Combining Deep Features and MSVM Characteristics for Enhanced Classification of Plant Diseases

Zahraa Jabbar Hussein
Department of Computer Science, College of Science for Women, University of Babylon, Iraq
zahraa.jabbar@uobabylon.edu.iq (corresponding author)

Received: 18 January 2025 | Revised: 18 April 2025, 23 April 2025, 14 May 2025, and 15 May 2025 | Accepted: 17 May 2025

DOI: https://doi.org/10.48084/etasr.10280

ABSTRACT

Plant diseases pose a significant global challenge, threatening food security and agricultural productivity. Accurate and timely diagnosis is essential for effective disease management and crop protection. This study utilized a comprehensive dataset of plant disease images collected from diverse agricultural regions to develop predictive models based on Convolutional Neural Networks (CNNs). The proposed CNN architectures were used to classify images after extracting important characteristics using several algorithms, such as Multi-class Support Vector Machine (MSVM), Decision Tree (DT), Neural Network (NN), and K-Nearest Neighbors (KNN). The results demonstrated that MSVM was the most accurate in identifying diseases that affect plant leaves, underscoring the role of deep learning techniques in rapid and precise disease detection and providing immediate intervention strategies to reduce agricultural losses. Furthermore, these models enable customized treatment recommendations for farmers, optimizing the use of pesticides and fungicides to reduce environmental impacts and enhance the economic sustainability of small farming communities. Experimental findings reveal that the deep learning model combined with MSVM significantly outperformed traditional methods, achieving an accuracy exceeding 99.1%.

Keywords-CNN; plant disease detection; feature extraction; MSVM; DT; KNN; DL

I. INTRODUCTION

Agriculture aims to exploit the available energies and resources and produce various agricultural products necessary to satisfy human needs and desires. Agriculture is important in the production of essential food commodities such as wheat, grains, and fruits. It is a source of raw material products, such as cotton and flax used in the clothing industry, in addition to supplying the industrial sector, such as oil seeds to extract oil, tomatoes to make sauces, and fruits and vegetables for making canned food [1]. Agriculture is considered a balance of trade, especially for agricultural countries that depend on its economy and the export of its surplus, thus obtaining foreign exchange necessary for import requirements [2]. In addition, agriculture is considered a consumer market for the products of other sectors used in agricultural production, such as fertilizers, improved seeds, machines, etc.

Environmental factors, such as drought, cold, and high temperatures, can expose plants to diseases, agricultural pests, fungi, bacteria, and parasites. These factors can cause distortions in the external appearance of the plant and cause physiological defects [3]. Depending on the type of disease and the extent to which the plant is damaged, these malformations usually appear on the stems, leaves, trunks, and branches. Early detection of plant diseases can reduce financial losses for farmers, benefiting the entire agricultural sector of a country [4]. Thus, many studies focus on discovering these diseases early to facilitate early treatment with minimal losses. Since these diseases have symptoms visible to the human eye, scientists can use deep learning and digital image processing techniques to diagnose them, rather than manually, which requires effort, manpower, experience, and a long time. In addition, the diagnosis can sometimes be inaccurate. Research in the diagnosis of plant diseases is crucial for several reasons:

- Crop Protection: It helps protect crops from diseases, ultimately contributing to food security [5].
- Economic Impact: As plant diseases can have a significant economic impact on agriculture, an accurate diagnosis can prevent losses.
- Environmental Impact: Proper disease management can reduce the need for chemical treatments, offering environmental benefits.
- Advancements in Technology: Plant disease diagnosis often involves using advanced technology such as computer vision and machine learning, contributing to the development of these fields.

II. RELATED WORKS

In [6], pre-trained models were used to accurately diagnose plant diseases, focusing on adjusting the parameters of famous models such as ResNet-50, DenseNet-121, Inception V4, and

VGG-16. The Plant Village dataset, which contains thousands of images showing many plant diseases divided into 38 classes, was used to conduct the experiments, where DenseNet-121 outperformed state-of-the-art models with a classification accuracy of 88%. In [7], a reliable classification method was proposed for plant diseases, using a custom CenterNet architecture with DenseNet-77 as the basic network. DenseNet-77-based deep key point extraction improved CenterNet by identifying and classifying 26 different plant diseases in 14 different plants, including tomatoes, apples, grapes, and other horticultural crops. In [8], deep feature extraction and DL models were used to identify plant diseases in the Plant Village dataset. SVM and KNN were used to extract features and GoogleNet, ResNet50, and VGG16 were chosen for classification. The results were compared using execution time and accuracy, showing that feature extraction efficiency can be much higher without transfer learning, offering better results and taking less time.

In [9], Few-Shot Learning (FSL) was used to identify and classify plant diseases. At first, the Inception V3 framework was used to calculate the key points, and then the collected features were used to train a multiclass SVM. This method was powerful for classifying plant diseases, but the results were presented on a modestly sized dataset, indicating that the model should be evaluated on broad and diversified cases. In [10], the efficacy of advanced plant disease classification was analyzed, using some models, such as DenseNet169, Xception, InceptionV3, MobileNetV2, and ResNet50V2, to process a dataset of 7623 training images and 1906 validation images. The method converted RGB images to grayscale and enhanced their quality using different techniques, such as Otsu thresholding, noise removal, and distance transformation. To find disease areas, contour characteristics were retrieved by computing morphological values. The results showed that, with loss values of 0.19 and 0.49, respectively, MobileNetV2 and ResNet50V2 had the best validation accuracy of 99.42%, while most models, except MobileNetV2 and InceptionV3, achieved optimal recall, precision, and F1 scores of 0.99. In [11], deep learning was used for image classification to identify crop diseases effectively, utilizing the Plant Village dataset comprising 50,000 RGB images of size 3×224×224, covering 38 diseases across 14 crops. This method enhanced images through pixel-based operations, extracted features, segmented images, and classified diseases using CNNs, achieving an accuracy of 90.4%.

In [12], the application of machine learning and deep learning techniques in precision agriculture was reviewed, with a particular focus on plant disease classification. A new classification scheme was introduced. This study also highlighted the available datasets for plant disease analysis, providing details on their classes, data, and whether they are appropriate for object detection or classification. Additionally, 18 classification algorithms were evaluated on the PlantDoc dataset in identifying disease presence. ResNet50 obtained a total accuracy of 61.01% after approximately 18 minutes of training, whereas MobileNetV2 acquired a total accuracy of 59.74% after approximately 16 minutes of training. These models provided the best balance between accuracy and training time for classification tasks.

In [13], an image classification model for plant diseases used a Vision Transformer (ViT). ViT divides images into patches and feeds them to the transformer encoder after applying linear embedding. This study combined CNN and ViT, leveraging their strengths: CNN extracted local features such as texture histograms, which are useful for recognizing diseases, while ViT captured global features of the entire image to detect and differentiate leaf types. The model's overall performance was evaluated through training, validation, and testing, achieving accuracy in the range of 95% to 96%.

III. DATASET

This study used the publicly available Plant Village dataset [14], which contains labeled images of healthy and diseased plant leaves. The original dataset consists of 38 folders corresponding to different plant species and conditions. However, only 33 folders were included in this work, as five folders were excluded for the following reasons: some folders (e.g., Background_without_leaves) contained irrelevant images such as vehicles and bicycles, while others (e.g., Blueberry healthy, Soybean healthy, Squash Powdery_mildew) included only healthy leaves without corresponding diseased samples. The final dataset consisted of approximately 44,500 images distributed across 33 folders, each representing a unique class of either healthy or diseased plant leaves. A detailed summary of the selected folders, their condition (healthy or diseased), and the number of images in each is provided in Figure 1 [14].

IV. METHODOLOGY

The first stage was to improve the data and remove noise through a linear median filter and then resize the images from 256×256 to 200×200. The data were augmented to improve diagnostic accuracy. The third stage involved deep feature extraction using a CNN. The dataset was divided into 80:10:10 for training, validation, and testing. The extracted features were fed into a Multi-class Support Vector Machine (MSVM) for classification. Figure 2 shows the process architecture. The main contributions of this study are:

- Develops a deep learning model to diagnose different diseases in plants.
- Chooses the most effective transfer learning strategy to obtain the most accurate classification of plant diseases with the best possible recognition accuracy across multiple classes.
- Elaborates on the CNN model to resolve labeling and class problems in plant disease identification using multi-class multi-label transfer learning.
- Uses data augmentation techniques to solve the overfitting issue.
- Uses batch-normalization to stabilize and accelerate training by normalizing layer outputs, reducing internal covariate shifts, and acting as a form of regularization to prevent overfitting. This ensures consistent activation distributions, improving performance and enabling faster convergence.

Apple-scab: 1000 images

Apple-black-rot: 1000 images

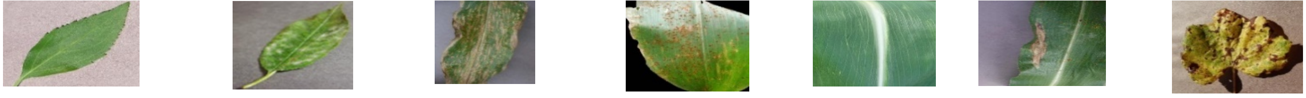
Apple-cedar-rust: 1000 images

Apple-healthy: 1645 images

Grape-healthy: 1000 images

Grape-black-rot: 1180 images

Grape-esca-black-measles: 1383 images



Cherry-healthy: 1000 images

Cherry-powdery-mildew 1052 images

Corn-cercospora-leaf-spot-gray-leaf-spot: 1000 images

Corn-common-rust: 1192 images

Corn-healthy: 1162 images

Corn-northern-leaf-blight: 1000 images

Grape-leaf-blight-isariopsis-leaf-spot: 1076 images

Peach-healthy: 1000 images

Peach-bacterial-spot: 2297 images

Pepper-bell-healthy: 1478 images

Pepper-bell-Bacterial-spot: 1000 images

Potato-healthy: 1000 images

Potato-early-blight: 1000 images

Potato-late-blight: 1000 images

Strawberry-healthy: 1000 images

Strawberry-leaf-scorch: 1109 images

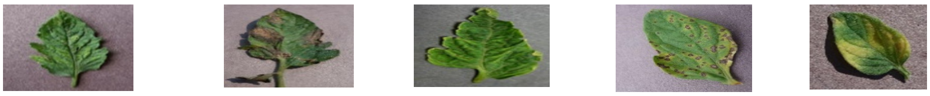
Tomato-healthy= 1591 images

Tomato-bacterial-spot: 2127 images

Tomato-early-blight: 1000 images

Tomato-spider-mites-two spotted-spider-mite: 1676 images

Tomato-target-spot: 1404 images



Tomato-mosaic-virus: 1000 images

Tomato-late-blight: 1000 images

Tomato-yellow-leaf-curl-virus: 5357 images

Tomato-septoria-leaf-spot: 1771 images

Tomato-leaf-mold: 1000 images

Fig. 1. Sample for the 33 classes of healthy and diseased leaves in the Plant Village dataset [14].

A. *Proposed CNN for Feature Extraction*

Deep learning, a subfield of Artificial Intelligence (AI), utilizes multiple hidden layers to extract features from raw data. It has shown great success in various applications, such as image recognition, handwriting detection, and facial recognition. CNNs, a type of deep learning model inspired by the human visual system, are widely used in data mining, natural language processing, computer vision, and gaming [15]. CNNs are more efficient in feature extraction than fully connected networks due to their hierarchical and feedforward structure. They are capable of learning abstract features and achieving near-optimal classification performance as the dataset grows. CNNs are especially effective when dealing with limited or complex training data. Among the top techniques in data mining, CNNs offer a powerful solution for pattern recognition tasks.

1) *The Convolutional Layer*

One of the fundamental components of a CNN is the convolutional layer. A collection of teachable filters makes up the parameters of each convolutional layer. Each filter performs a convolution on the input volume while employing forward propagation. At every given position, it calculates the products of the dots between the input and the filter elements. Feature maps are the outputs produced as a result.

TABLE I. PROPOSED CNN

Layer (Type)	Output Shape	Param
Input Layer	200, 200, 3	0
Conv2D	200, 200, 16	448
BatchNormalization	200, 200, 16	64
Activation (ReLU)	200, 200, 16	0
MaxPooling2D	100, 100, 16	0
Conv2D	100, 100, 32	4640
BatchNormalization	100, 100, 32	128
Activation (ReLU)	100, 100, 32	0
MaxPooling2D	50, 50, 32	0
Conv2D	50, 50, 64	18,496
BatchNormalization	50, 50, 64	256
Activation (ReLU)	50, 50, 64	0
MaxPooling2D	25, 25, 64	0
Conv2D	25, 25, 128	73,856
BatchNormalization	25, 25, 128	512
Activation (ReLU)	25, 25, 128	0
MaxPooling2D	12, 12, 128	0
Flatten	18432	0

2) *The Pooling Layer*

Following the convolution layer is the pooling layer (sub-sampling). Translation invariance and a large decrease in the number of trainable parameters are two major benefits of applying the pooling technique. To perform the pooling action, the input items from the selected window are subsequently transferred via a pooling function [14].

3) The Fully Connected Layer

This layer serves as an integrator that combines the image features in the feature maps through numerous convolutional and pooling layers. Data from the convolutional and pooling layers are transmitted to the fully connected layer. There is a connection between each neuron in a layer and the layer beneath it, but there is no connection between the neurons inside the same layer. The fully connected layer's role is to improve nonlinear mapping capabilities.

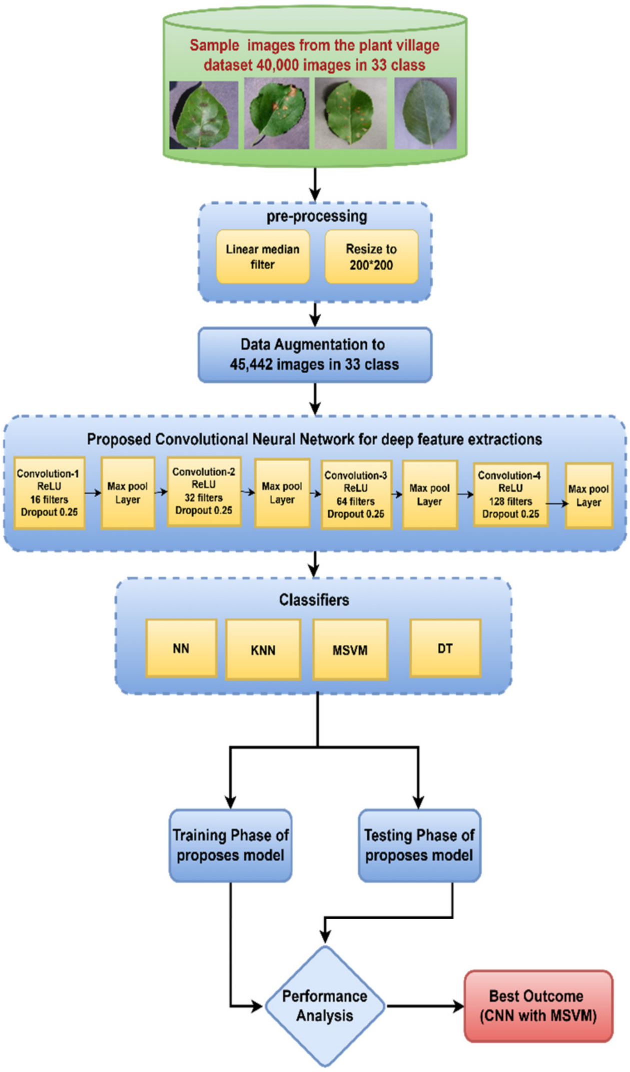


Fig. 2. Process flowchart illustration.

B. Classification Techniques

1) Multi-class SVM (MSVM)

The features extracted from the CNN are then classified using an MSVM, which resolves the multiclass problem by combining binary classifiers. Two popular implementations of MSVM include the One-versus-One (OvO) and the One-versus-All (OvA) methods. OvO trains a binary classifier for every pair of classes, while OvA trains one classifier per class against all remaining classes. Combining CNN with MSVM can significantly enhance classification performance by leveraging the strength of deep feature extraction in CNN and the robust classification mechanism in MSVM. This can be expressed as a classification process:

$$f(x) = argmax_c + \left(SVM_c\left(CNN(x)\right)\right) \tag{1}$$

where $CNN(x)$ extracts deep features from the input image, SVM_c represents the score or confidence that the input belongs to class c, with $c \in \{1,2,\ldots\ldots 33\}$, and $argmax_c$ selects the class c with the highest output score from the SVM.

2) K-Nearest Neighbor (KNN)

KNN is a straightforward yet effective algorithm used for both classification and regression tasks. Its simplicity makes it easy to implement and interpret [16]. KNN works by calculating the distance (typically Euclidean) between a given query point and all other points in the dataset and then selects the *K* closest data points (neighbors) to determine the output.

- In classification, the query point is assigned the most frequent class label among the *K* neighbors.
- In regression, the algorithm calculates the mean value of the K nearest neighbors.

However, KNN comes with a few limitations:

- Scalability: As the dataset grows larger, the time required to calculate distances to all points increases significantly, which can make the algorithm computationally expensive.
- Choice of *K*: The accuracy of KNN heavily depends on selecting an optimal value for *K*. A small *K* can make the model sensitive to noise, while a large *K* can cause underfitting.
- Feature Importance: KNN treats all features equally unless some form of weighting or normalization is applied. This makes it sensitive to irrelevant or redundant features.

To enhance the performance of KNN, deep features extracted using a CNN can be used instead of raw input features. These deep features are more abstract, discriminative, and compact, leading to better similarity comparisons in the KNN process. In this context, CNN extracts powerful features, and KNN acts as a lightweight, non-parametric classifier that uses them for decision-making. The effectiveness of this approach depends on two key factors: the quality of features extracted by the CNN and the optimal choice of *K* in the KNN classifier. Any bias or poor selection in either may affect the final classification accuracy.

3) Decision Tree (DT)

DT uses branching to illustrate a series of steps and potential results. As it does not require any assumptions about the distribution of the data or the structure of the classifier, it can be used for categories and numerical variables. Large-dataset classifications may be obtained quickly and accurately with DT.

V. RESULTS AND DISCUSSION

Features are extracted using the proposed CNN. Each classification model responded differently, as shown in Table II. The most important things to examine are the size of the dataset and the parameters utilized to launch the network. In addition, a classifier must be placed above the layers. According to the results, the MSVM classifier was the best and fastest for diagnosing plant diseases from leaf images stored in a database, as shown in Tables III and IV.

TABLE II. COMPARISON OF DIFFERENT CLASSIFIERS WITH CNN

Performances	Acc %	Precision %	Recall %	F1-score %
CNN+MSVM	99.1	98.1	98.9	98.8
CNN+ KNN	95.5	95.1	94.4	94.4
CNN+ DT	94.4	94.1	91.1	92.6
CNN+NN	96.1	95.1	95.2	95.1

Table III compares the impact of augmented data on CNN performance, highlighting how data augmentation and advanced preprocessing improve accuracy and F1 scores, showcasing the importance of effective data preparation.

TABLE III. DATA AUGMENTATION IMPORTANCE

Technique	Number of images	Accuracy	Precision	F1-score
With augmentation	45409	99.1	95.3	98.1
Without augmentation	26311	95.3	92.1	94.1

TABLE IV. COMPARISON OF CNN HYPERPARAMETERS

Optimizer	Batch size	Learning rate	Accuracy (%)	Optimizer
Adam	16	0.001	95.1	Adam
Adam	32	0.001	97.1	Adam
Adam	64	0.001	99.1	Adam
SGD	16	0.0001	93.01	SGD
SGD	32	0.0001	94.1	SGD
SGD	64	0.0001	94.3	SGD
RMSProp	16	0.01	93.0	RMSProp
RMSProp	32	0.01	93.1	RMSProp
RMSProp	64	0.01	94.1	RMSProp

Figure 3 shows the training and validation accuracy over epochs, with both metrics steadily increasing and converging after approximately 20 epochs. Figure 4 depicts the training and validation losses, which decrease rapidly at the start and stabilize as epochs progress, indicating model stability and avoidance of overfitting.

Table V presents a performance comparison between several well-known CNN architectures for plant disease classification, including LeNet-5, AlexNet, VGG-16, ResNet-50, DenseNet-121, and MobileNet. The comparison is based on four key evaluation metrics: Accuracy, Precision, Recall, and F1-score. The proposed hybrid model (CNN+MSVM) outperformed all the other architectures, achieving the highest accuracy of 99.1%, along with superior values in the remaining metrics. This highlights the effectiveness of combining deep feature extraction using CNN with the classification capabilities of MSVM, leading to a significant improvement in overall classification performance.

TABLE V. COMPARISON BETWEEN DIFFERENT CNN ARCHITECTURES FOR IMAGE CLASSIFICATION

CNN architecture	Accuracy %	Precision %	Recall %	F1-score %
LeNet-5	85.2	84.7	85.0	84.8
AlexNet	88.9	88.5	88.0	88.2
VGG-16	91.5	91.2	91.0	91.1
ResNet-50	94.3	94.1	94.4	94.2
DenseNet-121	96.2	96.0	96.3	96.1
MobileNet	90.4	90.0	90.2	90.1
CNN + Softmax	97.2	96.9	96.8	96.8
Proposed CNN + MSVM	99.1	98.1	98.9	98.8

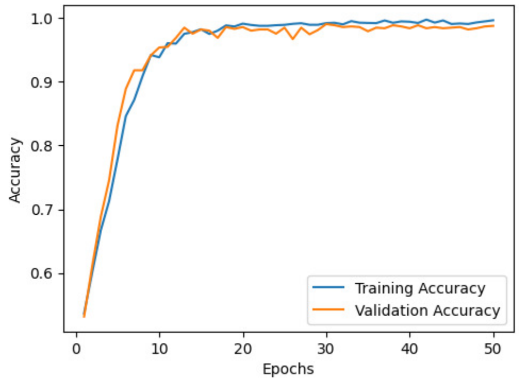


Fig. 3. Performance of the CNN-MSVM hybrid model using training and validation accuracy for 50 epochs.

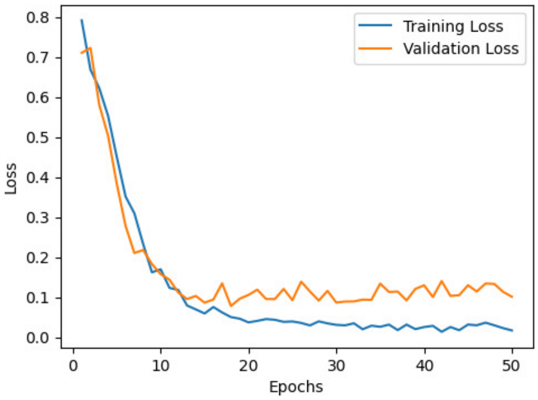
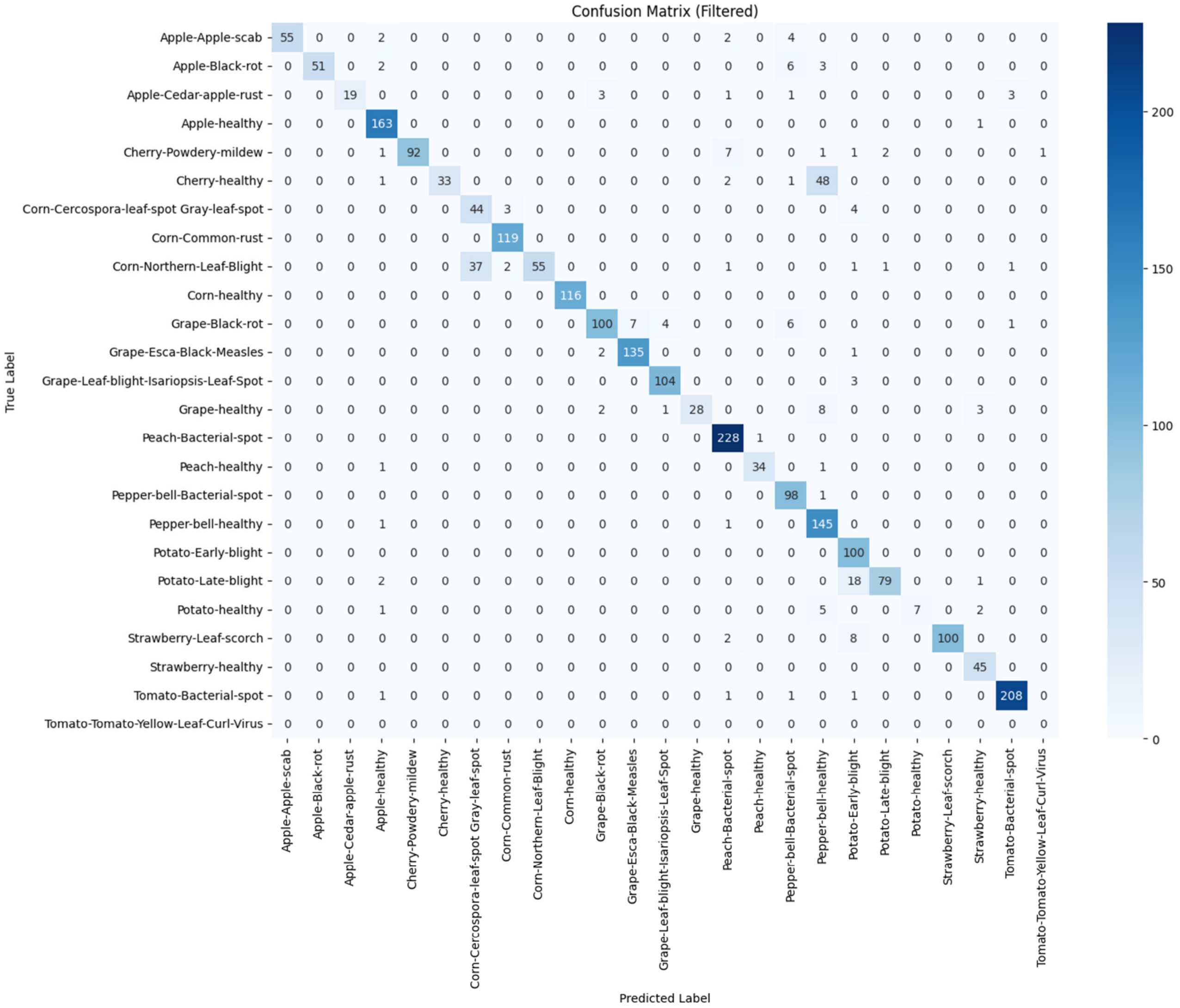


Fig. 4. Performance of the CNN-MSVM hybrid model using training and validation loss for 50 epochs.

The confusion matrix in Figure 5 shows the classification performance for 25 out of 33 plant disease categories using leaf and fruit images. Rows represent true labels, and columns represent predicted labels. The color gradient reflects prediction density, highlighting classification accuracy and confusion between classes. The matrix focuses on 25 key classes that were most representative and statistically significant in the testing phase. As the remaining eight classes were either underrepresented or had insufficient samples in the test set, they were excluded to ensure clarity and avoid skewed interpretation. This selective visualization allows for more meaningful performance analysis and highlights the model's effectiveness on the dominant and diagnostically important categories.



	Apple-Apple-scab	Apple-Black-rot	Apple-Cedar-apple-rust	Apple-healthy	Cherry-Powdery-mildew	Cherry-healthy	Corn-Cercospora-leaf-spot Gray-leaf-spot	Corn-Common-rust	Corn-Northern-Leaf-Blight	Corn-healthy	Grape-Black-rot	Grape-Esca-Black-Measles	Grape-Leaf-blight-Isariopsis-Leaf-Spot	Grape-healthy	Peach-Bacterial-spot	Peach-healthy	Pepper-bell-Bacterial-spot	Pepper-bell-healthy	Potato-Early-blight	Potato-Late-blight	Potato-healthy	Strawberry-Leaf-scorch	Strawberry-healthy	Tomato-Bacterial-spot	Tomato-Tomato-Yellow-Leaf-Curl-Virus
Apple-Apple-scab	55	0	0	2	0	0	0	0	0	0	0	0	0	0	2	0	4	0	0	0	0	0	0	0	0
Apple-Black-rot	0	51	0	2	0	0	0	0	0	0	0	0	0	0	0	0	6	3	0	0	0	0	0	0	0
Apple-Cedar-apple-rust	0	0	19	0	0	0	0	0	0	0	3	0	0	0	1	0	1	0	0	0	0	0	0	3	0
Apple-healthy	0	0	0	163	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Cherry-Powdery-mildew	0	0	0	1	92	0	0	0	0	0	0	0	0	0	7	0	0	1	1	2	0	0	0	0	1
Cherry-healthy	0	0	0	1	0	33	0	0	0	0	0	0	0	0	2	0	1	48	0	0	0	0	0	0	0
Corn-Cercospora-leaf-spot Gray-leaf-spot	0	0	0	0	0	0	44	3	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0
Corn-Common-rust	0	0	0	0	0	0	0	119	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Corn-Northern-Leaf-Blight	0	0	0	0	0	0	37	2	55	0	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0
Corn-healthy	0	0	0	0	0	0	0	0	0	116	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Grape-Black-rot	0	0	0	0	0	0	0	0	0	0	100	7	4	0	0	0	6	0	0	0	0	0	0	1	0
Grape-Esca-Black-Measles	0	0	0	0	0	0	0	0	0	0	2	135	0	0	0	0	0	0	1	0	0	0	0	0	0
Grape-Leaf-blight-Isariopsis-Leaf-Spot	0	0	0	0	0	0	0	0	0	0	0	0	104	0	0	0	0	0	3	0	0	0	0	0	0
Grape-healthy	0	0	0	0	0	0	0	0	0	0	2	0	1	28	0	0	0	8	0	0	0	0	3	0	0
Peach-Bacterial-spot	0	0	0	0	0	0	0	0	0	0	0	0	0	0	228	1	0	0	0	0	0	0	0	0	0
Peach-healthy	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	34	0	1	0	0	0	0	0	0	0
Pepper-bell-Bacterial-spot	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	98	1	0	0	0	0	0	0	0
Pepper-bell-healthy	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	145	0	0	0	0	0	0	0
Potato-Early-blight	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0
Potato-Late-blight	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	79	0	0	1	0	0
Potato-healthy	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	7	0	2	0	0
Strawberry-Leaf-scorch	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	8	0	0	0	100	0	0	0
Strawberry-healthy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	45	0	0
Tomato-Bacterial-spot	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	208	0
Tomato-Tomato-Yellow-Leaf-Curl-Virus	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Fig. 5. Confusion matrix of the proposed plant disease classification system.

VI. CONCLUSION

The use of CNNs for disease diagnosis in the Plant Village dataset is an important milestone in agriculture and global food security. CNNs have proven time and again their efficiency in classifying plant diseases with considerable accuracy and speed, making early action possible and reducing crop losses. They even exceed the precision of human specialists, making them an inevitable tool in modern agriculture. This study presented a hybrid approach that combines deep feature extraction with MSVM. The approach employs a multilevel and multiscale CNN to particularly improve prediction accuracy in plant leaf disease diagnosis. This multiscale structure allows for the detection of fine-grained visual features at various abstraction levels, thereby enhancing the discrimination ability of the model between various plant diseases even in the presence of large class variation. The hybrid model also demonstrated superior generalization performance and helped reduce the risk of overfitting, thereby being accurate and robust for real-world applications.

To increase the performance of the proposed hybrid model, pre-trained CNNs were fine-tuned with the Adam optimizer, batch size 64, and a learning rate of 0.001. This significantly enhanced model convergence and training accuracy compared to training the models from scratch. After deep features were extracted from the CNN, they were classified using MSVM, combining the strong feature representation of CNNs with the strong classification power of the MSVM. The results obtained were impressive, with an average precision of 98.1%, an average recall of 98.9%, an average F1 score of 98.8%, and an overall accuracy of 99.1%. These results support the notion that combining deep feature extraction with MSVM not only reduces training time but also significantly boosts classification performance, making it a potential contender for plant disease diagnosis at scale. The results obtained were compared with algorithms proposed in previous studies, as shown in Table V, demonstrating the superiority of the proposed method in terms of accuracy and performance.

REFERENCES

[1] N. S. Zaini, R. Karim, A. F. Abdull Razis, and N. Zawawi, "Utilizing Nutritional and Polyphenolic Compounds in Underutilized Plant Seeds for Health Application," *Molecules*, vol. 27, no. 20, Oct. 2022, Art. no. 6813, https://doi.org/10.3390/molecules27206813.

[2] D. R. F. R. Anandhi and S. Sathiamoorthy, "Enhanced Sea Horse Optimization with Deep Learning-based Multimodal Fusion Technique for Rice Plant Disease Segmentation and Classification," *Engineering, Technology & Applied Science Research*, vol. 13, no. 5, pp. 11959–11964, Oct. 2023, https://doi.org/10.48084/etasr.6324.

[3] S. Alqethami, B. Almtanni, W. Alzhrani, and M. Alghamdi, "Disease Detection in Apple Leaves Using Image Processing Techniques," *Engineering, Technology & Applied Science Research*, vol. 12, no. 2, pp. 8335–8341, Apr. 2022, https://doi.org/10.48084/etasr.4721.

[4] S. Anwar, S. R. Soomro, S. K. Baloch, A. A. Patoli, and A. R. Kolachi, "Performance Analysis of Deep Transfer Learning Models for the Automated Detection of Cotton Plant Diseases," *Engineering, Technology & Applied Science Research*, vol. 13, no. 5, pp. 11561–11567, Oct. 2023, https://doi.org/10.48084/etasr.6187.

[5] M. Nunes da Silva, M. G. Santos, M. W. Vasconcelos, and S. M. P. Carvalho, "Mitigation of Emergent Bacterial Pathogens Using Pseudomonas syringae pv. actinidiae as a Case Study—From Orchard to Gene and Everything in Between," *Crops*, vol. 2, no. 4, pp. 351–377, Dec. 2022, https://doi.org/10.3390/crops2040025.

[6] A. J., J. Eunice, D. E. Popescu, M. K. Chowdary, and J. Hemanth, "Deep Learning-Based Leaf Disease Detection in Crops Using Images for Agricultural Applications," *Agronomy*, vol. 12, no. 10, Oct. 2022, Art. no. 2395, https://doi.org/10.3390/agronomy12102395.

[7] 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, Feb. 2022, https://doi.org/10.1007/s40747-021-00536-1.

[8] F. Mohameth, C. Bingcai, and K. A. Sada, "Plant Disease Detection with Deep Learning and Feature Extraction Using Plant Village," *Journal of Computer and Communications*, vol. 08, no. 06, pp. 10–22, 2020, https://doi.org/10.4236/jcc.2020.86002.

[9] D. Argüeso *et al.*, "Few-Shot Learning approach for plant disease classification using images taken in the field," *Computers and Electronics in Agriculture*, vol. 175, Aug. 2020, Art. no. 105542, https://doi.org/10.1016/j.compag.2020.105542.

[10] K. Kaur and K. Bansal, "Enhancing Plant Disease Detection using Advanced Deep Learning Models," *Indian Journal of Science and Technology*, vol. 17, no. 17, pp. 1755–1766, Apr. 2024, https://doi.org/10.17485/IJST/v17i17.536.

[11] A. V. Panchal, S. C. Patel, K. Bagyalakshmi, P. Kumar, I. R. Khan, and M. Soni, "Image-based Plant Diseases Detection using Deep Learning," *Materials Today: Proceedings*, vol. 80, pp. 3500–3506, Jan. 2023, https://doi.org/10.1016/j.matpr.2021.07.281.

[12] V. Balafas, E. Karantoumanis, M. Louta, and N. Ploskas, "Machine Learning and Deep Learning for Plant Disease Classification and Detection," *IEEE Access*, vol. 11, pp. 114352–114377, 2023, https://doi.org/10.1109/ACCESS.2023.3324722.

[13] W. C. Yong, K. W. Ng, S. C. Haw, P. Naveen, and S. B. Ng, "Leaf Condition Analysis Using Convolutional Neural Network and Vision Transformer," *International Journal of Computing and Digital Systems*, vol. 16, no. 1, pp. 1685–1695, Oct. 2024, https://doi.org/10.12785/ijcds/1601125.

[14] A. J. Pandian and G. Geetharamani, "Data for: Identification of Plant Leaf Diseases Using a 9-layer Deep Convolutional Neural Network." Mendeley Data, Apr. 18, 2019, https://doi.org/10.17632/tywbtsjrjv.1.

[15] Z. J. Hussein, A. M. Hussein, G. I. Maki, and H. Q. Gheni, "Improved Model for Skin Illnesses Classification Utilizing Gray-Level Co-occurrence Matrix and Convolution Neural Network," *Journal of Advances in Information Technology*, vol. 14, no. 6, pp. 1273–1279, 2023, https://doi.org/10.12720/jait.14.6.1273-1279.

[16] A. Lu, Z. Fang, N. Farahpour, and L. Shannon, "CHIP-KNN: A Configurable and High-Performance K-Nearest Neighbors Accelerator on Cloud FPGAs," in *2020 International Conference on Field-Programmable Technology (ICFPT)*, Maui, HI, USA, Dec. 2020, pp. 139–147, https://doi.org/10.1109/ICFPT51103.2020.00027.