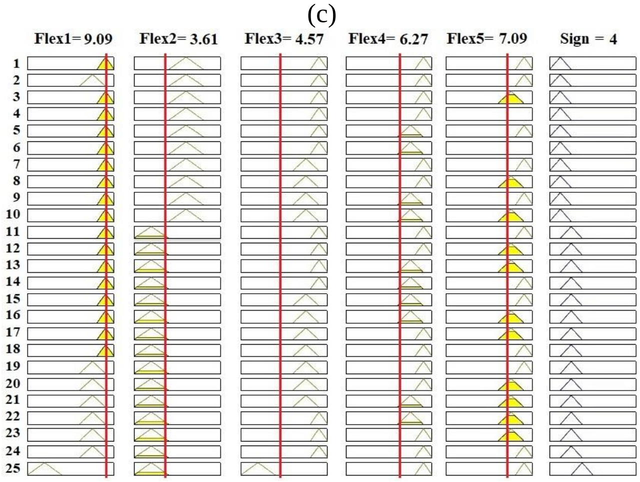
(a)



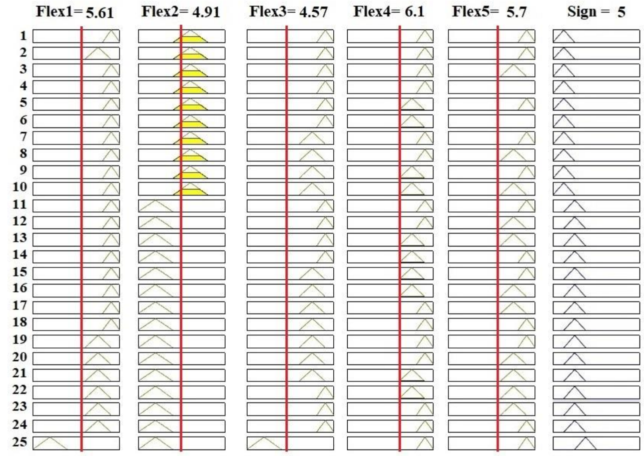
(b)



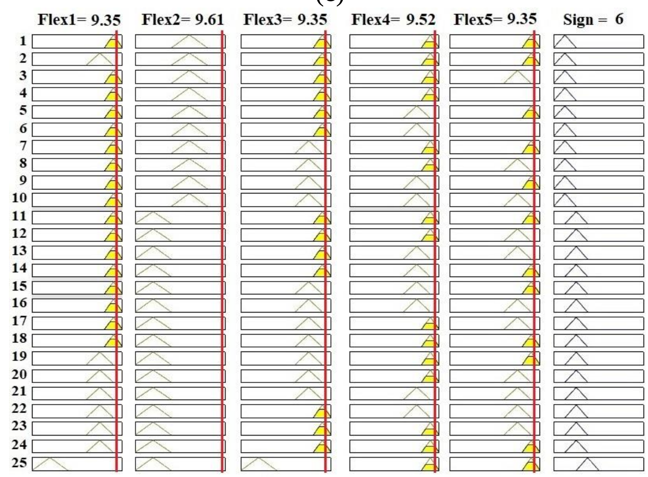
(c)



(d)



(e)



(f)

Fig.9. S1-S6 motion detection based on fuzzy inputs respectively. (Sign=1 (a), Sign =2 (b), Sign =3 (c), Sign =4 (d), Sign =5 (e), Sign =6 (f))

To test the accuracy of the measurement results, the glove was worn by different individuals, who were asked to perform 5 different movements. Accordingly, the real values and the results obtained by the FLS are comparably presented in Table 4. It was observed that the developed fuzzy model successfully identified all movements.

TABLE IV
THE COMPARISON OF RESULTS

Sign No.	*Flex1*	*Flex2*	*Flex3*	*Flex4*	*Flex5*	*FLS*	*Real*
1	***8.99***	***5.99***	***7.8***	***8.23***	***8.06***	***1***	***1***
2	***8.85***	***1.15***	***6.01***	***8.72***	***5.91***	***2.02***	***2***
3	***2.77***	***3.72***	***5.07***	***8.72***	***8.72***	***2.98***	***3***
4	***8.85***	***3.58***	***4.53***	***6.96***	***7.5***	***4.04***	***4***
5	***6.69***	***5.47***	***4.53***	***7.23***	***7.64***	***5***	***5***

According to the results, the movements (97.5%) were detected successfully. Table 5 presents the confusion matrix for the FLS across all movements.

TABLE V
THE CONFUSION MATRIX

n= 15.030	*Predicted: NO*	*Predicted: YES*
Actual: NO	*13.044*	*183*
Actual: YES	*124*	*1.679*

The results have been statistically analyzed and the results are shared in Table 6.

TABLE VI
DESCRIPTIVE STATISTICS RESULTS

	Measured	**Predicted**
Mean	232.61	202.23
Standard error	1.812	1.482
Median	332	284
Mode	354	305
Standard deviation	119.52	124.63
Sample Variance	15330.18	14222.48
Kurtosis	1.59	1.59
Skewness	0.502132	0.50223
Range	342	322
Maximum	372	319
Minimum	7	-3
Sum	128425	106452
Confidence Level (97.0%)	3.3592	3.3344

IV. DISCUSSIONS

Hearing impairment can be congenital or acquired later in life and may lead to significant communication challenges due to its impact on a large number of people worldwide. Individuals in this situation often face difficulties in social isolation, socialization, and participation in the workforce. This study presents a comprehensive framework using a new fuzzy logic model based on flexible sensors for recognizing hand movements with a prototype hand model. An intelligent hand movement classification system, based on the architecture of the fuzzy logic system, has been developed to assist deaf individuals in accurately recognizing various hand movements. The proposed community architecture achieves an impressive accuracy of 97.5%. Future work will emphasize improving recognition performance, particularly to achieve more accurate categorization results. Additionally, it is aimed to enable remote access to physiological information related to certain neurological diseases through system enhancements. Furthermore, the developed system is intended to be implemented in hospitals to address potential practical usage side effects, making it a valuable tool in clinical environments.

ACKNOWLEDGMENT

The research presented in this study was supported by the Coordinatorship of Scientific Research Projects at Ondokuz Mayis University in Samsun, Turkey. The project was identified by the project number PYO.YMY.1908.23.001.

REFERENCES

[1] Y. Zhang at all., “Mechanomyography signals pattern recognition in hand movements using swarm intelligence algorithm optimized support vector machine based on acceleration sensors”, Medical Engineering and Physics 124 (2024) 104060,

[2] M.A., Rahaman, M.H. Ali, and M. Hasanuzzaman, “Real-time computer vision-based gestures recognition system for bangla sign language using multiple linguistic features analysis”. Multimed Tools Appl 83, 22261–22294 (2024). https://doi.org/10.1007/s11042-023-15583-8

[3] World Health Organization (2021) Deafness and hearing loss. https://www.who.int/es/news-room/fact-sheets/detail/deafness-and-hearing-loss, 2019 Online; Accessed July 19

[4] H. Yakut, “İşaret dili harflerinin görüntü işleme yöntemleriyle taninmasi için bir uygulama”, Yüksek Lisans Tezi, Firat Üniversitesi, OCAK-2013

[5] F. ShengChen, C. MingFu, C. LinHuang, “Hand gesture recognition using a real-time tracking method and hidden Markov models”, Image and Vision Computing, 21(8), 1 August 2003, Pages 745-758

[6] Wu, D., Pigou, L., Kindermans, P. J., Le, N. D. H., Shao, L., Dambre, J., & Odobez, J. M., “Deep dynamic neural networks for multimodal gesture segmentation and recognition”, IEEE transactions on pattern analysis and machine intelligence, 2016, 38(8), 1583-1597.

[7] O. Güler, & I. Yucedag,. “Derin öğrenme ile el hareketi tanıma üzerine yapılan çalışmaların incelenmesi”, 21. Akademik Bilişim Konferansı, 2019, Ordu/TURKEY

[8] G. Devineau, F. Moutarde, W. Xi, & J. Yang, “Deep Learning for Hand Gesture Recognition on Skeletal Data”, In Automatic Face & Gesture Recognition (FG 2018), 2018, 13th IEEE International Conference on (pp. 106-113), IEEE.

[9] F. Başçiftçi, H. Sümbül, “Designing an expert system for detection of tuberculosis disease with logic simplification method”, e-Journal of New World Sciences Academy, NWSA Engineering Sciences, 1A0097, ISSN: 1308-7231,5(3), pp. 463-471, 2010.

[10] B. Ahmet, K. Yavuz, K. Adem, “Bulanik uzman sistem yaklaşimi ile yeşil kart başvuru değerlendirme sistemi”, Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi, Cilt: 24, Sayı: 1, 63-76, 2010

[11] H., Sümbül, A. H., Yüzer, “Design of a fuzzy input expert system visual information interface for classification of apnea and hypopnea”, Multimedia Tools and Applications, 83(7), pp.21133-21152,2024 https://dx.doi.org/10.1007/s11042-023-16152-9

[12] A.B. Sargano, X. Gu, P. Angelov et al., “Human action recognition using deep rule-based classifier”. Multimed Tools Appl., 2020, 30653–30667, https://doi.org/10.1007/s11042-020-09381-9

[13] B. Choi, S. Kang, K. Jun et al. “Rule-based soft computing for edge detection” Multimed Tools Appl, 24819–24831, 2017, https://doi.org/10.1007/s11042-016-4329-7.

[14] W. Bi, F. Gao, A. Zhang et al. “A framework for extended belief rule base reduction and training with the greedy strategy and parameter learning”. Multimed Tools Appl, 11127–11143, 2022, https://doi.org/10.1007/s11042-022-12232-4.

[15] A. Biswas, A. Adan, P. Haldar, D. Majumder, V. Natale, C. Randler, L. Tonetti, S. Sahu, “Exploration of transcultural properties of the reduced version of the Morningness–Eveningness Questionnaire (rMEQ) using adaptive neuro-fuzzy inference system”. Biological Rhythm Research. 45:6, 955-968, 2014, DOI:10.1080/09291016.2014.939442

[16] E. Allam, H.F. Elbab, M.A. Hady, S. Abouel-Seoud, “Vibration control of active vehicle suspension system using fuzzy logic algorithm”. Fuzzy Information and Engineering, 2(4):361-387, 2010, https://doi.org/10.1007/s12543-010-0056-3

[17] Ö. M. Soysal, S. Shirzad and K. Sekeroglu, "Facial action unit recognition using data mining integrated deep learning," 2017 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 2017, pp. 437-443, doi: 10.1109/CSCI.2017.74.

BIOGRAPHIES

Harun SÜMBÜL received the B.S. degree in Electronic Sciences from the Selçuk University, Konya, Turkey, in 2008, the M.S. degree from the Selçuk University from Electrical and Electronics Engineering, Konya, Turkey, in 2011, and received the Ph.D. degree in electrical and electronics engineering from Karabük University, Karabük, Turkey. His primary areas of research are biomedical engineering and biomedical signal processing. He is currently a Lecturer at the Yeşilyurt Demir Çelik Vocational School, Department of Biomedical Device Technologies at Ondokuz Mays University, Turkey.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 1, March 2025
Research Article

Development of a Smart Activity Recognition System with Transfer Learning Based Deep Learning Models for Elderly Care

Mehmet Ilyas Bayindir and Fahri Cihan Attila

***Abstract*—In recent years, smart healthcare services have become popular in recent scientific research trends. Under elderly care topic, fall detection and activity recognition of elderly person living alone in their house or in a nursing home are vitally important solutions. Because falls are primary cause of most of the injuries, traumas, need of care and even deaths. To find a solution to this issue, scientists are start to use Artificial Intelligence. In this study, an intelligent activity recognition and fall detection system based on Convolutional Neural Network was developed. To develop this system an original dataset was created. By the proposed system, time distributions and classes of the activities are observed. When a fall is detected, the system gives an alert and warns relevant persons. The performances of used different models were compared using the dataset we created. To evaluate the performance of the systems, accuracy, precision, recall and F1 score metrics was used. For the ResNet101, these metrics are obtained as 98.66%, 98.54%, 98.78%, 98.66% respectively, that is the best of all scores. This results show that trained ResNet101 system can be used to help elderly persons and can be integrated to the other IoT systems.**

Index Terms— **Deep Learning, Elderly Care, Fall Detection, Human Activity Recognition, Internet of Things.**

I. INTRODUCTION

THE HEALTH conditions of elderly individuals are fragile, and they have a challenge for self-care due to underlying illnesses. The population of these individuals is growing, for instance, in the UK, it was identified that there were 2376 individuals aged 75 and above living alone [1], [2]. As cities turn to smart solutions under increasing population pressures, ensuring the safety and well-being of these individuals becomes a battleground.

Mehmet İlyas Bayındır, is with Department of Ecoinformatics, Institute of Natural and Applied Sciences, Fırat University, Elazığ, Turkey, (e-mail: mbayindir@firat.edu.tr)
https://orcid.org/0000-0003-1999-014X

Fahri Cihan Attila, is with Department of Ecoinformatics, Institute of Natural and Applied Sciences, Fırat University, Elazığ, Turkey, (e-mail: fahricihanattila@gmail.com)
https://orcid.org/0009-0002-8680-2591

Manuscript received Oct 24, 2024; accepted Jan 31, 2025.
DOI: 10.17694/bajece.1572976

The conditions of the pandemic have emphasized the significance of minimum contact while monitoring quarantined elderly and/or sick individuals. Elderly people living alone or receiving healthcare in care facilities can be monitored by using artificial intelligence to analyze camera stream. If camera surveillance is provided for solitary or care-dependent elderly individuals, monitoring can be carried out by analyzing activities from video. When any falls occur, emergency alerts are triggered, supportation is provided to healthcare security and reducing the workload of caregivers. Employing deep learning solutions within live video monitoring processes will swiftly and effectively support the maintenance of a healthy state and facilitate easier caregiving.

As stated in many studies of the World Health Organization [3], falls are common in the initial stages of most of the injuries, traumas that cause need of care, and even deaths [4]. This is a danger case not only for the elderly but also for individuals in working life [5]. If the fall event is determined with early warning and intervened early, the individual can receive more effective healthcare and the cost of the health insurance system can be reduced. If scene tracking and human activity recognition solutions used in elderly care, a caregiver can serve more patients.

Internet of Things (IoT) technologies are especially used by smart city services. Deep Learning (DL) solutions, particularly Convolutional Neural Network (CNN) architectures, can effectively process large datasets to make predictions, forecasts, and insights without the need for preprocessing tasks like feature extraction [6] . When real-time videos are classified by a deep learning network trained on a comprehensive dataset, vital contributions can be provided for smart healthcare and elderly care.

Three main classifications can be made for assessing systems about falls detection and monitoring activities of humans, as regarding the types of data and sensors employed. As illustrated in Fig. 1, these classifications contain (I) systems depending on wearable sensors, (II) camera based systems, and (III) systems incorporating environmental sensors. For the first class, examples include wearable gyroscopes, smartwatches, wristbands, belt clips, motion detectors, and similar devices. The second category predominantly employs RGB cameras and depth cameras and infrared cameras can also be mentioned. Lastly, the third category entails acoustic sensors, pressure and floor sensors, presence detectors, inertial sensors, microphones etc. [7].

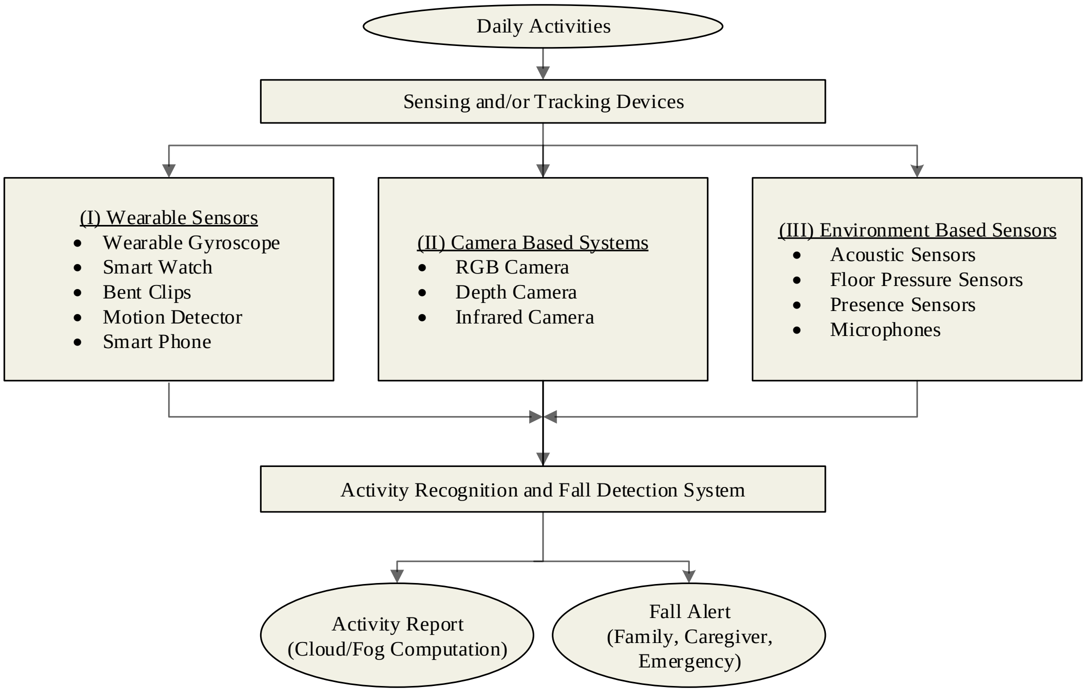


Fig. 1. Classification of fall detection approaches based on sensors

Perumal et al.[8] studied on fall detection systems developed for using in smart / non-smart home environment. They analyzed the most dangerous fall situation seen on wet floors. Different sensors, gateways and classification methods such as machine learning, rule-based, deep learning are investigated classification. They presented a scope especially for the researchers studying on the interoperability aspect of smart homes and activity recognition, especially fall detection systems in bathrooms/ toilets. Islam et al. [9] proposed a review of recent advances about fall detection systems based on deep learning. They are categorized as: CNN based systems, LSTM based systems, and Auto-encoder based systems. The most used method is CNN and 3D CNN or CNN with 10-fold cross-validation performed the best among the evaluated systems. Hardware technology have two categories: sensor based systems and vision based systems. The goal of the reviewed fall detection systems was to detect cases of elderly people falling using the best deep learning methods and to inform a nearby nurse or support/medical staff within a short time. Purwar, A. and Chawla [10] reviewed papers published in 2017-2023 on fall detection systems to protect elderly people. The papers contain both wearable, non-wearable systems and hybrid systems as regarding new technologies such as deep learning, computer vision, Internet of Things (IoT) and big data. The drawback of wearable sensors is that elderly people can forget to wear these devices. Zhou et al. [11] suggested using data from radar and optical cameras in a similar manner. Features extracted from these data by using Short-Time Fourier transform. They also used three different CNN models for feature extracting from radar signal. These are Alex-Net, a single-shot detector (SSD-Net) and a second SSD-Net. Maitre and Bouchard [12] introduced a new system to detect falls by using UWB radars and a CNN-LSTM architecture to use for indoor environment. They suggested that real-time fall detection performance should be studied in future works by using proper devices and communication protocol. Santos et al. [13] employed fog computing-enabled CNN to build a human fall detection system. They gathered training data from smartwatches and phones. The proposed CNN model was operated on a local fog device, capable of extracting appropriate features from the acquired data. Torti and colleagues [14] presented embedded software utilizing a Recurrent Neural Network (RNN) on the SensorTile wearable device, which incorporates a 3D accelerometer, gyroscope, magnetometer, and barometer. This software can detect the falls and then transmit alerts to a surveillance system via a wireless communication. Mauldin et al. [15] created an Android application. They used data from a smartwatch worn by the user. Their system is based on RNN model. This intelligent application is installed on a smartphone. In order to recognition instantaneously, computations are carried out directly on the smartphone. Another a CNN and LSTM architecture combination is also used for fall detection. The "ConvLSTM" model was developed by Nait Aicha et al. [16] by combining convolutional and recurrent models. Lie et al. [17] proposed the utilization of a Long Short-Term Memory (LSTM) classifier on human body joints for the characterization of various movements by using RGB sequence. Shojaei-Hashemi et al. [18] used transfer learning for human fall detection. This technique can identify human falls when they occur, extract features from camera depth maps, and send out an alert to notify families. Lu et al. [19] used 3D CNN and LSTM techniques. While the LSTM model is used to identify the interesting section of a frame, the 3D CNN is used to extract motion features from the temporal sequence of video sequences. Min et al. [20] introduced a deep CNN technique called Regions with CNN (R-CNN), which proves to be a more precise and expedient method for object detection in categorizing falling incidents. The fundamental concept of this approach is to determine the distances between individuals and furniture within the scene. This method is implemented by extracting various characteristics like human body's center of mass, shape parameters, and velocity of motion. Adhikari et al. [21] presented a vision-based monitoring method to detect falls utilizing an CNN model. Additionally, they employed recurrent networks with time-sequential and mobile data to rapidly and accurately capture falls when they occur. Jiang et al. [22] proposed a real-time human fall detection method by using an infrared array sensors and a lightweight deep learning network. They used RetinexNet to enhance picture contrast and remove noise information in the dataset. Some improvements are made in the original YOLOv5 to reduce the complexity of model calculations and the calculation time. Their suggestion about future work is to improve the robustness of the model by adding additional images of various home scenarios to a larger dataset. Mobsite et al. [23] are introduced a system for activity recognition and fall detection with monocular depth and motion analysis. They used the Xception network to reduce the false alerts caused by non-human motion. Zi et al. [24] proposed a vision-based fall detection system. The proposed method integrates dual illumination estimation to the YOLOv7 + Deep SORT tracking algorithm to enhance fall detection performance under suboptimal lighting conditions. The performance of the proposed method is validated on two fall detection datasets, namely, Le2i FDD and UR-Fall datasets. In a recent study, Khekan et al. [25] are investigated fall detection applications based on deep learning especially for YOLOv1-YOLOv8. They provide a benchmark for future studies based on YOLO for human fall detection. The literature studies considered in this study generally used processed datasets. However, in our study, the proposed system's performance was tested on videos which are obtained from realized experimental setup.

In this study, an intelligent activity recognition and fall detection system based on Convolutional Neural Network was developed. To develop this system an original dataset was created. It is aimed to observe time distributions of the daily activities by the proposed system. Five class of activity are defined to recognize by the system. When the system detects a fall, the system gives an alert and warns caregivers or insurance services. These activities are classified by analyzing video frames captured by a general purpose camera. Transfer learned successful ten different CNN models employed as classifiers in this system. These models trained by created dataset. At the end of training process, proposed system is separately tested as regarding employed different CNN models. To evaluate the performance of the systems, accuracy, precision, recall and F1 score metrics was used. The performances of proposed systems were compared using the dataset we created. The best of all scores are obtained for the ResNet101. This results show that trained ResNet101 system can be used to help elderly persons and can be integrated to IoT systems.

The remainder of the paper is organized as follows: In Section 2, the proposed method and its theoretical background are explained. Section 3 provides a detailed discussion of the dataset used in the study, the experimental setup and evaluation metrics, as well as the experimental results. Finally, a general summary of the study is presented in Section 4, the conclusion.

II. Material and Methods

A. The Created Dataset

This study is aimed to develop an smart system based on a deep learning network that can operate in real-time through simple general-purpose cameras for elderly care in smart healthcare services. In order to train the network, an original dataset was created. To avoid bandwidth overload when used in cloud computing, the video format is based on the 320x240 QVGA basic resolution. The dataset is created from videos captured while subjects of different ages and genders perform natural movements in a setting resembling an elderly care/resting room. Six types of classifications commonly accepted in the literature are identified: Empty room, standing, sitting, sleeping, bending, and falling. The used labels and the number of frames in the dataset are listed in Table 1 sorted by numbers. Fig. 2 shows sample images from the dataset with six different classes.

B. Transfer Learning

Transfer learning involves reusing a pre-trained primary deep learning network, which has been successfully trained on a large dataset [26], by retraining it for a specific classification purpose. When there is not enough large data available, the transfer learning process can be highly beneficial. In this process, by default, the weights of the convolutional layers before the fully connected layer in the primary network structure, which goes from general to specific, are kept untouched, and training with the new dataset is started from this stage.

TABLE I.
LABELS AND NUMBERS OF DATA IN THE DATA SET

Class	Number of Images
Standing	5349
Empty room	2022
Falling	2835
Bending	2328
Sitting	3765
Sleeping	1173
TOTAL	17472

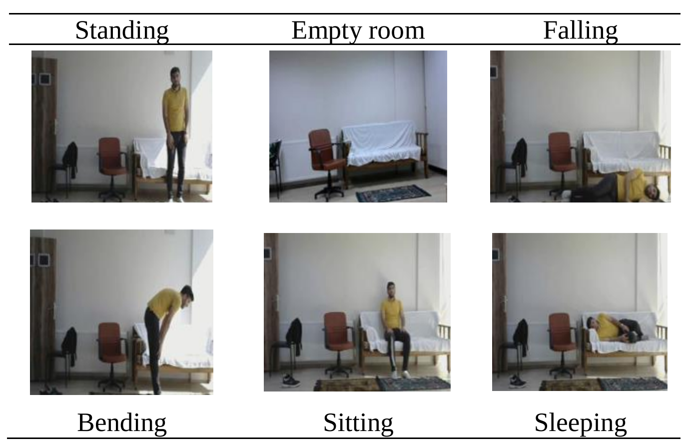


Fig. 2. Sample images of the categories in the data set

These weights can be frozen or re-adjusted during the training process. If the new dataset is not comprehensive enough and the weights are frozen, the problem of overfitting may be encountered. In this study, ten different deep learning architectures based on transfer learning are developed by the ImageNet dataset. These include ResNet50-ResNet101 [27], Xception [28], InceptionV3 [29], MobileNet [30], MobileNetV2 [31], DenseNet121-DenseNet201 [32] , EfficientNetB0 [33] and ConvNeXtTiny [34] methods.

1. Residual Network (ResNet): ResNet is a family of deep convolutional neural network architectures introduced by Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun in their 2015 paper titled 'Deep Residual Learning for Image Recognition' [27]. ResNet is known for its innovation in addressing the vanishing gradient problem, which frequently arises in very deep neural networks. The core idea behind ResNet is the use of residual connections, or skip connections, which allow information to flow more easily through the network. These skip connections essentially skip one or more layers, making it possible to train very deep networks. There are various variants of the ResNet architecture, primarily differing in depth, meaning the number of layers they contain. These include ResNet50, ResNet101, and ResNet152, described as follows:

ResNet50: A relatively shallow variant of the ResNet architecture. It has a total of 50 layers, including convolutional layers, batch normalization, and fully connected layers. The '50' refers to the total number of layers in the network. This model is widely used for a variety of computer vision tasks and is suitable for scenarios where computational capacity is limited.

ResNet101: A deeper variant of ResNet compared to ResNet50. It has a total of 101 layers, providing more depth and, in some cases, better performance. Deeper networks can capture more complex features and tend to perform better on challenging tasks, provided sufficient computational resources are available.

ResNet152: Even deeper than ResNet101, with a total of 152 layers. It is the deepest among the three variants mentioned. The increased depth allows ResNet152 to capture more complex and fine-grained features from input data. However, this also requires more computational resources and takes longer to train.

All these ResNet variants produce highly effective results in tasks like image classification, object detection, and other computer vision tasks. They are often used as a foundation for transfer learning, where pre-trained models on large image datasets are fine-tuned for specific applications. Researchers can take pre-trained ResNet models and adapt them to various image-related tasks, saving time and resources compared to training from scratch. The choice between ResNet50, ResNet101, and ResNet152 depends on the specific requirements of the task, available computational resources, and the desired level of model depth.

2. Xception: Xception is a deep convolutional neural network architecture introduced by François Chollet, the creator of the Keras deep learning library, in a research paper published in 2017 titled 'Deep Learning with Depthwise Separable Convolutions.' The name 'Xception' stands for 'Extreme Inception,' indicating its relationship with the Inception neural network architecture family [28]. Xception is particularly known for its innovative approach to convolutions, especially through the use of depthwise separable convolutions. Unlike traditional convolutions, which apply a single convolutional filter to the entire input image, Xception employs the depthwise separable convolution process. This process breaks down the traditional convolution into two steps: depthwise convolution and pointwise convolution. In depthwise convolution, a separate convolution operation is applied to each channel of the input image independently. This significantly reduces computational cost compared to traditional convolutions, where the same filter is applied to all input channels. After depthwise convolution, pointwise convolution, also known as 1x1 convolution, is applied to combine the output channels. This step helps capture complex patterns by allowing linear combinations of the output channels. The use of depthwise separable convolutions in Xception enables more efficient use of model parameters and computational resources. This architectural choice results in a highly impressive and efficient deep learning model. Xception also uses skip connections, similar to ResNet architectures, to mitigate the vanishing gradient problem.

The Xception neural network is an ESA architecture with a depth of 71 layers. The Xception architecture contains a total of 36 convolutional layers, which form the core of the feature extraction process of the network. These convolutional layers are grouped into 14 modules, making up a total of 36 layers. Except for the first and last modules, all modules have linear residual connections surrounding them. The Xception architecture consists of a linear stack of depthwise separable convolutional layers with residual connections. The Xception model achieved 94.5% accuracy on ImageNet, a dataset with over 14 million images across more than 1000 classes.

*3. InceptionV3:*InceptionV3 is a deep convolutional neural network architecture introduced as part of the Inception family, specifically in the 2015 paper "Rethinking the Inception Architecture for Computer Vision" by Google researchers [29]. It is a refined and more efficient version of its predecessors, such as InceptionV1 and InceptionV2, and is known for its effective balance between computational cost and accuracy in large-scale image classification tasks. One of the key innovations in InceptionV3 is the use of factorized convolutions, where larger convolutional filters are broken down into smaller, more efficient operations (for example, a 5x5 filter is factorized into a series of 3x3 convolutions). This reduces computational demands while preserving high performance. InceptionV3 also introduces a more efficient method for grid size reduction, replacing pooling layers with strided convolutions, which allows for better information flow in deeper layers while minimizing computational load. The model includes auxiliary classifiers during training, acting as regularizers to help the model converge faster and prevent overfitting by adding additional supervision to the middle layers. Asymmetric convolutions further boost efficiency by decomposing large convolutions, such as a 7x7, into smaller operations like a 1x7 convolution followed by a 7x1 convolution, improving feature extraction with fewer parameters. Extensive use of batch normalization in InceptionV3 enhances training stability and accelerates convergence, while reducing overfitting risks. The model achieved remarkable success on the ImageNet dataset, with a top-5 error rate of 3.46%, making it one of the top-performing architectures for image classification. InceptionV3 is widely used in various computer vision tasks such as image classification, object detection, and image generation. It is also popular in transfer learning, where pre-trained weights are fine-tuned for specific applications. Overall, InceptionV3 stands out for its computational efficiency and accuracy, making it highly suitable for both research and real-world applications in computer vision.

4. MobileNet: MobileNet is a family of deep convolutional neural network architectures specifically designed for mobile and embedded vision applications, aimed at providing high performance with low computational cost. Introduced by Google researchers, MobileNet employs depthwise separable convolutions, which separate the convolutional operation into two stages: depthwise convolution, where individual filters are applied to each input channel, and pointwise convolution, which combines the outputs. This innovative approach significantly reduces the number of parameters and computational complexity compared to traditional convolutions. Additionally, MobileNet incorporates width and resolution multipliers that allow for fine-tuning the model's size and resource usage according to specific application requirements. The architecture is optimized for real-time performance on devices with limited processing power, making it suitable for tasks such as image classification, object detection, and face recognition. MobileNet models are also widely used in transfer learning, as pre-trained versions can be easily adapted for various specific tasks, ensuring a good balance between accuracy and efficiency in diverse mobile and embedded environments [30].

*5. MobileNetV2:*MobileNetV2 is an advanced version of the MobileNet architecture designed to further improve efficiency and performance for mobile and embedded vision applications. One of its key innovations is the introduction of inverted residuals and linear bottlenecks, which enhance feature extraction and reduce information loss. Inverted residuals allow for a lightweight convolutional operation where the input is first expanded into a higher-dimensional space using a pointwise convolution, followed by a depthwise separable convolution, and then reduced back to a lower-dimensional space with another pointwise convolution. This process helps in capturing more complex features while maintaining a small model size. MobileNetV2 also utilizes linear activation in the bottleneck layers instead of ReLU, which helps preserve information, especially in deeper layers. The architecture comprises an initial standard convolution layer, followed by a series of inverted residual blocks that alternate depthwise separable convolutions and pointwise convolutions, all concluded with a global average pooling layer and a fully connected layer for classification. This structure not only enhances the model’s ability to learn intricate patterns but also maintains high efficiency, making MobileNetV2 highly suitable for real-time applications in constrained environments [31], [35].

6.DenseNet121-DenseNet201: DenseNet (Densely Connected Convolutional Networks) is a deep learning architecture that emphasizes feature reuse and efficient gradient flow through its unique connectivity pattern. Unlike traditional convolutional networks that connect each layer to its preceding layer, DenseNet connects each layer to every other layer in a feed-forward manner. This means that each layer receives feature maps from all preceding layers, promoting the flow of information and gradients throughout the network, which helps alleviate the vanishing gradient problem and reduces the number of parameters required. Key features of DenseNet include dense connections, which improve feature propagation, and bottleneck layers, which reduce dimensionality before applying convolution, thereby enhancing computational efficiency [32] .

DenseNet121: This model consists of 121 layers and is organized into four dense blocks, each followed by a transition layer that performs down sampling. The first block has a convolution layer followed by a batch normalization and ReLU activation. Each dense block consists of multiple convolutional layers where each layer takes input from all previous layers, culminating in a global average pooling layer and a fully connected layer for classification [32] [36] .

DenseNet201: Similar to DenseNet121, this model comprises 201 layers and follows the same architecture principles with more layers in each dense block, allowing it to capture more complex features. The additional layers enable deeper representations while maintaining the benefits of dense connectivity. Like DenseNet121, it concludes with a global average pooling layer and a fully connected layer, enhancing its performance in image classification tasks [32], [37].

Overall, DenseNet’s innovative approach to layer connectivity significantly improves feature extraction and model efficiency, making it highly effective for various computer vision applications.

7. EfficientNetB0: EfficientNetB0 is the foundational model of the EfficientNet family, designed to balance accuracy and efficiency in image classification tasks. It introduces a novel compound scaling method, which systematically scales up the model's depth, width, and input resolution in a way that optimizes performance without excessive computational cost. This approach allows EfficientNetB0 to outperform many previous architectures while using significantly fewer parameters. Key features of EfficientNetB0 include its use of depthwise separable convolutions, which reduce the number of parameters and computational load by separating the spatial and channel-wise convolutions. Additionally, it employs the swish activation function, which provides better performance compared to traditional activation functions by enabling smoother gradients and improving convergence during training [33].

EfficientNetB0 comprises several essential layers structured to maximize efficiency and feature extraction. The model begins with a stem block that consists of a standard convolutional layer. This layer captures initial features and reduces the spatial dimensions of the input image. The core of EfficientNetB0 consists of multiple MBConv blocks (Mobile Inverted Bottleneck Convolution). Each MBConv block utilizes depthwise separable convolutions, where a depthwise convolution is applied first to each channel, followed by a pointwise convolution to mix the outputs. This arrangement allows the model to efficiently capture both local and global features while keeping the computational requirements low. The blocks are arranged in a hierarchical fashion, with the number of filters and layers increasing as the network progresses, which enhances its capacity to learn complex patterns. Following the MBConv blocks, the architecture includes a global average pooling layer that aggregates the features across the spatial dimensions, effectively reducing the output to a single vector per feature map. This step significantly decreases the model's parameter count while retaining essential information. Finally, a fully connected layer processes the output from the global average pooling layer, producing class probabilities through a softmax activation function, making it suitable for multi-class classification tasks [33] .

In summary, EfficientNetB0 is a pioneering model that effectively balances performance and efficiency through its innovative architecture, making it highly suitable for various applications in computer vision. Its combination of depthwise separable convolutions, compound scaling, and an efficient layer structure allows it to achieve superior accuracy with reduced resource requirements.

8. ConvNeXtTiny: ConvNeXtTiny is a compact yet powerful convolutional neural network architecture that combines the principles of traditional convolutional networks with modern design elements inspired by transformer architectures. Key features of ConvNeXtTiny include its emphasis on simplicity and efficiency while maintaining high performance across various computer vision tasks. It employs a streamlined design that focuses on improved normalization techniques, such as layer normalization, and a novel architecture that utilizes depthwise separable convolutions to enhance computational efficiency. This model benefits from fewer parameters compared to larger models while still achieving competitive accuracy, making it ideal for resource-constrained environments [34].

The ConvNeXtTiny architecture consists of several key layers structured to optimize feature extraction and efficiency. The model begins with a stem block that processes the input image through a standard convolutional layer, reducing the spatial dimensions while extracting initial features. The core of ConvNeXtTiny consists of multiple ConvNeXt blocks, which are designed to efficiently learn hierarchical representations. Each block incorporates depthwise convolutions followed by pointwise convolutions, allowing for efficient computation. These blocks also feature layer normalization and an innovative activation function, such as GELU (Gaussian Error Linear Unit), which helps to model complex relationships within the data. After processing through the ConvNeXt blocks, the model uses a global average pooling layer that aggregates the features across the spatial dimensions, transforming the feature maps into a compact representation. The architecture concludes with a fully connected layer that outputs the final class probabilities through a softmax activation function, enabling the model to perform multi-class classification effectively. ConvNeXtTiny is a modern convolutional network that integrates elements from both traditional CNNs and transformer architectures to create a compact, efficient, and high-performing model for various vision tasks. Its thoughtful architecture, characterized by depthwise separable convolutions and improved normalization techniques, enables it to achieve strong performance while maintaining a lightweight footprint, making it suitable for deployment in applications with limited computational resources [34], [38] .

C. *Proposed Method*

The schema of the proposed method using pre-trained models for transfer learning is as shown in Fig.3. As shown in Fig.3, the input images are of size 320x240x3. Using these input images, training was carried out separately with ten different pre-trained network models based on transfer learning. The pre-trained network models used in this study are as follows: ResNet50, ResNet101, Xception, InceptionV3, MobileNet, MobileNetV2, EfficientNetB0, DenseNet121, DenseNet201, and ConvNeXtTiny.

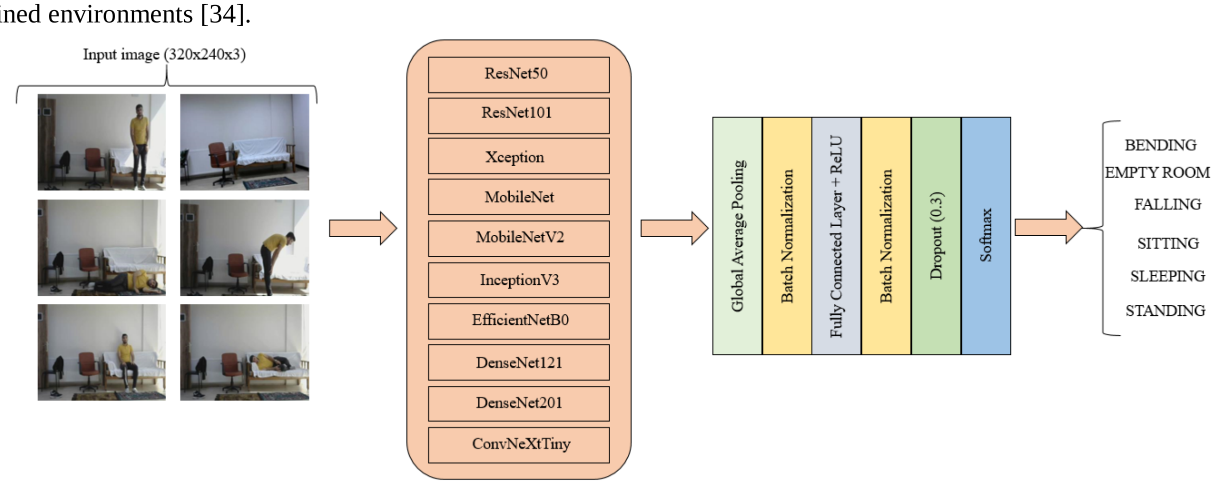


Fig. 3. The proposed method

After each pre-trained model, the output feature map is subjected to global average pooling (GAP), batch normalization, a fully connected layer with 128 neurons, batch normalization, and a dropout layer with a dropout rate of 0.3.

GAP is a technique commonly used in CNNs to reduce the spatial dimensions of feature maps before the final classification layer (softmax). Instead of using fully connected layers, GAP computes the average of all values in the feature map to obtain a single value for each feature map. The advantages of using GAP are as follows: (i) By reducing the number of parameters in the network, GAP helps to prevent overfitting, which is especially important when working with a limited amount of training data. It also acts as a regularization function. (ii) GAP is more efficient in terms of computation compared to fully connected layers, which require a large number of parameters and computation resources, particularly when dealing with high-resolution feature maps. GAP significantly reduces computational complexity. Additionally, GAP is simple to implement and does not require learning any additional parameters, which can simplify the training process. (iii) GAP requires less memory compared to fully connected layers, making it a good choice for resource-constrained applications such as mobile or embedded devices. (iv) Replacing fully connected layers with GAP layers is a common practice when fine-tuning pre-trained models, allowing the model to be adapted to a different task while preserving the knowledge learned during pre-training. When batch normalization is used in pre-trained network models, (i) it helps maintain the stability and performance of the model during fine-tuning on new tasks or datasets. (ii) It accelerates the training process by reducing the number of training iterations required for model convergence, leading to faster training. Thus, batch normalization is used to speed up and simplify the training process. The advantages of using a dropout layer in the proposed model are as follows: Dropout serves as an effective regularization technique to prevent overfitting. By randomly deactivating a portion of the neurons during training, it helps the model generalize better to new tasks or datasets. This randomness caused by dropout reduces the dependency on certain learned features, making the model more adaptable and robust during fine-tuning. As a result, dropout can improve the model's ability to generalize across different data distributions and enhance its performance in transfer learning scenarios. After the GAP, batch normalization, fully connected and dropout layers, the resulting feature map is passed to the input of the softmax classifier for classification. Softmax is used as the final layer in the neural network for classification tasks. The softmax classifier takes a vector of arbitrary real-valued scores as input and converts them into a probability distribution over multiple classes. Softmax transforms the input scores into positive values by exponentiating them and then normalizes these values by dividing them by the sum of all the exponentiated scores. This normalization ensures that the resulting values are between 0 and 1 and that their sum equals 1, representing probabilities. The class with the highest probability is then predicted as the output class.

III. EXPERIMENTAL RESULTS AND DISCUSSION

A. *Experimental Setup*

In the experimental studies, the Python programming language was utilized. All Python code was implemented on the Kaggle platform. The GPU P100 was used as the execution environment for training, validating, and testing the transfer learning models on the Kaggle platform. The hyperparameters used are as follows: a batch size of 64, a learning rate of 0.001, and a total of 50 epochs. The Adam optimization method was employed. The input image dimensions were set to 320x240x3. Sparse Categorical Crossentropy was used as the loss function during the training of the transfer learning models. Sparse Categorical Crossentropy is a highly advantageous loss function in multi-class classification tasks, particularly when target labels are represented as integers rather than one-hot encoded vectors. One of its primary benefits is reduced memory usage, as it eliminates the need to create a one-hot encoding of the target labels, which can be particularly advantageous in scenarios with a large number of classes, where one-hot encoding would require significant additional memory. This efficiency simplifies the data preprocessing pipeline, allowing for direct use of integer labels, which streamlines the workflow and reduces the potential for errors during data transformation. Furthermore, Sparse Categorical Crossentropy enhances interpretability, as it maintains the original labeling scheme without necessitating any transformation into binary formats, making it easier for practitioners to understand and debug model predictions. This loss function is also well-suited for applications with large class sets, as it effectively manages extensive label spaces without the overhead associated with one-hot encoding, ultimately leading to improved computational performance. Overall, Sparse Categorical Crossentropy offers a memory-efficient, simplified, and intuitive approach to handling multi-class classification tasks, making it a popular choice in modern deep learning frameworks. Training-testing separation was set to 70-30% in all models. In other words, a total of 17472 images were used, with 12231 images for training and 5241 images for testing.

B. *Performance Metrics of Classification*

Confusion matrix allows comparison of predicted classifications versus true classifications (Fig. 4). At the end of the training process, a confusion matrix is created with the classification results obtained from the test data. Performance evaluation is carried out using standard formulations based on this matrix. In the confusion matrix, the correctly predicted classifications of the trained network on the test data (not used in training) are defined as True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). TP is the number of inputs that are actually positive and predicted as positive, FN is the number of inputs that are actually positive but predicted as negative, TN is the number of inputs that are actually negative and predicted as negative, and FP is the number of inputs that are actually negative but predicted as positive. After creating the confusion matrix, the TP value is found as the sum of the numbers in the diagonal cells.

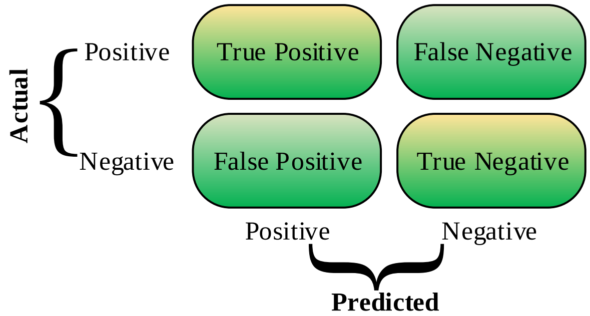


Fig. 4. Basic confusion matrix structure

TABLE II.
PERFORMANCE METRICS AND THEIR EXPRESSIONS

Performance metric	Formula
Sensivity (Recall)	SEN = TP / (TP + FN)
Precision	PRE = TP / (TP + FP)
Accuracy	ACC = (TP + TN) / (P + N)
F1 Score	F1 = 2TP / (2TP + FP + FN)

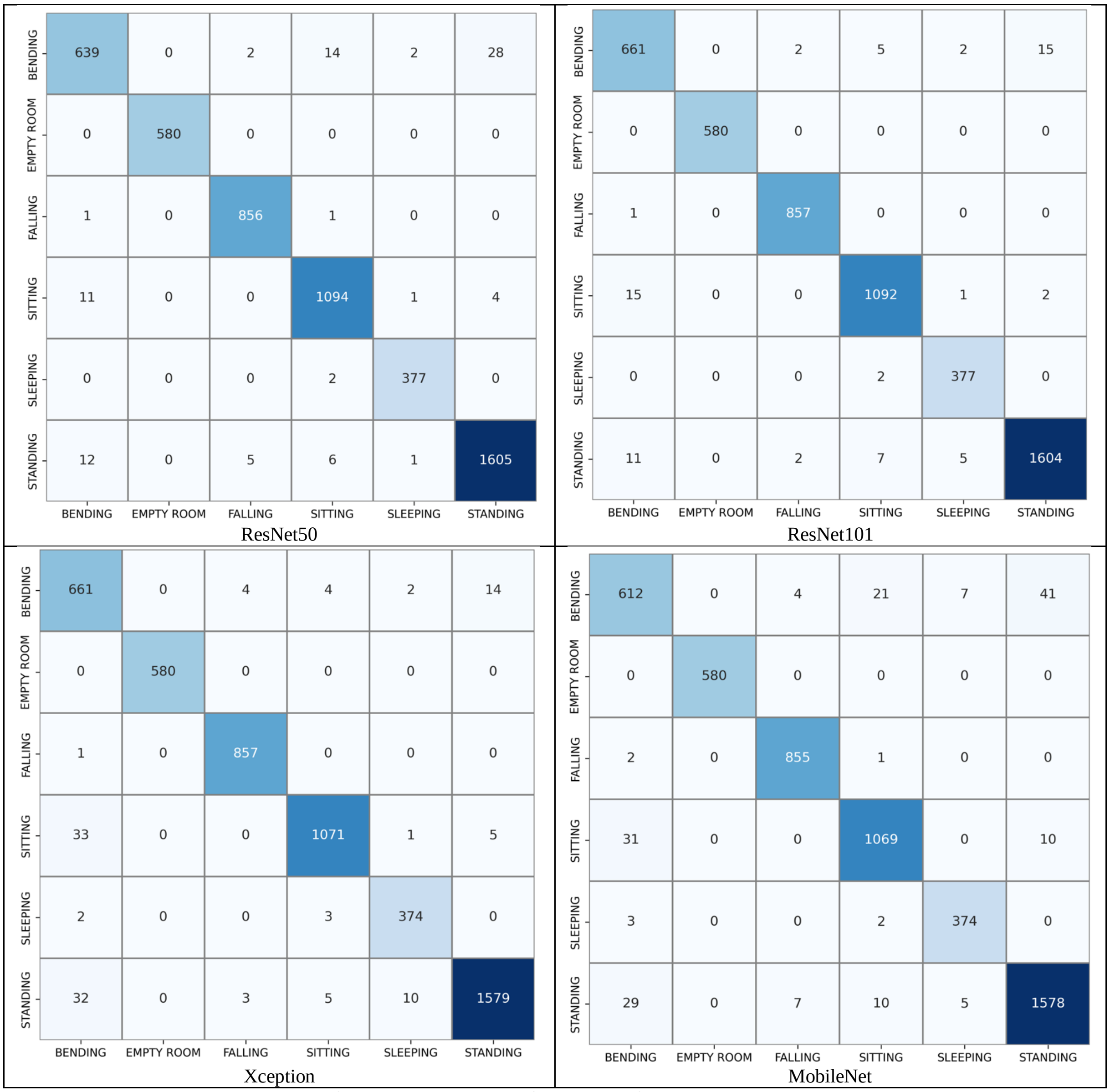
The TN value is the sum of the numbers in cells other than the row and column for that classification. The FP value is the sum of the numbers along the row, excluding the diagonal. The FN value is the sum of the numbers along the column, excluding the diagonal. Then, performance metrics such as sensitivity or recall, precision or positive predictive values, accuracy, and F1 Score can be calculated. Formulations of these metrics presented in Table 2.

C. Results of Experimental Studies

The confusion matrices obtained for all models are given in Fig.5 and Fig.6. The confusion matrix for ResNet50 shows strong classification performance, with minimal misclassifications across categories. The majority of instances are correctly classified, which is reflected in its high accuracy and balanced precision, recall, and F1-score values (Accuracy: 98.28%, F1: 98.34%). The confusion matrix for ResNet101 indicates even fewer misclassifications compared to ResNet50, suggesting that this deeper architecture is able to distinguish between classes more effectively. This is supported by its higher recall (98.78%) and accuracy (98.66%). The Xception model's confusion matrix shows slightly more misclassifications compared to ResNet models. Its lower precision (97.37%) could be due to false positives in certain classes, affecting its ability to maintain a balance between precision and recall (F1: 97.70%). The confusion matrices for MobileNet and MobileNetV2 show more noticeable errors. MobileNet's accuracy (96.70%) and MobileNetV2's lower accuracy (95.92%) suggest higher rates of misclassification. InceptionV3's confusion matrix reflects its relatively good performance, with accuracy of 97.99%. EfficientNetB0's confusion matrix is expected to show minimal misclassifications, reflecting its high accuracy (98.32%) and recall (98.60%). This model is effective at capturing relevant features while maintaining high precision. The confusion matrices for DenseNet121 and DenseNet201 are expected to be similar, with both models showing high accuracy and minimal errors. DenseNet121 performs slightly better, as indicated by its higher F1-score (98.25%) and accuracy (98.17%). ConvNeXtTiny's confusion matrix is among the best model, with few errors across all categories, reflected in its high accuracy (98.35%) and strong balance across precision (98.17%), recall (98.46%), and F1-score (98.31%). Overall, the confusion matrices indicate that deeper models like ResNet101, EfficientNetB0, and ConvNeXtTiny provide superior classification performance with fewer errors. Shallower or lighter models such as MobileNet and MobileNetV2 show more misclassifications, likely due to their reduced capacity to capture complex patterns.

In Table 3, we compare the performance of various deep learning models based on transfer learning for the classification task. The models compared include ResNet50, ResNet101, Xception, MobileNet, MobileNetV2, InceptionV3, EfficientNetB0, DenseNet121, DenseNet201, and ConvNeXtTiny. The classification performance metrics used are accuracy, precision, recall, and F1-score, as shown in Table 3. ResNet50 achieved a high classification accuracy of 98.28%, with precision and recall values of 98.42% and 98.27%, respectively. The F1-score, which balances precision and recall, is 98.34%. ResNet101 performed slightly better, with an accuracy of 98.66%, precision of 98.54%, recall of 98.78%, and an F1-score of 98.66%. The higher recall in ResNet101 indicates better detection of relevant instances compared to ResNet50, which slightly improves overall performance. The Xception model demonstrated slightly lower performance with an accuracy of 97.73%. Precision (97.37%) and recall (98.08%) values were slightly unbalanced, leading to an F1-score of 97.70%. The model still performs effectively but falls short compared to some of the other models, particularly in terms of precision. MobileNet achieved an accuracy of 96.70%, with precision and recall values of 96.63% and 96.81%, respectively, resulting in an F1-score of 96.72%. MobileNetV2, however, showed slightly lower accuracy (95.92%), with similar precision (96.12%) and recall (95.77%). The F1-score of MobileNetV2 was 95.93%, suggesting that this model has slightly lower performance compared to MobileNet. Despite their lightweight architecture, MobileNet-based models deliver decent performance but lag behind the deeper networks in this study. InceptionV3 provided a balanced performance with an accuracy of 97.99%, precision of 97.92%, and recall of 98.07%, resulting in an F1-score of 97.99%. The high recall value reflects the model's ability to correctly classify most of the relevant instances, positioning it competitively among the other models. EfficientNetB0 performed well across all metrics, achieving an accuracy of 98.32%, precision of 98.23%, recall of 98.60%, and an F1-score of 98.39%. Its high recall indicates that the model effectively captured relevant instances, while maintaining a strong balance with precision, making it one of the best-performing models. DenseNet121 and DenseNet201 performed similarly, with DenseNet121 having a slightly higher accuracy (98.17%) compared to DenseNet201 (98.13%). The F1-scores for both models are also close, with 98.25% for DenseNet121 and 98.16% for DenseNet201. DenseNet121 exhibited a marginally better balance between precision (97.97%) and recall (98.56%), indicating that it might generalize slightly better than DenseNet201.

ConvNeXtTiny achieved one of the highest performances among all models, with an accuracy of 98.35%, precision of 98.17%, recall of 98.46%, and an F1-score of 98.31%. The model's strong results across all metrics demonstrate its potential in delivering accurate and robust classifications. Overall, the results show that ResNet101, EfficientNetB0, and ConvNeXtTiny consistently outperform other models in terms of accuracy, precision, recall, and F1-score. ResNet101, in particular, achieved the highest recall, accuracy, precision and F1-score, making it the most reliable model for this task. EfficientNetB0 also stands out due to its efficient architecture and high performance across all metrics. ConvNeXtTiny, with its strong balance of metrics, proves to be a competitive model as well. Models like Xception and MobileNet series, while still effective, lag behind the top-performing models in terms of classification accuracy.



ResNet50

	BENDING	EMPTY ROOM	FALLING	SITTING	SLEEPING	STANDING
BENDING	639	0	2	14	2	28
EMPTY ROOM	0	580	0	0	0	0
FALLING	1	0	856	1	0	0
SITTING	11	0	0	1094	1	4
SLEEPING	0	0	0	2	377	0
STANDING	12	0	5	6	1	1605

ResNet101

	BENDING	EMPTY ROOM	FALLING	SITTING	SLEEPING	STANDING
BENDING	661	0	2	5	2	15
EMPTY ROOM	0	580	0	0	0	0
FALLING	1	0	857	0	0	0
SITTING	15	0	0	1092	1	2
SLEEPING	0	0	0	2	377	0
STANDING	11	0	2	7	5	1604

Xception

	BENDING	EMPTY ROOM	FALLING	SITTING	SLEEPING	STANDING
BENDING	661	0	4	4	2	14
EMPTY ROOM	0	580	0	0	0	0
FALLING	1	0	857	0	0	0
SITTING	33	0	0	1071	1	5
SLEEPING	2	0	0	3	374	0
STANDING	32	0	3	5	10	1579

MobileNet

	BENDING	EMPTY ROOM	FALLING	SITTING	SLEEPING	STANDING
BENDING	612	0	4	21	7	41
EMPTY ROOM	0	580	0	0	0	0
FALLING	2	0	855	1	0	0
SITTING	31	0	0	1069	0	10
SLEEPING	3	0	0	2	374	0
STANDING	29	0	7	10	5	1578

Fig. 5. The confusion matrices obtained for each transfer learning-based deep learning model

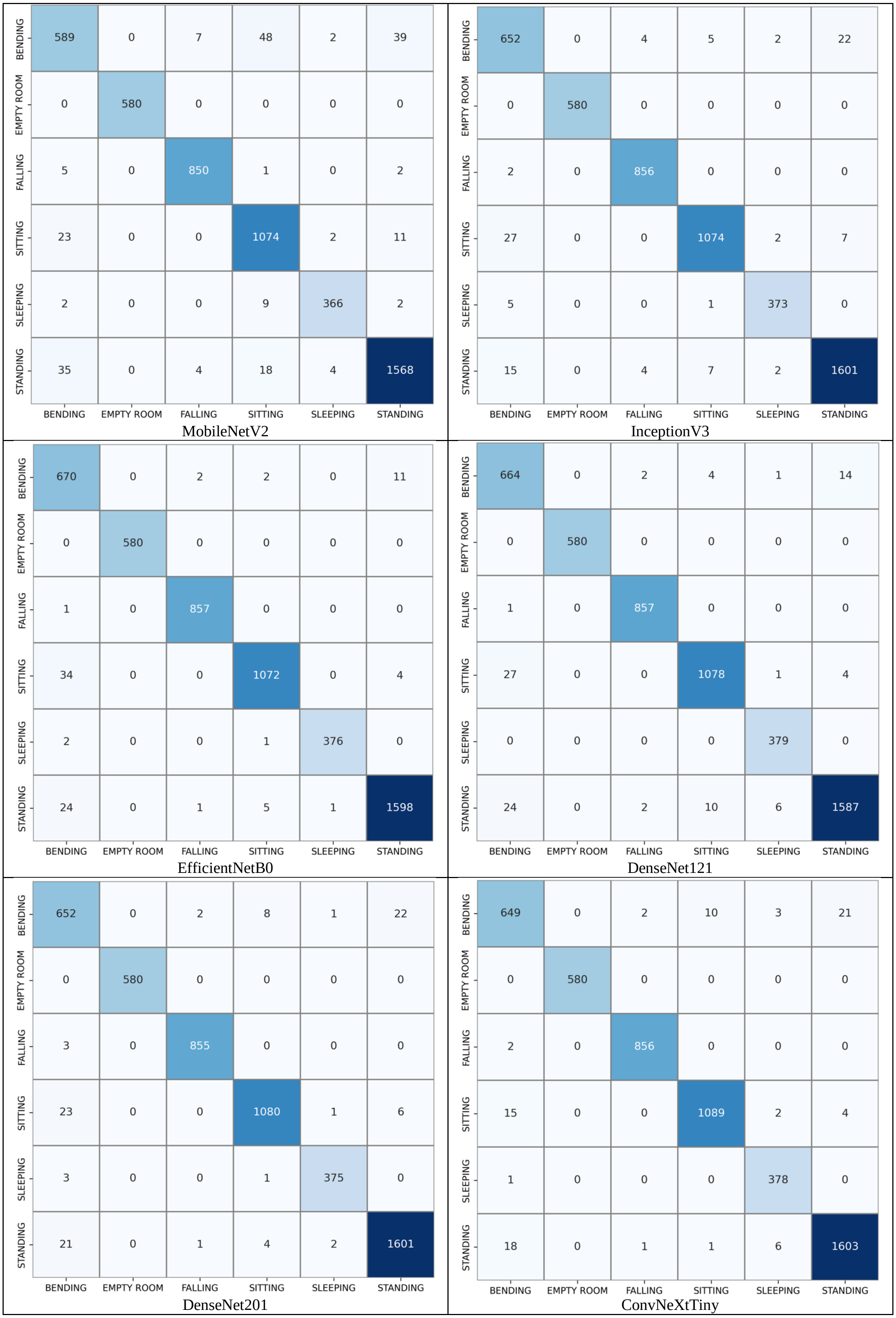


Fig. 6. The confusion matrices obtained for each transfer learning-based deep learning model (continued)

TABLE III.
CLASSIFICATION RESULTS FOR EACH MODEL (%)

Model	Accuracy	Precision	Recall	F1-score
ResNet50	98.28	98.42	98.27	98.34
ResNet101	**98.66**	**98.54**	**98.78**	**98.66**
Xception	97.73	97.37	98.08	97.70
MobileNet	96.70	96.63	96.81	96.72
MobileNetV2	95.92	96.12	95.77	95.93
InceptionV3	97.99	97.92	98.07	97.99
EfficientNetB0	98.32	98.23	98.60	98.39
DenseNet121	98.17	97.97	98.56	98.25
DenseNet201	98.13	98.09	98.23	98.16
ConvNeXtTiny	98.35	98.17	98.46	98.31

IV. Conclusion

A smart system for activity recognizing and fall detection has been developed for use as an elderly care application within the scope of smart healthcare services. This intelligent activity recognition and fall detection system is based on CNN type of deep learning. The system is aimed to make smart healthcare services more efficient through big data analytics. An original dataset is created with frame snapshots selected from videos recorded in an experimental environment. With the purpose of real-time video processing, this CNN based system used for reporting the duration of specific activities and whether a person in a room has fallen during live video capture.

The transfer learned networks are expected to yield higher performances as they have been previously trained on a very large dataset and have more advanced architectures. Ten successful CNN model used for transfer learning: ResNet50, ResNet101, Xception, InceptionV3, MobileNet, MobileNetV2, DenseNet121-DenseNet201, EfficientNetB0 and ConvNeXtTiny. The training performance metrics of the networks regarding test data are found very successful. These CNN models employed as classifiers in this system. After training process using by created dataset, proposed system is separately tested for each employed different CNN models. To evaluate the performance of the systems, accuracy, precision, recall and F1 score metrics was used. For the ResNet101, these metrics are obtained as 98.66%, 98.54%, 98.78%, 98.66% respectively, that is the best of all scores. This results show that trained ResNet101 system can be used to help elderly persons and can be integrated to the other IoT systems.

Application performance of the designed networks can be raised by training with a larger dataset. In an extended dataset that includes different features, it is suggested to add falling and semi-falling situations as additional classification categories. Additionally, to enhance practical performance, the addition of a logical error elimination process or a Long Short-Term Memory (LSTM) classifier for the flow of successive activities is recommended to distinguish between intentional sitting and unintentional falling events.

Funding and Ethics Permission Disclosure: This study was supported by the Scientific Research Projects Coordination Unit of Fırat University (FÜBAP) with project number ADEP.22.06. The study was conducted with the approval of the Ethics Committee for Social and Human Sciences Research at Fırat University, at 03/03/2023 with approval number 9.

Data availability: The dataset created in this study and experiment videos are available in the GitHub - fahricihan/FALLDETECTION: FALL DETECTION MATLAB

References

[1] S. B. Atitallah, M. Driss, W. Boulila, and H. B. Ghézala, "Leveraging Deep Learning and IoT big data analytics to support the smart cities development: Review and future directions," Comput. Sci. Rev., vol. 38, p. 100303, Nov. 2020, doi: 10.1016/j.cosrev.2020.100303.

[2] "Statista, UK: people living alone, 2019, https://www.statista.com/statistics/281616/people-living-alone-in-the-united-kingdom-uk-by- age-and-gender/ (Accessed 5 December 2023).".

[3] "World Health Organization, Falls." [Online]. Available: World Health Organization https://www.who.int/news-room/fact-sheets/detail/falls

[4] "Global Pharma News & Resources." [Online]. Available: https://www.pharmiweb.com/press-release/2020-05-05/fall-detection-system-market-to-register-cagr-4-growth-in-revenue-during-the-forecast-period-2019-t

[5] Bureau of Labor Statistics (2023) https://data.bls.gov/timeseries/FWU00X4XXXXX8EN00

[6] M. A. Khan, F. Algarni, and M. T. Quasim, Smart Cities: A Data Analytics Perspective. Springer, 2021.

[7] N. Zerrouki, F. Harrou, Y. Sun, A. Z. A. Djafer, and H. Amrane, "A Survey on Recent Advances in Fall Detection Systems Using Machine Learning Formalisms," in 2022 7th International Conference on Frontiers of Signal Processing (ICFSP), IEEE, 2022, pp. 35–39.

[8] R. E, T. Perumal, and S. K, "A review on fall detection systems in bathrooms: challenges and opportunities," Multimed. Tools Appl., Jan. 2024, doi: 10.1007/s11042-023-18088-6.

[9] Md. M. Islam et al., "Deep Learning Based Systems Developed for Fall Detection: A Review," IEEE Access, vol. 8, pp. 166117–166137, 2020, doi: 10.1109/ACCESS.2020.3021943.

[10] A. Purwar and I. Chawla, "A systematic review on fall detection systems for elderly healthcare," Multimed. Tools Appl., vol. 83, no. 14, pp. 43277–43302, Oct. 2023, doi: 10.1007/s11042-023-17190-z.

[11] X. Zhou, L.-C. Qian, P.-J. You, Z.-G. Ding, and Y.-Q. Han, "Fall detection using convolutional neural network with multi-sensor fusion," in 2018 IEEE international conference on Multimedia & Expo Workshops (ICMEW), IEEE, 2018, pp. 1–5.

[12] J. Maitre, K. Bouchard, and S. Gaboury, "Fall Detection With UWB Radars and CNN-LSTM Architecture," IEEE J. Biomed. Health Inform., vol. 25, no. 4, pp. 1273–1283, Apr. 2021, doi: 10.1109/JBHI.2020.3027967.

[13] G. L. Santos, P. T. Endo, K. H. de C. Monteiro, E. da S. Rocha, I. Silva, and T. Lynn, "Accelerometer-based human fall detection using convolutional neural networks," Sensors, vol. 19, no. 7, p. 1644, 2019.

[14] E. Torti et al., “Embedded real-time fall detection with deep learning on wearable devices,” in 2018 21st euromicro conference on digital system design (DSD), IEEE, 2018, pp. 405–412.
[15] T. R. Mauldin, M. E. Canby, V. Metsis, A. H. Ngu, and C. C. Rivera, “SmartFall: A smartwatch-based fall detection system using deep learning,” Sensors, vol. 18, no. 10, p. 3363, 2018.
[16] A. Nait Aicha, G. Englebienne, K. S. Van Schooten, M. Pijnappels, and B. Kröse, “Deep learning to predict falls in older adults based on daily-life trunk accelerometry,” Sensors, vol. 18, no. 5, p. 1654, 2018.
[17] W.-N. Lie, A. T. Le, and G.-H. Lin, “Human fall-down event detection based on 2D skeletons and deep learning approach,” in 2018 International workshop on advanced image technology (IWAIT), IEEE, 2018, pp. 1–4.
[18] A. Shojaei-Hashemi, P. Nasiopoulos, J. J. Little, and M. T. Pourazad, “Video-based human fall detection in smart homes using deep learning,” in 2018 IEEE International Symposium on Circuits and Systems (ISCAS), IEEE, 2018, pp. 1–5.
[19] N. Lu, Y. Wu, L. Feng, and J. Song, “Deep learning for fall detection: Three-dimensional CNN combined with LSTM on video kinematic data,” IEEE J. Biomed. Health Inform., vol. 23, no. 1, pp. 314–323, 2018.
[20] W. Min, H. Cui, H. Rao, Z. Li, and L. Yao, “Detection of human falls on furniture using scene analysis based on deep learning and activity characteristics,” IEEE Access, vol. 6, pp. 9324–9335, 2018.
[21] K. Adhikari, H. Bouchachia, and H. Nait-Charif, “Deep learning based fall detection using simplified human posture,” Int J Comput Syst Eng, vol. 13, no. 5, pp. 255–260, 2019.
[22] Y. Jiang, T. Gong, L. He, S. Yan, X. Wu, and J. Liu, “Fall detection on embedded platform using infrared array sensor for healthcare applications,” Neural Comput. Appl., vol. 36, no. 9, pp. 5093–5108, Mar. 2024, doi: 10.1007/s00521-023-09334-x.
[23] S. Mobsite, N. Alaoui, M. Boulmalf, and M. Ghogho, “Activity Classification and Fall Detection Using Monocular Depth and Motion Analysis,” IEEE Access, vol. 12, pp. 1525–1541, 2024, doi: 10.1109/ACCESS.2023.3348413.
[24] X. Zi, K. Chaturvedi, A. Braytee, J. Li, and M. Prasad, “Detecting Human Falls in Poor Lighting: Object Detection and Tracking Approach for Indoor Safety,” Electronics, vol. 12, no. 5, p. 1259, Mar. 2023, doi: 10.3390/electronics12051259.
[25] A. R. Khekan, H. S. Aghdasi, and P. Salehpour, “The Impact of YOLO Algorithms Within Fall Detection Application: A Review,” IEEE Access, vol. 13, pp. 6793–6809, 2025, doi: 10.1109/ACCESS.2024.3496823.
[26] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in 2009 IEEE conference on computer vision and pattern recognition, Ieee, 2009, pp. 248–255.
[27] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
[28] F. Chollet, “Xception: Deep Learning with Depthwise Separable Convolutions,” in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 1800–1807. doi: 10.1109/CVPR.2017.195.
[29] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 2818–2826.
[30] A. G. Howard et al., “MobileNets: efficient convolutional neural networks for mobile vision applications (2017),” ArXiv Prepr. ArXiv170404861, vol. 126, 2017.
[31] 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.
[32] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4700–4708.
[33] M. Tan and Q. Le, “Efficientnet: Rethinking model scaling for convolutional neural networks,” in International conference on machine learning, PMLR, 2019, pp. 6105–6114.
[34] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, “A convnet for the 2020s,” in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022, pp. 11976–11986.
[35] H. Fırat and H. Üzen, “Detection of Pneumonia Using A Hybrid Approach Consisting of MobileNetV2 and Squeeze-and-Excitation Network,” Türk Doğa Ve Fen Derg., vol. 13, no. 1, pp. 54–61, 2024.
[36] A. Bello, S.-C. Ng, and M.-F. Leung, “Skin Cancer Classification Using Fine-Tuned Transfer Learning of DENSENET-121,” Appl. Sci., vol. 14, no. 17, p. 7707, 2024.
[37] Ü. Hüseyin and H. FIRAT, “ÖZNİTELİK ENTEGRASYONUNA DAYALI ESA MİMARİSİ KULLANILARAK ENDOSKOPİK GÖRÜNTÜLERİN SINIFLANDIRILMASI,” Kahramanmaraş Sütçü İmam Üniversitesi Mühendis. Bilim. Derg., vol. 27, no. 1, pp. 121–132, 2024.
[38] A. R. KP and S. Gowrishankar, “Convnext-based mango leaf disease detection: Differentiating pathogens and pests for improved accuracy,” Int. J. Adv. Comput. Sci. Appl., vol. 14, no. 6, 2023.

BIOGRAPHIES

Mehmet İlyas Bayındır received the B.S., the M.S. and the Ph.D. degrees in electrical and electronics engineering from Fırat University, Elazığ, Turkey, in 1992, 1996, 2003 respectively. He worked in Şanlıurfa Vocational High School in Harran University as lecturer in 1993-1998 years. He has been worked in Technical Sciences Vocational High School as assistant professor since 1998. He continues his research studies at the Department of Ecoinformatics, Institute of Natural and Applied Sciences, Fırat University, Elazığ, Türkiye.

Fahri Cihan Attila was born in Elazığ in 1998. He received his bachelor's degree from Fırat University, Department of Electrical and Electronics Engineering in 2016. He received Master's degree from Ecoinformatics department in the Institute of Natural and Applied Sciences at Fırat University in 2023. In 2021, he worked as a Telecommunications engineer at Geoplus Solutions Engineering Inc in Elazığ. He works as a Senior Fixed Network Operations and Infrastructure Planning Specialist at Turkcell Communication Services. He specialized in developing and optimizing telecommunications infrastructures.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 1, March 2025

Research Article

Comparative Analysis of Analog and FPGA Realizations Based on Matsuda Method for Fractional Order Integral Operator

Omer Pektas and Murat Koseoglu

Abstract— **Realization of fractional order (FO) transfer functions is essential for real-time applications such as communication systems, video, and digital signal processing. In general, in both implementation methods, the FO transfer function including FO integral and derivative operators is transformed to an integer order approximate transfer function by one of the approximation methods such as Oustaloup, Matsuda, Continued Fraction Expansion (CFE), Modified Stability Boundary Locus (MSBL), etc. Then, the integer order approximate transfer function can be implemented using analog circuit elements such as opamps, resistors, capacitors, or digitally with field-programmable gate arrays (FPGA). In this study, integer order approximate continuous time transfer function obtained for FO integral operator by Matsuda's approximation method is converted to a discrete time function, and that function is digitally implemented by FPGA with Xilinx System Generator. The results obtained are analyzed in comparison with analog circuit implementation results presented in a former study. The study emphasizes the growing importance of fractional calculus in providing accurate models for real-world systems and the challenges posed by the long memory effect in digital implementations. Simulation and experimental results, including sinusoidal waveform, step response and impulse response analysis, reveal the pros and cons of FPGA implementation. Considering these issues, conclusions are made on the effectiveness, efficiency and potential of the FPGA implementation for real-time applications in control systems and signal processing.**

Index Terms—**Fractional Order (FO) Integral, Matsuda's Method, Field Programmable Gate Array (FPGA), Digital Circuit Implementation.**

I. Introduction

THERE HAS been a growing interest among engineers in fractional calculus due to its potential for providing more accurate representations of real-world systems. Over the past two decades, Fractional Order (FO) modeling techniques have been applied to a wide range of problems in various scientific and engineering disciplines. Researchers have implemented FO system models with the accuracy required for their specific applications as well as possible. In the field of control systems, approximate implementation methods have enabled the simulation and experimental realization of FO controllers. However, a significant challenge in the digital implementation of near-ideal FO elements is the long memory effect, which arises from the non-local nature of FO integrals and derivatives. FO operators inherently involve a backward memory, requiring consideration of all past values of a function when processing current values. This leads to a continuously increasing computational load in terms of memory and processing power when handling continuous data streams. Even for non-ideal experimental realizations of fractional elements, achieving real-time computation may necessitate high-speed digital hardware, such as FPGAs [1].

Analog realizations of FO systems also have some advantages. There are two main approaches for the analog implementation of FO systems: circuit-based implementation and implementation using Field-Programmable Analog Arrays (FPAA). The study conducted by A. Hassanein and his colleagues examines the design process of FO analog filters, focusing on both circuit-based and FPAA methods. The study demonstrates how classical integer-order filters can be generalized to FO elements. Simulation and experimental results were consistent with theoretical analysis, showing that FO systems offer greater flexibility and design freedom [2]. In another study conducted by Nako and colleagues, the design of generalized 1+α order Butterworth filters and their implementation using FPAA were presented. The FO filters were generalized from integer-order filters and made programmable in various types, such as low-pass, high-pass, band-pass, and band-stop filters. The study developed a method using MATLAB to model the filter transfer functions as integer-order and prepared these transfer functions for implementation on an FPAA device. Results showed that the proposed method allowed the filters to be easily implemented, with simulation and experimental findings aligning with theoretical expectations [3]. In addition, another study conducted by M.A. George and colleagues presented the design of complex-order (complex-order) PI/PID speed controllers and their implementation using FPAA. Due to the challenges of implementing complex-order controllers in digital environments, these controllers were made practically

Ömer Pektaş, is with Department of Vocational School of Technical Sciences, Electric and Energy, University of Karamanoglu Mehmetbey University, Karaman, Turkey,(e-mail: omerpektas@kmu.edu.tr).
https://orcid.org/0000-0002-6927-6529

Murat Köseoğlu, is with Department of Electrical & Electronics Engineering University of Inonu, Malatya, Turkey, (e-mail: murat.koseoglu@inonu.edu.tr).
https://orcid.org/0000-0003-3774-1083

Manuscript received Oct 01, 2024; accepted Dec 28, 2024.
DOI: 10.17694/bajece.1577733

implementable using FPAA. Simulation and experimental results demonstrated that complex-order controllers exhibited better stability and speed-tracking performance than traditional technologies [4]. As noted, despite the analog implementations having certain advantages, FO systems, which were transformed to high integer order systems, are difficult to implement due to their complex structures. These systems can be digitally realized using FPGAs, which provide a more efficient, faster, and cost-effective platform for digital system implementation. The study by A. Ali and colleagues is a comprehensive review of this subject. It examines the implementation of FO differentiators and integrators, exploring different digital approaches and software tools for digital system applications. Future research directions are discussed, addressing open problems and proposing potential solutions [5]. Considering the advantages of digital systems and the drawbacks of analog implementations, this study emphasizes the preference for FPGA-based digital realizations for fractional-order operators.

Various computational platforms can be utilized to design or calculate FO systems. For many years, digital signal processors (DSPs) and microprocessors have been widely used in low-speed applications where power and efficiency are not critical concerns. However, due to technological advancements in recent years, FPGAs have gained popularity as platforms for digital signal-processing applications. As known, FPGAs are integrated circuits composed of tens of thousands of programmable logic cells interconnected by programmable switches and wires. The most significant advantage of FPGAs over DSPs and microprocessors is their parallel architecture and versatility. Since an FPGA can be adapted to execute computations in maximum parallelism, it outperforms microprocessors and DSPs, which must execute computations serially [6].

FPGAs offer inherent advantages in terms of performance, flexibility, and efficiency. FPGAs enable parallel processing, which allows for high-speed data handling and low-latency operation, essential for real-time applications such as communication systems, video processing, and digital signal processing. The reconfigurability of FPGAs allows designers to customize and optimize filters for specific applications, ensuring better performance compared to fixed-function hardware [7]. Additionally, the deterministic nature of FPGA operations provides consistent and reliable performance, which is critical for applications requiring precise timing. Overall, implementing filters on FPGAs combines the benefits of high performance, adaptability, and efficiency, making them an ideal choice for complex and demanding signal processing tasks [8].

In a former study [1], the authors present a low-cost analog circuit realization of Matsuda's approximate FO integral operators for industrial electronics using operational amplifiers (op-amps), resistors, and capacitors. The proposed method utilizes Matsuda's approximation and partial fraction expansion to decompose FO integral models into low-pass filter forms, facilitating their implementation with standard electronic components. The designed circuit's performance is validated through experimental comparisons with exact analytical solutions and alternative FO circuit designs, demonstrating satisfactory accuracy in time and frequency responses. The results highlight the practicality and efficiency of the analog realization for FO systems, emphasizing its potential for industrial applications where cost and simplicity are critical. However, in this former study, it was seen that it is difficult to readjust a capacitor or resistor value when a parameter or the degree changed in the transfer function. The increase in the degree of the transfer function may increase the complexity and the sensitivity of the analog circuit which is composed of different analog components having different tolerance values.

Considering the encountered difficulties in analog realization, this study aims to examine the advantages and disadvantages of the FPGA digital realization compared to the analog realization. In this study, an integer 4^{th} order approximate continuous time transfer function, which was obtained by the transform of FO integral operator $s^{-0.5}$ with Matsuda's approximation method in $\omega \in [0.1\ 10] rad/sec$, was considered. This function is decomposed by the partial fraction expansion method and expressed as the sum of first order low-pass filter forms in s domain [1], [9]. The sum of these filter functions can be digitally implemented with FPGA after the conversion of continuous time function to discrete time function. The obtained function is converted to a discrete time function using MATLAB's 'c2d' function for the implementation of the function in the FPGA. The FPGA realization results, obtained with the aid of Xilinx System Generator, are compared with analog realization results obtained in the former study [1]. Simulation results, including the responses for square waveform, sinusoidal waveform, 100 s step waveform and 1 s impulse waveform, demonstrate the effectiveness and efficiency of the FPGA implementation, highlighting its potential for real-time applications in control systems and signal processing. The FPGA digital realization results obtained via Xilinx System Generator were analyzed in comparison with the analog application results obtained using conventional electronic components. Simulations and experimental outcomes, encompassing sine and step response analyses, exhibit the conformity and efficiency of the FPGA implementation. These findings underscore the practicality of FPGAs for real-time control systems and signal processing applications.

II. MATERIALS AND METHODS

In the study, FO integral operator $s^{-0.5}$ was realized digitally by FPGA. Since an FO operator or an FO transfer function can not be realized directly, the FO integral operator $s^{-0.5}$ was transformed to an integer (4^{th}) order approximate transfer function by Matsuda's approximation method, which is one of the frequently used approximation methods along with Oustaloup, Continued Fraction Expansion (CFE), Modified Stability Boundary Locus (MSBL) and El-Khazali methods [1]. Matsuda's approximation technique is fundamentally grounded in the CFE method and seeks to align an integer-order approximate model with FO derivative operators. This alignment is specifically targeted at sampled frequencies that are distributed evenly on a logarithmic scale. In other words, Matsuda's approximation method was proposed as an improvement of the CFE approximation method because the CFE method does not allow adjustment of operating frequency ranges, where the transfer function model of the CFE method behaves as a FO operator. This frequency adjustability improvement provides versatility advantages in the practical realization of FO elements for the real-world applications [1], [10]. In this study, to reduce the complexity of the realization process, the FO integral operator was transformed to a 4th integer order approximate transfer function rather than a higher order approximate transfer function by Matsuda approximation method as follows [1]:

$$T(s) = \frac{0.1132s^4+3.439s^3+5.853s^2+1.068s+0.01778}{s^4+6.007s^3+3.291s^2+0.1934s+0.0006366} \quad (1)$$

Then, by using partial fraction expansion method, Equation (1) can be expressed as [1]:

$$T(s) = \frac{2.0010}{s+5.4049} + \frac{0.4943}{s+0.5360} + \frac{0.1806}{s+0.0628} + \frac{0.0828}{s+0.0035} + 0.1132 \quad (2)$$

This transfer function was obtained by using Matsuda approximation method for the 4 th order realization of fractional order integral operator $s^{-0.5}$ in the interval of $\omega \in [0.1,10)$ rad/s. The reason why this transfer function was used is that Multisim analog responses were obtained by using this transfer function in the author's former study [1]. To make a proper comparison of analog and digital (FPGA) responses, the same transfer function was considered.

A. Digital Implementation

To implement the continuous time function with FPGA in digital form, it is needed to convert the s-domain continuous time transfer function to z-domain discrete time transfer function. For this purpose, there are many conversion methods such as zero-order hold (ZOH), first-order hold (FOH), impulse-invariant mapping, Tustin approximation (bilinear transform), zero-pole matching equivalents, least squares. Each of these methods has its own strengths and weaknesses, and the choice of method depends on the specific requirements of the system being modeled and the nature of the input signals. All the mentioned methods have been tested in MATLAB, and it has been found for this study that Tustin method with the sample time of 0.01 s gives relatively better results for the handled transfer function when error rates are considered in a defined frequency range [11], [12], [13], [14].

Using the Tustin method with a small sampling time like Ts=0.01 yields better results because it minimizes frequency warping, accurately captures high-frequency behavior, reduces discretization error, and ensures the stability and dynamic characteristics of the continuous system are well-preserved in the discrete domain. Also, any different Ts value could be chosen. Ts value of 0.01 s yields a sufficiently proper discrete time signal. On the other hand, in a former study where analog implementation was realized, the frequency range was taken in the interval of [0.1, 10) rad/s to approximate the $s^{-0.5}$ integral operator by Matsuda method. This operating frequency range was selected hypothetically, considering any frequency range that may be encountered in systems in real-life applications, another frequency range could also have been chosen.

When the continuous time transfer function given in Equation (2), which expresses the main transfer function as the sum of first order low-pass filter forms, is converted to a discrete time transfer function by using Tustin method in MATLAB, the following function is obtained in z-domain:

$$\begin{aligned} T_m(z) = & \frac{0.009742 + 0.009742\, z^{-1}}{1 - 0.9474\, z^{-1}} \\ & + \frac{0.002464 + 0.002465\, z^{-1}}{1 - 0.9947\, z^{-1}} \\ & + \frac{0.0009027 + 0.0009027\, z^{-1}}{1 - 0.9994 z^{-1}} \\ & + \frac{0.000414 + 0.000414\, z^{-1}}{1 - z^{-1}} \\ & + 0.1132 \end{aligned} \quad (3)$$

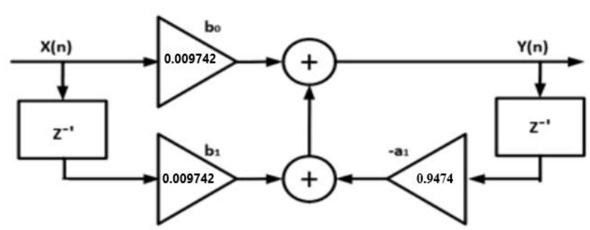
Each rational part of Equation (3) can be represented as a difference equation as shown in the following example, in which the first rational component of Equation (3) was written as a difference equation:

$$\frac{Y(z)}{X(z)} = \frac{0.009742 + 0.009742\, z^{-1}}{1 - 0.947\, z^{-1}} \quad (4)$$

$$X(z) * (0.009742 + 0.009742\, z^{-1}) = Y(z) * (1 - 0.947\, z^{-1}), \quad (5)$$

$$Y(n) = 0.009742 * X(n) + 0.009742 * X(n-1) + 0.947 * Y(n-1). \quad (6)$$

Since each rational term depends on the present and past inputs as well as the past outputs, the transfer function is realized as the sum of Infinite Impulse Response (IIR) filters. IIR filter is a digital filter with an impulse response that theoretically continues indefinitely. IIR filters are efficient regarding computational resources but can be less stable and introduce phase distortion if not designed carefully. Direct Form I structure was used to realize each first-order IIR filter as shown below, and the design example for the first rational part of the right side of Equation (3) is shown in Fig. 1.



$$\frac{Y(z)}{X(z)} = \frac{b_0 + b_1 z^{-1}}{1 + a_1 z^{-1}} => \frac{0.009742 + 0.009742 z^{-1}}{1 - 0.9474 z^{-1}}$$

Fig.1. Direct Form 1 Structure for a First Order IIR Filter

Equation (3) expresses the sum of four separate IIR filter equations and a static gain of 0.1132. The frequency response obtained for discrete time transfer function given by Equation (3), approximate continuous time transfer function given by Equation (1) and exact output are given in Fig. 2, which shows both the magnitude and phase responses in the operating frequency range of [0.1,10] rad/s. The frequency response of the discrete-time transfer function $T_m(z)$ and continuous time transfer function T(s) obtained by Matsuda approximation method are almost identical to each other, and they are very close to exact response in the defined frequency range.

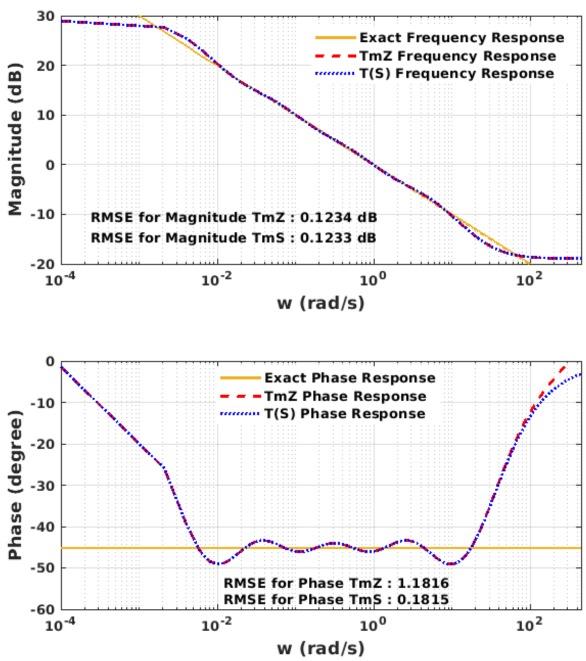


Fig.2. Bode Diagram of the Discrete Time Transfer Function

Then, Equation (3) is implemented as the sum of IIR filters on Simulink using the Xilinx System Generator tool. The Xilinx System Generator generates the IP for the function. The generated IP is used on Vivado IDE to synthesize, get Register Transfer Level (RTL) analysis, and implement it on FPGA, as seen with the block diagram in Fig. 3.

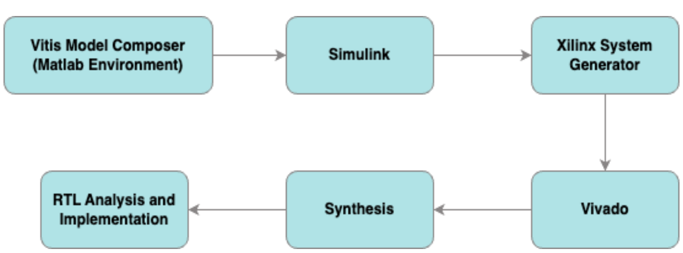


Fig.3. Simulink Block diagram based on Xilinx System Generator Tool

Xilinx System Generator design tool is provided by Xilinx, one of the leading companies producing FPGA, for implementing FPGA applications using a model-based approach in MATLAB Simulink [15]. The working principle of this tool can be explained as follows: Xilinx System Generator includes a DSP block set specific to Xilinx within Simulink. This block set allows for creating a model based on design specifications. Xilinx System Generator then uses the model designed using these block sets to generate a Register Transfer Level (RTL) netlist with the help of Xilinx components. The netlist represents the implementation or the connections of a specific logic design, which can be depicted through diagrams as a visual representation. The RTL netlist is transferred to the Xilinx Vivado design suite to generate the bitstream and application. Hardware Co-Simulation can be performed after the bitstream file is interfaced with the Xilinx System Generator and downloaded onto the FPGA device [16], [17], [18]. In this study, subsystems have been created for filters that are part of the transfer function. Pre-built add blocks were utilized to combine these systems or filters. Gateway blocks serve as interfaces between Simulink blocks and FPGA blocks [19].

III. RESULTS AND DISCUSSIONS

For FPGA simulation, IIR filters are designed using Xilinx System Generator. The discrete time transfer function consisting of a static gain block and four IIR filter blocks is implemented in two separate ways, shown in parallel with each other in Fig. 4, as FPGA (IIR multiple subsystem) and Simulink discrete filter subsystem, for comparison. The outputs of these subsystems and the input signals are compared by displaying them with a time scope.

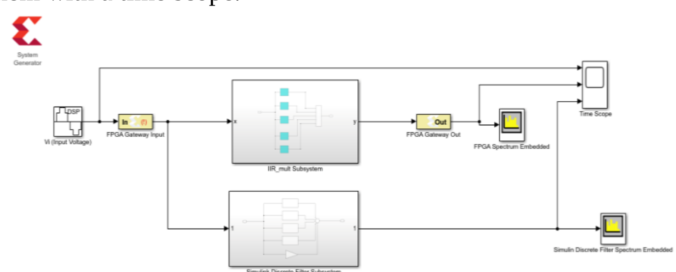


Fig.4. Implementation and comparison of filter in Xilinx System Generator at Vitis Model Composer

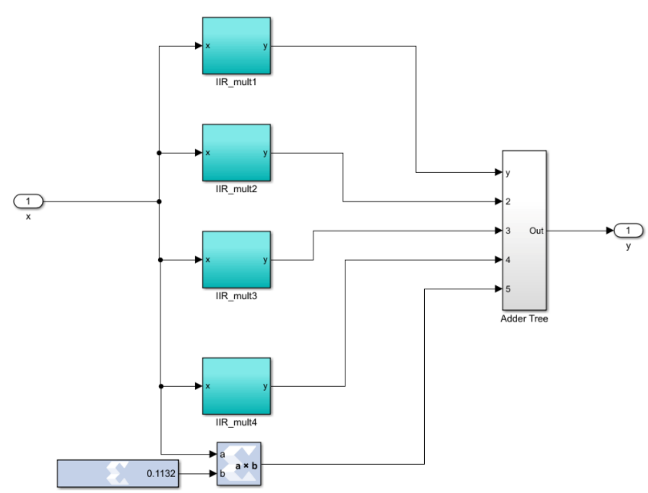


Fig.5. Implementation of IIR_mult_Subsystem Block

Each first-order IIR filter was represented by a subsystem which was called IIR_mult (see in Fig. 5). These subsystems were combined using Adder Tree from the Xilinx DSP library to obtain the output of the implemented function.

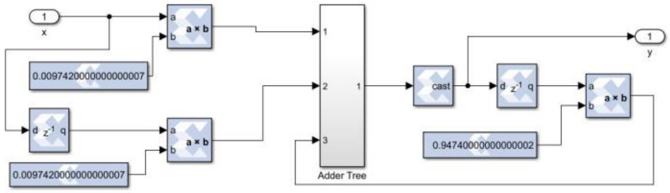


Fig.6. Implementation of IIR_mult1 Block

In these IIR_mult multiplexer subsystems seen in Fig. 6, the Direct Form I implementation of the IIR filter, which is the first rational component of $T_m(z)$, is achieved using three multipliers, two registers, one convert block and an adder tree. Although full resolution is maintained through the multipliers and the adder tree, the data path cannot expand indefinitely. Therefore, a quantization block has been placed at the output of the adder tree to reduce data width back to its input width [15].

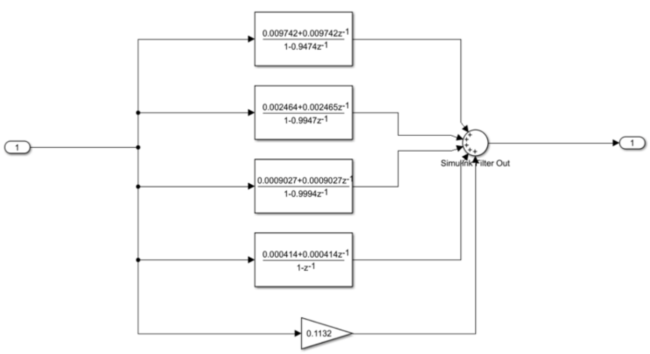


Fig.7. Design of Simulink Discrete Filter Subsystem

On the other hand, a Simulink Discrete Filter subsystem is implemented to simulate the filter function. This subsystem contains four Discrete Filter blocks and a static gain block as seen in Fig. 7, similar to IIR_mult subsystems. The outputs of the blocks are combined to obtain the final output of the filter.

Then, different waveforms were applied to the input of the system to check the accuracy, consistency and practicability for FPGA implementation. At first, sine waveform of 0.2 V and 0.5 Hz was applied as input. The obtained output waveforms and the Root Mean Square Error (RMSE) for Multisim analog and FPGA responses were presented in Fig. 8. This figure indicates that the model, when implemented in simulation, performs consistently, suggesting accurate system behavior for sine input. The relatively low RMSE values for both analog and digital realization outputs suggest that the FPGA and Multisim implementations are highly accurate and closely follow the analytical response in defined frequency range. The relatively higher error in the FPGA implementation can be attributed to the discrete time nature of the system, which includes quantization and discretization effects. However, these small errors are omittable for practical applications. On the other hand, the coherence seen at the output signals reveals negligible phase or amplitude errors between FPGA, Multisim, Simulink and analytical outputs. Thus, one can conclude that FPGA can reliably emulate the theoretical characteristic of a transfer function in the real world in a defined frequency range. This comparison shows the effectiveness of the FPGA design and its suitability for real-time signal processing applications.

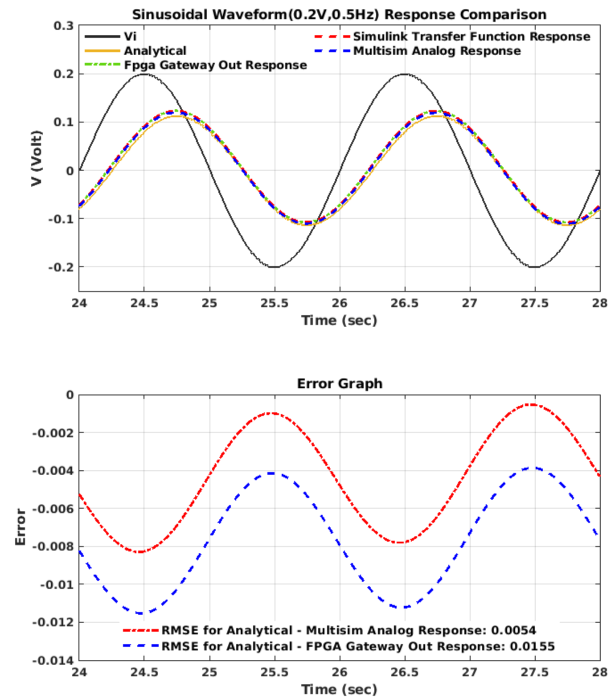


Fig.8. Sinusoidal Waveform Responses for Discrete Time Filter and Its Digital and Analog Circuit Realization and the Error Graph

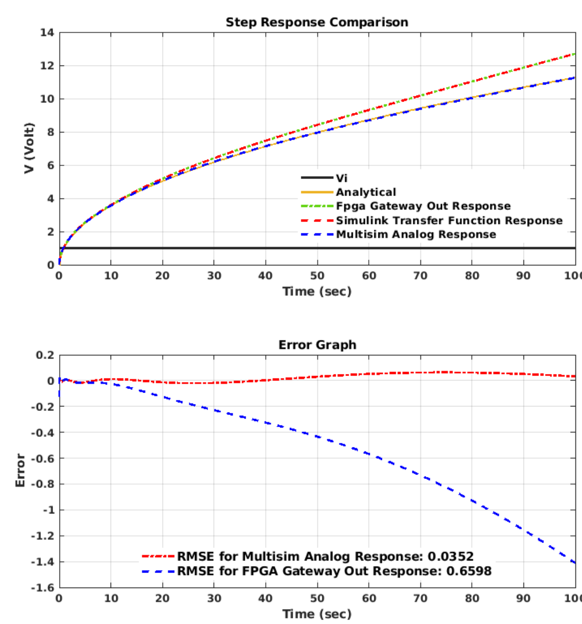


Fig.9. Step Responses for Discrete Time Filter and Its Digital and Analog Realization and the Error Graph

Fig. 9 shows 100 seconds step response and error graph for FPGA gateway out, Simulink transfer function, analytical response and Multisim analog output. The error graph was obtained by calculating the difference between analytical and realization responses. Also, calculated RMSE values for both Multisim Analog response and FPGA gateway out were added to the graph. The FPGA and Simulink responses closely follow the Multisim analog circuit response and analytical response for the first 30-35 seconds. After 35 seconds a deviation between FPGA response and analytical or Multisim analog response is observed. This deviation may arise from a variety of factors resulting from the nature of digital implementation besides the characteristics of the system being simulated. This case can be explained by numerical precision and quantization errors, sampling and discretization issues, accumulated errors from numerical computations and the closeness of the poles of the discrete-time transfer function to the unit circle. More comprehensively, the differentiations observed especially in time domain responses can be explained as follows: The divergence over time can be attributed to fundamental differences in how analog and digital systems process signals. Analog systems operate on continuous signals with theoretically infinite precision, while digital systems rely on sampled and quantized representations, introducing small errors that accumulate over time. In digital systems, numerical integration, frequency warping during discretization, and finite precision arithmetic can lead to inaccuracies, especially in long-duration responses. Feedback delays and numerical limitations in digital computations further contribute to the observed discrepancies. These factors together cause more deviations in digital response compared to the analog responses as time progresses. Nonetheless, the FPGA may be used particularly for real-time applications requiring flexibility and high-speed processing.

Fig. 10 shows continuous time analytical response, discrete time response obtained from FPGA Gateway out and Simulink for 1 second pulse input voltage. Also, the calculated RMSE value for FPGA Gateway out was added to the graph. As seen, the responses obtained for MATLAB Xilinx system generator realization and Simulink discrete model agree with each other. The low RMSE value confirms the consistency, the reliability and the effectiveness of the FPGA implementation in accurately modeling the dynamic behavior of the system for a 1-second pulse input.

Fig. 11 shows the 1 V square wave responses for the FPGA, Simulink Transfer Function, and Multisim Analog Circuit. A remarkable consistency is observed between the FPGA, Simulink, and Multisim responses, indicating that digital simulations can be as accurate as analog implementations.

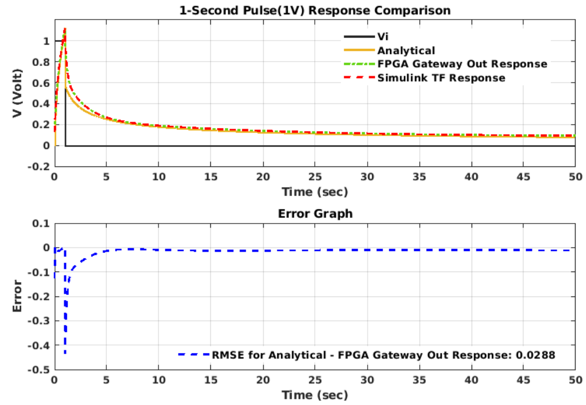


Fig.10. Responses of the Discrete Time Function Obtained for 1-Second Pulse (1 V) and Analytical-FPGA Error Graph

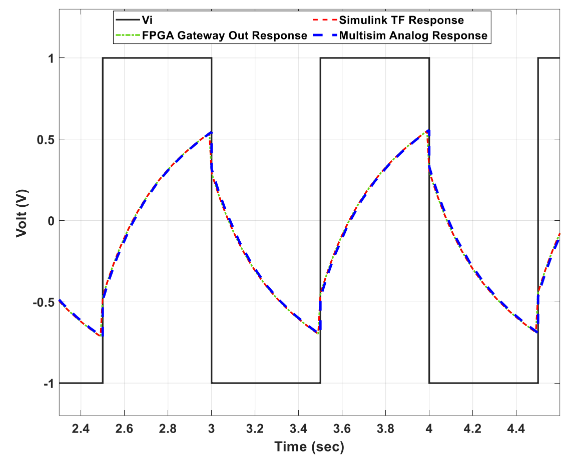


Fig.11. Square Wave Response for FPGA, Simulink Transfer Function and Multisim Analog Circuit

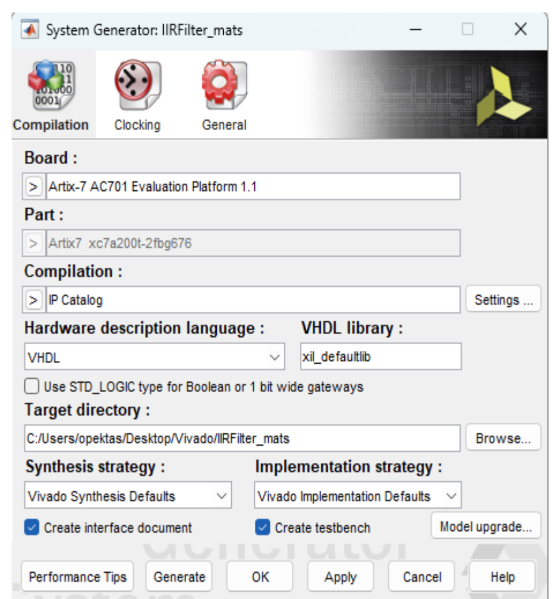


Fig.12. Xilinx System Generator Settings

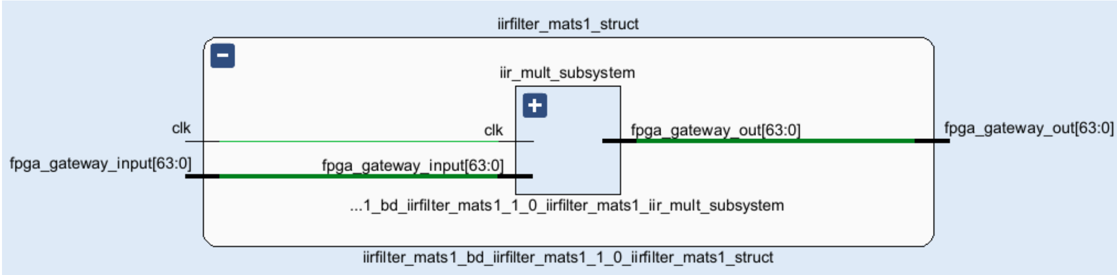


Fig.13. Filter's Elaborated Design Structure obtained with RTL Analysis at Vivado

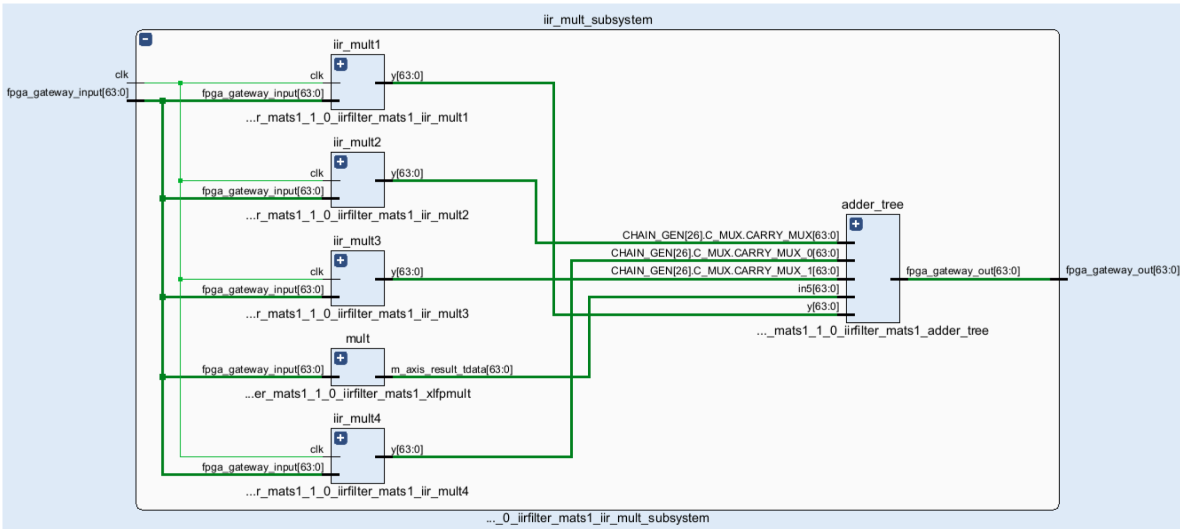


Fig.14. IIR_mult_subsystem Structure at Filter's Elaborated Design

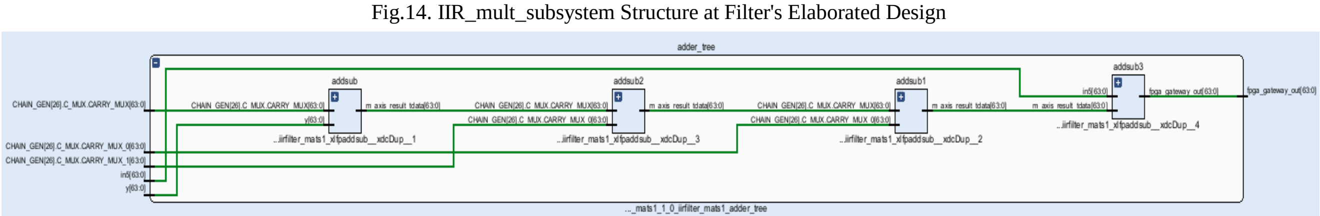


Fig.15. The Adder Tree Structure at the Design of IIR_mult_subsystem

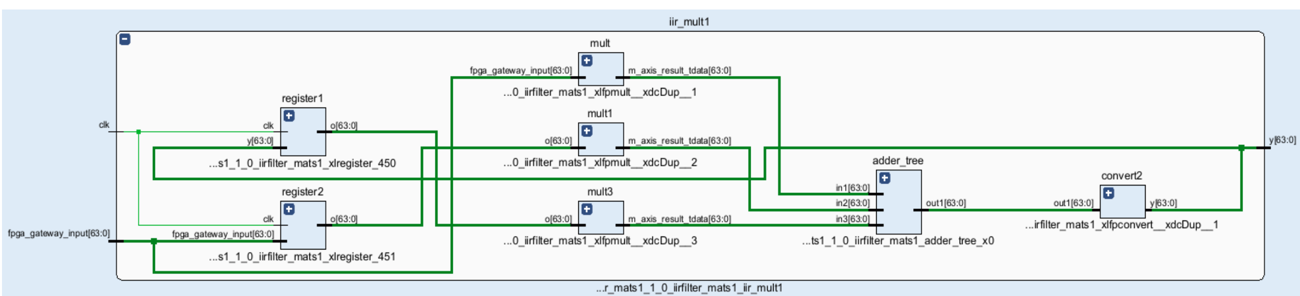


Fig.16. IIR_mult1 multiplexer Structure at the Design

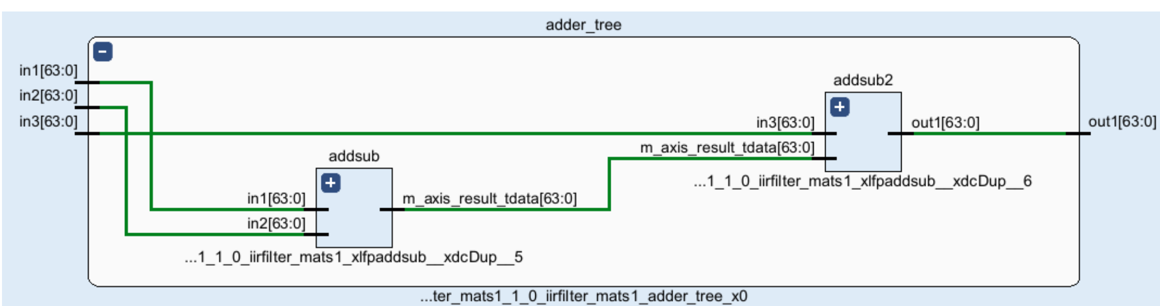


Fig.17. The Adder Tree Structure at the Design of IIR_mult1

After obtaining the simulation result graph, we generate the filter design and obtain power data using the Xilinx System Generator block shown in Fig. 12. Artix-7 AC701 board is used to synthesis and RTL analysis. Since the synthesis will be performed via Vivado, we select the default Vivado Synthesis option in the Synthesis Strategy section. Then, we choose the target directory for saving the output and click the generate button to complete the system generation.

To open the generated system, launch the Vivado application and select the .xpr file from the saved location via the Open Project option. Once the project is open, access the Elaborated Design to view the detailed design of the filter. This will display the design shown in the image below. The iirfilter_mats_bd block in the image represents our filter design, which includes the clk and FPGA_gateway_input inputs and the FPGA_gateway_output output. By double-clicking on this block, we can access the subsystems that comprise it. The iir_mult_subsystem in the design is similar to our design in MATLAB Simulink. As shown in the image, each iir_multx system consists of three multiplexers, three registers, and one adder tree, just like the MATLAB Simulink design. Since our filter is of fourth order, a total of four multiplexers were used, and their results were combined using adders to obtain the filter output (see in Fig. 13,14,15,16 and 17).

Fig. 13, 14, 15, 16 and 17 show the Elaborated design created on Vivado as a result of the RTL analysis of the filters. As can be seen, a design similar to the filter design we designed on MATLAB Simulink has been created. In this design, the filter was designed using the collectors and multipliers in Vivado. Thus, the filter was synthesized, and the necessary working data was obtained by performing RTL analysis.

TABLE I
SYNTHESIS REPORT RESULTS OF THE FILTER AT VIVADO

Resource	Utilization	Available	Utilization
LUT	11589	134600	8.61%
FF	1348	269200	0.50%
DSP	165	740	22.30%
IO	129	400	32.25%

Table 1 shows the synthesis report results which are obtained on Vivado. The utilization of LUTs is quite low, indicating that the FPGA still has a significant number of available resources for further implementation or additional functionality. The 8.61% utilization shows that the design is efficient in terms of LUT usage. Flip-Flop (FF) utilization is minimal, similar to LUTs, with only 0.5% of the available FFs being used. This suggests that the design does not heavily rely on sequential logic elements, leaving ample room for further expansion. DSP utilization is higher than LUTs and FFs but still reasonable at 22.3%. This level of usage indicates that the design includes some memory-intensive components but remains within a comfortable range for resource availability. The IO utilization is relatively higher at 32.25%, which is expected for designs requiring substantial external interfacing. However, it remains within a manageable range, suggesting that the design is efficient regarding IO resource allocation. Global buffers (BUFG) are minimal, indicating that the design's clock distribution network is not heavily taxing the available resources. This is a good sign of efficient clock management.

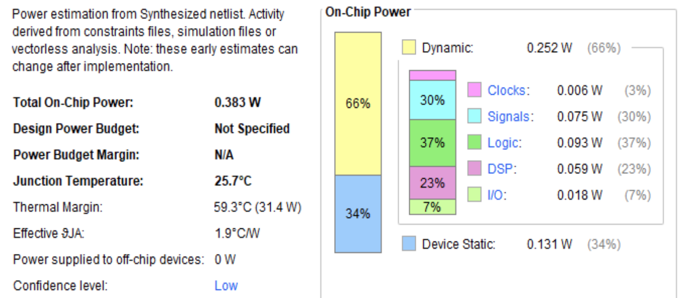


Fig.18. Power Analysis of Filter

Figure 18 shows the Power analysis obtained by implementing the filter to FPGA. As is seen, the total power consumption of the FPGA is moderate at 0.383 W, which suggests an efficient design with manageable thermal characteristics. The junction temperature is well within safe operating limits, with a substantial thermal margin indicating that the FPGA runs cool and efficiently. On the other hand, the dynamic power analysis shows the usage of clocks, signals, logic, DSP, and input-output(I/O) percentages on-chip power. The clock distribution is minimal, indicating efficient clock gating and distribution mechanisms. The most considerable portion is consumed by signal switching, which is expected in designs with significant data movement and processing. The logic power consumption is moderate, reflecting the computational activities within the FPGA. The DSP blocks consume substantial power, indicating the design's reliance on digital signal processing capabilities. Lastly, the I/O power consumption is relatively low, which is advantageous for minimizing overall power consumption. Also, the static power analysis highlights the importance of leakage power in the overall power budget. This is typical for modern FPGAs.

This study shows that FPGAs offer significant advantages for implementing FO systems in digital signal processing, particularly in performance, flexibility, and energy efficiency. However, these benefits come with trade-offs in design complexity, development cost, and resource constraints. As FPGA technology continues to evolve, some of these challenges will likely be mitigated, further enhancing the suitability of FPGAs for fractional-order DSP applications [20], [21]. In summary, FPGA implementation shows efficient resource utilization and power consumption. The low LUT and FF usage percentages indicate that the FPGA has ample room for additional logic or future expansion. Power consumption is low, with dynamic power dominating due to active signal processing and logic operations. The design runs cool with a significant thermal margin, suggesting good thermal management. Overall, the FPGA design is resource-efficient and power-efficient, with room for further optimization and expansion.

IV. CONCLUSION

In this study, 4^{th} integer order approximate continuous time transfer function was obtained for FO integral operator $s^{-0.5}$ in the operating frequency range of [0.1,10] rad/s by Matsuda's approximation method. Then this function was converted to a discrete time function with Tustin method, and that function was digitally implemented by FPGA with Xilinx System Generator. The results obtained were analyzed in comparison with analog circuit implementation results presented in a former study [1]. The results obtained from FPGA design, executed with Xilinx System Generator, have demonstrated a superior coherence with the analog realization results and exact results when frequency response is considered.

For the sinusoidal waveform, the FPGA system exhibited a slightly higher RMSE value (0.0155) compared to that of Multisim analog response (0.0054), highlighting minor discrepancies due to digital quantization. For the 1-second pulse response, the FPGA system achieved an RMSE of 0.0288, demonstrating a strong agreement with the analytical model. Also, an excellent agreement between the Multisim analog response and FPGA Gateway out for the square wave input was observed. When the responses of FPGA realization, analog realization and exact analysis for sinusoidal input, square wave input and 1-second pulse input were compared and calculated RMSE values were considered, it was seen that the FPGA results mostly yielded a good agreement with exact and analog results, demonstrating a considerable accuracy and reliability in replicating the dynamic behavior of fractional-order systems especially in the predefined frequency ranges.

When a 100 second step time response was considered, a time dependent increasing deviation was observed especially after 35 seconds between analytical and FPGA Gateway out responses, resulting in the RMSE of 0.6598 for FPGA Gateway out and RMSE of 0.0352 for Multisim analog output. This deviation is supposed to be resulted from the discrete time nature of digital systems, which causes numerical precision and quantization errors, sampling and discretization issues, accumulated errors from numerical computations. This issue is predicted to be substantially solved by increasing precision in the FPGA realization, monitoring and mitigating accumulated numerical errors, validating sampling time and stability.

Compared to the analog realization, the FPGA design demonstrated significant advantages, including reconfigurability, low power consumption (0.383 W), and efficient resource utilization, with LUT usage at 8.61% and DSP utilization at 22.3%. These findings underscore the practicality of FPGA implementations for real-time applications, offering high-speed processing, adaptability, and precision while overcoming the limitations of analog circuits, such as inflexibility and sensitivity to component tolerances. However, it should be noticed that analog implementation, which includes many design and configuration difficulties, has given relatively better time response. This comparative analysis highlights the suitability of FPGA-based implementations for FO systems in control and signal processing applications when frequency response is considered.

REFERENCES

[1] M. Koseoglu, F. N. Deniz, B. B. Alagoz, A. Yuce, and N. Tan, "An experimental analog circuit realization of Matsuda's approximate fractional-order integral operators for industrial electronics," *Eng. Res. Express*, vol. 3, no. 4, p. 045041, Dec. 2021, doi: 10.1088/2631-8695/ac3e11.

[2] A. M. Hassanein, A. H. Madian, A. G. G. Radwan, and L. A. Said, "On the Design Flow of the Fractional-Order Analog Filters Between FPAA Implementation and Circuit Realization," *IEEE Access*, vol. 11, pp. 29199–29214, 2023, doi: 10.1109/ACCESS.2023.3260093.

[3] J. Nako, C. Psychalinos, and A. S. Elwakil, "A 1 + α Order Generalized Butterworth Filter Structure and Its Field Programmable Analog Array Implementation," *Electronics*, vol. 12, no. 5, p. 1225, Mar. 2023, doi: 10.3390/electronics12051225.

[4] M. A. George, A. S. Elwakil, A. Allagui, and C. Psychalinos, "Design of Complex-Order PI/PID Speed Controllers and its FPAA Realization," *IEEE Access*, vol. 11, pp. 118606–118614, 2023, doi: 10.1109/ACCESS.2023.3326446.

[5] A. Ali, K. Bingi, R. Ibrahim, P. A. M. Devan, and K. B. Devika, "A review on FPGA implementation of fractional-order systems and PID controllers," *AEU - Int. J. Electron. Commun.*, vol. 177, p. 155218, Apr. 2024, doi: 10.1016/j.aeue.2024.155218.

[6] C. X. Jiang, J. E. Carletta, and T. T. Hartley, "Implementation of Fractional-order Operators on Field Programmable Gate Arrays," in *Advances in Fractional Calculus: Theoretical Developments and Applications in Physics and Engineering*, J. Sabatier, O. P. Agrawal, and J. A. T. Machado, Eds., Dordrecht: Springer Netherlands, 2007, pp. 333–346. doi: 10.1007/978-1-4020-6042-7_23.

[7] W. Wolf, *FPGA-based system design*. Pearson education, 2004.

[8] U. Meyer-Baese and U. Meyer-Baese, *Digital signal processing with field programmable gate arrays*, vol. 65. Springer, 2007.

[9] M. Koseoglu, "Time response optimal rational approximation: Improvement of time responses of MSBL based approximate fractional order derivative operators by using gradient descent optimization," Eng. Sci. Technol. Int. J., vol. 35, p. 101167, Nov. 2022, doi: 10.1016/j.jestch.2022.101167.

[10] F. N. Deniz, B. B. Alagoz, N. Tan, and M. Koseoglu, "Revisiting four approximation methods for fractional order transfer function implementations: Stability preservation, time and frequency response matching analyses," *Annu. Rev. Control*, vol. 49, pp. 239–257, 2020, doi: 10.1016/j.arcontrol.2020.03.003.

[11] G. F. Franklin, J. D. Powell, and M. L. Workman, *Digital control of dynamic systems*, vol. 3. Addison-wesley Menlo Park, 1998.

[12] "Matlab Documentation 'c2d,'" c2d. [Online]. Available: https://www.mathworks.com/help/control/ref/dynamicsystem.c2d.html?s_tid=doc_ta

[13] K. Ogata, *Discrete-time control systems*. Prentice-Hall, Inc., 1995.

[14] J. G. Proakis, *Digital signal processing: principles, algorithms, and applications, 4/E*. Pearson Education India, 2007.

[15] "Xilinx System Generator for DSP Getting Started Guide." AMD. Accessed: Aug. 16, 2024. [Online]. Available: https://docs.amd.com/v/u/en-US/sysgen_gs

[16] A. Sharma and T. K. Rawat, "Design and FPGA implementation of lattice wave fractional order digital differentiator," *Microelectron. J.*, vol. 88, pp. 67–78, Jun. 2019, doi: 10.1016/j.mejo.2019.04.013.

[17] D. Datta and H. S. Dutta, "High performance IIR filter implementation on FPGA," *J. Electr. Syst. Inf. Technol.*, vol. 8, no. 1, p. 2, Dec. 2021, doi: 10.1186/s43067-020-00025-4.

[18] V. Dhillon, S. Nair, A. Pabarekar, M. Kumbhare, K. Thakur, and R. Krishnan, "Implementation of FIR Digital Filter on FPGA," in *2021 4th Biennial International Conference on Nascent Technologies in Engineering (ICNTE)*, NaviMumbai, India: IEEE, Jan. 2021, pp. 1–5. doi: 10.1109/ICNTE51185.2021.9487744.

[19] H. V. Dixit and D. V. Gupta, "IIR filters using Xilinx System Generator for FPGA implementation," *Int. J. Eng. Res. Appl.*, vol. 2, no. 5, pp. 303–307, 2012.

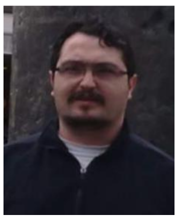
[20] P. Paz and M. Garrido, "Efficient Implementation of Complex Multipliers on FPGAs Using DSP Slices," *J. Signal Process. Syst.*, vol. 95, no. 4, pp. 543–550, Apr. 2023, doi: 10.1007/s11265-023-01867-7.

[21] P. Kwiatkowski, "Digital-to-time converter for test equipment implemented using FPGA DSP blocks," *Measurement*, vol. 177, p. 109267, Jun. 2021, doi: 10.1016/j.measurement.2021.109267.

BIOGRAPHIES

Ömer Pektaş Malatya, Turkey, in 1991. He received the B.S. and M.S. degrees from the Department of Electrical & Electronics Engineering, Inonu University, Malatya, Turkey in 2015 and 2018, respectively. He is a Ph.D. candidate in electrical-electronics engineering at Inonu University, Malatya.

He is working as a lecturer at Department of Vocational School of Technical Sciences, Electric and Energy, University of Karamanoglu Mehmetbey, Karaman, Turkey. His research interests include digital circuit implementation, FPGA design and programming, wearable sensors, object-oriented programming and mobile application development.



Murat Köseoğlu was born in Giresun, Turkey in 1978. He received the B.Sc., the M. Sc. and the Ph.D. degrees from the Department of Electrical & Electronics Engineering, Inonu University, Malatya, Turkey in 2000, 2003 and 2010, respectively. He worked as engineer in several power plants between 2000 and 2002. He worked as Research Assistant in the Department of Electrical & Electronics Engineering, Inonu University, Malatya, Turkey between 2002 and 2011. He worked as an Assistant Professor in the same department between 2011 and 2024. He has been working as an Associate Professor in the same department since 2024. His research interests include electronics, dielectrics, semiconductors and discharge phenomena.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 1, March 2025
Research Article

An Integrated Web Security Application: Integration of Nginx Reverse Proxy, Fail2ban, WAF, Postgresql and Laravel

Raif Sime, Necmettin Sezgin and Fikri Aggun

Abstract— **Recently, the increase in network-connected devices and the ability to run every application over the web has made web application security an issue that needs to be seriously considered. Although firewall solutions are used to protect networked systems and users, it seems that they are insufficient to ensure application security, especially in today's conditions. In this context, WAF (Web Application Firewall) systems have been developed and continue to be developed, especially to ensure the security of web applications. While the firewall filters traffic at the network layer, which is a lower layer, WAF protects at the application layer closest to the user. Network administrators intensively use WAF applications and the systems they create with new technologies integrated into these applications in order to maximize security.**

In this study, the WAF application, which is used together with Laravel, File2ban and Postgresql, is discussed, which we compiled and ran to protect the corporate network we manage from attacks and application vulnerabilities. In addition, it is thought that this study will guide other researchers working in this field and aims to open doors to produce more effective solutions.

Index Terms— **WAF, File2ban, Laravel, Postgresql.**

I. Introduction

THE INCREASING use of the web and the proliferation of web applications in every field makes it necessary to ensure application security and it is becoming increasingly difficult to ensure security. In addition, developers use risky tools and techniques to detect such attacks, threats and vulnerabilities present in the application [1]. In order to ensure security in corporate networks, problems arising from weak points should be solved as much as possible.

Raif Sime, Master Student, Graduate School of Education, Batman, University 72100 Batman, Turkey,(e-mail: raifsime@yahoo.com).

https://orcid.org/ 0009-0008-4292-2456

Necmettin Sezgin, Department of Computer Engineering, Batman University, Batman 72100, Turkey, (e-mail: necmettin.sezgin@batman.edu.tr).

https://orcid.org/0000-0002-4893-6014

Fikri Ağgün, Department of Computer Technologies, Adilcevaz Vocational School , Bitlis Eren University, Bitlis 13100, Turkey, (e-mail: faggun@gmail.com).

https://orcid.org/0000-0001-9550-1462

Manuscript received Sep 11, 2024; accepted March 06, 2025.
DOI: 10.17694/bajece.1547456

This is possible with security systems that complement each other and work interactively together. Such a structure is called multi-layered security and defense in depth [2].

In addition, even people who are authorized to use the systems may unknowingly damage the system as a result of unconscious use. For this reason, information systems must be protected against both internal threats and possible malicious threats from outside [3].

Firewalls have been used for general protection, port and protocol-based filtering, network layer security, etc. to ensure web security in the use of applications, but providing only network security is insufficient to prevent recent attacks. Web applications have recently become the primary target of attacks due to apathy, lack of awareness, lack of secure software development techniques and disregard for web application security. Although traditional firewalls successfully prevent network layer attacks, they are not effective in web-based attacks on web applications. Therefore, web applications need security in order to prevent information loss and vulnerability on the Internet, which is not a secure environment [4].

In this context, in addition to firewalls, which alone are insufficient to ensure the security of web applications, application layer protection, http-based filtering, web application security and detailed content review, applications called web application firewall (WAF) have started to be used extensively. WAF is likened to a security shield for applications accessed using HTTP [5]. In summary, a firewall is designed for general network security, while a WAF is designed to secure web applications. A firewall filters traffic at the network layer level, while a WAF filters traffic at the application layer level. When both security mechanisms are used together, it creates a holistic security strategy and protects your network and web applications more effectively. In addition to these features, they can bring additional protection in areas that were not taken into account when the software was developed [6]. The complexity of web applications, the increase in cyber-attacks, the difficulty of detecting unknown threats, the ability to perform security controls at the application level, the protection of sensitive data, and the adaptation of standards that must be followed for data security to web applications reveal the need for WAF applications. WAF can have two types of security models, positive or negative, depending on the type of policy. A positive security model only allows traffic that matches the policies to pass through. All other traffic is blocked. A negative security model allows all traffic through and only tries to block traffic represented by malicious rules. If the negative model alone is used in a WAF, hackers can bypass the inspection. Therefore, it is recommended to use a combination of both negative and positive model with better protection [7]. In teheir work, the authors emphasize the importance of enhancing web application security against SQL injection attacks and present their findings as a promising solution for improving the effectiveness of existing WAFs. The solution proposed in the article demonstrates robust detection capabilities, enhanced performance with higher generalization rates, and competitive advantages over existing solutions, indicating its potential to serve as a valuable tool in web application security [8]. In conclusion, WAF is a critical component to enhance the security of modern web applications and protect against cyber-attacks.

With the increasing security risks associated with web applications, the need for a WAF is also increasing. As shown in figure (Fig.1a), all SSL traffic passes directly through without content checking in networks with traditional firewalls, but in networks with WAF architecture (Fig.1b), all requests for user-to-server traffic are answered at the reverse proxy server, security checked, and if an attack is detected, it can be rejected before the request reaches the original servers.

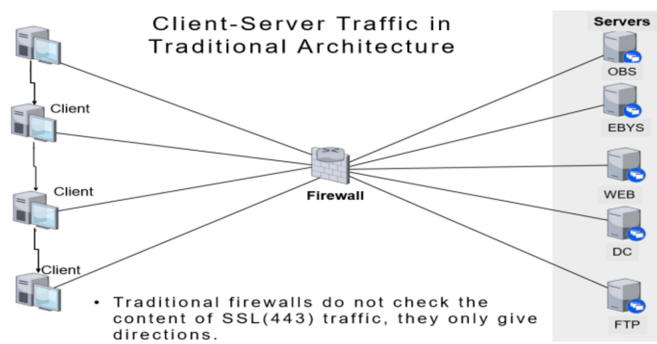


Fig.1a. Client-Server traffic in traditional architecture

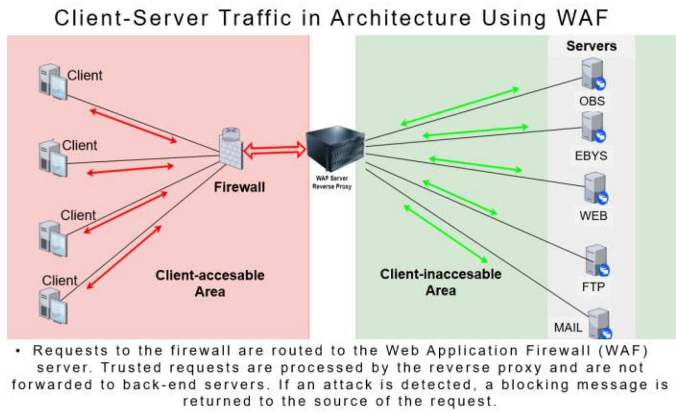


Fig.1b. Client-Server traffic in WAF-enabled architecture

Web applications, which enable many transactions to be performed, become the target of attacks because they contain various information such as personal, banking and corporate information. Ensuring the security of these environments is directly related to the security of web applications. Web application security refers to all measures taken to ensure the confidentiality, integrity and accessibility of the data it contains [9]. In their article the authors aim to develop an affordable and user-friendly WAF framework. They employ proxy techniques to block and sanitize malicious user requests and implement two machine learning models for detecting malicious requests and classifying attack types. As a result, they propose a more accessible, effective, and user-friendly WAF solution that could significantly benefit organizations with limited resources. The proposed framework addresses common security vulnerabilities and provides real-time monitoring, assisting organizations in enhancing their security measures. Additionally, focusing on interface development through user feedback, as expressed for future work, is among the outcomes of our study [10]. The security of web applications is of increasing importance today. Therefore, developers and system administrators try to protect web applications by implementing various security measures. This paper will focus on a web application security system integrated with Nginx reverse proxy, its structure, components and benefits.

II. MATERIALS AND METHODS

In our application, a complex and effective structure has been created to increase web application security. In this system, an efficient and effective network performance is achieved by combining different technologies such as Nginx reverse proxy, WAF, PostgreSQL, Laravel and Fail2ban. Nginx and WAF are used for traffic routing and filtering, PostgreSQL for logging and data storage, File2ban for automatic IP blocking mechanism, and Laravel for security rules management and application environment.

A. *Nginx Reverse Proxy and WAF Integration*

Nginx is an open source web server software that can be used as an HTTP and reverse proxy server. Nginx is stable, secure and very easy to configure, and has performance and efficiency advantages over Apache [11]. A reverse proxy takes requests and redirects them to another server. That is, clients make requests to the Nginx server and Nginx receives the request, forwards it to the destination server and forwards the answers to the clients. This is useful in situations where the target servers are not directly accessible and provides several advantages. The client (for example a browser) sends a request to the Nginx server, like an HTTP request to a URL. Nginx examines the incoming request and routes it according to certain rules. These rules ensure that certain URLs or domains are directed to certain destination servers. In the next step, Nginx forwards the request to the destination server for processing. The response from the destination server is received by Nginx and forwarded to the client. This is in response to the client's original request.

In this way, we can list the benefits of using Nginx as a reverse proxy as follows:

1. *Load Balancing*
   Nginx balances the load by distributing traffic between multiple destination servers. This improves the performance and availability of the target servers.
2. *Provide Cache Support*
   Nginx reduces the load on target servers and speeds up response times by caching static content.
3. *Security*
   Nginx provides security by filtering incoming requests, validating them, and applying firewall rules if necessary.
4. *URL Routing*
   Provides a flexible structure by routing specific URLs or domains to different destinations. This is especially useful when multiple web applications or microservices are used.
5. *SSL Termination*
   Nginx manages SSL/TLS certificates by managing HTTPS traffic and can forward HTTPS traffic as HTTP to destination servers. This does not require destination servers to deal with SSL/TLS and simplifies the structure.

Using Nginx as a reverse proxy makes web servers and applications more secure, efficient and scalable in many cases. For this reason, many large-scale websites and applications prefer Nginx.

B. PostgreSQL Integration with Laravel and Criteria Based Transaction Execution

Laravel is a PHP-based open source web application framework that aims to create fast and efficient web applications, is based on MVC (Model-View-Controller) architecture and offers developers a set of useful tools and libraries. In the development of the project, Laravel's features and components such as Eloquent ORM, Blade Template Engine, Migrations and Seeding, Routing and Controls, and Artisan Command Line Tool were used extensively for efficiency.

It is important to record the operation of the system in question and the movements during the operation process, both in order to identify problems that may arise in the system and for legal obligations. In order to keep the event records of the system developed for this purpose, integration with a powerful database such as PostgreSQL was provided. PostgreSQL is an open source and relational database management system. PostgreSQL is known for its strong ACID compatibility, wide data type support and advanced features. In the application, PostgreSQL's ACID Compatibility, Wide Data Types, Advanced Extension Support, JSON Support and high performance features were utilized to maximize the recording capability.

C. Mod Security Integration

Web application firewalls identify, monitor and block HTTP traffic to and from a web server. By controlling HTTP traffic, application vulnerabilities such as SQL injection, cross-site scripting (XSS), file injection and protection failures can be prevented [12]. ModSecurity is an open source WAF method. Version 3 (ModSecurity3) is the latest and most actively available version of this application. ModSecurity3 includes a number of improvements, new features and updates compared to previous versions. ModSecurity3 has abandoned the single-threaded architecture in favor of a Multi-Threaded Architecture. This provides better performance and scalability, allowing the WAF to work more effectively, especially in high-traffic environments. Another feature, the New Low-Level Engine, uses a more flexible and efficient engine compared to previous versions. This allows to process requests faster and apply more complex security rules. ModSecurity3 offers better performance and efficiency compared to previous versions. The multi-threaded architecture and new low-level engine enable ModSecurity to consume fewer resources and run faster. ModSecurity3 includes new security rules, predefined rules, cyber threat signatures and updates. This enables the WAF to more effectively protect against current threats. It improves compatibility with popular web servers such as Apache HTTP Server and Nginx. It also provides better integration with various security products and services. ModSecurity3 improves debugging and diagnostics. This enables faster detection and resolution of errors and problems, which makes the WAF work more reliably. In addition, ModSecurity3 provides enhanced features, performance and updates to protect web applications against security threats. In today's world of growing security awareness and increasing complexity of web applications, security tools like ModSecurity play an important role.

D. Fail2ban Integration and Automatic Blocking

PostgreSQL and Fail2ban integration is an effective method to increase the security of a database system. This integration ensures database security by automatically responding to malicious activity and allows system administrators to quickly respond to security threats. In our application, File2Ban and PostgreSQL integration works as follows: PostgreSQL offers various logging levels. One of the logging levels, which is important for security, enables logging of certain types of queries, failed login attempts or other critical events. Fail2ban monitors specific log files and detects certain patterns in those files (for example, failed login attempts) and temporarily or permanently blocks IP addresses that match those patterns. Fail2ban manages blocked IP addresses dynamically. IP addresses blocked for a certain period of time can be automatically released or permanently blocked according to a specific set of rules.

Thanks to PostgreSQL and Fail2ban Integration, automatic security is provided by monitoring and blocking events recorded for web security in the running system. In addition, thanks to Fail2ban, malicious activities are quickly detected and intervention is provided. Automatic blocking helps prevent threats without requiring manual intervention and with less hassle for system administrators. By recording and archiving events, system administrators can analyze the security status and take faster precautions against new vulnerabilities that may arise.

When these effects are analyzed, the PostgreSQL-Fail2ban integration is an effective method to increase the security of a system. This integration ensures database security by automatically responding to malicious activity and provides system administrators with a rapid response to security threats.

III. Results and Discussion

When the studies and the operation of the implemented system were analyzed, it was observed that the integrated Nginx reverse proxy, WAF, PostgreSQL, Laravel and Fail2ban system increased the security of web applications. Thanks to the enhanced security provided by the system, a more efficient structure has been achieved by ensuring that each job is done by a different module instead of uploading all jobs to a system.

In our study, as a security measure, the Nginx reverse proxy filters incoming requests and blocks malicious traffic, while the WAF detects vulnerabilities by performing more in-depth examinations. Thus, it provided a more effective protection for the web application against cyber-attacks.

Thanks to the Laravel application, security policies can be defined and customized more flexibly. Security checks are performed according to the criteria determined by the event logs stored in PostgreSQL, so that each application can determine and implement specific security requirements. This is an ideal solution that facilitates the implementation of security policies and determines the security level specific to each application. Event logs obtained as a result of the application provide an important data source for monitoring, analyzing the security situation and responding when necessary. Thus, it has been observed that security measures can be taken more easily and earlier. It is also clear that analyzing this data will enable continuous improvement and strengthening of security policies.

Another benefit emphasized in the system is the caching mechanism. This mechanism and the accompanying Nginx reverse proxy improve the performance of the web application running on the system. Thanks to this system, the necessary protection measures are taken for the security of the web application, while performance loss is minimized and the user experience is prevented from being exposed to negative effects.

As a result of our study, the system established with our application works in the active corporate network and responds to many requests as seen in the graph and faces a large number of attacks.

As can be seen from the graph (Fig.2), 2% of the requests directed to the servers are attack requests. Although this rate may seem low, it means a serious attack attempt in a network with very heavy traffic and can pose significant security risks.

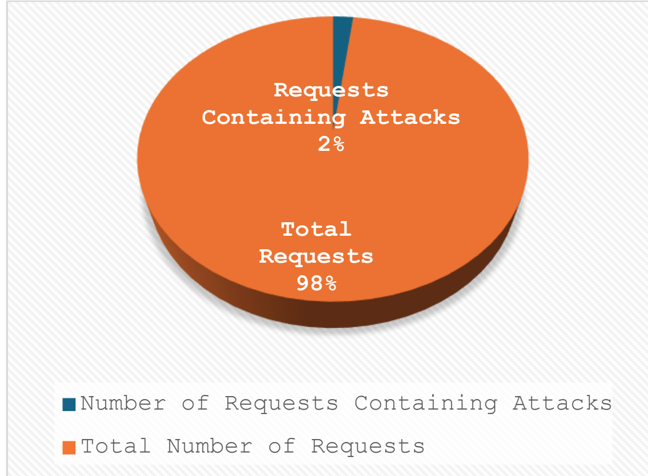


Fig.2. Total / Harmful Requests for corporate network traffic

For this reason, detecting and blocking these attack requests is of great importance in terms of ensuring system continuity.

The graph below shows the number of attacks on our organization's network for 12 days and the increases in these numbers on some days.

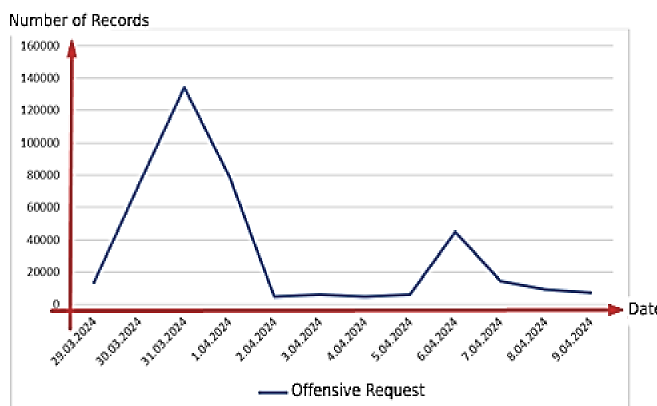


Fig.3. Number of traffic records containing attacks between 29.03.2024 - 09.04.2024

When the graph in Figure 3 is analyzed, it is seen that servers are attacked on some days and in this context, the number of attacks increases significantly. That is, the daily distribution of requests clearly shows that there are more attacks on certain days than others. This is an important consideration for the security of the servers and requires measures to be taken to detect and prevent attacks more effectively. The increases in the event logs generated by the application and analyzed with the help of PostgreSQL and Laravel show that attacks are being made at certain intervals and that there is a search for open doors to infiltrate the system. An example of attack record table that have been stored in PostgreSQL is showed in Figure 4.

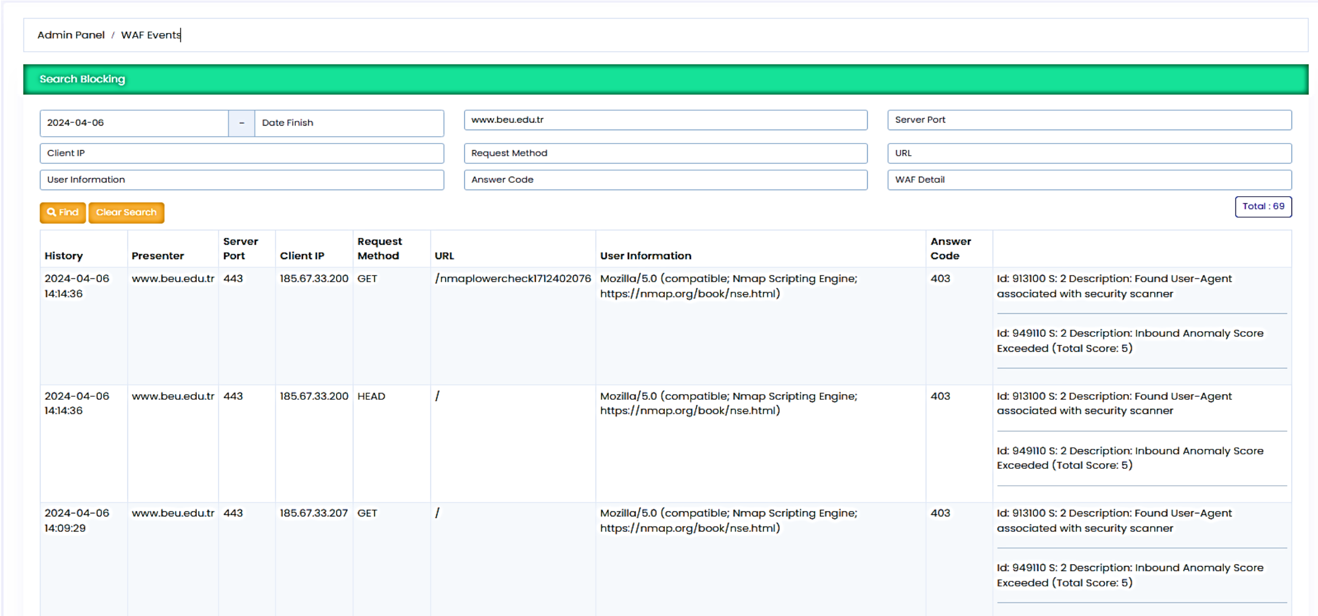


Fig.4. Attack records that stored in database

Some useful information can be obtained as a result of examining these attempts, which are perceived as attacks on systems, with reverse engineering methods and identifying the sources of attack.

When the data and IP addresses in the figure are checked, it is seen that the attempts that appear to be attacks are the penetration tests performed by the Information and Communication Technologies Authority on our organization's network (Fig.5). Since penetration tests are controlled attack and vulnerability scanning procedures performed to assess the security of an institution's or an organization's information systems, such tests help to take preventive steps to detect and correct vulnerabilities of information systems. Therefore, the identification that the IP addresses examined were used during the penetration tests shows that we have taken an important step to increase the effectiveness of security measures and strengthen the network's defenses.

Fig.5. Location analysis of an attack

When we examine the communication requests and attacks in network traffic on the same graph(Fig.6); Requests directed to servers and attacks are independent of each other.

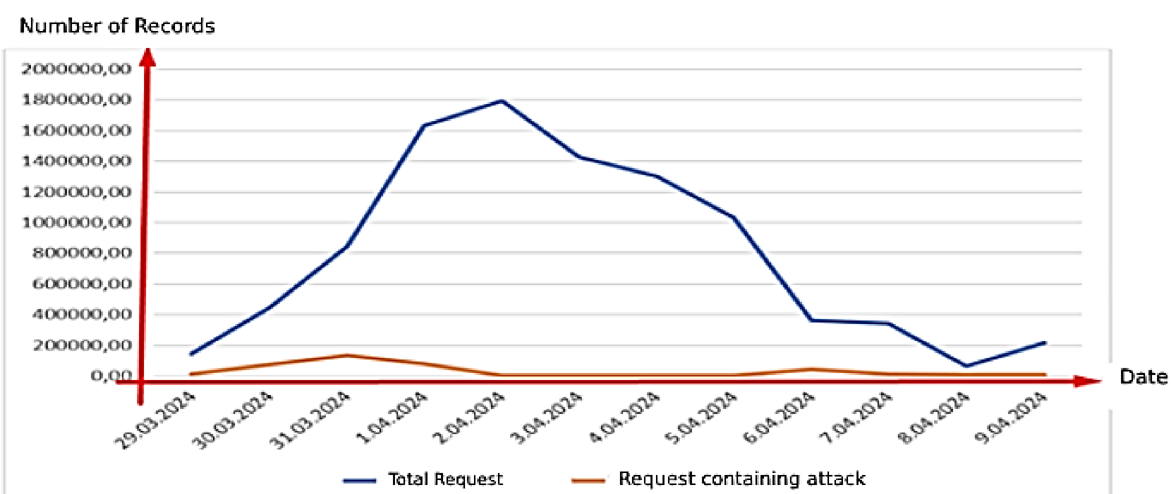


Fig.6. All requests and attacks

That is, there is no direct correlation between the number of requests directed to servers and the number of requests containing attacks. For example, on any day, the number of requests may increase, but this increase does not affect the number of requests containing attacks. This may indicate that attacks are carried out independently of the request traffic and that attackers are not focused on density at certain times. This detail is an important consideration in the design of security measures and intrusion detection mechanisms.

IV. Conclusion

Our web application security system, integrated with Nginx reverse proxy, offers an effective solution to increase the security of web applications. Considering the benefits, it provides, the security of web applications is increased while performance and efficiency are also ensured. This is a gain that increases both the security of users and the operational efficiency of the business. It is thought that our system will strengthen the hands of system developers and administrators in areas such as providing web security, cyber security measures, etc. and offers both fast and effective solutions. Since the system has an open source structure, it can be said that it is a system open to development. In addition, it is thought that the system is a work that can be a source of inspiration for system developers working in this field and experiencing security problems.

References

[1] D. Mairaj Inamdar and S. Gupta, "A Survey on Web Application Security," Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol., vol. 3307, pp. 223-228, 2020, doi: 10.32628/cseit206543.
[2] E. Karaarslan, T. Tuğlular, and H. Şengonca, "Enterprise web security structure," in Akademik Bilişim, 2008, pp. 1-9.
[3] M. Baykara, R. Daş, and G. Tuna, "Web-based log analysis platform for detection of web attacks from web server access logs," Firat University Engineering Sci. Derg., vol. 28, no. 2, pp. 291-302, 2016.
[4] A. Tekerek, C. Gemci, and O. F. Bay, "Development of a hybrid web application firewall to prevent web based attacks," in 8th IEEE International Conference on Application of Information and Communication Technologies, AICT 2014 - Conference Proceedings, 2014, pp. 1-4, doi: 10.1109/ICAICT.2014.7035910.
[5] R. A. Muzaki, O. C. Briliyant, M. A. Hasditama, and H. Ritchi, "Improving Security of Web-Based Application Using ModSecurity and Reverse Proxy in Web Application Firewall," 2020 Int. Work. Big Data Inf. Secur. IWBIS 2020, pp. 85-90, 2020, doi: 10.1109/IWBIS50925.2020.9255601.
[6] H. Tan and A. Z. Aktas, "An approach for an organization's information system security," in Network and Information Security Symposium, 2011, pp. 34-39.
[7] V. Clincy and H. Shahriar, "Web Application Firewall: Network Security Models and Configuration," in Proceedings - International Computer Software and Applications Conference, 2018, vol. 1, pp. 835-836, doi: 10.1109/COMPSAC.2018.00144.
[8] F. Omar, D. Ahmed, O. Elnakib, et al., "Towards a User-Friendly Web Application Firewall.," In: Proceedings - 11th IEEE International Conference on Intelligent Computing and Information Systems, ICICIS 2023. pp. 483–488. IEEE (2023).
[9] D. Aydogdu and M. Gündüz, "A Research on Web Application Security Vulnerabilities and Security Solutions," Uluslararası Uluslararası Bilgi Güvenliği Mühendisliği Dergisi, vol. 2, no. 1, pp. 1-7, 2016, doi: 10.18640/ubgmd.56836.
[10] A. Coscia, V. Dentamaro, S. Galantucci, A. Maci, and G. Pirlo, "PROGESI: A PROxy Grammar to Enhance Web Application Firewall for SQL Injection Prevention.," IEEE Access. vol. 12, no. August, pp. 107689–107703, 2024.
[11] Nginx: the High-Performance Web Server and Reverse Proxy, https://dl.acm.org/doi/fullHtml/10.5555/1412202.1412204.
[12] T. D. Sobola, P. Zavarsky, and S. Butakov, "Experimental Study of ModSecurity Web Application Firewalls," Proc. - 2020 IEEE 6th Intl Conf. Big Data Secur. Cloud, Big Data Security 2020, 2020 IEEE Intl Conf. High Perform. Smart Comput. HPSC 2020 2020 IEEE Intl Conf. Intell. Data Secur. IDS 2020, pp. 209-213, 2020, doi: 10.1109/BigDataSecurity-HPSC-IDS49724.2020.00045.

Biographies

Raif Sime was born in Ağrı in 1984. He graduated from Atatürk University Ağrı Vocational School, Department of Computer Technologies and Programming in 2006. He graduated from Anadolu University, Faculty of Business Administration and Ağrı İbrahim Çeçen University, Faculty of Education, Department of Religious Culture and Ethics Education in 2014. Since 2022, he has been working as the Branch Manager at Department of Information Technologies in Bitlis Eren University.

Necmettin Sezgin graduated from Hacettepe University, Faculty of Engineering, Department of Electrical and Electronics Engineering in 2001. He received MSc degree at Dicle University, Institute of Science, Department of Electrical and Electronics Engineering in 2003 and PhD at İnönü University, Institute of Science, Department of Electrical and Electronics Engineering in 2011. In 2011, he started working as an assistant professor at Batman University at Department of Electrical and Electronics Engineering. In 2014, he received the degree of associate professor in the field of Electrical and Electronics Engineering and started working as an associate professor at Batman University, Department of Electrical and Electronics Engineering in the same year. He was appointed as a Professor at Batman University at Department of Computer Engineering in 2019.

Fikri Ağgün was born in Ahlat, Bitlis in 1980. He received the B.S. in computer engineering from the University of Selçuk, Konya in 2001 and M.S. degree in in Biometry-Genetic from Yuzuncu Yil University, Van in 2011 and the Ph.D. degree in electrical and computer engineering from Ankara Yildirim Beyazit University, Ankara, in 2020. His research interests are wireless sensor networks, Vanets and Mac protocol design. He still works as an Assistant Professor at Bitlis Eren University.

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 1, March 2025

Research Article

Space Power Block Coding (SPBC): A Novel 6G Method for Providing High Throughput and Ultra-Reliability for Processing-limited Receivers

Jehad M. Hamamreh

***Abstract*—In this work, we introduce a new transmission method inspired by space-time block-coding (STBC) with orthogonal frequency division multiplexing (OFDM) to jointly improve spectral efficiency and reliability, while reducing receiver complexity, which is deemed as a critical requirement for processing-limited IoT devices. The newly developed transmission method, coined as space-power block-coding (SPBC), utilizes the superposition of specially precoded data streams to be transmitted to serve a single end-receiving device with two antennas, while keeping the receiver structure unchanged, thus making it compatible with any current or future modern smart end-devices. The proposed scheme objectives are to improve spectral efficiency, increase the link throughput, and enhance reliability for a single OFDM-based receiver. In particular, the user data streams are superimposed with intelligently designed pre-coders intending to cancel the intentionally introduced interference at the transmitter as well as the channel impact and effect at the receiver, while ensuring that the reception process is much simpler and way less complex. Moreover, we analyze the obtained comparative simulation results, where it is found that the performance of the proposed design is significantly higher than that of its conventional counterparts.**

***Index Terms*—MIMO, STBC, SPBC, OFDM, 5G, 6G, Superposition, Precoding, Reliability, Spectral Efficiency, IoT, V2X.**

I. INTRODUCTION

THE grand vision of vehicle-to-everything (V2X) communication supported by the sixth generation (6G) of wireless systems will be an instrumental element of future connected autonomous vehicles [1]. On the other hand, the Internet of Things (IoT) is a gigantic network of things ranging from simple sensors to smartphones, wearables, autonomous cars, and drones that are connected and used to collect and analyze data to perform actions. These low-cost sensing modules and more enhanced powerful devices are integral parts of IoT today. The diversity of such devices and associated applications result in heterogeneous data in various senses [2].

To improve the reception quality and meet the many requirements of IoT related applications in telecommunication systems, diversity techniques have widely been applied to wireless systems. A base station in a wireless network often serves tens to hundreds of users. It is, therefore, more economical to add equipment to base stations rather than the user devices. For this reason, transmit diversity schemes are very attractive [3]. In [3], S. M. Alamouti presented a simple transmit diversity scheme which improves the signal quality at the receiver on one side of the link by simple processing across two transmit antennas on the opposite side. The obtained diversity order is equal to the acquired diversity gain by applying maximal-ratio receiver combining (MRRC) with two antennas at the receiver.

In orthogonal space-time block codes (STBC), due to the decoding error probability being high, authors in [4] developed a linear transformation (precoding) solution of orthogonal STBC for reducing the probability of decoding error, when the transmitter knows channel covariance matrix. Particularly, the authors provided a closed form solution for multi-input single-output (MISO) systems and a numerical solution for the multi-input multi-output (MIMO) systems.

In practical wireless communication scenarios, the antennas available on the reception side are usually more restricted than the antennas on the transmitter side, whereas there is no such limitation in STBC. Therefore, in [5], a hybrid STBC scheme was proposed that exploits these two systems. The hybrid STBC technique involves two approaches: a pre-coding and a decision feedback detection method. By using the linear pre-coding matrix, performance can be enhanced and the interference can be canceled out by utilizing the decision feedback detection scheme.

In [6], a new scheme was investigated that integrates a linear pre-coder exploiting both the channel mean and transmit correlation with non-orthogonal space-time code systems. The precoder is based on pairwise error probability with minimized Chernoff bound on it. This precoder can be viewed as a multi-function beamformer and can converge to a single beam as the K factor increases. Usually, due to low minimum diversity and decoder's error propagation, there is a bit error rate (BER) performance degradation in vertical Bell-Labs layered space-time (VBLAST) MIMO system. On the other hand, even an STBC-integrated VBLAST system does not outperform the pre-coded VBLAST system [7]. Therefore, in [7], the authors developed a hybrid joint design that includes a pre-coder and its relevant decoder for the VBLAST-STBC system.

Conventionally, to obtain transmit diversity, multiple antennas are required, but due to possible physical constraints like in MIMO, diversity cannot be realized. However, cooperative communication solves this dilemma by enabling a single antenna to function as a virtual multiple antenna transmitter under sharing the available antennas [8]. To achieve high-

Jehad M. HAMAMREH is with the WISLAB, Department of Electrical and Electronics Engineering, Antalya Bilim University, Antalya, 07190, Turkey (e-mail: jehad.hamamreh@gmail.com).
https://orcid.org/0000-0003-4435-5500

Manuscript received May 03, 2024; accepted Jan 16, 2025.
DOI: 10.17694/bajece.1478007

performance authors in [8], utilize STBC with one transmit antenna, and pre-code the transmit symbols, thus, achieving high gain without loss in the transmission rate.

In [9], authors investigate unitary pre-coders using STBC for space-time coded orthogonal frequency division multiplexing (ST-OFDM) systems. Two-channel variations are assumed where in the case of LTI, optimal pre-coders are derived that are channel-dependent and greatly enhance the system performance, while in the time-varying case, channel variation causes inter-carrier interference (ICI) in the ST-OFDM system, ICI can be alleviated by properly designing the pre-coder which can average out the error variance.

Authors in [10] developed a multi-mode pre-coder design by varying the number of streams depending on the channel conditions and also proposed a design criterion of reducing the vector symbol error rate with an efficient algorithm to generate pre-coders. In [11], authors proposed a Zadoff-Chu matrix transform (ZCMT) based STBC MIMO-OFDM system with reduced peak-to-average power ratio (PAPR), to obtain low PAPR than Walsh-Hadamard transform (WHT) pre-coded STBC MIMO-OFDM systems and the conventional STBC MIMO-OFDM systems.

In [12], coordinated multi-point system STBC is implemented as a joint coding strategy with pre-coding user data comparative to channel state information (CSI) to improve spectrum utilization and edge user receptivity. The study in [13] presents an STBC dirty paper coding (DPC) system which includes maximum likelihood (ML), one-dimensional searching algorithm, and QR decomposition of channel matrix for gaining reliable performance. Channel on/off assignment while using the water filling algorithm to combat deep fading channel problem, and Tomlinson-Harashima precoding (THP) scheme for lowering the PAPR.

Authors in [14] proposed a combination of STBC, and spatial modulation with hybrid analog-digital beamforming for 60 GHz Millimeter-wave (mmWave) communications to take advantage of these schemes and avoid their drawbacks. In [15], [16], we proposed a modified STBC with a non-orthogonal superposition technique dedicated to serving multi-users by utilizing the inherent diversity and redundancy that exist in both time and space domains of STBC transmission scheme. In particular, purposely designed assistant signals are superimposed on top of the multi-user data during the two transmission time slots of STBC to intelligently cancel inter-user interference as well as channel fading at the receiving end device, while ensuring that the reception process is much simpler and less power-consuming than conventional systems.

As discussed above, the state-of-the-art in the light of pre-coding and classical STBC is not suitable for the requirements and challenges imposed by 6G. To address these necessities, in this paper, **we propose a novel transmission method, coined as Space-Power Block-Coding (SPBC), which can provide high throughput, ultra-reliability, elimination of inter-antenna and inter-symbol interference, physical layer security improvement, and diversity enhancement, while ensuring less complexity at the receiver as there is no need for doing equalization in the proposed method.**

The novelty and main contributions of the proposed method design can be explained as follows:

- Exploiting the STBC concept and the super positioning of user data streams in a novel manner to simultaneously complete the transmission of two symbols within just one single time slot rather than two as is the case with conventional STBC method.
- The completion of the transmission within just one time slot while serving the end user with the same transceiver resources enables achieving spectral efficiency enhancement by double compared to conventional STBC which requires at least two-time slots to complete the transmission. Consequently, the overall throughput gain is twice that of the conventional STBC.
- The design and utilization of special type of pre-coding matrices $\mathbf{P_1}$, $\mathbf{P_2}$, $\mathbf{P_3}$ and $\mathbf{P_4}$ to eliminate interference by the time the signal reaches the receiver without forcing the receiver to do any extra processing like data signal combination or successive interference cancellation (SIC). Consequently, the processing complexity gets reduced at the received as well as the delay, thus making the proposed technique very suitable for low complexity IoT devices as the receiver is not required to do any extra processing.

In the subsequent sections, we will discuss the proposed transmission method, SPBC, the explain the proposed mathematical derivation behind the method, present the simulation results, and finally draw the conclusion.

II. Proposed System Model

The proposed system design is displayed in Fig. 1, where a transmitting base station equipped with two antennas is serving a single OFDM receiver with two antennas $\mathbf{R_{x_1}}$ and $\mathbf{R_{x_2}}$. By exploiting the proposed SPBC transmission scheme, which will be explained in the next section, the superimposed data signals are simultaneously transmitted from two transmit antennas $\mathbf{A_{n_1}}$ and $\mathbf{A_{n_2}}$ in just one single time slot $\mathbf{t_1}$ instead of needing to complete the transmission within two-time slots $\mathbf{t_1}$ and $\mathbf{t_2}$ as is the case in the conventional STBC scheme.

Furthermore, we superimpose the specially designed pre-coders to data signals before transmission in each time slot at the transmitter. The channels between transmit and receive antennas are denoted by (H_{km}, which is the channel frequency response diagonal matrix with k_{th} receiver antennas and m_{th} transmit antennas. The channels are assumed to be multi-path Rayleigh fading with exponential decaying factor and are calculated by using channel sounding techniques exploiting the reciprocity od the channel in a time division duplex (TDD) system.

III. Proposed Transmission Method (SPBC) Design

According to the proposed system model as shown in Fig. 1, the mathematical algorithm is illustrated as follows: The two transmit antennas $\mathbf{A_{n_1}}$ and $\mathbf{A_{n_2}}$ simultaneously serve a single OFDM receiver with two antennas $\mathbf{R_{x_1}}$ and $\mathbf{R_{x_2}}$ during just one single time slot instead of two different time slots, as the case in conventional STBC systems. The fact that the proposed

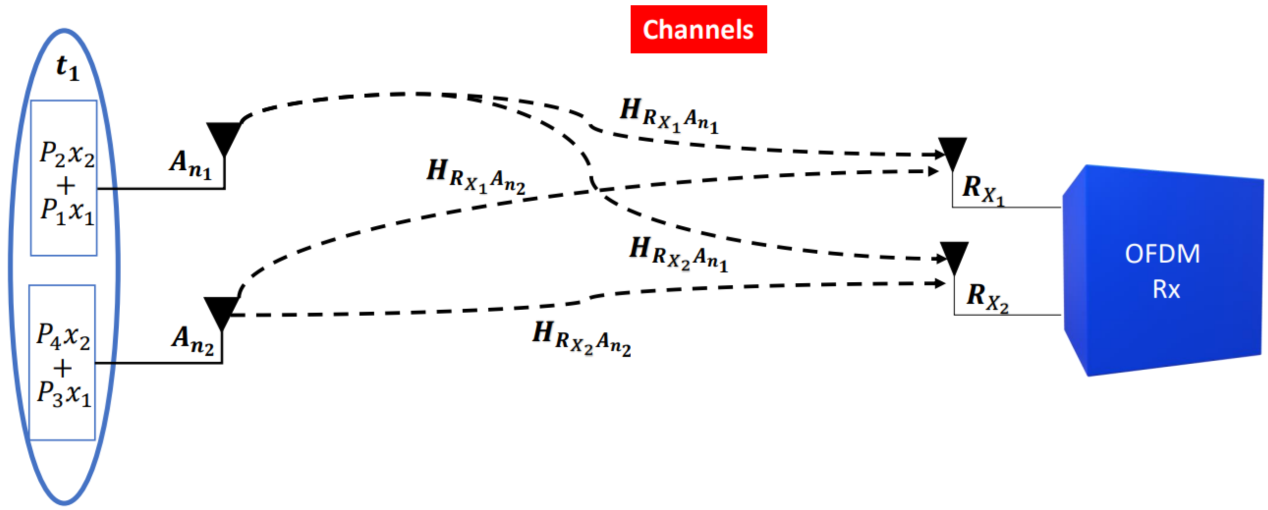


Fig. 1. Proposed SPBC system model serving a single OFDM receiver with same transceiver resources. When SPBC is used with OFDM, which is what is assumed in this paper, the precoding and superposition processes done on the transmitted signals will be implemented in the frequency domain before the IFFT block of an OFDM-based system to make the design less complex by exploiting the diagonal nature of the estimated channel frequency response for all the OFDM subcarriers

method can send two symbol vectors within one single time slot results in improving the spectral efficiency and throughput by double compared to traditional STBC systems. The total number of modulated symbols in one OFDM block is N_f and each data vector is $\mathbf{x_A} = [\mathbf{x_0}, \mathbf{x_1}, ..., \mathbf{x_{N-}}]^\mathbf{T} \in C^{[N\times1]}$ and $\mathbf{x_B} = [\mathbf{x_0}, \mathbf{x_1}, ..., \mathbf{x_{N-1}}]^\mathbf{T} \in C^{[N\times1]}$, respectively. The pre-coder matrices $\mathbf{P_1}$, $\mathbf{P_2}$, $\mathbf{P_3}$ and $\mathbf{P_4}$ are designed based on the receiver channels and superimposed with the data to be transmitted. $y_{km} \in C^{[N\times1]}$, $h_{km} \in C^{[N\times N]}$, and $n_{km} \in C^{[N\times1]}$ represent the received signal, the channel response, and the additive white Gaussian noise (AWGN) between k_{th} receiver antenna and m_{th} transmit antenna, respectively.

1) Transmission and Reception Process: The superimposed transmitted signal from $\mathbf{A_{n_1}}$ is represented as

$$\mathbf{s_1} = \mathbf{P_1x_1} + \mathbf{P_2x_2}, \tag{1}$$

and the transmitted superimposed signal from $\mathbf{A_{n_2}}$ is given as

$$\mathbf{s_2} = \mathbf{P_3x_1} + \mathbf{P_4x_2}, \tag{2}$$

where $\mathbf{x_1}$ and $\mathbf{x_2}$ are frequency domain data vectors designed for $\mathbf{R_{x_1}}$ and $\mathbf{R_{x_2}}$. Furthermore, $\mathbf{P_1}$, $\mathbf{P_2}$, $\mathbf{P_3}$ and $\mathbf{P_4}$ are the superimposed pre-coder matrices that are the function of the each receivers' channel.

In time slot $\mathbf{t_1}$, $\mathbf{R_{x_1}}$ and $\mathbf{R_{x_2}}$ receive the transmitted signals from antennas $\mathbf{A_{n_1}}$ and $\mathbf{A_{n_2}}$, which are given as

$$\mathbf{y^1_{R_{x_1}A_{n_1}}} = \mathbf{H_{R_{x_1}A_{n_1}}s_1}, \tag{3}$$

where $\mathbf{H_{R_{x_1}A_{n_1}}}$ represents the channel's frequency response between $\mathbf{R_{x_1}}$ and antenna $\mathbf{A_{n_1}}$. Similarly, the received signal at $\mathbf{R_{x_1}}$ transmitted from $\mathbf{A_{n_2}}$ is represented as

$$\mathbf{y^1_{R_{x_1}A_{n_2}}} = \mathbf{H_{R_{x_1}A_{n_2}}s_2}, \tag{4}$$

where $\mathbf{H_{R_{x_1}A_{n_2}}}$ is the channel frequency response between $\mathbf{R_{x_1}}$ and transmit antenna $\mathbf{A_{n_2}}$. The signal received at $\mathbf{R_{x_2}}$ from $\mathbf{A_{n_1}}$ is given as

$$\mathbf{y^1_{R_{x_2}A_{n_1}}} = \mathbf{H_{R_{x_2}A_{n_1}}s_1}, \tag{5}$$

where $\mathbf{H_{R_{x_2}}A_{n_1}}$ represents the channel's frequency response between $\mathbf{R_{x_2}}$ and antenna $\mathbf{A_{n_1}}$. Similarly, the received signal at $\mathbf{R_{x_2}}$ which is transmitted from $\mathbf{A_{n_2}}$ is represented as

$$\mathbf{y^1_{R_{x_2}A_{n_2}}} = \mathbf{H_{R_{x_2}A_{n_2}}s_2}, \tag{6}$$

where $\mathbf{H_{R_{x_2}A_{n_2}}}$ is the channel frequency response between $\mathbf{R_{x_2}}$ and transmit antenna $\mathbf{A_{n_2}}$.

The combined received signals at $\mathbf{R_{x_1}}$ transmitted from $\mathbf{A_{n_1}}$ and $\mathbf{A_{n_2}}$ are given as follows

$$\mathbf{y^1_{R_{x_1}}} = \mathbf{y^1_{R_{x_1}A_{n_1}}} + \mathbf{y^1_{R_{x_1}A_{n_2}}} + \mathbf{n^1_{R_{x_1}}}, \tag{7}$$

Now putting the values of $\mathbf{y^1_{R_{x_1}A_{n_1}}}$ and $\mathbf{y^1_{R_{x_1}A_{n_2}}}$ in (7) as

$$\mathbf{y^1_{R_{x_1}}} = \mathbf{H_{R_{x_1}A_{n_1}}s_1} + \mathbf{H_{R_{x_1}A_{n_2}}s_2} + \mathbf{n^1_{R_{x_1}}}, \tag{8}$$

In (8), $\mathbf{s_1}$ and $\mathbf{s_2}$ represent the superimposed transmitted signals and $\mathbf{n^1_{R_{x_1}}}$ is the AWGN at $\mathbf{R_{x_1}}$ during time slot $\mathbf{t_1}$. Substituting the values of $\mathbf{s_1}$ and $\mathbf{s_2}$ in (8) we get

$$\begin{aligned}\mathbf{y^1_{R_{x_1}}} &= \mathbf{H_{R_{x_1}A_{n_1}}}(\mathbf{P_1x_1} + \mathbf{P_2x_2}) \\ &\quad + \mathbf{H_{R_{x_1}A_{n_2}}}(\mathbf{P_3x_1} + \mathbf{P_4x_2}) + \mathbf{n^1_{R_{x_1}}}.\end{aligned} \tag{9}$$

Combining like terms after rearranging (9) we get

$$\begin{aligned}\mathbf{y^1_{R_{x_1}}} &= (\mathbf{H_{R_{x_1}A_{n_1}}P_1} + \mathbf{H_{R_{x_1}A_{n_2}}P_3})\mathbf{x_1} \\ &\quad + (\mathbf{H_{R_{x_1}A_{n_1}}P_2} + \mathbf{H_{R_{x_1}A_{n_2}}P_4})\mathbf{x_2} + \mathbf{n^1_{R_{x_1}}}.\end{aligned} \tag{10}$$

The first mathematical term of (10) is basically the desired term concerning and pertaining to $\mathbf{R_{x_1}}$, whereas the remaining other expressions are undesired and unwanted.

The combined received signals at $\mathbf{R_{x_2}}$ transmitted from $\mathbf{A_{n_1}}$ and $\mathbf{A_{n_2}}$ during time slot $\mathbf{t_1}$ are given as follows

$$\mathbf{y^1_{R_{x_2}}} = \mathbf{y^1_{R_{x_2}A_{n_1}}} + \mathbf{y^1_{R_{x_2}A_{n_2}}} + \mathbf{n^1_{R_{x_2}}}, \tag{11}$$

Now putting the values of $\mathbf{y^1_{R_{x_2}A_{n_1}}}$ and $\mathbf{y^1_{R_{x_2}A_{n_2}}}$ in (11) as

$$\mathbf{y^1_{R_{x_2}}} = \mathbf{H_{R_{x_2}A_{n_1}}s_1} + \mathbf{H_{R_{x_2}A_{n_2}}s_2} + \mathbf{n^1_{R_{x_2}}}, \tag{12}$$

In (12), $\mathbf{s_1}$ and $\mathbf{s_2}$ represent the superimposed transmitted signals and $\mathbf{n^1_{R_{x_2}}}$ is the AWGN at $\mathbf{R_{x_2}}$ in time slot $\mathbf{t_1}$. Substituting the values of $\mathbf{s_1}$ and $\mathbf{s_2}$ in (12), we get

$$\begin{aligned}\mathbf{y^1_{R_{x_2}}} &= \mathbf{H_{R_{x_2}A_{n_1}}}(\mathbf{P_1x_1} + \mathbf{P_2x_2}) \\ &\quad + \mathbf{H_{R_{x_2}A_{n_2}}}(\mathbf{P_3x_1} + \mathbf{P_4x_2}) + \mathbf{n^1_{R_{x_2}}}.\end{aligned} \tag{13}$$

Equation (13) can be rearranged and represented as

$$\begin{aligned}\mathbf{y^1_{R_{x_2}}} &= (\mathbf{H_{R_{x_2}A_{n_1}}P_1} + \mathbf{H_{R_{x_2}A_{n_2}}P_3})\mathbf{x_1} \\ &\quad + (\mathbf{H_{R_{x_2}A_{n_1}}P_2} + \mathbf{H_{R_{x_2}A_{n_2}}P_4})\mathbf{x_2} + \mathbf{n^1_{R_{x_2}}}.\end{aligned} \tag{14}$$

The second mathematical term of (14) is essentially the desired term pertaining to $\mathbf{R_{x_2}}$, whereas the other remaining expressions are the undesired and unwanted terms that need to be removed to get red of the interference.

Equations (10) and (14) represent the received signals at $\mathbf{R_{x_1}}$ and $\mathbf{R_{x_2}}$ in time slot $\mathbf{t_1}$ respectively.

2) Designing Pre-coders: The superimposed pre-coders $\mathbf{P_1}$, $\mathbf{P_2}$, $\mathbf{P_3}$, and $\mathbf{P_4}$ are inspired by the works [17], [18], [19], [20], [21]. The desired terms for $\mathbf{R_{x_1}}$ are $(\mathbf{H_{R_{x_1}A_{n_1}}P_1} + \mathbf{H_{R_{x_1}A_{n_2}}P_3})\mathbf{x_1}$. For $\mathbf{R_{x_2}}$, the desired expressions are $(\mathbf{H_{R_{x_2}A_{n_1}}P_2} + \mathbf{H_{R_{x_2}A_{n_2}}P_4})\mathbf{x_2}$. First we have to set up conditions according to the desired and unwanted terms at $\mathbf{R_{x_1}}$ and $\mathbf{R_{x_2}}$.

Condition for $\mathbf{P_1}$, $\mathbf{P_2}$, $\mathbf{P_3}$, and $\mathbf{P_4}$ during $\mathbf{t_1}$:

To formulate the conditions for the pre-coders $\mathbf{P_1}$, $\mathbf{P_2}$, $\mathbf{P_3}$, and $\mathbf{P_4}$, the first term in the equation (10) is equated to identity matrix $\mathbf{I}$ and also the unwanted second term is made equal to null matrix $\mathbf{O}$. Furthermore, the second expression in equation (14) is the required one, thus is equated to identity matrix $\mathbf{I}$, and the undesired first term is equated to null matrix $\mathbf{O}$ as given below

$$\mathbf{H_{R_{x_1}A_{n_1}}P_1} + \mathbf{H_{R_{x_1}A_{n_2}}P_3} = \mathbf{I}. \tag{15}$$

$$\mathbf{H_{R_{x_2}A_{n_1}}P_1} + \mathbf{H_{R_{x_2}A_{n_2}}P_3} = \mathbf{O}. \tag{16}$$

From the above-obtained two equations we can get a conditional system of mathematical equations for the model such as

$$\begin{cases}\mathbf{H_{R_{x_1}A_{n_1}}P_1} + \mathbf{H_{R_{x_1}A_{n_2}}P_3} = \mathbf{I}, \\ \mathbf{H_{R_{x_2}A_{n_1}}P_1} + \mathbf{H_{R_{x_2}A_{n_2}}P_3} = \mathbf{O}.\end{cases} \tag{17}$$

$$\mathbf{H_{R_{x_2}A_{n_1}}P_2} + \mathbf{H_{R_{x_2}A_{n_2}}P_4} = \mathbf{I}. \tag{18}$$

$$\mathbf{H_{R_{x_1}A_{n_1}}P_2} + \mathbf{H_{R_{x_1}A_{n_2}}P_4} = \mathbf{O}. \tag{19}$$

From the above-obtained two equations we can get a conditional system of mathematical equations for the model such as

$$\begin{cases}\mathbf{H_{R_{x_2}A_{n_1}}P_2} + \mathbf{H_{R_{x_2}A_{n_2}}P_4} = \mathbf{I}, \\ \mathbf{H_{R_{x_1}A_{n_1}}P_2} + \mathbf{H_{R_{x_1}A_{n_2}}P_4} = \mathbf{O}.\end{cases} \tag{20}$$

The above-mentioned two systems of mathematical equations can together be solved to get the values of the precoders that need to be used at the transmitter to make sure the interference completely gets eliminated by the time the signal reaches the receiver due to using $\mathbf{P_1}$, $\mathbf{P_2}$, $\mathbf{P_3}$, and $\mathbf{P_4}$ such as

$$\begin{cases}\mathbf{P_1} = -\mathbf{H_{R_{x_2}A_{n_2}}} \\ (\mathbf{H_{R_{x_1}A_{n_2}}H_{R_{x_2}A_{n_1}}} - \mathbf{H_{R_{x_1}A_{n_1}}H_{R_{x_2}A_{n_2}}})^{-1}, \\ \mathbf{P_3} = -\mathbf{H_{R_{x_2}A_{n_1}}} \\ (\mathbf{H_{R_{x_2}A_{n_1}}H_{R_{x_1}A_{n_2}}} - \mathbf{H_{R_{x_1}A_{n_1}}H_{R_{x_2}A_{n_2}}})^{-1}.\end{cases} \tag{21}$$

$$\begin{cases}\mathbf{P_2} = -\mathbf{H_{R_{x_1}A_{n_2}}} \\ (\mathbf{H_{R_{x_1}A_{n_1}}H_{R_{x_2}A_{n_2}}} - \mathbf{H_{R_{x_1}A_{n_2}}H_{R_{x_2}A_{n_1}}})^{-1}, \\ \mathbf{P_4} = -\mathbf{H_{R_{x_1}A_{n_1}}} \\ (\mathbf{H_{R_{x_1}A_{n_1}}H_{R_{x_2}A_{n_2}}} - \mathbf{H_{R_{x_1}A_{n_2}}H_{R_{x_2}A_{n_1}}})^{-1}.\end{cases} \tag{22}$$

Using the above derived system of mathematical equations shown in (21) and (22), the undesired/unwanted terms and the channels effects are eliminated during the reception of one single time slot instead of two time slots $\mathbf{t_1}$ and $\mathbf{t_2}$ as is the cas in conventional STBC Method. Therefore, through (21) and (22) the equations (10) and (14) can be modified as

$$\mathbf{y^1_{R_{x_1}}} = \mathbf{x_1} + \mathbf{n^1_{R_{x_1}}}. \tag{23}$$

$$\mathbf{y^1_{R_{x_2}}} = \mathbf{x_2} + \mathbf{n^1_{R_{x_2}}}. \tag{24}$$

IV. Differences of the Proposed Method (SPBC) from Conventional NOMA

One might look at our proposed method and say that there seem to be some similarities with the conventional NOMA proposed in the literature [22]. Therefore, we thought that it would be useful to dedicate a section to talk about that. Motivated by this, the differences can be summarized in the following main points:

- NOMA is mainly aimed and specifically designed for multi-user scenarios to improve the overall system spectral efficiency; however, our proposed method is specifically focused on the single user scenario to improve the link spectral efficiency per device.
- NOMA always needs to employ successive interference cancellation (SIC) to remove the interference of far user from the near users. However, our proposed method does not use SIC at all, but rather depend on a specially designed precoded utilized at the transmitter to make

the signal free of interference by the time it reaches the receiver, which does not need to do any extra processing to get rid of the interference.

- NOMA can only work when there is a pathloss power difference between the near and far users, but completely fails when both receivers are in the same range. However, our proposed method can work in all scenarios including the worst-case one, where both receivers have more or less the same distance from the base station transmitter.
- NOMA needs power allocations to separate the far user signal from that of the near user; however, this is not a condition for our proposed techniques because it does not require a pathloss difference between users to make the transmission successful. In the contrary to that, our proposed technique can work without the need for power allocations that is needed to make the SIC successful in decoding and separating the signals, rather it depends on deploying a specially designed precoders that can cancel interference intelligently. This is so because NOMA operates without precoding, while our techniques utilizes precoders to cancel interference.
- NOMA is not inherently secure at the physical layer because the transmitted signals are not designed based on the channel variations of the receiver, aside from the fact that the near user has to decode the far user data signal before it decodes its own, which jeopardize the security of the system and make it vulnerable to the eavesdropping type of attacks. contrarily, our propsoed method is inherently secure at the physical layer because the signal data precoding are designed to be dependent on the channel specifications of the receiver.

V. Physical Layer Security Notion for SPBC

In present wireless communication systems, providing security is a must due to the ever-increasing number of users and IoT devices. However, security shouldn't be the cause of degradation of the system's performance. Authors in [23], propose a practical precoded orthogonal space-time block coding (POSTBC) method. An optimum matrix is used to precode the space-time codewords to minimize the error rate only for the legitimate user. According to the acquired results, there is and existing security gap region in the resulting BER performance as a consequence of using POSTBC. The authors also propose precoding along with a partial pre-equalizing (PCPPE) hybrid method, in which a new precoder is proposed, as a result of the original precoder, and a new unitary matrix that can map Bob's channel amplitudes into a 2D orthonormal matrix.

In [24], the authors have proposed physical layer security (PLS) scheme that can further enhance the security of STBC-OFDM via signals space diversity (SSD). This scheme works under the impression that the legitimate user has less error than the eavesdropper, even when the worst-case scenario occurs where the unauthorized receiver can somehow capture the interleaving strategy compromised by the transmitter and legitimate receiver. In the study [25], authors exploit the chaos communication technique and present an improved chaos MIMO scheme where PLS is realized with additional channel coding gain. In particular, a chaotic modulation symbol is multiplied by the data to be transmitted, and at the receiver, joint detection and chaos decoding is performed by maximum likelihood decoding (MLD).

On the contrary, our proposed SPBC method provides the PLS security solution against internal and external security vulnerabilities with less complexity. The systems of equations presented by (21), and (22) illustrate the values of pre-coders $\mathbf{P_1}$, $\mathbf{P_2}$, $\mathbf{P_3}$, and $\mathbf{P_4}$. In addition, these pre-coders are basically function of the wireless channel variations and characteristics thus yielding the desired outcome for security. Accordingly, the internal and external physical layer security threats can be avoided while securing the transmission of legitimate data to the intended receiver.

VI. Simulated Results

In this section, we evaluate and investigate the performance of the proposed SPBC method by adopting metrics such as bit error rate (BER), link throughput (TER), and peak to average power ratio (PAPR). The parameters considered in the computer simulation analysis of SPBC by using MATLAB software are given in Table I.

TABLE I
MATLAB Simulation Parameters

Channel	Multipath Rayleigh Fading Channel
Channel Length	9
Cyclic Prefix (CP)	9
IFFT/FFT Size	64
Number of Iterations	10000
Number of Symbols	9000000
Modulation Type	BPSK
Signal to noise ratio in dB	0-25 dB
Channel Delay	[0 3 5 6 8]
Channel Power Profile	[0 -8 -17 -21 -25]
Guard Interval Length	16

It is assumed that we have an OFDM system consisting of N_f=64 subcarriers with a CP size equals to the channel delay spread length L to eliminate inter-symbol interference (ISI). As shown in Table I, the channels between the transmitting and receiving antennas are assumed to be multi-path Rayleigh fading channels with a number of taps L=9.

As seen in Fig. 2, the BER performances depicted by Rx1-BER and Rx2-BER are much less (approximately at 15 dB SNR) than the conventional OFDM, theoretical STBC (with nTX=2, nRx=1), conventional Alamouti STBC (with nTX=2, nRx=1), and the transmit diversity with pre-coding (with nTX=2, nRx=1). The proposed design can ensure the best performance in comparison to conventional designs, and the case of PLS due to different wireless channels, the legitimate receiver can easily and securely receive and decode the transmitted signals.

In Fig. 3, the illustrations of the TER performances for SPBC, conventional STBC (with nTX=2, nRx=1), and transmit diversity with pre-coding (with nTX=2, nRx=1) are presented. The individual TER performances for the Rx1 and Rx2 are greater than the conventional STBC and similar to

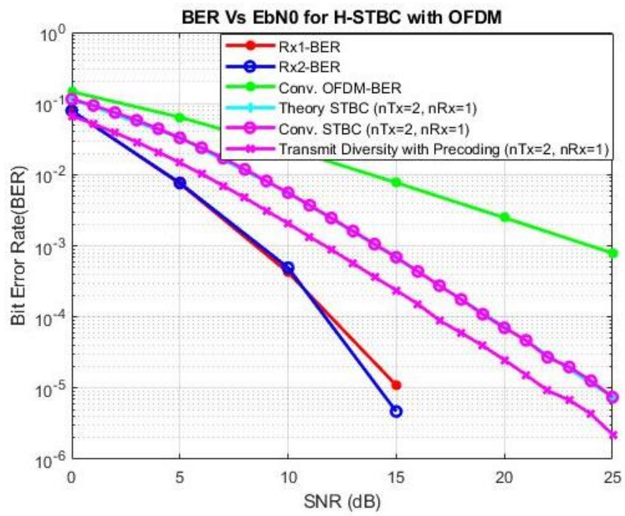


Fig. 2. Bit error rate (BER) performance of the proposed SPBC method compared to the conventional scheme.

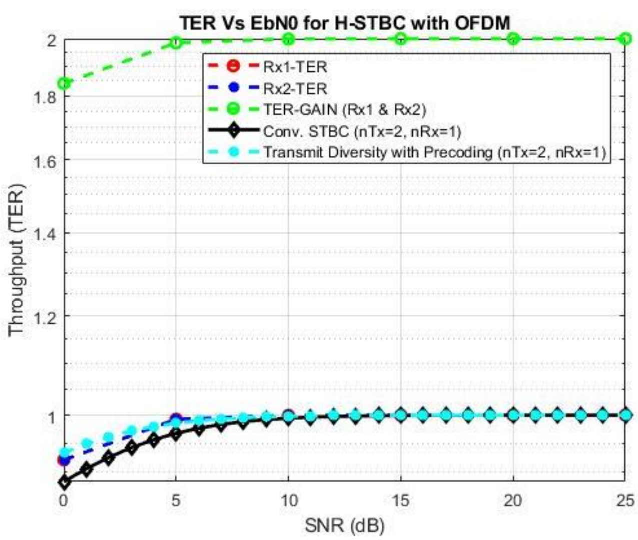


Fig. 3. Throughput (TER) performance of the proposed SPBC method compared to the conventional system.

transmit diversity. Furthermore, the overall performance gain for the SPBC denoted by TER-GAIN (Rx1 & Rx2) is doubled that of the conventional STBC and transmit diversity.

VII. Conclusion

In this article, we investigated the reliability performance of the introduced SPBC technique with the achievable double throughput for a single user. The results obtained during the computer simulations depicted that the proposed design outperforms the conventional comparable wireless systems. In addition, due to the resulting BER performance as a consequence of SPBC, the security against internal and external threats is realized. Moreover, the PAPR performance illustrates that the proposed design can also achieve better results in power saving compared to the conventional OFDM. Considering future works, this scheme can be extended to multiple users, and the security aspect can be further enhanced.

References

[1] M. Noor-A-Rahim, Z. Liu, H. Lee, M. O. Khyam, J. He, D. Pesch, K. Moessner, W. Saad, and H. V. Poor, "6G for vehicle-to-everything (V2X) communications: Enabling technologies, challenges, and opportunities," *Proceedings of the IEEE*, 2022.

[2] M. Vaezi, A. Azari, S. R. Khosravirad, M. Shirvanimoghaddam, M. M. Azari, D. Chasaki, and P. Popovski, "Cellular, wide-area, and non-terrestrial IoT: A survey on 5G advances and the road towards 6G," *IEEE Communications Surveys & Tutorials*, 2022.

[3] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on selected areas in communications*, vol. 16, no. 8, pp. 1451–1458, 1998.

[4] Y. Zhao, R. Adve, and T. J. Lim, "Precoding of orthogonal STBC with channel covariance feedback for minimum error probability," in *2004 IEEE 15th International Symposium on Personal, Indoor and Mobile Radio Communications (IEEE Cat. No. 04TH8754)*, vol. 1. IEEE, 2004, pp. 503–507.

[5] C.-K. Sung, I. Lee, and B. D. Choi, "An improved hybrid STBC scheme with precoding and decision feedback detection," in *IEEE 60th Vehicular Technology Conference, 2004. VTC2004-Fall. 2004*, vol. 2. IEEE, 2004, pp. 1328–1332.

[6] M. Vu and A. Paulraj, "Linear precoding for MIMO wireless correlated channels with non-zero means: K factor analysis, extension to non-orthogonal STBC," in *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, vol. 3. IEEE, 2005, pp. iii–1113.

[7] C. Meng and J. Tuqan, "PRECODED STBC-VBLAST for MIMO wireless communication systems," in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*, vol. 3. IEEE, 2007, pp. III–337.

[8] H.-J. An, J.-H. Kim, and H.-K. Song, "Cooperative transmission scheme to increase gain by using STBC," *Eng. Lett.*, vol. 15, no. 1, pp. 135–139, 2007.

[9] Y.-H. Chung and S.-M. Phoong, "Unitary precoders for ST-OFDM systems using alamouti STBC," *IEEE Transactions on circuits and Systems I: Regular Papers*, vol. 55, no. 9, pp. 2860–2869, 2008.

[10] R. Ran, J. Yang, and D. Kim, "Multimode precoder design for STBC with limited feedback in MIMO based wireless communication system," *IEICE electronics express*, vol. 6, no. 3, pp. 173–179, 2009.

[11] I. Baig, V. Jeoti, and M. Drieberg, "A ZCMT precoding based STBC MIMO-OFDM system with reduced PAPR," in *2011 National Postgraduate Conference*. IEEE, 2011, pp. 1–5.

[12] J. Yang, Z. Zhang, Z. Jiang, X. Jiang, and M. He, "Study on joint-coding based on pre-coding and STBC in coordinated multi-point system," in *2012 Spring Congress on Engineering and Technology*. IEEE, 2012, pp. 1–4.

[13] J.-H. Deng, S.-C. Jhan, and S.-Y. Huang, "A low complexity precoding transceiver design for double STBC system," *IEICE transactions on communications*, vol. 96, no. 4, pp. 1075–1080, 2013.

[14] T. Y. Elganimi and A. A. Elghariani, "Space-time block coded spatial modulation aided mmWave MIMO with hybrid precoding," in *2018 26th Signal Processing and Communications Applications Conference (SIU)*. IEEE, 2018, pp. 1–6.

[15] M. F. Khan, D. Pesch, A. Iqbal, S. Iqbal, and J. M. Hamamreh, "A novel multi-user space-time block coding based superposition transmission for future generation wireless networks," in *2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*, 2023, pp. 1–7.

[16] S. Iqbal and J. M. Hamamreh, "Modified MU-STBC with Round Transmission for Improving the Date Rate and Reliability of Future 6g and Beyond Wireless Networks," *RS Open Journal on Innovative Communication Technologies*, vol. 3, no. 8, may 31 2023, https://rs-ojict.pubpub.org/pub/ev4icpzz.

[17] J. P. Lemayian and J. M. Hamamreh, "Hybrid MIMO: a new transmission method for simultaneously achieving spatial multiplexing and diversity gains in MIMO systems," *RS Open Journal on Innovative Communication Technologies*, vol. 2, no. 4, 2021.

[18] ——, "Physical layer security analysis of hybrid MIMO technology," *RS Open Journal on Innovative Communication Technologies*, vol. 2, no. 5, p. 8, 2021.

[19] M. F. Zia, H. M. Furqan, and J. M. Hamamreh, “Multi-cell, Multi-user, and Multi-carrier Secure Communication Using Non-Orthogonal Signals’ Superposition with Dual-Transmission for IoT in 6g and Beyond,” *RS Open Journal on Innovative Communication Technologies*, vol. 2, no. 3, apr 6 2021, https://rs-ojict.pubpub.org/pub/om0x27nx.

[20] M. F. Zia and J. M. Hamamreh, “An advanced non-orthogonal multiple access security technique for future wireless communication networks,” *RS Open Journal on Innovative Communication Technologies*, vol. 1, no. 2, 2020.

[21] S. Iqbal and J. M. Hamamreh, “Precoded universal mimo superposition transmission for achieving optimal coverage and high throughput in 6g and beyond networks,” *Balkan Journal of Electrical and Computer Engineering*, vol. 11, no. 1, p. 25â€“34, 2023.

[22] H. M. Furqan, J. Hamamreh, H. Arslan *et al.*, “Physical layer security for noma: Requirements, merits, challenges, and recommendations,” *arXiv preprint arXiv:1905.05064*, 2019.

[23] J. M. Hamamreh, E. Guvenkaya, T. Baykas, and H. Arslan, “A practical physical-layer security method for precoded OSTBC-based systems,” in *2016 IEEE Wireless Communications and Networking Conference*. IEEE, 2016, pp. 1–6.

[24] M. A. Reşat, M. C. Karakoç, and S. Özyurt, “Improving physical layer security in alamouti OFDM systems with subcarrier coordinate interleaving,” *IET Communications*, vol. 14, no. 16, pp. 2687–2693, 2020.

[25] E. Okamoto, “A chaos MIMO transmission scheme for channel coding and physical-layer security,” *IEICE transactions on communications*, vol. 95, no. 4, pp. 1384–1392, 2012.

BIOGRAPHIES

Jehad M. Hamamreh is the Founder and Director of WISLAB (wislabi.com), and A. Professor at the Electrical and Electronics Engineering Department, Antalya Bilim University. He received his Ph.D. degree in telecommunication engineering and cyber systems from Istanbul Medipol University, Turkey, in 2018. Previously, he worked as a Researcher at the Department of Electrical and Computer Engineering at Texas AM University. He is the inventor of more than 20+ patents and an author of more than 85+ peer-reviewed scientific papers along with several book chapters. His innovative patented works won gold, silver, and bronze medals at numerous international invention contests and fairs. His current research interests include wireless physical and MAC layer security, orthogonal frequency-division multiplexing and multiple-input multiple-output systems, advanced waveform design, multidimensional modulation techniques, and orthogonal/non-orthogonal multiple access schemes for future wireless systems. He is a serial referee for various scientific journals as well as a TPC member for several international conferences. He is an Editor at Researcherstore, RS-OJICT journal, and Frontiers in Communications and Networks. Email: jehad.hamamreh@gmail.com.

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
An International Peer Reviewed, Indexed and Open Access Journal

ISSN: 2147- 284X
Year: March 2025
Volume: 13
Issue: 1

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