SAKARYA UNIVERSITY JOURNAL OF COMPUTER AND INFORMATION SCIENCES
http://saucis.sakarya.edu.tr/
e-ISSN: 2636-8129
Publisher: Sakarya University
Vol. 8, No. 4, 606-620, 2025
DOI: 10.35377/saucis...1634387

RESEARCH ARTICLE

Classification of Plant Diseases with ResNet-GAN Integration: Comparative Analysis of Machine Learning and Deep Learning Methods

Buse Çalışır[1*], Bihter Daş[1]

[1] Firat University, Faculty of Technology, Department of Software Engineering, Elazig, Türkiye, ror.org/ 05teb7b63

Corresponding author:
Buse Çalışır, Fırat University, Faculty of Technology, Department of Software Engineering, Elazıg, Türkiye,
bcalsr2709@outlook.com

Article History:
Received: 06.02.2025
Revised: 03.08.2025
Accepted: 31.08.2025
Published Online: 13.10.2025

ABSTRACT
Accurate and effective classification of plant diseases is critical for increasing yield and quality in agricultural production, minimizing economic losses through early detection of diseases, and implementing sustainable agriculture approaches. This study presents an approach for detecting and classifying plant leaf diseases. We compare the performance of machine learning and deep learning-based models, and we use GAN-based data synthesis methods on a dataset we created to improve the model performance. ResNet-based feature extraction is performed for machine learning methods, and XGBoost, Random Forest, SVM, and InceptionV3 models are evaluated. In contrast, AlexNet, VGG16, VGG19, DenseNet, and ResNet models are examined within the scope of deep learning. The study was analyzed in three classes: Phytophthora Infestans, Potassium Deficiency, and Healthy, and tested on data obtained from 21 different plant species. According to the model performances obtained, the deep learning-based ResNet model showed the highest success in all performance metrics and achieved 98% accuracy, showing superior performance compared to other methods. In the study, a comprehensive evaluation of multiple classification, GAN-based data synthesis, machine learning, and deep learning models was carried out. A valuable contribution was made to the existing studies in the literature.

Keywords: Plant disease classification, Data synthesis, Deep learning, Machine learning, Image processing

1. Introduction

Agriculture is the foundation of the economic structure of many countries and is recognized as an important factor in national income and employment [1]. Plant leaf diseases are a serious threat that negatively affects agricultural productivity, and research in this field is considered important for the sustainability of agricultural production. Additionally, it is one of the primary factors directly affecting food production, negatively impacting gains in both crop productivity and economic efficiency [2]. Plant diseases are among the factors that cause yield and quality loss in agricultural production. Plant foliar diseases are a serious agricultural problem caused by various factors, including fungi, bacteria, viruses, and nutrient deficiencies, which cause significant damage to agricultural production. Symptoms such as root and fruit rot, leaf blight and spots, wilting and shriveling are observed in plants affected by diseases [3].

The agricultural sector plays a critical role in economic development by ensuring food security, sustaining the food supply, creating employment, promoting economic growth, and supporting rural development. However, plant diseases threaten agricultural production, severely reducing crop yields and quality and causing significant economic losses. Biotic and abiotic factors, such as viruses, nutrient deficiencies, fungi, and bacteria, can hinder the healthy growth of plants and result in production losses. Therefore, early detection of diseases is crucial for preventing their spread and minimizing economic losses. Especially when examining leaf images to determine whether a disease is present in the plant or its type, this can provide important information for early diagnosis.

Traditional methods of detecting plant diseases are often based on visual assessments. Experts diagnose the disease by examining the symptoms on the leaf. However, this manual diagnosis process is time-consuming and labor-intensive. It also has disadvantages, such as requiring expertise and increasing costs. Considering these drawbacks, machine learning and deep learning methods can automatically detect plant diseases, resulting in faster, reliable and cost-effective disease diagnosis. Artificial intelligence-based systems analyze leaf images to identify disease symptoms, enabling effective, highly accurate, and cost-effective diagnosis. Machine learning and deep learning models can accurately detect disease symptoms when trained on large datasets. This study considers two different plant diseases: Phytophthora Infestans and potassium deficiency. Phytophthora Infestans is a serious plant disease caused by a fungus-like microorganism, particularly on plant leaves.

Cite as: B. Çalışır and B. Daş, Classification of plant diseases with ResNet-GAN integration: Comparative analysis of machine learning and deep learning methods. *Sakarya University Journal of Computer and Information Sciences*, vol. 8, no. 4, pp. 606–620, 2025. doi: 10.35377/saucis...1634387

Symptoms of the disease include yellowing of the leaf surface, formation of a white fungal layer on the undersurface, defoliation, and, in later stages of the disease, plant rot. Potassium deficiency is a nutritional disorder resulting from a lack of potassium, a vital plant nutrient. Potassium is a vital macronutrient for the healthy growth and development of plants. It helps cells maintain water balance, increases the efficiency of the photosynthesis process and ensures that basic metabolic functions continue to run smoothly. Symptoms of potassium deficiency include yellowing or bruising of leaf margins, slowed growth, curling and drying of leaves. These diseases are characterized by poor fruit quality, low yield, and economic damage. Automatic diagnosis using artificial intelligence-based methods detects the symptoms of these diseases, allowing the farmer to intervene early and effectively. Thus, agricultural damages are minimized, chemical intervention is optimized, and crop productivity is increased. Early diagnosis and correct classification of both diseases are important for protecting agricultural productivity.

The main motivation of this paper is to minimize losses in agricultural production through early and accurate diagnosis of plant diseases. It also aims to help farmers take more informed and effective interventions against diseases. It can accurately classify disease symptoms using machine learning and deep learning methods. In this context, the dataset used in artificial intelligence-based methods for plant diseases comprises 21 species and is categorized into three classes. The results of the methods used are compared, and the methods that provide the highest accuracy in disease classification are analyzed.

1.1 Problem Statement

The study investigates the detection and classification of plant leaf diseases using artificial intelligence-based methods. Diseases are detected and classified using machine learning and deep learning models. Plant diseases are categorized into three classes and 21 species. The methods used were tested on our data, and their performances were compared.

1.2 Main Contributions

Our main contributions in this study are the following:

- General plant images were examined, and the disease type was identified.
- Multiple methods were compared using deep learning and machine learning methods.
- Disease detection is not limited to a single plant species but is evaluated by examining 21 species.
- The study examined three classes (Phytophthora Infestans, Potassium Deficiency and Healthy).
- A new dataset was created for the study, and data augmentation was performed using a Generative Adversarial Network (GAN) based data synthesis method.

These contributions are crucial for enhancing model performance and generalizability in agricultural disease detection processes.

The rest of the paper is structured as follows. Section 2 reviews the related work in the literature, emphasizing previous research. Section 3 provides detailed information about the used dataset, data preprocessing steps, data augmentation techniques, the applied methods and the feature extraction process. In Section 4, the experimental results are analyzed and discussed. Finally, Section 5 summarizes the study's results and presents the findings.

2. Related Works

In this section, machine learning, deep learning and other methods used in analyzing and classifying plant diseases are discussed by reviewing the studies in the literature in detail.

Upadhyay et al. [4] utilized deep learning techniques to detect and classify diseases in rice plants. In this study, a CNN model was designed by analyzing the features such as size, shape, and hue of the damage on the leaf surface. The Otsu thresholding method was used in image processing to improve data quality and reduce noise.

Yao et al. [5] aimed to detect foliar diseases early to reduce pesticide use in tropical regions. They investigated the automatic detection of diseases in detail using deep learning methods, including YOLO, SSD, CNN, and ResNet. They also address attention mechanisms, such as leaf segmentation and disease classification, and offer solutions to the challenges faced in tropical regions, including leaf overlap and multi-scale detection.

Li et al. [6] demonstrated that deep learning models, such as Convolutional Neural Networks (CNNs), yield faster, more accurate, and objective results in plant disease detection compared to traditional methods. In addition, it was emphasized that effective and successful results can be obtained even with limited data sets thanks to transfer learning.

Lu et al. [7] in their study discussed the use of CNN-based deep learning models in the classification of plant leaf diseases. In the study, the advantages and disadvantages of these models are analyzed, and solutions to the challenges encountered are presented. The paper focuses on problems such as insufficient data and the diversity of disease patterns, and shows that data augmentation and transfer learning methods effectively address these issues. In the study, a comparison of CNN-based models is made, and suggestions for developing more efficient models are examined in response to these models.

In detail, Saleem et al. [8] examined deep learning models and optimization algorithms for plant disease classification. They evaluated 18 different Convolutional Neural Network (CNN) models on the PlantVillage dataset and compared them on performance metrics. This study emphasizes that the accuracy and generalization ability of deep learning models can be improved by using transfer learning and data augmentation techniques to analyze limited datasets.

Shrivastava et al. [9] developed a machine learning model that detects diseases in rice plants using only color features. They utilized 14 color spaces and extracted four statistical features for each color channel, resulting in a total of 172 features. A dataset with four different leaf states was used, and tests were performed with seven different classifiers. This study demonstrates that color features can be an effective method for detecting plant diseases. It is also stated that this approach can be improved with larger datasets and deep learning models.

Borhani et al. [10] developed lightweight and real-time deep learning models for plant disease detection by combining Vision Transformer (ViT) and Convolutional Neural Network (CNN) based approaches. In the study, the proposed methods were tested on three different datasets. As a result of these tests, it was observed that ViT-based models provide high accuracy but are computationally costly. However, it is stated that ViT+CNN-based models provide both speed and high accuracy values.

Argüeso et al. [11] developed a deep learning model for classifying plant diseases using Few-Shot Learning (FSL) methods using plant images collected from a specific area. The study compares the Triplet loss method and classical transfer learning of FSL-based Siamese networks. It is demonstrated that the Triplet loss-based approach is more effective in class separation and outperforms classical methods in low-data scenarios.

Zhao et al. [12] propose RIC-Net, a deep learning model optimized with Convolutional Block Attention Module (CBAM), which integrates Inception and residual structures, to classify plant leaf diseases. The study focuses on maize, potato and tomato crops, and shows a high success rate in detecting diseases in these crops. Moreover, the model's compactness, its ability to be trained quickly, and its high recognition accuracy make it more efficient compared to existing deep learning models. Thanks to its fast operation and low computational requirements, it is emphasized that it can be effectively applied even in systems with limited resources.

Haridasan et al. [13] developed a system that combines deep learning and image processing techniques to detect and classify five main diseases in paddy plants. Various image processing and segmentation methods were applied to identify the diseases. Then, Support Vector Machines (SVM) and Convolutional Neural Networks (CNN) were used in the classification phase to compare the results. In the evaluations, it was determined that the SVM model provides higher accuracy and has the potential to be used in disease treatment.

Rao et al. [14] proposed a more effective and comprehensive classification method by combining the powerful image processing capabilities of CNN with the capacity of GCN to model structural relationships in plant disease classification. Among these models, the CNN-based model enables the extraction of features of image data. In contrast, the GCN-based model increases the classification success by analyzing the connectional data between plants. The study used a dataset with 31 classes obtained from Kaggle, and the results were evaluated using the CNN-GCN model. As a result of this comparison, it was observed that the proposed model had a higher accuracy rate.

Topcu et al. [15] evaluated the performance of the ResNet model in detecting plant diseases and found that the model can classify different types of diseases with high accuracy. The ResNet-based model demonstrated high accuracy in disease detection, achieving successful results even under varying outdoor conditions. The study demonstrated that ResNet provides accurate diagnoses and offers an effective solution for sustainable agriculture.

Applalanaidu et al. [16] compared machine learning algorithms and deep learning models by analyzing articles published between 2017 and 2020. The authors examined the factors affecting classification accuracy by analyzing the diversity of data processing techniques and datasets. The study shows that machine learning models can work with more limited data, while deep learning models provide more accurate results thanks to automatically extracting features.

Zhao et al. [17] developed a model called MAFDE-DN4, which incorporates a multi-level attention mechanism, to classify plant diseases. Based on the Few-Shot Learning (FSL) approach, this model aims to detect new disease types with high accuracy with limited data. To enable the model to perform a more detailed analysis on complex images, a bidirectional weighted feature integration module (BWFM) and an episodic attention module (EA) that generates context-based attention maps are integrated.

C.K. et al. [18] developed a model that integrates the MFFN network and the ACSPAM attention mechanism to detect and classify tomato diseases. The model was tested on 10 classes of tomato leaf diseases from the PlantVillage dataset, utilizing a ResNet50-based method to address the imbalanced data problem. By optimizing both local and global features, the ACSPAM-MFFN method has been adapted for real-time applications, providing an innovative solution in agricultural technologies.

In addition to these studies, details on the datasets used in the literature and the articles utilizing these sets are presented in Table 1. Upon analyzing Table 1, it becomes apparent that Kaggle and PlantVillage datasets are the most frequently used in the studies.

Table 1. Plant Disease Detection Studies in the Literature

Ref	Data set	Constraint	Contribution	Models
[4]	Kaggle	A single plant species was studied.	Performance is improved by noise reduction using the Otsu thresholding method.	CNN
[5]	Tomato Disease Rice Disease Apple Leaf Plant Pathology 2021	Complex backgrounds in images, limited data sets and computational intensity on devices	A comprehensive review of how deep learning can be applied in tropical plant diseases has been conducted, and a mobile application has been developed.	YOLO, SSD, CNN, MaxViT, ResNet, MobileNet, DenseNet
[6]	PlantVillage Plant Pathology Challenge Self-Created	Hyperspectral images are costly, and real data are limited.	Using Few-Shot Learning and transfer learning, hyperspectral applications have been made.	AlexNet, GoogLeNet, VGG16, ResNet50, FSL
[7]	Kaggle, BIFROST	Insufficient data sets and complex images	Studies on classification with CNN are summarized.	AlexNet, VGGNet, ResNet, GoogLeNet, MobileNet, EfficientNet
[8]	PlantVillage	Data imbalance and long training period	18 Different CNN architectures are compared.	AlexNet, VGG-16, ResNet-50, DenseNet, Inception-v4, MobileNet, Xception, Improved GoogLeNet, Modified MobileNet, Reduced MobileNet, LeafNet, MLCNN, Cascaded AlexNet & GoogLeNet, Hybrid AlexNet & VGG
[9]	Problem-specific dataset	Limited capture of complex features, as prediction is based on color features	A total of 172 color features are extracted in 14 different color spaces, and a wide analysis is performed. Classification is made according to color features.	SVM, Naïve Bayes, DT, RF, LR, DC
[10]	Wheat Rust Classification Rice Leaf Disease PlantVillage	Limited data and data imbalance	Comparison of ViT, CNN and hybrid models	ViT, CNN, EfficientNet, ResNet, Inception
[11]	PlantVillage	Some plant species have few images.	By using FSL, new classes are learned with less data. With Triplet Loss, class boundaries are learned more efficiently.	Siamese networks (Contrastive and Triplet Loss), InceptionV3, SVM

Table 1 (Continued)				
[12]	PlantVillage	Single disease images were used, and images were created in the laboratory.	With RIC-Net, fewer parameters, faster training, and higher accuracy were achieved, and with CBAM, they developed an attention mechanism.	RIC-Net, InceptionV3, Xception, ResNet50, Inception-ResNetV2
[13]	Problem-specific dataset	The images are from the Kerala region of India, and only images of the rice plant were taken.	A low-cost and fast system for the automatic detection and early diagnosis of rice leaf diseases is proposed, utilizing SVM and CNN models.	SVM, CNN
[14]	Kaggle Plant Leaf Image	Data imbalance between classes	A new hybrid model (CNN+GCN) is proposed, and accuracy is improved by integrating visual and relational information using a combined CNN and GCN model.	CNN (ResNet18), GCN, Hybrid CNN+GCN model
[15]	Plant Diseases	Imbalance and limited diversity of data	A model with a high accuracy rate is proposed, enabling early and accurate diagnosis of plant diseases.	ResNet
[16]	PlantVillage various specialized data sets	Most models are trained on data from a laboratory environment, which can degrade their performance in a real-world setting.	Suggestions for accurate and rapid classification of plant diseases are presented.	SVM, KNN, Naive Bayes, AlexNet, GoogLeNet, VGGNet, InceptionV3, ResNet, XceptionNet, etc.
[17]	PlantVillage various specialized data sets	Intra-class diversity and inter-class similarity problems exist	A DN4-based MAFDE-DN4 model was developed, utilizing multi-scale feature fusion with BWFM and scene-specific attention maps via the EA module, and class distinction was enhanced by incorporating Cosine margin.	DN4, BWFM, EA
[18]	PlantVillage	The similarity between some classes in the dataset is high.	Multilayer feature fusion is implemented with MFFN, and an Adaptive attention mechanism is integrated.	MFFN, ACSPAM

This study creates and analyzes a three-class dataset by collecting images from 21 plant species. In the literature, it is often observed that studies are conducted with ready-made datasets or datasets created under laboratory conditions. Unlike other studies, this study aims to enhance the model's generalization capacity by utilizing a dataset collected under real environmental conditions. Classification of plant diseases has been performed using machine learning and deep learning models, and the performance metrics of these models have been thoroughly compared. This study aims to address a significant gap in the literature with its unique dataset and methodological approach.

3. Materials and Methods

This study evaluates machine learning and deep learning models for the detection and classification of plant leaf diseases. In the preprocessing step, data augmentation, GAN-based data synthesis, and image scaling techniques are applied to enhance the accuracy and generalization capabilities of the models. ResNet-based feature extraction was performed during the machine learning stage, and models including XGBoost, InceptionV3, SVM, and Random Forest were utilized. In contrast, in the deep learning stage, AlexNet, VGG16, VGG19, DenseNet and ResNet models were used to analyze and classify the image data directly. The methods were applied to diagnose diseases belonging to Phytophthora Infestans, Potassium Deficiency and Healthy classes. Figure1. shows the system architecture. The proposed method enables reliable and successful classification by automating agricultural disease detection processes, thereby contributing to informed decision-making.

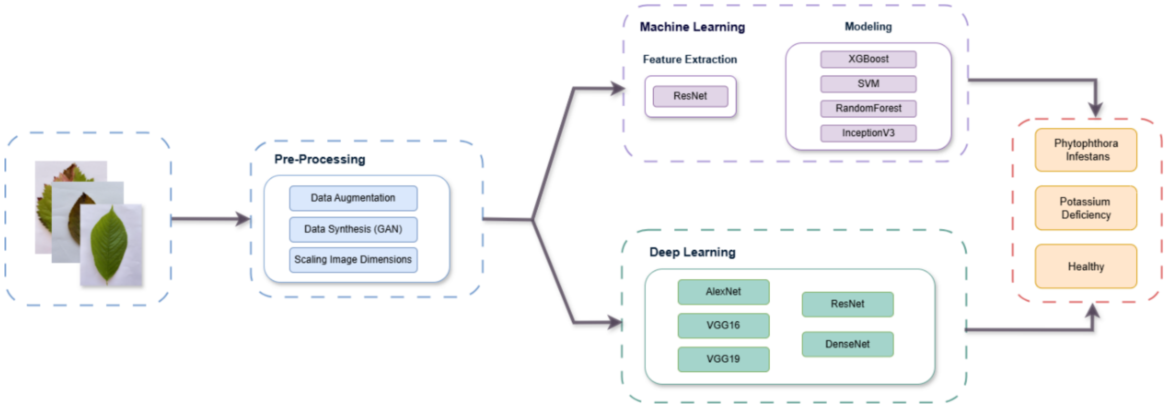

Figure 1. System architecture

3.1 Datasets

Our dataset consists of images taken from various Ankara plant growers and gardens in Elazığ. The dataset comprises images of 21 different plant species, captured with a camera featuring a resolution of 4284 x 5712 pixels and an aperture of f/1.6. The study analyzes the images in three classes: 'Phytophthora Infestans', 'Potassium Deficiency' and 'Healthy'. The dataset consists of 200 images, including 50 samples of Phytophthora Infestans, 50 samples of Potassium Deficiency, and 100 Healthy samples. This dataset was augmented using the Generative Adversarial Network (GAN) model, and the number of data points for each class was expanded to 250 for Phytophthora Infestans, 250 for Potassium Deficiency, and 500 for Healthy. Figure 2 shows some examples from our dataset.

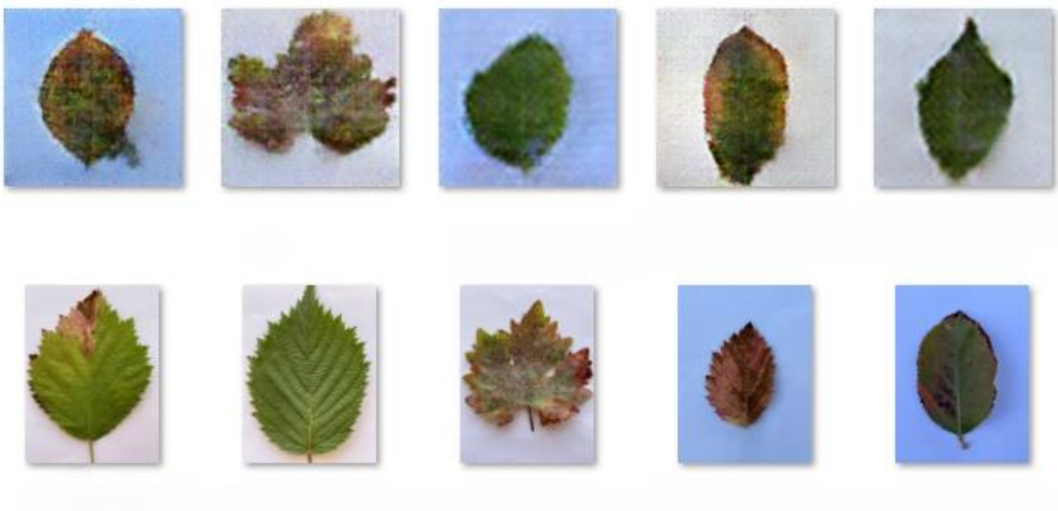

Figure 2. Plant Leaves Dataset

3.2 Preprocessing and Data Augmentation

In this study, preprocessing steps were performed to enhance the image data and bring it into a format suitable for classification. Within the scope of data augmentation, rotation, brightness adjustment, exposure operations, and Generative Adversarial Network (GAN)-based data augmentation methods were employed. All images were then scaled to a 640x640 pixel size. The data was first converted from one-hot vector format to integer encoding format for labeling. The dataset was divided into approximately 80% training and 20% testing sets in each layer using the 5-fold cross-validation method to train and evaluate the models. Different samples were included in the test set in each layer to evaluate the model's generalization performance. Synthetic images produced with GAN were only used in the training dataset and were not included in the testing phase.

GAN:

Generative Adversarial Networks (GANs) are a deep learning model developed by Ian Goodfellow in 2014 [19]. GANs play a crucial role in generative AI by offering an efficient method for generating synthetic images during the data augmentation process. As shown in Figure 3, the GAN architecture comprises two neural network components: the generator, which generates data, and the discriminator, which determines whether the generated data is real or not. While the generator increases data diversity by generating synthetic images, the discriminator is trained to distinguish these images from real data [20].

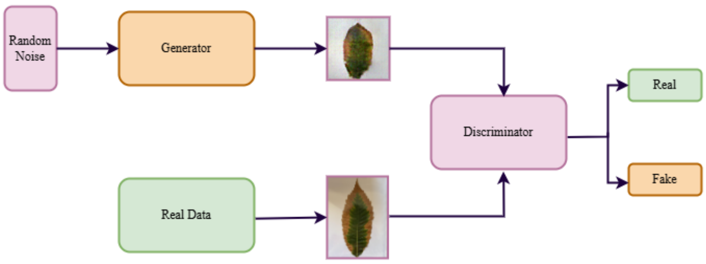

Figure 3. Traditional GAN Architecture

GANs are built on a game-theoretic learning model in which two neural networks, the generator and the discriminator, compete [21]. In this structure, the generator generates fake image samples using a random input vector z (G(z)), while the discriminator attempts to distinguish between real and fake images. Working in constant competition, these two networks improve each other over time to produce more realistic images. The mathematical function of GAN is defined as in Equation 1 [22].

$$\min_G \max_D V\,(D,G) = E_{x\sim P_{data(x)}}[\log D(x)] + E_{z\sim P_{data(z)}}\left[\log\left(1 - D\left(G(z)\right)\right)\right] \quad (1)$$

In Equation (1), Ex~ pdata (x) represents the real data, and Ex~ pdata (z) represents the expected value of the generated fake data. The function is based on a competitive process in which the discriminator aims, as far as possible, to obtain correct results from the real image data (D(x)) and to detect fake data. At the same time, the generator tries to pass off the fake image data as real data. This structure increases the generator's ability to generate increasingly realistic image data.

3.3 Methods and Feature Extraction

For deep learning methods, the following models were used for classification.

ResNet:

It is a convolutional neural network (CNN) model widely used in image classification and feature extraction, introduced by Kaiming He and colleagues at Microsoft Research Asia in 2015 [23]. The basic approach of ResNet is to overcome the gradient loss problem in deep networks by using residual learning connections to facilitate the learning process and support the learning of deeper features through multilayer structures [24]. ResNet addresses issues such as gradient loss and model distortion by utilizing residual structures, allowing for more efficient training of deep network models without the need for Dropout [25]. The pre-learned weights of the ResNet model on the ImageNet dataset can be easily downloaded by transfer learning or retraining techniques. They can be adapted to different problems by changing the final layers. Additionally, the original datasets can be enriched with data augmentation techniques to enhance the model's accuracy and overall performance [26].

VGG16/VGG19:

VGG16 and VGG19 are convolutional neural network (CNN) models developed by the Visual Geometry Group (VGG), which are widely used in computer vision. These models perform particularly well in tasks such as image classification, object detection and feature extraction [27]. While VGG16 consists of 13 convolutional layers, 3 fully connected layers, 5 Max-Pooling layers, ReLU, Dropout, and Softmax layers, the difference between the VGG19 model is that it includes 16 convolutional layers [28]. Using 3x3 convolution filters and Max Pooling layers in the VGG16 and VGG19 models, the computational load is balanced, and the model runs efficiently. Due to their simple applicability and good performance, these models are widely used in deep learning [29].

DenseNet:

DenseNet employs an enhanced design featuring cross-layer connections to address ResNet's issue of insufficient data transmission between layers. It provides the ability to clearly distinguish between new and existing data, as it gradually incorporates important features in the last layer [30].

AlexNet:

AlexNet is a convolutional neural network model designed by Alex Krizhevsky, Ilya Sutskever and Geoffrey Hinton in 2012 [31]. The AlexNet architecture has been a milestone in deep learning with its success in ImageNet competitions. It has gained popularity due to its success in visual data processing and has attracted considerable attention in studies [32]. This architecture consists of layers that come together sequentially, moving away from convolution and pooling layers. This is because increasing the number of layers and filters increases memory, computational cost and the possibility of memorizing the model. AlexNet was developed to solve this problem by creating parallel connected components [33].

Regulatory Techniques Against Overfitting Risk:

Various regulations and control techniques were applied to prevent overfitting during the training process of deep learning models. In particular, by incorporating Dropout and Batch Normalization layers into the models, the network's generalization ability was enhanced, and the impact of unnecessary weights during learning was mitigated. With the Early Stopping strategy, the risk of overlearning is minimized by terminating the training process when the validation loss fails to improve for a specified number of epochs. In addition, the ReduceLROnPlateau (ReduceLROnPlateau) mechanism makes the model more sensitive to changes in the verification loss. To evaluate the model's performance more reliably, the 5-fold cross-validation method was applied, and the model's generalization capability was measured using different data samples at each fold. The models performed well against both the training data and previously unseen test data, thanks to these comprehensive strategies.

Feature Extraction:

In this study, feature extraction was performed using the ResNet50 model for machine learning applications. Classification was performed using different deep learning models presented in Table 2, and the related data was analyzed. As a result of the evaluations, the ResNet method was identified as the most successful technique. A detailed application study was conducted in Section 4, in line with the results obtained.

Table 2. Hyperparameters Used for Deep Learning Methods

Methods	**Activation Function**	**Epoch**	**Optimization Function**	**Learning Rate**
AlexNet	ReLU, Softmax	100	Adam	0.00005
VGG16	ReLU, Softmax	100	Adam	0.0001
VGG19	ReLU, Softmax	100	Adam	0.00005
DenseNet	ReLU, Softmax	100	Adam	0.00005
ResNet	ReLU, Softmax	100	Adam	0.00005

4. Experimental Results and Discussion

4.1. Experimental Results

This study aims to accurately and efficiently classify plant leaf diseases. For this purpose, a large-scale analysis was performed using different deep learning and machine learning models. The models used include Convolutional Neural Network (CNN)- based deep learning models (VGG16, VGG19, DenseNet, ResNet, and AlexNet), as well as traditional machine learning algorithms (XGBoost, SVM, Random Forest, and InceptionV3). The performance of each model is evaluated in detail with metrics such as accuracy, precision, recall, F1 score and AUC. Additionally, the total time spent training and testing the models was also compared and included in the analysis. The results of this evaluation are shown in Table 3. These metrics reveal both the overall success level of the models and their ability to discriminate between classes.

Table 3. Performance Evaluation Table of Models

Methods	Precision	Recall	F1-Score	Accuracy	AUC	Time (seconds)
XGBoost	0,82	0,82	0,81	0,82	0,93	3105.23
SVM	0,82	0,80	0,79	0,80	0,92	1123.00
RandomForest	0,79	0,77	0,75	0,77	0,90	1067.08
InceptionV3	0,92	0,92	0,92	0,92	0,97	6367.52
VGG16	0,90	0.89	0.89	0.89	0,97	2383.78
VGG19	0.88	0.87	0.87	0.87	0,96	2760.99
AlexNet	0,97	0,96	0,96	0,96	0.99	2399.23
DenseNet	0.93	0.93	0.93	0.93	0.98	2716.87
ResNet	0,98	0,98	0,98	0,98	0,99	2082.62

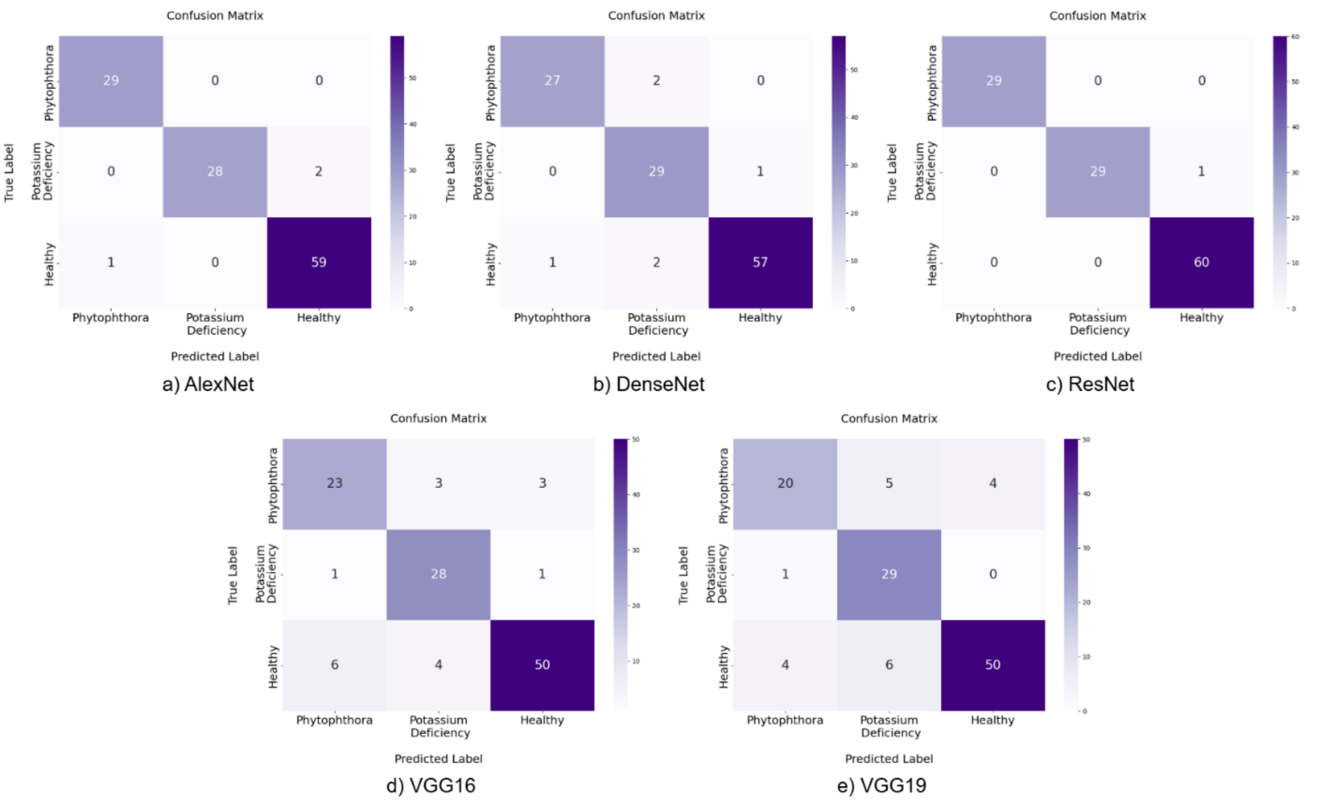

Figure 4. Deep Learning Models Confusion Matrices

The study presents the complexity matrices of the models, as well as the correct and incorrect predictions of each model across the classes. To evaluate the classification success of deep learning methods for Phytophthora Infestans, Potassium Deficiency, and Healthy Samples, Figure 4 shows the confusion matrices. The DenseNet model demonstrated exceptionally high accuracy in the Healthy class and achieved a highly successful performance in the Potassium Deficiency class. The

ResNet model performed the best among all classes, maximizing the correct classification performance and significantly reducing the error rate. In particular, it correctly predicted all instances in the Healthy and Potassium Deficiency classes, misclassifying only one instance in the Phytophthora Infestans class, resulting in very high overall accuracy. The DenseNet model demonstrated high accuracy in the Potassium Deficiency and Healthy classes, while experiencing limited confusion in the Phytophthora Infestans class, misclassifying only two instances. The AlexNet model made a few incorrect predictions in the Phytophthora Infestans and Potassium Deficiency classes and showed consistent success in the Healthy class. The VGG16 model exhibited significant confusion in the Phytophthora Infestans class, while achieving acceptable accuracy in the Potassium Deficiency and Healthy classes. The VGG19 model again experienced the most confusion in the Phytophthora Infestans class, but achieved good accuracy in the Potassium Deficiency class and reasonable success for the Healthy class. The matrices in Figure 4 serve as a crucial reference for evaluating the models' performance in real environmental conditions.

According to the overall accuracy analysis across classes, ResNet and DenseNet models have the highest success rate compared to other models.

In Figures 5 and 6, the accuracy and loss plots are analyzed during the training and validation processes to examine the models' learning process in detail. The accuracy plots in Figure 5 show the learning ability of the models over time. The AlexNet model exhibited a steady increase in training accuracy during the first phase, but the validation accuracy plateaued and increased marginally after approximately epoch 15. In the DenseNet model, the training and validation curves initially increased rapidly. However, in the later epoch, the validation accuracy slightly lagged behind the training accuracy; however, high accuracy rates were still achieved. The ResNet model demonstrated the best performance in the validation accuracy curve, and both training and validation accuracy stabilized relatively early, reaching 98% around epoch 10. The accuracy curves of the VGG16 and VGG19 models generally increased in parallel, but the verification accuracy of the VGG19 model fluctuated across epochs. In this study, the parallel training and validation accuracy graphs indicate that the models do not overfit and exhibit high generalization ability.

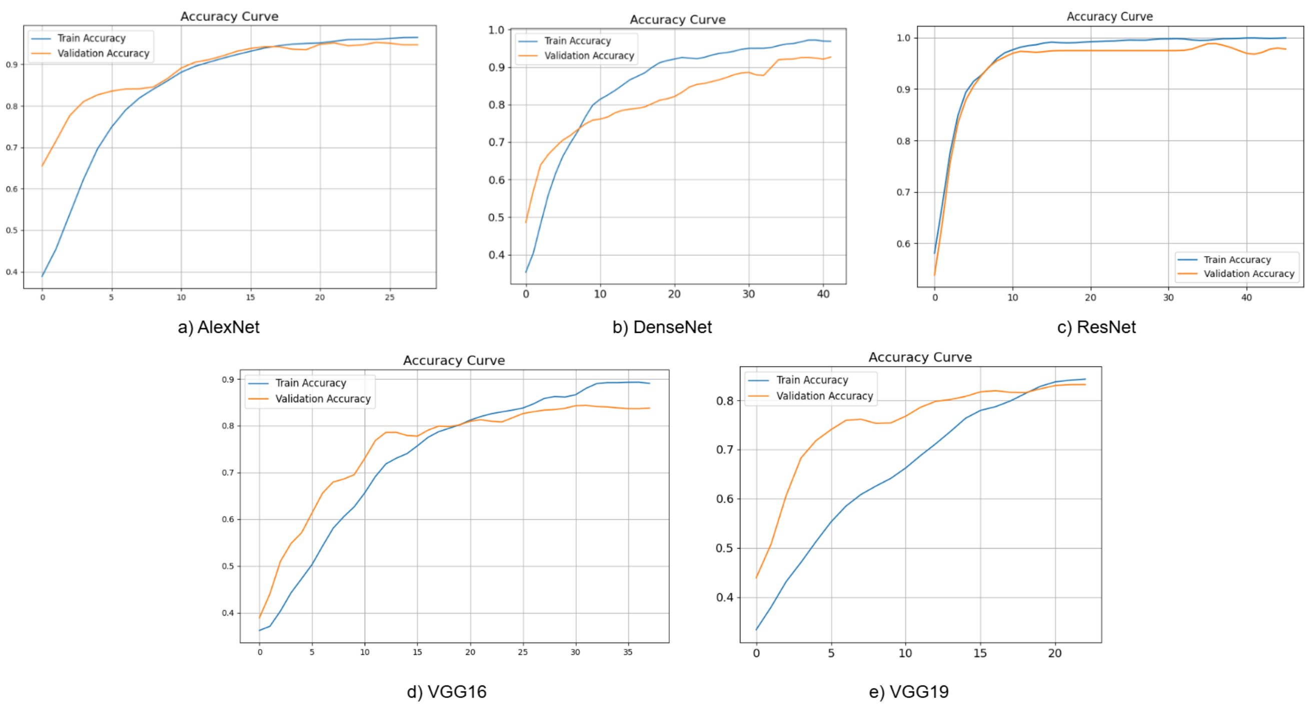

Figure 5. Accuracy Comparison of Deep Learning Models for Classification of Plant Leaf Diseases

The loss graphs shown in Figure 6 are examined to understand and evaluate the optimization performance of the deep learning models used in the learning process. In the AlexNet model, training and validation losses decreased consistently, reaching minimum levels around epoch 25. In the DenseNet model, the loss curves decreased in parallel, with the training loss being lower than the verification loss, and the trend continued to decrease steadily. In the ResNet model, the validation loss decreased to a very low value after epoch 10 and stabilized alongside the training loss, indicating strong learning performance. The verification loss generally decreased in the VGG16 and VGG19 models; however, in the VGG19 model, significant fluctuations were observed from the early epochs, and the verification loss increased occasionally. The balanced progress of accuracy and loss values in the training and validation processes in the study demonstrated that the DenseNet and ResNet models exhibited the best learning performance during the learning processes. The other models also performed well in generalization, albeit with minimal fluctuations in accuracy and loss value.

Figure 7. ROC curves visualize the discrimination levels between classes by evaluating the classification performance of the models at different thresholds. The AUC values calculated for each model reflect the model's ability to discriminate between classes accurately. The AlexNet model showed strong discrimination performance, achieving an AUC value of 1.00 for the Phytophthora and Potassium Deficiency classes and 0.96 for the Healthy class. The DenseNet model demonstrated high accuracy, with an AUC value of 1.00 for the Phytophthora and Healthy classes, and 0.99 for the Potassium Deficiency class. The ResNet model achieved ideal classification performance with an AUC value of 1.00 in all classes. The VGG16 model demonstrated satisfactory discrimination performance, with an AUC value of 0.93 for all three classes. The VGG19 model maintained its discrimination power between classes, with an AUC of 0.96 for the Phytophthora and Potassium Deficiency classes and 0.95 for the Healthy class. However, it showed lower performance than the other models. In line with these results, the ROC curves and AUC values indicate that the ResNet model exhibits the highest discrimination across all classes.

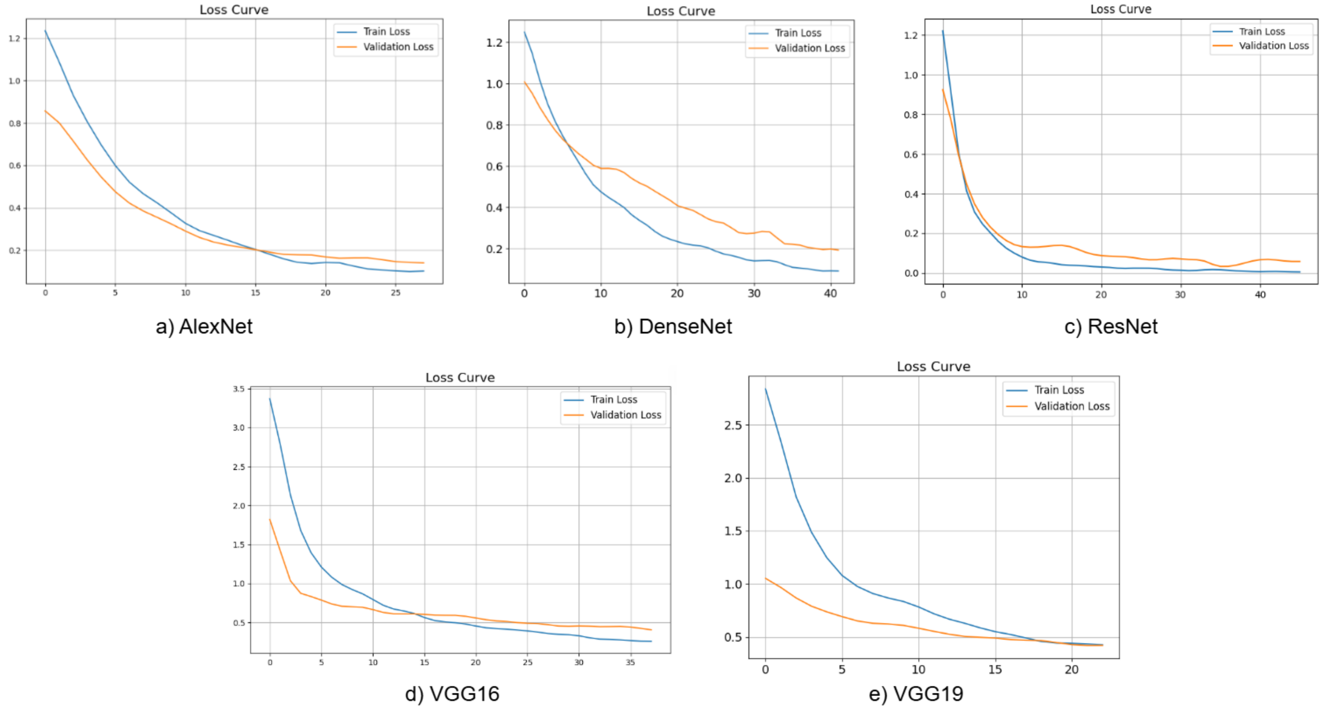

Figure 6. Loss Curves of Deep Learning Models Used in Plant Disease Classification

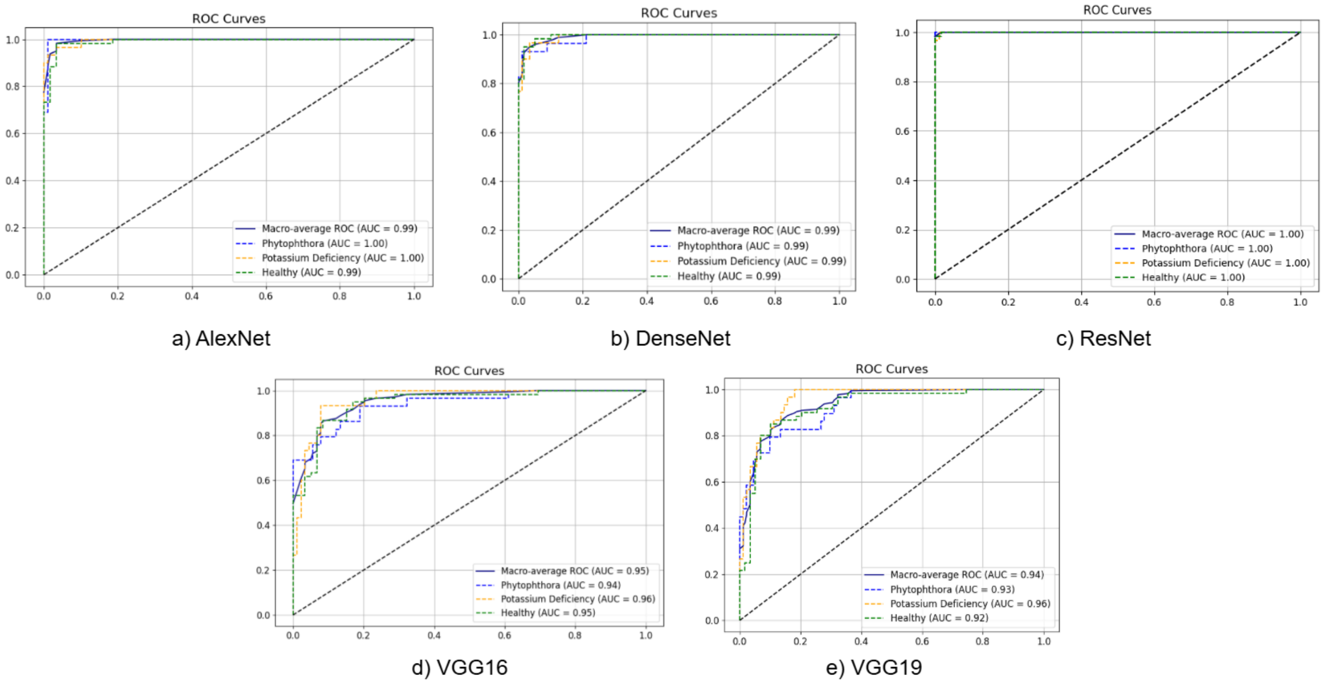

Figure 7. Comparison of ROC Curves and AUC Values of Deep Learning Models for Plant Leaf Disease Classification

In contrast, the DenseNet model performs quite closely behind it. The convergence of the ROC curves from the linear (0.5) line to the upper left corner indicates that the models' overall accuracy and classification performance are quite high.

Table 4. Performance Comparison of the Proposed Method with the Methods in the Literature

Author	Method	Dataset	Accuracy
Argüeso et al. [11]	FSL, Siamese Ağlar, Triplet Loss	PlantVillage	%90
Borhani et al. [10]	ViT, CNN	Wheat Rust Classification Rice Leaf Disease PlantVillage	%97,99
Rao et al. [14]	CNN + GCN	Kaggle	%94,23
Zhao et al. [17]	MAFD - DN4	Plant Pathology 2021	%81,41
C. K. et al. [18]	MFFN	PlantVillage(Tomato)	%99,83
Proposed study	ResNet + GAN	New Dataset(21 Plants Type)	%98

Table 4 compares the methods, datasets, and accuracy rates of different studies. While most studies used ready-made datasets, such as PlantVillage and Kaggle, for classification, the proposed method was tested on a new dataset created by our team. Thanks to GAN-based data augmentation and ResNet integration, our model achieves 98% accuracy, outperforming other studies. These results show that, unlike the methods in the existing literature, we achieved high success on a new dataset created with real-world data, and the proposed method provides a more generalizable model.

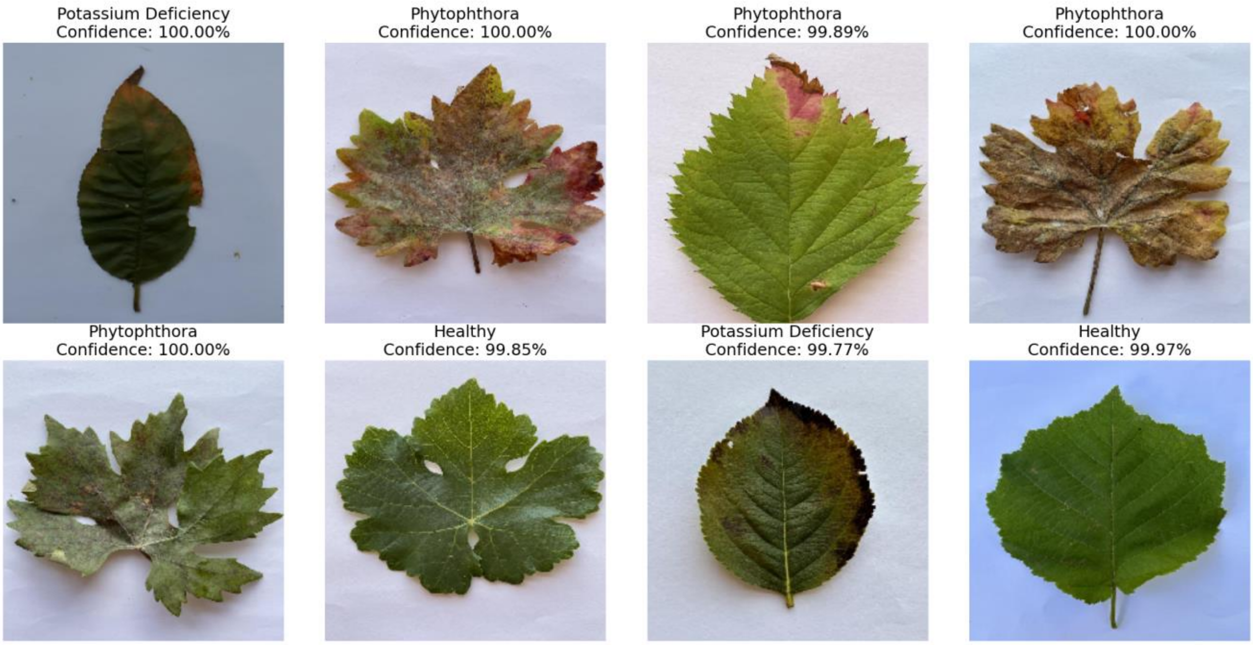

Figure 7. Classification Predictions of Plant Leaf Diseases

The test images of the model, shown in Figure 7, demonstrate that it performs successfully in classifying plant leaf diseases with high accuracy rates. Diseases such as potassium deficiency and Phytophthora Infestans were correctly detected with 99.77% and 100% accuracy rates. Healthy leaves were also classified with 99.97% accuracy, demonstrating the model's strong generalization capacity. The results demonstrate that the model can be utilized as a reliable AI-based system for detecting agricultural diseases.

4.2. Discussion

The ResNet deep learning architecture, supported by the GAN-based data augmentation method presented in this study, has demonstrated more successful results than many existing approaches in terms of accuracy and generalization ability in classifying plant leaf diseases. In particular, the 98% accuracy rate achieved in tests conducted with real images collected from the field demonstrates that the model is effective on both controlled laboratory data and complex, diverse images representing real-world conditions. When examining studies frequently cited in the literature, it is observed that most studies utilize the PlantVillage dataset, which is typically created under fixed lighting conditions and simple backgrounds, resulting in a dataset with limited environmental diversity. However, the original field data used in this study were prepared to include leaf overlaps, various disease symptoms, and a wide range of plant species. Thus, the model was trained and tested under more realistic conditions. Comparative analyses with current approaches in the literature also support the model's success. The Few-Shot Learning-based model developed by Argüeso et al. [11] achieved partial success with 90% accuracy under low data conditions. Borhani et al. [10] combined ViT and CNN architectures to achieve an accuracy rate of 97.99%, but the high computational requirements of the models were considered a disadvantage for field applications. The MFFN+ACSPAM model, developed by C.K. et al. [18], achieved 99.83% accuracy on the PlantVillage dataset; however, its generalization ability is limited due to its focus on tomato leaves and the limited diversity of the data. Many approaches in the literature focus only on specific visual features and may be insufficient for distinguishing between similar diseases. While Shrivastava et al. [9] performed classification based solely on color spaces, Upadhyay et al. [4] developed a classical CNN model based on the shape, color, and size characteristics of lesions on the leaf surface. Yao et al. [5] compared models such as YOLO, SSD, CNN, and ResNet, highlighting field-specific challenges, including leaf overlap and multi-scale detection, and emphasizing the importance of attention mechanisms and segmentation. Similarly, Lu et al. [7] and Saleem et al. [8] emphasized the necessity of data augmentation and transfer learning strategies to achieve high success rates with limited datasets. The CBAM-supported RIC-Net model, developed by Zhao et al. [12], offers high accuracy with low resource requirements; however, it is limited to only three plant species. Additionally, Topcu et al. [15] demonstrated that the ResNet architecture achieves high accuracy under different environmental conditions, highlighting its importance for sustainable agriculture. In this context, the proposed model differs from similar methods in the literature. It has been trained with a broader dataset that covers various plant species and reflects natural conditions, thereby providing a significant advantage in classification accuracy and generalization ability. The images generated by GAN were used only during the training phase and were separated from the test set, eliminating the risk of data leakage during the evaluation process. Instead of a fixed training/testing split, 5-fold cross-validation was applied to objectively measure the model's performance on different data subsets. During the training process, the accuracy and loss graphs showed that the ResNet architecture exhibited a stable learning curve from the early stages and avoided overfitting. The classification capabilities of the models, evaluated by their ROC curves and AUC values, also demonstrated the success of the ResNet architecture. Although the DenseNet model showed high performance with an AUC of 0.99 in the potassium deficiency class, the ResNet model achieved maximum classification capability with an AUC of 1.00 in all classes. The training and testing times presented in Table 3 demonstrate that the proposed model is accurate and computationally efficient. This provides an important advantage for integrating the model with low-power devices in agricultural settings. The study results are not limited to experimental success but also offer practical implications for the widespread adoption of AI-based diagnostic systems in agriculture. Models trained on images collected from the field, which encompass a wide variety of data, provide much more reliable and generalizable outputs than traditional methods. However, the study also has some limitations. In particular, the usage rate of data generated by GAN and the long-term impact of these synthetic examples on the model should be investigated in more detail. The model should be tested under various environmental conditions, including seasonal variations, different plant growth stages, and diverse soil types. The study's results are not limited to experimental success but also provide practical implications for the widespread use of artificial intelligence-based diagnostic systems in agriculture. Models trained on images collected from the field, which contain a wide variety of data, have provided much more reliable and generalizable outputs than traditional methods. However, the study also has some limitations. Despite its good classification performance, the study has some limitations. First, although the dataset covers 21 plant species, the absolute sample size is 1000 images. While the GAN-based approach helps reduce class imbalance, synthetic images may not fully reflect the complexity of naturally occurring variation. Second, the computational requirements of deep models such as ResNet, DenseNet, and InceptionV3 may limit their use in low-power environments. While training is feasible on high-performance computing platforms, real-time inference on edge devices remains to be tested.

5. Conclusion

Early detection of plant diseases in the agricultural sector is crucial for increasing crop productivity and minimizing economic losses. Traditional disease detection methods are time-consuming and costly processes that require expert knowledge. Additionally, artificial intelligence-based models offer a reliable diagnostic process, enabling the quick and accurate detection of diseases through image processing. These approaches provide early intervention by optimizing disease management in agricultural production and support sustainable agriculture by reducing the use of chemicals. This study analyzed leaf images of 21 plant species and classified them into three categories: Phytophthora Infestans, Potassium Deficiency, and Healthy. As a result of the classification process, the ResNet model showed the highest performance among the deep learning models. Experimental findings revealed that the ResNet model was the most successful method, reaching 98% accuracy. In addition, the model also achieved 98% success in precision, recall and F1-score values, indicating superior performance in the classification task. The results demonstrate that the ResNet model is a powerful method for plant disease classification, enabling high accuracy in real-world applications. In future studies, the model's generalization capability can be enhanced and its performance improved by utilizing larger datasets and exploring different hyperparameter optimizations. Furthermore, the model can be tested in different environmental conditions to ensure its integration into field applications, and its potential for practical use with mobile or drone-based imaging systems can be evaluated. In addition, computational efficiency can be optimized to increase the model's real-time analysis capacity, and its applicability on low-power devices can be examined.

References

[1] T. Saranya, C. Deisy, S. Sridevi, and K. S. M. Anbananthen, "A comparative study of deep learning and Internet of Things for precision agriculture," *Eng. Appl. Artif. Intell.*, vol. 122, p. 106034, 2023.

[2] S. N. Benli, *Derin evrişimli yapay sinir ağı kullanarak meyve yaprağı hastalık tespiti*, M.S. thesis, Fen Bilimleri Enstitüsü, Bilecik Şeyh Edebali Üniversitesi, Bilecik, Turkey, 2021.

[3] R. Thangaraj, S. Anandamurugan, P. Pandiyan, and V. K. Kaliappan, "Artificial intelligence in tomato leaf disease detection: A comprehensive review and discussion," *J. Plant Dis. Prot.*, vol. 129, no. 3, pp. 469–488, 2022.

[4] S. K. Upadhyay and A. Kumar, "A novel approach for rice plant diseases classification with a deep convolutional neural network," *Int. J. Inf. Technol.*, vol. 14, no. 1, pp. 185–199, 2022.

[5] Z. Yao and M. Huang, "Deep learning in tropical leaf disease detection: Advantages and applications," *Trop. Plants*, vol. tp-0024-0018, pp. 1–11, 2024.

[6] L. Li, S. Zhang, and B. Wang, "Plant disease detection and classification by deep learning—A review," *IEEE Access*, vol. 9, pp. 56683–56698, 2021.

[7] J. Lu, L. Tan, and H. Jiang, "Review on convolutional neural network (CNN) applied to plant leaf disease classification," *Agriculture*, vol. 11, no. 8, p. 707, 2021.

[8] M. H. Saleem, J. Potgieter, and K. M. Arif, "Plant disease classification: A comparative evaluation of convolutional neural networks and deep learning optimizers," *Plants*, vol. 9, no. 10, p. 1319, 2020.

[9] V. K. Shrivastava and M. K. Pradhan, "Rice plant disease classification using color features: A machine learning paradigm," *J. Plant Pathol.*, vol. 103, no. 1, pp. 17–26, 2021.

[10] Y. Borhani, J. Khoramdel, and E. Najafi, "A deep learning-based approach for automated plant disease classification using vision transformer," *Sci. Rep.*, vol. 12, no. 1, p. 11554, 2022.

[11] D. Argüeso, A. Picon, U. Irusta, A. Medela, M. G. San-Emeterio, A. Bereciartua, and A. Alvarez-Gila, "Few-shot learning approach for plant disease classification using images taken in the field," *Comput. Electron. Agric.*, vol. 175, p. 105542, 2020.

[12] Y. Zhao, C. Sun, X. Xu, and J. Chen, "RIC-Net: A plant disease classification model based on the fusion of inception and residual structure and embedded attention mechanism," *Comput. Electron. Agric.*, vol. 193, p. 106644, 2022.

[13] A. Haridasan, J. Thomas, and E. D. Raj, "Deep learning system for paddy plant disease detection and classification," *Environ. Monit. Assess.*, vol. 195, no. 1, p. 120, 2023.

[14] S. U. M. Rao, K. Sreekala, P. S. Rao, N. Shirisha, G. Srinivas, and E. Sreedevi, "Plant disease classification using novel integration of deep learning CNN and graph convolutional networks," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 36, no. 3, pp. 1721–1730, 2024.

[15] C. Topçu and P. Güneş, "Bitki hastalıklarını tespitte derin öğrenme: ResNet modelinin etkinliği," *Anadolu Bil. Meslek Yüksekokulu Derg.*, vol. 19, no. 69, pp. 31–65, 2024.

[16] M. V. Applalanaidu and G. Kumaravelan, "A review of machine learning approaches in plant leaf disease detection and classification," in *Proc. 3rd Int. Conf. Intell. Commun. Technol. Virtual Mobile Network (ICICV)*, Feb. 2021, pp. 716–724.

[17] Y. Zhao, Z. Zhang, N. Wu, Z. Zhang, and X. Xu, "MAFDE-DN4: Improved few-shot plant disease classification method based on deep nearest neighbor neural network," *Comput. Electron. Agric.*, vol. 226, p. 109373, 2024.

[18] C. K. Sunil, C. D. Jaidhar, and N. Patil, "Tomato plant disease classification using multilevel feature fusion with adaptive channel spatial and pixel attention mechanism," *Expert Syst. Appl.*, vol. 228, p. 120381, 2023.

[19] G. Iglesias, E. Talavera, and A. Díaz-Álvarez, "A survey on GANs for computer vision: Recent research, analysis and taxonomy," *Comput. Sci. Rev.*, vol. 48, p. 100553, 2023.

[20] B. T. Weldemikael, G. Woldetinsae, and G. Neshir, "Generating land gravity anomalies from satellite gravity observations using pix2pix GAN image translation," *Appl. Comput. Geosci.*, p. 100205, 2024.

[21] E. Kardal and V. Nabiyev, "İnsan yüzü modifikasyonu için farklı bir GAN modeli," *Karadeniz Fen Bilimleri Derg.*, vol. 14, no. 2, pp. 403–418, 2024.

[22] A. Abdollahi, B. Pradhan, S. Gite, and A. Alamri, "Building footprint extraction from high resolution aerial images using generative adversarial network (GAN) architecture," *IEEE Access*, vol. 8, pp. 209517–209527, 2020.

[23] M. Arun, M. Deenadhayalan, R. Deepak, S. Balaraman, and D. Deepak, "Lung cancer detection using ResNet-50 CNN architecture," in *Proc. 2nd Int. Conf. Adv. Inf. Technol. (ICAIT)*, vol. 1, Jul. 2024, pp. 1–6.

[24] Z. Nie, W. Gao, H. Jiang, J. Lu, Z. Lu, and X. Jiang, “Predicting critical flame quenching thickness using machine learning approach with ResNet and ANN,” *J. Loss Prev. Process Ind.*, vol. 92, p. 105448, 2024.

[25] C. Guo, Y. Chen, and J. Li, “Radiographic imaging and diagnosis of spinal bone tumors: AlexNet and ResNet for the classification of tumor malignancy,” *J. Bone Oncol.*, vol. 48, p. 100629, 2024.

[26] B. Koonce, *Convolutional Neural Networks with Swift for TensorFlow: Image Recognition and Dataset Categorization*. Berkeley, CA, USA: Apress, 2021. [Online]. Available: https://link.springer.com/content/pdf/10.1007/978-1-4842-6168-2.pdf. Accessed: Dec. 15, 2024.

[27] M. R. Hasan, M. I. Fatemi, M. M. Khan, M. Kaur, and A. Zaguia, “Comparative analysis of skin cancer (benign vs. malignant) detection using convolutional neural networks,” *J. Healthc. Eng.*, vol. 2021, no. 1, p. 5895156, 2021.

[28] S. Toraman and B. Daş, “Evrişimsel sinir ağları kullanılarak normal ve göğüs kanseri hücreleri içeren genomların sınıflandırılması,” *Dicle Univ. Muhendis. Fak. Muhendis. Derg.*, vol. 11, no. 1, pp. 81–90, 2020.

[29] Y. Lu, Y. Qiu, Q. Gao, and D. Sun, “Infrared and visible image fusion based on tight frame learning via VGG19 network,” *Digit. Signal Process.*, vol. 131, p. 103745, 2022.

[30] S. Gaba, I. Budhiraja, V. Kumar, S. Garg, G. Kaddoum, and M. M. Hassan, “A federated calibration scheme for convolutional neural networks: Models, applications and challenges,” *Comput. Commun.*, vol. 192, pp. 144–162, 2022.

[31] K. Kılıç and U. Özcan, “AlexNet architecture optimized for wood defect detection,” *Bozok J. Eng. Archit.*, vol. 2, no. 2, pp. 20–28, 2023.

[32] W. Tang, J. Sun, S. Wang, and Y. Zhang, “Review of AlexNet for medical image classification,” *EAI Endorsed Trans. e-Learning*, vol. 9, Dec. 2023. doi: 10.4108/eetel.4389.

[33] B. Taşcı, “An overview on COVID-19 detection algorithm using deep learning,” in *Proc. Innovations Power Adv. Comput. Technol. (i-PACT)*, 2021, pp. 1–5.

Article Information Form

Authors Contributions

Buse Calisir: Data collection, implementation of methods, analysis, and writing.

Bihter Das: Supervision, validation, writing, review and editing.

Conflict of Interest Notice

The authors declare that they have no conflict of interest regarding the publication of this paper.

Artificial Intelligence Statement

No artificial intelligence tools were used while writing this article.

Plagiarism Statement

This article has been scanned by iThenticate ™.