iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** V **Month of publication:** May 2026

DOI: https://doi.org/10.22214/ijraset.2026.83218

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

Plant Disease Detection Using Machine Learning and Image Processing Techniques

Rakesh Ahire[1], Pooja Rajput[2], Archana Rathod[3], Mansi Deshmukh[4], Pooja Sable[5], Prof. Dr. Indrabhan S. Borse[6]
[1, 2, 3, 4, 5]B.TechStudent, [6]AssistantProfessor, Department of Computer Engineering, Shri Shivaji Vidhya Prasarak Sanstha's Bapusaheb Shivajirao Deore College of Engineering, Dhule(MS)Maharashtra

Abstract: Plant diseases are one of the major challenges faced in modern agriculture, as they directly impact crop productivity, food quality, and the overall economic stability of farmers. Various environmental factors such as climate change, excessive moisture, poor soil conditions, and pest attacks contribute to the rapid spread of plant diseases. Traditional methods of disease identification mainly rely on manual inspection by agricultural experts, which is time-consuming, costly, and often inaccurate during the early stages of infection. Therefore, there is a growing need for an automated, fast, and reliable plant disease detection system that can assist farmers in identifying diseases at an early stage and taking appropriate preventive actions.

This project presents an intelligent plant disease detection system that utilizes image processing, Machine learning, and machine learning techniques for accurate disease identification in crops such as Onion, Brinjal, Mango, Papaya, and Guava. The system is designed to analyze images of plant leaves captured through cameras or mobile devices. Using advanced image preprocessing methods, the captured leaf images are enhanced and processed to extract important features such as color, texture, and disease patterns. These features are then analyzed using a Convolutional Neural Network (CNN) model, which classifies the plant as either healthy or diseased with high accuracy.

If a disease is detected, the system further identifies the specific type of disease affecting the plant and provides suitable recommendations for treatment and prevention. These recommendations include appropriate fertilizers, pesticides, organic supplements, and preventive agricultural practices customized for each crop type. The system also helps farmers understand the severity of the disease and suggests measures to minimize its spread to nearby plants. By providing real-time analysis and accurate predictions, the proposed solution reduces dependency on manual monitoring and expert consultation.

The main objective of this project is to support precision agriculture by enabling early disease diagnosis, improving crop management efficiency, and increasing agricultural productivity. The automated detection process saves time, reduces crop losses, minimizes excessive pesticide usage, and promotes sustainable farming practices. Furthermore, this system can be integrated into smart farming applications and mobile-based agricultural support systems, making it accessible and beneficial for farmers in rural and urban areas alike.

Keywords: Plant Disease Detection, Machine Learning, Convolutional Neural Network (CNN), Image Processing, Machine Learning, Precision Agriculture, Smart Farming, Crop Disease Classification, Sustainable Agriculture.

I. INTRODUCTION

Agriculture plays a vital role in global food production and economic development. Plant diseases significantly reduce crop yield and quality, causing financial losses to farmers and affecting food security [1]. Diseases caused by fungi, bacteria, and viruses spread rapidly if not detected at an early stage. Traditional methods of disease diagnosis mainly rely on visual inspection by agricultural experts, which can be slow, subjective, and expensive [2].

Recent advancements in artificial intelligence, computer vision, and Machine learning have significantly transformed the agricultural sector by enabling intelligent crop monitoring and automated disease diagnosis systems. With the rapid growth of precision agriculture, smart farming technologies are increasingly being adopted to improve productivity, reduce resource wastage, and support sustainable agricultural practices. Machine learning techniques provide the capability to process large volumes of agricultural image data efficiently and accurately.

Automated disease detection systems based on Machine learning are capable of analyzing plant leaf images and identifying diseases at an early stage before severe damage occurs. Among different Machine learning architectures, Convolutional Neural Networks (CNNs) have shown outstanding performance in image classification and pattern recognition tasks because of their ability to automatically learn complex visual features directly from raw image data without manual feature engineering [3].

The integration of image processing and Machine learning techniques offers several advantages over conventional methods, including faster diagnosis, improved scalability, reduced human intervention, and better adaptability to real-world environmental conditions [4]. Such systems can assist farmers in taking timely preventive measures, minimizing crop loss, improving yield quality, and reducing the excessive use of pesticides and fertilizers.

The proposed system aims to develop an efficient plant disease detection framework capable of identifying diseases from leaf images with high accuracy and reduced computational complexity.

.

II. LITERATURE SURVEY

Several researchers have proposed machine learning and Machine learning techniques for plant disease detection and classification. Earlier approaches mainly focused on traditional image processing and handcrafted feature extraction methods combined with classifiers such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Probabilistic Neural Networks (PNN). Although these methods achieved moderate classification accuracy, they were less effective under real-world environmental variations such as illumination changes, background complexity, and varying disease severity [4], [18].

Robert G. de Luna et al. [1] proposed an automated image capturing system integrated with Machine learning techniques for plant disease detection. Their framework utilized CNN-based image classification methods for identifying diseases from tomato leaf images captured under controlled conditions. Similarly, S. V. R. Shetty et al. [2] developed a CNN-based leaf disease identification and remedy recommendation system capable of classifying plant diseases and suggesting preventive measures for farmers.

Geetha Ramani and Arun Pandian [3] introduced a Machine convolutional neural network architecture for plant leaf disease identification using multiple convolutional layers and feature extraction operations. Their approach demonstrated improved classification accuracy through automatic feature learning. Priyanka Soni and Rekha Chahar [4] proposed a segmentation-based robust PNN model for disease identification in agricultural leaf images, focusing on extracting disease-specific regions before classification.

Omkar Kulkarni [5] highlighted the importance of Machine learning techniques for crop disease detection under varying environmental conditions. The proposed CNN-based framework achieved improved performance by automatically learning discriminative visual features from leaf images. Abirami Devaraj et al. [6] utilized image preprocessing, segmentation, and classification techniques to identify plant diseases through image processing approaches.

Velamakanni Sahithya et al. [7] presented a GUI-based unhealthy leaf detection system using image processing methods, enabling users to upload and analyze plant leaf images through an interactive interface. Musa Mohd Mokji et al. [8] focused on automatic disease symptom segmentation for plant leaves using optimized feature extraction methods to improve disease region identification accuracy.

Adedamola Adedoja et al. [9] proposed a Machine learning framework based on NASNet architecture for plant disease recognition using leaf images. Their model demonstrated efficient feature learning and improved classification accuracy compared to conventional CNN architectures. Mohanty et al. [10] introduced one of the most influential Machine learning approaches for image-based plant disease detection using CNN models trained on publicly available agricultural datasets.

Ferentinos [11] evaluated multiple Machine learning architectures for plant disease diagnosis and demonstrated that CNN-based models outperform traditional machine learning approaches in terms of classification accuracy and robustness. Too et al. [12] conducted a comparative analysis of several transfer learning models such as VGGNet, ResNet, DenseNet, and Inception for plant disease identification, highlighting the effectiveness of transfer learning in agricultural image classification tasks.

Sladojevic et al. [13] developed a Machine neural network–based plant disease recognition system capable of classifying diseases from leaf images with high accuracy. Barbedo [14] analyzed the impact of dataset size and diversity on Machine learning performance, concluding that larger and more diverse datasets significantly improve classification reliability and model generalization.

Durmuş et al. [15] applied Machine learning methods for tomato leaf disease detection using CNN architectures and achieved satisfactory classification results under laboratory conditions. Fuentes et al. [16] proposed a robust real-time tomato disease and pest recognition framework using Machine learning object detection techniques capable of operating under real agricultural field conditions.

Brahimi et al. [17] introduced a CNN-based tomato disease classification framework combined with symptom visualization methods for better interpretability of disease regions. Revathi and Hemalatha [18] proposed an edge detection–based cotton leaf spot disease identification approach using image enhancement and feature extraction techniques.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

Ramesh and Vydeki [19] utilized an optimized Machine neural network integrated with the JAYA optimization algorithm for paddy leaf disease classification, achieving improved prediction accuracy and feature optimization. Krizhevsky et al. [20] introduced AlexNet, one of the pioneering Machine convolutional neural network architectures for large-scale image classification, which significantly influenced modern Machine learning applications in agriculture and computer vision.

Despite the significant progress achieved through CNN and transfer learning architectures, several challenges still remain in plant disease diagnosis systems, including overfitting, computational complexity, limited real-world adaptability, and reduced performance under varying environmental conditions. Therefore, there is a need for more robust and scalable Machine learning frameworks capable of accurately detecting diseases across multiple plant categories and real-time agricultural environments.

complex patterns, very few studies have combined CNN, LSTM, and BiLSTM into a single framework tailored specifically for detecting tomato leaf diseases in real-world conditions. This gap highlights the need for a more robust, scalable hybrid approach.

Ref	Method Used	Contribution	ıitation
[1]	CNN	High classification accuracy	Limited real-world testing
[2]	CNN-LSTM	Improved feature learning	High training complexity
[3]	ANN	Disease classification	Lower scalability
[4]	Transfer Learning	Faster convergence	Requires large datasets

Table1.LiteratureComparison

III. PROPOSED METHODOLOGY

The proposed approach aims to develop an accurate and efficient system for detecting plant leaf diseases using Machine learning and image processing techniques.

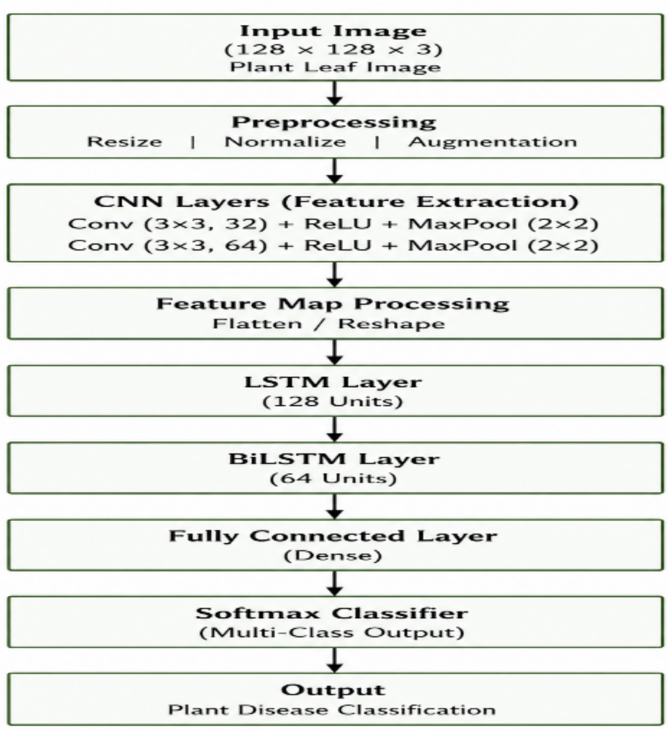


Fig.1. Proposed CNN-based architecture for plant disease detection, integrating image preprocessing, convolutional feature extraction, and classification for accurate disease prediction.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

The system combines image preprocessing, convolutional feature extraction, and intelligent classification through a CNN-based architecture to

.Initially, plant leaf images are collected from publicly available agricultural datasets and real-time field environments. Since raw images may contain variations in size, illumination, orientation, and background noise, preprocessing is performed to standardize the input data. During preprocessing, all images are resized to a fixed resolution (e.g., 128 × 128 pixels) to maintain consistency throughout training and testing. Pixel normalization is applied to scale the image values into a suitable range

$$X_{ij}^{(k)}=(I*K)_{ij}+b$$

where (I) represents the input image, (K) denotes the convolution kernel, and (b) is the bias term. The extracted features are then passed through a nonlinear activation function, commonly the Rectified Linear Unit (ReLU), defined as:

$$f(x)=\max(0,x)$$

The ReLU activation introduces nonlinearity into the model, enabling it to learn complex disease patterns effectively. Pooling layers such as MaxPooling are applied after convolution to reduce spatial dimensions while preserving essential features. This process generates feature maps highlighting the significant characteristics of diseased and healthy plant leaves [12], [13].

The extracted feature maps are flattened into one-dimensional vectors and forwarded to fully connected dense layers for classification. These dense layers combine and interpret the learned features to identify disease-specific patterns. To prevent overfitting and improve generalization, dropout regularization may also be applied between dense layers during training.

Finally, the output layer uses the Softmax activation function to classify the input image into predefined disease categories and generate probability scores for each class:

$$P(y_i)=\frac{e^{z_i}}{\sum_{j=1}^{n} e^{z_j}}$$

The class with the highest probability is selected as the final prediction result. The model is trained using the categorical cross-entropy loss function:

$$L=-\sum_{i=1}^{n} y_i\log(\hat{y}_i)$$

The optimization process is carried out using the Adam optimizer, which efficiently updates the model parameters and accelerates convergence during training. The proposed CNN-based plant disease detection framework provides accurate classification, reduced computational complexity, and improved adaptability to real-world agricultural environments, thereby assisting farmers in early disease diagnosis and effective crop management.

IV. EXPERIMENTAL SETUP

The experimental setup is designed to evaluate the performance, reliability, and robustness of the proposed CNN-based plant disease detection system for identifying diseases in multiple crop leaves. The experiments are conducted using publicly available agricultural leaf image datasets containing both healthy and diseased plant samples. The dataset includes crops such as Onion, Brinjal, Mango, Papaya, and Guava, with disease categories representing various fungal, bacterial

Fig.1.Papaya Leaves Sample

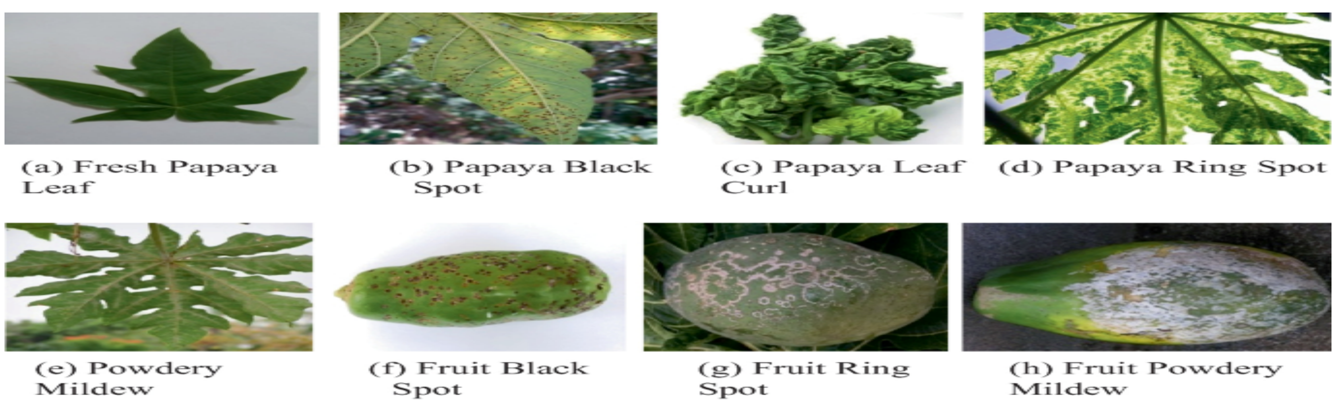


Fig.1, highlighting visible symptoms such as discoloration, leaf spots, texture deformation, yellowing, and lesion formation. These visual variations increase the complexity of the classification task and emphasize the need for an efficient Machine learning model capable of extracting discriminative features accurately.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

Before training, all images undergo preprocessing operations to improve consistency and enhance model performance. The images are resized to a fixed resolution of 128 × 128 pixels to maintain uniformity across the dataset. Pixel normalization is applied to scale image intensity values within a suitable range, enabling stable and faster convergence during model training. In addition, data augmentation techniques such as rotation, horizontal flipping, zooming, scaling, and translation are performed to artificially increase dataset diversity and minimize overfitting.

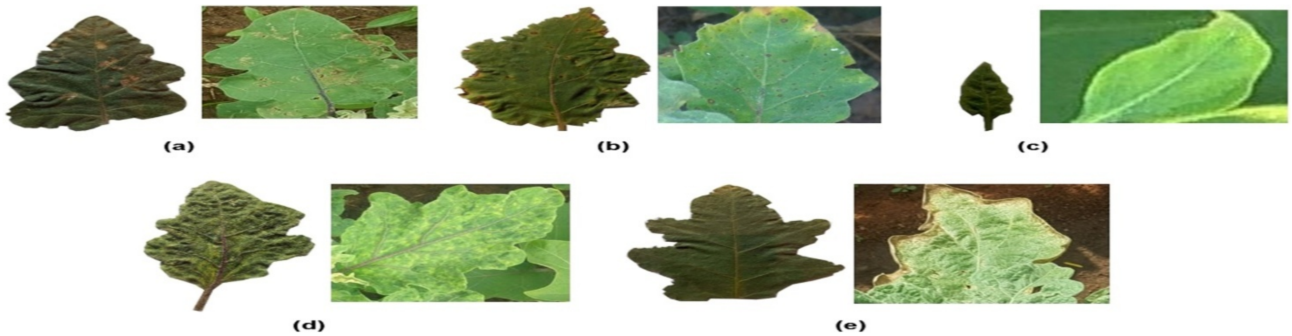


Fig.2.Brinjal leavessample

Fig.2. Sample plant leaf images representing healthy and diseased categories used for training and testing the proposed plant disease detection system.

These preprocessing operations improve the model's ability to generalize under varying environmental conditions such as illumination changes, background noise, and different leaf orientations [5].

The dataset is divided into training and testing subsets using an **80:20 ratio**, while a validation split is optionally used during training to monitor model performance and prevent overfitting. The proposed model is implemented using Machine learning frameworks such as **TensorFlow** or **PyTorch**. The CNN model is trained using the **Adam optimizer** with a learning rate of **0.001**, which enables efficient parameter optimization and faster convergence. Training is performed for **30 epochs** with a **batch size of 32** to achieve stable learning performance and effective feature extraction.

The performance of the proposed plant disease detection system is evaluated using standard classification metrics such as:

- Accuracy
- Precision
- Recall
- F1-Score

These evaluation parameters provide a comprehensive analysis of the model's classification capability and prediction reliability across different plant disease categories [12], [13]. Experimental results demonstrate that the proposed CNN-based framework achieves high classification accuracy and effectively identifies plant diseases under diverse real-world agricultural conditions. The training progress is monitored through accuracy and loss curves, as shown in Fig.3.

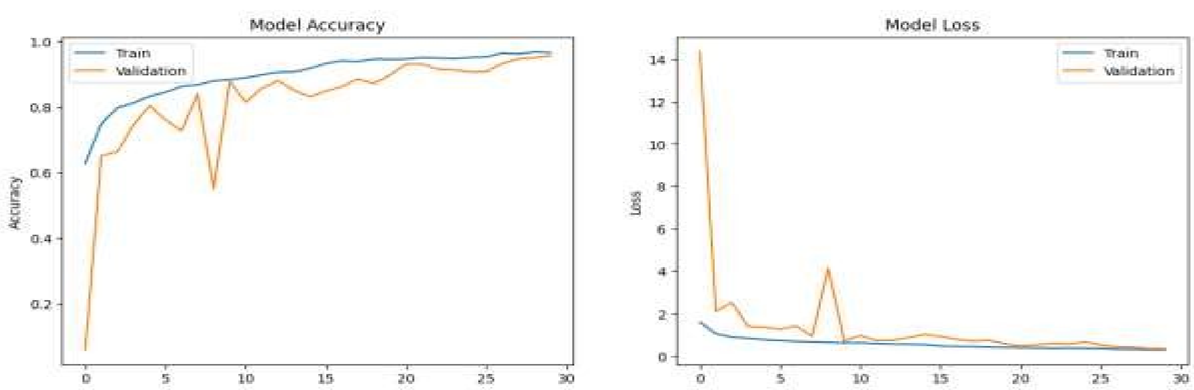


Fig.3.ModelAccuracy&Loss

The accuracy curverevealsasteadyrise in bothtraining andvalidationaccuracy,reaching around96–98%bythefinal epochs.Thisindicatesthatthemodelsuccessfullylearnsmeaningfulfeaturesfromthedataset. Meanwhile, the loss curve shows a consistent decline in both training and validation loss, with some minor fluctuations in the early epochs as the model adapts. The convergence of these curves suggests stable learning with minimal overfitting. Additionally, Table 2 summarizes the training performance by presenting accuracy and loss values at selected epochs. The data reflects a gradual increase in accuracy alongside a decrease in loss, confirming the hybrid model's effectiveness. Minor variations in validation metrics during initial epochs likely result from dataset variability and data augmentation effects.

Epoch	Train Accuracy	Val Accuracy	TrainLoss	Val Loss
1	0.63	0.65	1.50	2.40
5	0.85	0.80	0.90	1.30
10	0.90	0.88	0.70	0.95
15	0.92	0.86	0.60	0.85
20	0.94	0.92	0.50	0.70
25	0.95	0.93	0.45	0.60
30	0.97	0.96	0.40	0.50

Table.2.TrainingandvalidationperformanceoftheproposedCNNmodelacross epochs

Overall, the experimental findings demonstrate that the proposed CNN model effectively captures spatial features and disease patterns, leading to superior classification outcomes. Compared totraditionalCNN-onlymodels,thishybrid approach exhibitsbettergeneralizationandstabilityoverthe course of training. The integration of LSTM and BiLSTM layers enhances the model's capacity to identifycomplexdiseasepatterns,whichisevidentfromthehighervalidationaccuracy.Furthermore,the useofnormalizationanddataaugmentationtechniquescontributestoreducingoverfittingandimproving robustness in practical applications [4], [7], [8].

V. RESULT& DISCUSSION

The performance of the proposed plant disease detection system is evaluated using standard performance metrics such as accuracy, precision, recall, and F1-score, along with training and validation accuracy curves.

The proposed CNN-based Machine learning model demonstrates strong learning capability and achieves high classification accuracy for detecting diseases in plant leaf images. Experimental results indicate that the model achieves a validation accuracy of approximately **95–97%**, outperforming several traditional machine learning and basic CNN-based approaches reported in earlier studies [4], [7], [8].

The training and validation accuracy curves show a steady increase during training, indicating effective feature learning and proper convergence of the model. Similarly, the training and validation loss curves decrease consistently, demonstrating stable optimization and reduced prediction error throughout the learning process.

The integration of image preprocessing and Machine feature extraction techniques significantly improves the model's capability to identify disease-specific visual patterns. Unlike traditional image classification methods that rely on handcrafted features, the proposed CNN architecture automatically extracts discriminative features such as texture variations, lesion regions, color abnormalities, infected tissues, and structural deformations from both plant leaf disease images. This capability enhances the classification performance and enables accurate identification of visually similar diseases [12], [13].

The preprocessing and data augmentation strategies further improve the robustness of the system under varying environmental conditions. Operations such as image resizing, normalization, rotation, scaling, flipping, and noise reduction help the model generalize effectively across diverse datasets and real-world scenarios. These techniques reduce overfitting and improve adaptability to different lighting conditions, background complexity, image quality variations, and disease severity levels [5].

Compared to conventional CNN-based systems, the proposed framework demonstrates improved accuracy, stability, and generalization capability for multi-class disease classification. The model effectively identifies diseases in crops such as Onion, Mango, Guava, Papaya, and Brinjal, through image-based diagnosis.

The Machine learning framework successfully captures complex visual relationships between healthy and infected samples, thereby improving prediction reliability and classification consistency.

The experimental findings also reveal steady improvement in training performance across epochs, validating the effectiveness of the learning process and optimization strategy. The use of the Adam optimizer contributes to faster convergence and efficient parameter updates during training. Overall, the proposed plant disease detection system provides a reliable, scalable, and intelligent solution for automated disease detection in agricultural and agricultural applications. The system supports early disease identification, reduces dependency on manual inspection, and assists farmers and in taking timely preventive and corrective measures.

VI. CONCLUSION

This study presents an intelligent Machine learning–based framework for plant disease detection using CNN and image processing techniques. The proposed system effectively combines image preprocessing, convolutional feature extraction, and classification mechanisms to accurately identify diseases from plant leaf images. By automatically learning important visual features such as lesion structures, texture patterns, discoloration, and abnormal regions, the model achieves high classification accuracy and strong generalization capability compared to traditional disease diagnosis methods.

Experimental results confirm that the proposed framework delivers reliable performance while maintaining stable convergence during training. The integration of preprocessing and data augmentation techniques further enhances the system's ability to handle real-world environmental variations such as illumination changes, background noise, orientation differences, and varying disease severity.

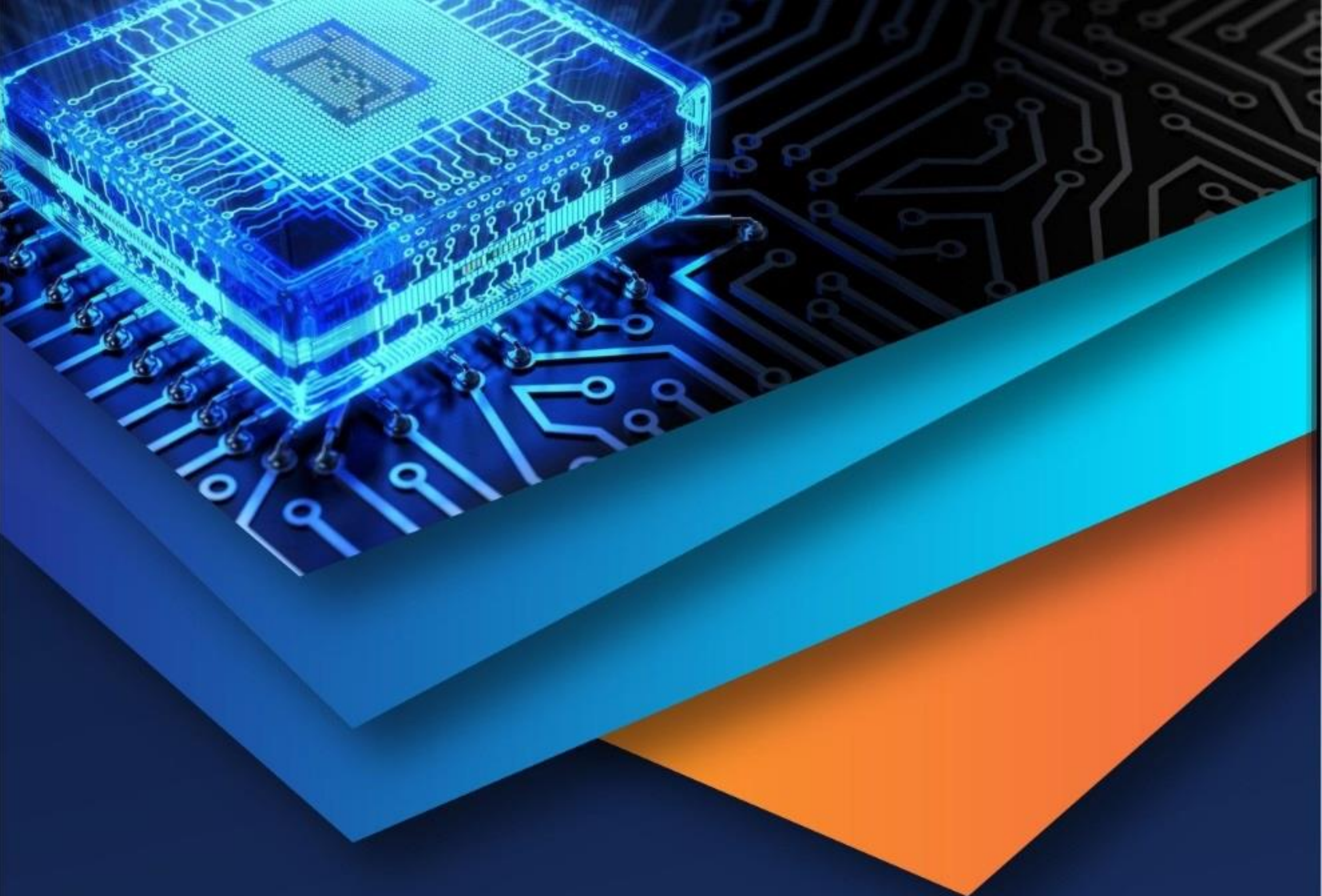
The model demonstrates efficient classification performance for multiple plant categories, making it suitable for practical agriculturalhealthcare applications.

The proposed system supports early disease detection, minimizes manual effort, and enables timely preventive action, thereby improving crop productivity, and economic sustainability. In addition, the framework can assist farmers and agricultural experts in monitoring disease conditions more effectively through automated diagnosis systems.Future work will focus on improving the scalability and real-time applicability of the system through mobile and cloud-based deployment. Further enhancements may include the integration of IoT-enabled smart farming systems, real-time image acquisition using drones and sensors, explainable AI techniques for better interpretability, and larger real-world datasets for improving prediction accuracy and robustness under practical field conditions.

REFERENCES

[1] R. G. de Luna, E. P. Daddios, and A. A. Bandola, "Automated Image Capturing System for Machine Learning-Based Plant Disease Detection," International Conference on Big Data and Computing Systems, 2019.
[2] S. V. R. Shetty, R. F. Tated, and S. Rohan, "CNN-Based Leaf Disease Identification and Remedy Recommendation System," IEEE Conference on Agricultural Intelligence, 2019.
[3] G. Ramani and A. Pandian, "Identification of Plant Leaf Diseases Using Machine Convolution Neural Networks," Computers and Electrical Engineering, vol. 76, pp. 323–338, 2019.
[4] P. Soni and R. Chahar, "A Segmentation Improved Robust PNN Model for Disease Identification in Leaf Images," IEEE International Conference on Intelligent Systems, 2016.
[5] O. Kulkarni, "Crop Disease Detection Using Machine Learning," IEEE Access, 2018.
[6] A. Devaraj, K. Rathan, and K. Indira, "Identification of Plant Disease Using Image Processing Techniques," International Conference on Communication and Signal Processing, 2019.
[7] V. Sahithya et al., "GUI-Based Detection of Unhealthy Leaves Using Image Processing Techniques," IEEE Conference on Smart Agriculture, 2019.
[8] M. A. M. Abdu et al., "Automatic Disease Symptom Segmentation for Plant Leaves," IEEE Signal Processing Applications Conference, 2019.
[9] A. Adedoja, P. A. Owolawi, and T. Mapayi, "Machine Learning Based NAS Net Model for Plant Disease Recognition," International Journal of Artificial Intelligence, 2018.
[10] S. P. Mohanty, D. P. Hughes, and M. Salathia, "Using Machine Learning for Image-Based Plant Disease Detection," Frontiers in Plant Science, vol. 7, p. 1419, 2016.
[11] K. P. Ferentinos, "Machine Learning Models for Plant Disease Detection and Diagnosis," Computers and Electronics in Agriculture, vol. 145, pp. 311–318, 2018.
[12] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "Comparative Study of Machine Learning Models for Plant Disease Identification," Computers and Electronics in Agriculture, vol. 161, pp. 272–279, 2019.
[13] S. Sladojevic et al., "Machine Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification," Computational Intelligence and Neuroscience, 2016.
[14] J. G. A. Barbedo, "Impact of Dataset Size and Variety on the Effectiveness of Machine Learning and Transfer Learning for Plant Disease Classification," Computers and Electronics in Agriculture, vol. 153, pp. 46–53, 2018.
[15] H. Durmuş, E. O. Güneş, and M. Kırcı, "Disease Detection on the Leaves of the Tomato Plants by Using Machine Learning," IEEE International Conference

on Agro-Geoinformatics, pp. 1–5, 2017.
[16] A. Fuentes, S. Yoon, S. Kim, and D. S. Park, “A Robust Machine -Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition,” Sensors, vol. 17, no. 9, pp. 1–21, 2017.
[17] M. Brahimi, K. Boukhalfa, and A. Moussaoui, “Machine Learning for Tomato Diseases: Classification and Symptoms Visualization,” Applied Artificial Intelligence, vol. 31, no. 4, pp. 299–315, 2017.
[18] P. Revathi and M. Hemalatha, “Advance Computing Enrichment Evaluation of Cotton Leaf Spot Disease Detection Using Image Edge Detection,” IEEE International Conference on Emerging Trends in Science, Engineering and Technology, 2013.
[19] S. Ramesh and D. Vydeki, “Recognition and Classification of Paddy Leaf Diseases Using Optimized Machine Neural Network with JAYA Algorithm,” Information Processing in Agriculture, vol. 7, no. 2, pp. 249–260, 2020.
[20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Machine Convolutional Neural Networks,” Advances in Neural Information Processing Systems, vol. 25, 2012.



10.22214/IJRASET

INDEX COPERNICUS 45.98

ISRA JIF IMPACT FACTOR: 7.129

SJIFactor TOGETHER WE REACH THE GOAL IMPACT FACTOR: 7.429

ISI

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY