Intelligent Methods in Engineering Sciences 5(1): 1-9, 2026

INTELLIGENT METHODS
IN ENGINEERING SCIENCES

https://www.imiens.org

International Open Access

April, 2026

e-ISSN 2979-9236

Research Article

https://doi.org/10.58190/imiens.2026.166

# Plant Disease Detection and Classification Using Deep Learning and Image Processing

***Oguzhan BASPINAR*** [a,*], ***Ali YASAR*** [a], ***Yavuz Selim TASPINAR*** [a]

[a] *Technology Faculty, Mechatronics Engineering, Selcuk University, Konya, Türkiye*

ARTICLE INFO

*Article history:*
Received 21 December 2025
Accepted 20 April 2026

*Keywords:*
Classification,
Comparison of SqueezeNet and InceptionV3,
Image processing,
Performance analysis,
Plant diseases detection

ABSTRACT

In agriculture, plants play a vital role in sustaining both human life and the ecosystem. However, plant diseases significantly affect crop yield and quality, making early detection essential. In this study, a dataset consisting of 458 healthy and 435 diseased plant leaf images was used for classification. SqueezeNet and InceptionV3 deep learning architectures were employed for feature extraction, and machine learning models including Artificial Neural Network (ANN), K-Nearest Neighbors (KNN), Support Vector Machine (SVM) and Logistic Regression (LR) were used for classification. Model performance was evaluated using accuracy, precision, recall, F1-score, and confusion matrix metrics. The experimental results show that the InceptionV3-based models achieved higher classification performance compared to SqueezeNet with the LR classifier providing the best overall accuracy of 98.48%. This study provides a comprehensive comparison of deep learning architectures combined with traditional machine learning classifiers and demonstrates the effectiveness of hybrid approaches for plant disease detection. The findings contribute to the development of accurate and efficient systems for early plant disease diagnosis.



## 1. INTRODUCTION

Plant diseases are various diseases that negatively affect the growth, development, and yield of plants. They can have different causes; diseases caused by microorganisms, parasites, fungi, bacteria, and viruses affecting plants can manifest in the leaves, stems, roots, or fruits of plants. These symptoms can generally appear in various ways, such as leaf spots, fungal infections, drying, wilting, browning, rotting, or complete plant death. One of the most important factors affecting yield in agricultural production is the diseases that can occur in plants. Detecting and treating these diseases with artificial intelligence (AI) is of serious importance, as it will both positively or negatively affect the yield of the plant and directly affect the profit of the person harvesting. For this reason, AI-based approaches enable faster and more efficient plant disease detection compared to traditional methods. With AI algorithms, faster, efficient, and ultimately more efficient detection can be achieved. The aim of this study is to classify plant disease images with the help of AI algorithms. In this context, a performance comparison was made using four different classification models: Artificial Neural Network (ANN), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Logistic Regression (LR). In recent years, AI, Machine Learning (ML), and Deep Learning (DL) methods have become popular in agriculture for the automatic classification of plant diseases. Many studies on this subject are being conducted in the literature. The relevant studies are listed in order.

### *1.1. Literature Review*

Jinzhu Lu et al. performed plant leaf disease classification using Convolutional Neural Network (CNN) based Deep Learning (DL) methods with data consisting of approximately 500 plant leaves. They achieved a satisfactory performance with 95% accuracy and 94.5% recall rate with the CNN model VGG16. They suggested that it would be a good resource to determine the future development direction in plant disease classification in the literature [1].

Lili Li et al. collected 54,309 plant leaf disease images and achieved a performance of 96.8% accuracy in combination with the CNN SVM model using data consisting of 3,651 high-quality annotated RGB images of 865 healthy apple leaves, 1,200 apple scab and 1,399 cedar apple rust symptoms and 187 complex disease models for the ‘Plant Pathology Challenge’ for CVPR 2020-FGVC7. They presented a source to the literature that can identify plant leaf diseases with high accuracy [2].

Jie Hang et al. used a dataset with 10 types of diseased

*** Corresponding Author:** oguzhan190572@gmail.com

leaf images. They performed the classification of plant leaf diseases using Global Mean Pooling (GAP) and CNN model with VGG16 convolutional layers and Squeeze & Stimulate (SE) module. They achieved a maximum accuracy of 91.7% in the test set. They proposed an advanced convolutional neural network structure for the identification and classification of a large dataset of plant leaf diseases to the literature [3].

Demilie used the PlantVillage dataset containing 38 classes and 54,305 images of 14 different plant species, 12 of which were healthy and 26 were diseased. By performing the classification of plant leaf diseases using CNNs with DL preference, they achieved an accuracy of 99.35%. The literature has contributed to the need to use more datasets to achieve better results and increase the accuracy percentage in the fields of ML, DL and image processing [4].

Albattah et al. used the PlantVillage dataset containing 54,305 images. Using DenseNet-77, a branch of Deep Learning (DL), they performed the classification of plant leaf diseases with a high accuracy rate. They contributed to the literature by automatically detecting and classifying diseased and healthy plants [5].

Sofuoğlu et al. used the PlantVillage dataset containing 54,305 images and achieved 96.82% accuracy in the training set and 99.48% accuracy in the validation set using a CNN model. They contributed to the literature by classifying healthy, early and late blighted leaves as plant diseases [6].

Altan used the PlantVillage dataset containing 54,305 images. Using the Capsule Network (CapsNET) model, a Machine Learning (ML) system, they achieved performance of 95.76%, 96.37%, and 97.49% in terms of accuracy, precision, and specificity, respectively. They contributed to the literature by stating that CapsNet has great potential, but training large datasets is still a time-consuming process for CapsNET [7].

Yaşar et al. used a dataset consisting of 40 different plants and a total of 340 data samples from 2014. Using an Artificial Neural Network (ANN) model, they achieved a 92% classification success rate. Silva et al. contributed to the literature by improving their work [8].

Akman et al. used the New Plant Diseases Dataset, which contains healthy and diseased leaves divided into 38 different classes. Using the Tensorflow Lite library, they achieved a 92.1% performance in the MobileNetV2 model. They have contributed to the literature not only by identifying and understanding plant diseases but also by discovering unknown or rare diseases [9].

Irmak and Saygılı used a dataset of 8,345 diseased and healthy tomato leaves. Using the CNN model, they achieved accuracy rates of 99.5%, 98.50%, and 97.0% for 2-class, 6-class, and 10-class classifications, respectively. They have contributed to the literature on the use of computer-aided recognition and detection systems in increasing agricultural productivity [10].

Atila et al. used the PlantVillage dataset, which contains 38 classes and 54,305 images of 14 different plant species, 12 of which were healthy and 26 were diseased. They performed the classification of plant leaf diseases using the EfficientNet deep learning architecture and CNN models. The B5 and B4 models of the EfficientNet architecture achieved performances of 99.91% and 99.97%, respectively. They contributed to the literature on the classification of plant diseases using deep learning [11].

Yin Min Oo and Nay Chi Htun used a dataset containing a database of 560 diseased and healthy image samples. Using SVM Classifier, KNN Classifier, and Ensemble Classifier methods, they achieved accuracy rates of 98.2%, 80.2%, and 84.6%, respectively. They contributed to the literature on increasing the efficiency of automatically detecting plant diseases [12].

Dandawate and Kokare used a total of 120 images in a database consisting of healthy, diseased, and pest-infected images. They created a model with Support Vector Machines (SVM) that provided an average accuracy of 93.79%. They contributed to the literature on automated plant disease classification based on image processing in leaves [13].

Hari et al. used the PlantVillage dataset containing 54,305 images. They performed the classification of plant leaf diseases using Convolutional Neural Networks (CNN) and achieved 86% accuracy. They contributed to the literature by creating a model for identifying diseased and healthy plants [14].

Arivazhagan et al. used a dataset containing approximately 500 plant leaf images. They performed the classification of plant leaf diseases using Support Vector Machines (SVM) and achieved 94.74% accuracy. They contributed to the literature by stating that diseased leaves can be detected with very little computational power [15].

Islam et al. performed classification operations using image processing and Naive Bayes methods on rice leaf images. Successful classification of rice brown spot, bacterial blight, and blight diseases was obtained using the Naive Bayes method. No device was developed; the method was used in agriculture. It contributed to the literature by providing rapid disease detection with low computational cost [16].

Al Hiary et al. addressed the automated detection and classification of plant leaf diseases. Classification was performed using image processing, K-means clustering, and Artificial Neural Networks methods with leaf images. An average accuracy of 94% was obtained from the Artificial Neural Networks method. Contributions to the literature include rapid and automated disease detection and improvement in computation time [17].

Sharma et al. created a dataset of plant leaf images consisting of 20,000 images classified into 19 different categories from online sources such as Github and Kaggle.

Using this dataset, classification was performed using image processing methods including K-means clustering, Convolutional Neural Network, Logistic Regression, and Support Vector Machine. The Convolutional Neural Network method yielded the highest performance with an average accuracy rate of 98%. A dataset containing over 20,000 images from 19 different categories has been presented to the literature. The model is considered suitable for future use on Android and iOS platforms [18].

Akila and Deepan used leaf image data from different plant species to detect and classify plant diseases using Faster R-CNN, R-FCN, and SSD deep learning methods. High-accuracy disease recognition results were obtained from deep learning-based object detection methods. An automated disease detection system was developed and used in the field of early diagnosis of agricultural and plant diseases. Contributions to the literature were made, such as deep learning-based multi-model comparison and real-time detection [19].

Hossain et al. performed analysis using image processing and machine learning methods on a dataset consisting of leaf images for the detection and classification of plant leaf diseases. In this context, color space transformation, K-nearest neighbor (KNN) based segmentation, morphological processing, and GLCM-based feature extraction methods were applied. As a result of the applied methods, an accuracy rate of 96.76% was obtained in the classification of Alternaria alternata, anthracnose, bacterial blight, leaf spot, and canker diseases. The study provides a low-cost and feasible automated disease detection system, contributing to the literature with its high accuracy rate and effective segmentation approach [20].

Literature review indicates that although many studies have achieved high accuracy using deep learning models, there is still a need for comprehensive comparative analyses that combine deep learning architectures with traditional machine learning classifiers on the same dataset. In addition, limited studies evaluate the effectiveness of hybrid approaches using multiple classifiers and architectures in a unified framework.

The main contributions of this study can be summarized as follows:

A comprehensive comparative analysis of traditional machine learning models (KNN, SVM, LR, ANN) and deep learning architectures (SqueezeNet and InceptionV3) is performed on the same dataset.

The effectiveness of hybrid approaches combining deep feature extraction and machine learning classifiers is systematically evaluated.

The impact of different deep learning architectures on classification performance is analyzed comparatively.

The proposed framework achieves high classification accuracy, demonstrating its effectiveness in plant disease detection.

The study provides a practical and comparative benchmark for future research in plant disease classification.

The rest of the article is structured as follows: Section 2 presents the materials and methods, including the dataset and the models used. Section 3 presents the experimental results. Section 4 presents the result. Section 5 contains the discussion. Finally, Section 6 concludes the study.

## 2. MATERIALS AND METHODS

### *2.1. Plant Disease Prediction Dataset (PDPD)*

The Plant Disease Prediction Dataset (PDPD) used in this study was obtained from the Kaggle platform [21]. The dataset contains a total of 892 images from 2 classes. There are 458 images in the Healthy class and 435 images in the Disease class. Sample images of multiple plant varieties consisting of diseased and healthy leaves from the dataset are given in Figure 1.


**Figure 1.** Sample images of healthy and diseased plants in PDPD

The images in the dataset used in this study have the .jpg extension and are all in RGB format. The images of diseased leaves show disease symptoms such as spotting, yellowing, and deformations on the leaf surface. Healthy leaves, on the other hand, exhibit a homogeneous color distribution and a smooth texture with no fading or deformation. The images were collected under different environmental conditions, including variations in lighting, background, and orientation; this increases the diversity and confusion of the dataset. The dataset is divided into two folders representing class labels (Healthy and Diseased), making it suitable for binary classification tasks. This dataset is widely used in the development and evaluation of machine learning and deep learning models for automated plant disease detection. The dataset was divided into 80% training and 20% testing for model creation. 10% of the training data was used as the validation set.

### *2.2. Convolutional Neural Networks (CNN)*

CNN is a deep learning architecture that can achieve a high success rate in studies such as image processing, object recognition, and classification thanks to its multi-layered structure. It can be used as a classifier in image processing studies and as a classification system by extracting features from images in machine learning. Therefore, it can be preferred in this study to detect diseases in plant leaves. CNN images consist of various layers. These layers are formed by processing five layers in order: Convolution Layer, Non-Linearity Layer, Pooling Layer, Flattening Layer, and Fully-Connected Layer [22]. Figure 2 visually illustrates how the CNN architecture works through the plant leaf layers.

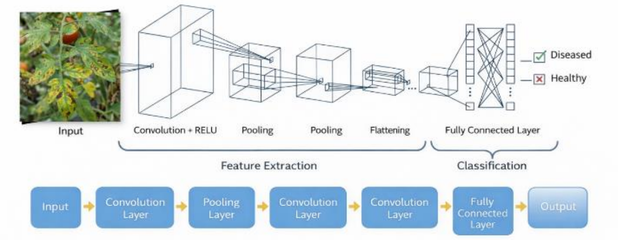


**Figure 2**. Layers and implementation stages of the CNN architecture

In this study, the classification process was performed by sequentially processing the input images through these layers. Feature extraction (distinguishing features) from the image was carried out through the layers, and in the final stage, the model classified the image as "healthy" or "diseased".

### *2.3. SqueezeNet and InceptionV3 Architecture*

In this study, classification and performance comparison was made using two CNN architectures with different algorithm architectures, SqueezeNet and InceptionV3. Both architectures are used in the field of image classification. The differences between them are the number of parameters, computations, and accuracy performance.

SqueezeNet aims to achieve a high accuracy rate with fewer parameters and lower computational costs. This architecture uses a special structure called Fire Module. The Fire module consists of a compression (squeeze) and an expansion (expand) layer. The number of parameters in SqueezeNet has been reduced by approximately 50 times compared to traditional CNNs. In the SqueezeNet papers, the authors showed that a model compression technique called deep compression can be applied to SqueezeNet to reduce the size of the parameter file from 5 MB to 500 KB or even more [23]. Thanks to these features, SqueezeNet is suitable for working environments where the hardware is not large enough and resources are insufficient, such as embedded systems, mobile devices, and real-time agriculture applications. InceptionV3, on the other hand, is a more complex and deeper CNN architecture. Inception combines convolution kernels of different sizes in the same layer. This allows the model to learn from the smallest details to large-scale features. The computational cost is significantly reduced by factoring large convolution filters into smaller filters (e.g., dividing 5×5 into two 3×3s, and 3×3 into 1×3 and 3×1). This allows the model to maintain high accuracy performance while increasing parameter efficiency [24].

In conclusion, while SqueezeNet is a CNN architecture that should be used in limited environments and resources due to its low number of parameters, fast training, and low memory usage, InceptionV3 is a high-performance CNN architecture with good generalization developed for deeper and more complex systems. In this study, the performance of the two architectures was compared in terms of accuracy, training time, and parameter values by training them separately.

### *2.4. Artificial Neural Network (ANN)*

An Artificial Neural Network (ANN) is an algorithmic system developed by taking inspiration from the human brain, i.e., neurons. Advanced artificial neural networks incorporate deep learning. Neural networks possess learning, remembering, and generalization capabilities, and can be used for a variety of tasks, including recognizing and classifying complex data patterns. Artificial neural networks have input, hidden, and output layers. These networks can be used to address a wide range of problems involving supervised learning techniques. They receive information to process, then perform the processing, teach, and make predictions [25]. ANN models automate decision-making processes by learning the relationships between input data and target outputs through supervised learning. In this study, the ANN model used features obtained from images. Using these features, the plants were classified as healthy or diseased. ANNs are small, have low parameters, can be trained quickly, and prevent overlearning.

### *2.5. K-Nearest Neighbors (KNN)*

The KNN model is a supervised learning method where measurements are taken at the nearest point to a sample and the results of these measurements are compared to determine which class the sample belongs to. KNN can work well with linear and nonlinear data and can be used for multi-class classifications. When using this model, the choice of the K parameter is very important for the success of the model. Choosing a small K value can cause noise and overlearning in the model. Choosing a large K value can blur the boundaries and reduce the generalization ability. Therefore, the K value is usually optimized with the cross-validation method. The most commonly used distance measure is the Euclidean distance. Alternatively, Manhattan or Minkowski distances can also be used. In this study, it was applied to feature vectors obtained after feature extraction from images. These features were given as input to KNN and classified as healthy and sick. The most important advantages of the KNN algorithm are its simple structure, no parameter requirement, and high accuracy rate in small datasets [26].

### *2.6. Support Vector Machine (SVM)*

SVM is a supervised learning system developed to solve binary classification problems. In this system, hyperplanes are used to perform classification. Classification is done by dividing classes into two with hyperplanes, but for classifications made with datasets containing many classes, more than one hyperplane is required. SVM classifiers with many hyperplanes are called multi-class SVMs [22]. In this study, there are also two classes. Features obtained from the images are given as input to the SVM. The model determined the most suitable hyperplane that separates the healthy and diseased classes based on these features. Since SVM will provide high accuracy and prevent over-learning in small datasets, choosing this method will yield a successful result.

### *2.7. Logistic Regression (LR)*

Logistic regression is a machine learning method used to classify categorical or numerical data. LR estimates probabilities and labels these probabilities as classes. In this study, the necessary classification was made by predicting whether the plant is sick or healthy. LR predicts the probability, not the results themselves. By passing the result through a logit or sigmoid function, a probability value between 0 and 1 is produced, and this value is compared with a certain threshold value to determine which class it belongs to [27]. It is one of the most efficient models for distinguishing between two classes. In this study, the logistic regression model was used to predict whether the plant leaves are diseased or healthy. If the output is 0, the plant is classified as sick, and if it is 1, it is classified as healthy.

### *2.8. Confusion Matrix*

The confusion matrix is used to evaluate the prediction performance of training and test data. The values in the matrix are used to calculate the performance metrics of classification problems. Figure 3. shows the meanings of the row and column values of the confusion matrix.

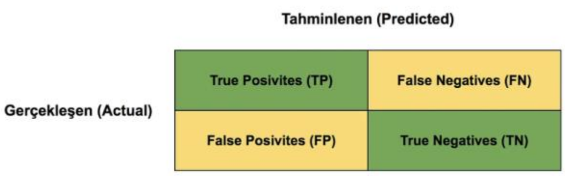


**Figure 3.** Confusion Matrix of Classification Results

Explanation of the values in the Confusion Matrix given in Figure 3:

TP (True Positive): Number of samples correctly predicted as positive,

FP (False Positive): Number of samples incorrectly predicted as positive,

FN (False Negative): Number of samples incorrectly predicted as negative,

TN (True Negative): Number of samples correctly predicted as negative.

## 2.9. Performance Metrics

The Confusion Matrix method was used to evaluate the performance of the classification models used in the study. As a result of this method, Accuracy, Precision, Recall, and F1-Score values were obtained. To obtain these values, TP, FP, FN, and TN values were found in the confusion matrix.

**Table 1.** Performance Metrics, Formulas, and Explanation

| Performance Metrics | Math Formulas | Explanation |
|---|---|---|
| Accuracy | $\frac{TP + TN}{TP + TN + FP + FN}$ | Overall accuracy of the model |
| Precision | $\frac{TP}{TP + FP}$ | The result of how many positive predictions were correct. |
| Recall | $\frac{TP}{TP + FN}$ | The result of how many of the true positives were accurately detected. |
| F1 Score | $\frac{2 \times (Precision \times Recall)}{Precision + Recall}$ | Balanced average of Precision and Recall. |

In this table, the Accuracy value, Sensitivity value, Recall value, and F1 Score value are found using formulas and the performance metrics of the models are calculated using mathematical formulas. These formulas are given in Table 1.

## 3. EXPERIMENTAL RESULTS

In this study, ANN, KNN, SVM, and LR methods were used to classify the Plant Disease Prediction Dataset (PDPD) by identifying diseased and healthy plant leaves separately from crops in the field using image processing and machine learning methods. A computer with an AMD Ryzen™ 9 7845HX 3.00 GHz processor, NVIDIA GeForce RTX 4060 graphics card, and 16 GB of RAM was used. Python programming language was used in this study.

**Table 2.** The main hyperparameters used for ANN, KNN, SVM, and LR models.

| Model | Parameter | Value |
|---|---|---|
| ANN | Hidden Layers | 100 |
| | Activation Function | ReLU |
| | Regularization | 0.0001 |
| | Iterations | 200 |
| KNN | Number of Neighbors (k) | 5 |
| | Distance Metric | Euclidean |
| | Weight Function | Uniform |
| SVM | Kernel | RBF |
| | Cost (C) | 1 |
| | Epsilon | 0.10 |
| LR | Regularization Type | L2 (Ridge) |

The main hyperparameters used for ANN, KNN, SVM, and LR models are summarized in Table 2. The same parameter settings were applied for both SqueezeNet and InceptionV3 architectures. In addition, 10-fold cross-validation was used for model evaluation. The flowchart of the procedures performed in the study is shown in Figure 4.

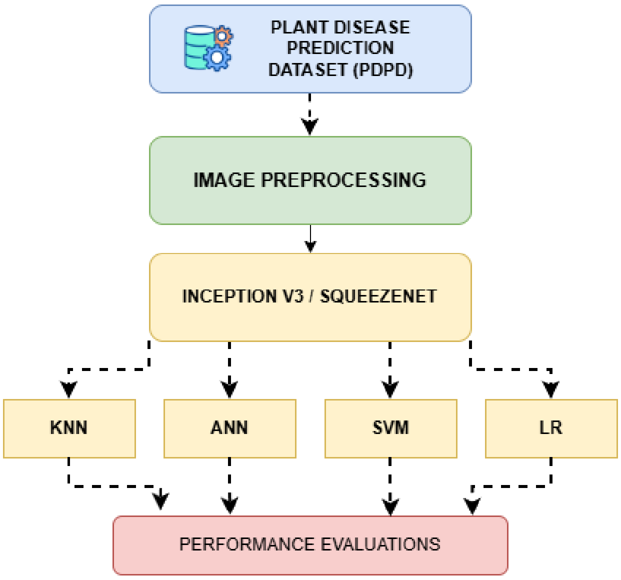


Figure 4. Steps in the process of classifying plants

1000 features obtained from the SqueezeNet model and 2048 features obtained from the Inception V3 model were provided as an introduction to machine learning methods and their training was carried out. These features were obtained from the fully connected layer, which is the layer immediately preceding the classification layer. The confusion matrices obtained from the machine learning methods are given in order. The confusion matrices obtained from the SVM, LR, KNN, and ANN methods of the SqueezeNet and Inception V3 architectures are shown in Figures 5 and 6.

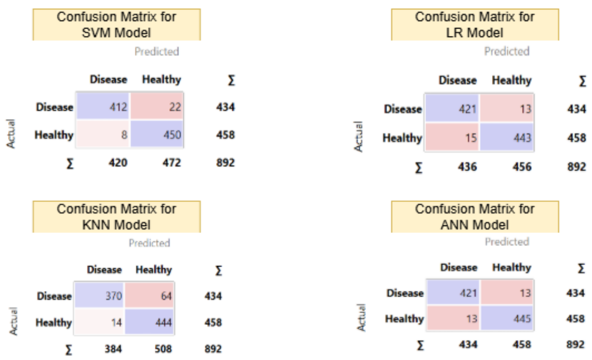


**Figure 5.** Confusion matrices of SVM, LR, KNN, and ANN models resulting from the Squeezenet architecture

According to the confusion matrix results given in Figure 5, all models generally demonstrated high classification performance. The SVM model achieved the most balanced results, classifying the Disease and Healthy classes with high accuracy respectively. The Logistic Regression (LR) model performed similarly to SVM, exhibiting only a small amount of misclassification. In contrast, the KNN model showed a higher error rate, particularly in the Disease class. The ANN model on the other hand offered balanced performance in both classes achieving a high overall accuracy value. The results reveal that the SVM and ANN models in particular demonstrated more successful classification performance compared to other methods.

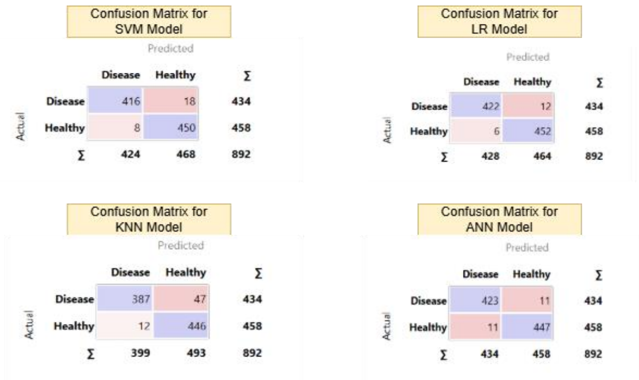


**Figure 6.** Confusion matrices of SVM, LR, KNN, and ANN models resulting from the Inception V3 architecture

According to Figure 6, the confusion matrix results of the Inception V3 architecture indicate that all models achieved high classification performance for both Disease and Healthy classes. Among the evaluated methods, SVM and LR models demonstrated the most balanced performance with lower misclassification rates compared to KNN and ANN. In particular, KNN showed a relatively higher number of misclassifications in the Disease class, while ANN maintained a balanced performance across both classes. The results show that LR and SVM generally outperform the other models in terms of classification accuracy and consistency.

Performance metrics for all models were calculated based on the corresponding confusion matrices, and the results for the Inception V3 and Squeezenet architectures are presented in Table 3 and Table 4.

**Table 3.** Performance metrics of the Inception V3 architecture

| Class | AUC | Precision | Recall | F-1 Score | Average Accuracy (%) |
|---|---|---|---|---|---|
| ANN | 0.998 | 0.975 | 0.975 | 0.975 | 98.08 |
| SVM | 0.996 | 0.971 | 0.971 | 0.971 | 97.73 |
| LR | 0.999 | 0.980 | 0.980 | 0.980 | 98.48 |
| KNN | 0.982 | 0.936 | 0.934 | 0.934 | 94.65 |

According to Table 3, the LR model achieved the highest overall performance in the Inception V3 architecture with an average accuracy of 98.48%, followed closely by the ANN and SVM models. In contrast, the KNN model obtained the lowest performance with an average accuracy of 94.45%. Similar trends were observed in other evaluation metrics, including precision, recall, and F1-score, where LR consistently outperformed the other models. Overall, all models achieved satisfactory classification performance, with accuracy values exceeding 94%, indicating the effectiveness of the proposed approaches.

**Table 4.** Performance metrics of the Squeezenet architecture

| Class | AUC | Precision | Recall | F-1 Score | Average Accuracy (%) |
|---|---|---|---|---|---|
| ANN | 0.997 | 0.971 | 0.971 | 0.971 | 97.05 |
| SVM | 0.993 | 0.967 | 0.966 | 0.966 | 96.52 |
| LR | 0.996 | 0.969 | 0.969 | 0.969 | 96.82 |
| KNN | 0.971 | 0.918 | 0.913 | 0.912 | 90.95 |

According to Table 4, the ANN model achieved the highest overall performance in the Squeezenet architecture, while the KNN model showed the lowest classification performance. The LR and SVM models exhibited similar and competitive results, closely following the ANN model. A consistent trend was observed across all evaluation metrics, including precision, recall, and F1-score, which aligned with the accuracy results. Overall, although performance levels varied among the models, all classifiers demonstrated acceptable classification performance within the Squeezenet framework, with accuracy values exceeding 90%.

**Figure 7.** Misclassified images in the Dataset

When the prediction results were examined, it was seen that the ANN, SVM, LR and KNN models in the Squeezenet architecture predicted the plant as healthy even though the image in Image-a was diseased. In the same image in Inception V3, only the LR model made the correct prediction. Although the image in Image-b in figure-6 is diseased, the KNN and SVM models in the Squeezenet architecture made incorrect predictions, while the LR and ANN models made correct predictions. All models predicted correctly on the same image in Inception V3. Even though the Image-c image in Figure 7 is a healthy example, only the LR model in the Squeezenet architecture made the correct prediction. In Inception V3, all models predicted correctly. Reasons for these inaccurate predictions include problems with the images themselves, such as the model not being able to identify a particular image.

## 4. RESULTS

In the study on plant diseases, classification was carried out using a data set consisting of 892 images of diseased and healthy plants. The highest accuracy rate was obtained by the InceptionV3-based LR model with 98.48%, while the lowest performance was obtained by the SqueezeNet-based KNN model with 90.95%. In the Inception V3 architecture, the average accuracy rates were 98.08% for ANN, 97.73% for SVM, 98.48% for LR, and 94.65% for KNN. In the Squeezenet architecture, the average accuracy rates were 97.05% for ANN, 96.52% for SVM, 96.82% for LR, and 90.95% for KNN.

## 5. DISCUSSION

Classification results demonstrate that the proposed machine learning and deep learning models are effective for plant disease detection. In particular, LR, SVM, and ANN models achieved more balanced and higher performance compared to KNN, which shows lower performance due to its sensitivity to data distribution and neighborhood dependency.

When compared with related studies in the literature, the proposed approach demonstrates competitive or superior performance. For instance, previous studies using CNN-based models or transfer learning approaches reported accuracy values ranging between approximately 90% and 97%, depending on the dataset and preprocessing strategies. In contrast, the models used in this study, particularly InceptionV3-based LR and Squeezenet-based approaches, achieved higher accuracy levels, reaching up to 98.48%.

Furthermore, the use of deep feature extraction through architectures such as InceptionV3 and SqueezeNet significantly improves classification performance when combined with traditional machine learning classifiers. This highlights the effectiveness of hybrid approaches that integrate deep learning feature extraction with classical classifiers.

Overall, the results indicate that both model selection and feature representation play a critical role in classification performance. The proposed framework provides a robust and efficient solution for plant disease detection and can serve as a benchmark for future studies.

## 6. CONCLUSIONS

This study demonstrates that machine learning and deep learning approaches can effectively be used for plant disease classification. The results indicate that the proposed models achieve high classification performance and can support early disease detection in agricultural applications. By evaluating the performance of hybrid approaches that combine deep feature extraction with machine learning classifiers, insights into the effectiveness of different classification strategies have been provided. Future studies may improve performance by expanding the dataset and including more disease classes. A comparison of both traditional machine learning models (SVM, KNN, LR, ANN) and deep learning-based architectures (InceptionV3 and SqueezeNet) was performed using the same plant disease dataset.

## Declaration of Ethical Standards

The article does not contain any studies with human or animal subjects.

## Credit Authorship Contribution Statement

Authors individually were responsible for the ideation, modeling, analysis, and writing of this article.

## Declaration of Competing Interest

Author claims that there are no conflicts of interest.

## Funding / Acknowledgements

No funding was received from any organization for this research.

## Data Availability

https://www.kaggle.com/datasets/dittakavinikhita/plant-disease-prediction-disease-and-healthy

## References

[1] J. Lu, L. Tan, and H. Jiang, "Review on Convolutional Neural Network (CNN) Applied to Plant Leaf Disease Classification," *Agriculture,* vol. 11, no. 8, 2021, doi: 10.3390/agriculture11080707.

[2] 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, doi: 10.1109/access.2021.3069646.

[3] J. Hang, D. Zhang, P. Chen, J. Zhang, and B. Wang, "Classification of Plant Leaf Diseases Based on Improved Convolutional Neural Network," *Sensors (Basel),* vol. 19, no. 19, Sep 25 2019, doi: 10.3390/s19194161.

[4] W. B. Demilie, "Plant disease detection and classification techniques: a comparative study of the performances," *Journal of Big Data,* vol. 11, no. 1, 2024, doi: 10.1186/s40537-023-00863-9.

[5] W. Albattah, M. Nawaz, A. Javed, M. Masood, and S. Albahli, "A novel deep learning method for detection and classification of plant diseases," *Complex & Intelligent Systems,* vol. 8, no. 1, pp. 507–524, 2021, doi: 10.1007/s40747-021-00536-1.

[6] C. İ. SofuoĞLu and D. Birant, "Potato Plant Leaf Disease Detection Using Deep Learning Method," *Tarım Bilimleri Dergisi,* 2023, doi: 10.15832/ankutbd.1276722.

[7] G. Altan, "Performance Evaluation of Capsule Networks for Classification of Plant Leaf Diseases," *International Journal of Applied Mathematics Electronics and Computers,* vol. 8, no. 3, pp. 57–63, 2020, doi: 10.18100/ijamec.797392.

[8] A. Yasar, I. Saritas, M. A. Sahman, and A. O. Dundar, "Classification of Leaf Type Using Artificial Neural Networks," *International Journal of Intelligent Systems and Applications in Engineering,* vol. 3, no. 4, 2015, doi: 10.18201/ijisae.49279.

[9] M. Akman, O. Çam, B. Cinar, and D. Özdemir, "Derin Öğrenme Tabanlı Mobil Uygulama ile Bitki Hastalığı Tespiti," *Eskişehir Türk Dünyası Uygulama ve Araştırma Merkezi Bilişim Dergisi,* vol. 6, no. 1, pp. 20–27, 2025, doi: 10.53608/estudambilisim.1531231.

[10] G. Irmak and A. Saygili, "A novel approach for tomato leaf disease classification with deep convolutional neural networks," *Tarım Bilimleri Dergisi,* 2023, doi: 10.15832/ankutbd.1332675.

[11] Ü. Atila, M. Uçar, K. Akyol, and E. Uçar, "Plant leaf disease classification using EfficientNet deep learning model," *Ecological Informatics,* vol. 61, p. 101182, 2021.

[12] N. C. H. Yun Min Oo, Mainkar, Prakash, "Plant leaf disease detection and classification using image processing," *International Journal of Research and Engineering,* 2018.

[13] Y. Dandawate and R. Kokare, "An automated approach for classification of plant diseases towards development of futuristic Decision Support System in Indian perspective," in *2015 International conference on advances in computing, communications and informatics (ICACCI)*, 2015: IEEE, pp. 794–799.

[14] S. S. Hari, M. Sivakumar, P. Renuga, and S. Suriya, "Detection of plant disease by leaf image using convolutional neural network," in *2019 International conference on vision towards emerging trends in communication and networking (ViTECoN)*, 2019: IEEE, pp. 1–5.

[15] S. Arivazhagan, R. N. Shebiah, S. Ananthi, and S. V. Varthini, "Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features," *Agricultural Engineering International: CIGR Journal,* vol. 15, no. 1, pp. 211–217, 2013.

[16] T. Islam, M. Sah, S. Baral, and R. R. Choudhury, "A faster technique on rice disease detectionusing image processing of affected area in agro-field," in *2018 second international conference on inventive communication and computational technologies (ICICCT)*, 2018: IEEE, pp. 62–66, doi: 10.1109/ICICCT.2018.8473322.

[17] H. Al Hiary, S. B. Ahmad, M. Reyalat, M. Braik, and Z. Alrahamneh, "Fast and accurate detection and classification of plant diseases," *International journal of computer applications,* vol. 17, no. 1, pp. 31–38, 2011.

[18] P. Sharma, P. Hans, and S. C. Gupta, "Classification of plant leaf diseases using machine learning and image preprocessing techniques," in *2020 10th international conference on cloud computing, data science & engineering (Confluence)*, 2020: IEEE, pp. 480–484, doi: 10.1109/Confluence47617.2020.9057889.

[19] M. Akila and P. Deepan, "Detection and classification of plant leaf diseases by using deep learning algorithm," *International Journal of Engineering Research & Technology (IJERT),* vol. 6, no. 7, pp. 1–5, 2018.

[20] E. Hossain, M. F. Hossain, and M. A. Rahaman, "A color and texture based approach for the detection and classification of plant leaf disease using KNN classifier," in *2019 international conference on electrical, computer and communication engineering (ECCE)*, 2019: IEEE, pp. 1–6, doi: 10.1109/ECACE.2019.8679247.

[21] Kaggle. *Images of healthy and diseased plants for deep learning-based disease detection.* [Online]. Available: https://www.kaggle.com/datasets/dittakavinikhita/plant-disease-prediction-disease-and-healthy

[22] Y. S. Taspinar, I. Cinar, and M. Koklu, "Classification by a stacking model using CNN features for COVID-19 infection diagnosis," *Journal of X-Ray Science and Technology,* vol. 30, no. 1, pp. 73–88, 2022, doi: 10.3233/xst-211031.

[23] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size," *arXiv preprint arXiv:1602.07360,* 2016.

[24] A. Ghantiwala, "Understanding Architecture Of Inception Network & Applying It To A Real-World Dataset," May 21, 2022. [Online]. Available: https://gghantiwala.medium.com/understanding-the-architecture-of-the-inception-network-and-applying-it-to-a-real-world-dataset-169874795540.

[25] M. B. Yildiz, E. T. Yasin, and M. Koklu, "Fisheye freshness detection using common deep learning algorithms and machine

learning methods with a developed mobile application," *European Food Research and Technology,* vol. 250, no. 7, pp. 1919–1932, 2024.

[26] X. Wu and V. Kumar, *The top ten algorithms in data mining*. CRC press, 2009.

[27] S. Bulut, E. Maşazade, and D. E. Barkana, "A Logistic Regression and Neural Network Based Approach for Disease Detection from Physiological Sensors," in *2025 33rd Signal Processing and Communications Applications Conference (SIU)*, 2025: IEEE, pp. 1–4.