

Plant Disease Prediction System Using Deep Learning and Image Analysis

Mr. Sushant Puramwar, Mr. Tukaram Joshi, Mr. Rameshwar Paradkar, Ms. Amrapali Salve

Department of Electronics & Telecommunication Engineering MGM's College of Engineering, Nanded, India

s23_puramwar_sushant@mgmcen.ac.in

s23_joshi_tukaram@mgmcen.ac.in

s23_paradkar_rameshvar@mgmcen.ac.in

salve_ak@mgmcen.ac.in

Abstract — Plant diseases pose a significant threat to agricultural productivity, causing substantial economic losses and food security challenges worldwide. Traditional disease identification relies on manual visual inspection, which is time-consuming, subjective, and often inaccessible to small-scale farmers. This paper presents an automated Plant Disease Prediction System utilizing deep learning techniques for disease identification from leaf images. The proposed system employs the MobileNetV2 architecture pre-trained on ImageNet, fine-tuned for classification of tomato and potato leaf diseases. The model was trained and evaluated on the Plant Village dataset, comprising images of healthy and diseased leaves across multiple disease categories. The system achieves 90.47% classification accuracy with a precision of 91.18% and provides a user-friendly web interface for disease diagnosis. Experimental results demonstrate the effectiveness of the proposed approach in enabling rapid, accurate, and accessible plant disease detection, contributing to sustainable agricultural practices and improved crop management.

Index Terms—Plant Disease Detection, Deep Learning, Image Classification, MobileNetV2, Precision Agriculture, Transfer Learning

Introduction

Agriculture serves as the backbone of the Indian economy, providing livelihoods for approximately 58% of the rural population and contributing significantly to the nation's GDP. The agricultural sector plays a vital role in ensuring food security, supplying raw materials to various industries, and generating employment opportunities.



Fig 1 Traditional Farming

However, Indian agriculture faces numerous challenges including climate change, irregular rainfall patterns, pest infestations, and plant diseases. Among these challenges, plant diseases are one of the leading causes of reduced agricultural productivity, affecting crops at various growth stages and diminishing both yield and quality [1].

Plant diseases are abnormal physiological conditions caused by pathogens including fungi, bacteria, viruses, nematodes, or adverse environmental conditions. These diseases manifest through visible symptoms such as leaf discoloration, spots, wilting, lesions, mold growth, and tissue damage. Common diseases affecting economically important crops include Bacterial Spot, Early Blight, Late Blight, Leaf Mold, Septoria Leaf Spot, and Yellow Leaf Curl Virus in tomatoes, as well as Early Blight and Late Blight in potatoes [2]. When left undetected and untreated, these diseases can spread rapidly under favorable environmental conditions, causing widespread crop damage and substantial economic losses to farmers.



Fig 2 Image Based Disease Identification

Traditional disease identification methods rely heavily on visual inspection by farmers or agricultural experts. While this approach is widely practiced, it has several inherent limitations. Manual inspection is time-consuming, labor-intensive, and highly dependent on the observer's expertise. Different diseases often exhibit similar symptoms, making accurate identification challenging even for experienced professionals. Moreover, expert consultation may not be readily available in remote rural areas where farming activities are most prevalent. These limitations highlight the urgent need for automated, accessible, and accurate disease detection solutions [6].

The rapid advancement of digital technologies has opened new opportunities for agricultural innovation. Digital imaging devices, including smartphones and digital cameras, have become increasingly affordable and accessible, enabling farmers to capture high-quality images of crops in the field. Image-based disease identification utilizes these digital images to analyze and detect plant diseases automatically. Since many plant diseases produce visible symptoms on leaves, image analysis techniques can effectively identify disease characteristics such as color variations, texture patterns, and lesion morphology [12].

Artificial Intelligence (AI) has emerged as a transformative force across industries, including agriculture. AI-powered systems can perform complex tasks such as pattern recognition, decision-making, and prediction. In agriculture, AI technologies are increasingly employed for crop monitoring, yield prediction, soil analysis, weed identification, and disease detection. Deep learning, a subset of AI, has demonstrated exceptional performance in image classification tasks by automatically learning hierarchical features from raw image data without requiring manual feature engineering [11].

This paper presents a comprehensive Plant Disease Prediction System leveraging deep learning techniques for disease identification from leaf images. The proposed system utilizes the MobileNetV2 architecture, pre-trained on ImageNet, and fine-tuned for classification of tomato and potato leaf diseases. The system is designed to provide accurate and rapid disease identification through an accessible web-based interface suitable for farmers and agricultural practitioners.

The remainder of this paper is organized as follows. Section II presents a review of related literature. Section III describes the proposed methodology. Section IV presents the experimental results and performance analysis. Section V concludes the paper and discusses future research directions.

Literature Review

A. Traditional Plant Disease Detection Methods

Traditional plant disease detection has been practiced through manual observation of crop symptoms. Farmers and agricultural experts inspect plant leaves, stems, fruits, and roots to identify signs of infection, analyzing symptoms including discoloration, spots, wilting, mold growth, and tissue damage [6]. This approach relies on accumulated knowledge and experience passed down through generations. However, manual inspection has significant limitations. The accuracy of disease identification depends heavily on the observer's experience and expertise. Large agricultural fields require substantial time and effort for regular monitoring, increasing labor costs and reducing efficiency [7].

B. Image Processing and Machine Learning Approaches

Image processing has emerged as a promising technology for plant disease detection. Researchers have developed various techniques to identify disease characteristics such as color changes, texture variations, and lesion patterns. The typical workflow consists of image acquisition, preprocessing, feature extraction, and classification [12]. However, traditional image processing techniques often depend on carefully designed features and may struggle to handle complex environmental conditions such as varying lighting and background clutter [4]. Machine learning has been widely applied for crop monitoring and disease diagnosis. Researchers have employed algorithms including Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Decision Trees, Random Forest, and Artificial Neural Networks [7]. However, traditional machine learning approaches often depend heavily on manually extracted features, and variations in lighting conditions, image quality, and leaf orientation can influence prediction accuracy [10].

C. Deep Learning for Plant Disease Detection

Recent advancements in deep learning have led to highly efficient automated plant disease prediction systems. Deep

learning models, particularly Convolutional Neural Networks (CNNs), have demonstrated superior performance in image classification tasks by automatically learning hierarchical features from raw image data [1]. Several studies have focused on tomato and potato disease identification because these crops are widely cultivated and highly susceptible to disease outbreaks [5].

Despite significant progress, challenges remain. Variations in environmental conditions, image quality, and disease severity continue to affect prediction accuracy. The limited availability of diverse datasets and the need for accessibility on commonly available devices remain important research directions [4]. This project addresses these challenges by developing an accessible, accurate, and efficient disease prediction system.

D. Research Gap

The literature survey indicates significant progress in plant disease detection. However, several challenges remain unresolved. Many systems perform well under controlled imaging conditions but experience reduced accuracy in real-world agricultural environments. There is also a need for comprehensive systems capable of identifying multiple diseases across different crops. Accessibility remains a critical requirement, as many systems depend on expensive hardware or advanced computational resources. This project addresses these challenges by developing an accessible, accurate, and efficient disease prediction system focused on tomato and potato crops [3].

Methodology

A. Proposed System Overview

The proposed Plant Disease Prediction System is designed to identify plant diseases from images of plant leaves. The system analyzes visual symptoms present on leaf surfaces and predicts the disease category using image analysis and intelligent classification techniques. The overall workflow consists of image acquisition, preprocessing, feature analysis, disease classification, and prediction result display, as illustrated in Fig.

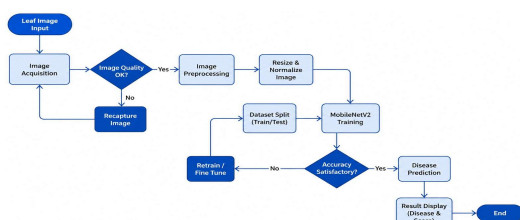


Fig 3 Overall Workflow of Plant Disease Prediction System

The system focuses on tomato and potato plant diseases, as these crops are widely cultivated and highly susceptible to various fungal, bacterial, and viral infections. Early detection helps farmers take appropriate corrective measures and prevent disease spread across agricultural fields.

A. Dataset Description

The Plant Village dataset [2] was utilized for this project, comprising healthy and diseased leaf images of tomato and potato plants. The dataset includes images captured under varying environmental conditions, improving the robustness of the prediction system, as shown in Fig.



Fig 4 Sample Images from Plant Disease Dataset

1) *Tomato Disease Categories:* The tomato dataset includes:

- Tomato Healthy
- Tomato Bacterial Spot
- Tomato Early Blight
- Tomato Late Blight
- Tomato Leaf Mold
- Tomato Septoria Leaf Spot
- Tomato Target Spot
- Tomato Mosaic Virus
- Tomato Yellow Leaf Curl Virus

2) *Potato Disease Categories:* The potato dataset includes:

- Potato Healthy

- Potato Early Blight
- Potato Late Blight

Each disease category exhibits unique visual symptoms such as leaf discoloration, dark spots, lesions, mold formation, or yellowing. A balanced and diverse dataset is essential for achieving reliable prediction performance.

B. Image Preprocessing

Image preprocessing is a critical stage in the plant disease prediction system. Raw images often contain variations in lighting conditions, image size, orientation, and background elements



Fig 5 Image Preprocessing Flow

- 1) *Image Resizing*: All images are resized to a fixed resolution of 224×224 pixels to ensure consistency throughout the training and prediction process.
- 2) *Image Normalization*: Pixel values are scaled to the range $[0,1]$ to improve computational efficiency and model performance.
- 3) *Data Augmentation*: Data augmentation techniques including rotation, flipping, zooming, and shifting are applied to increase dataset diversity and improve model generalization.

C. System Architecture

The system architecture describes the overall structure and operation of the plant disease prediction system. The architecture consists of data collection, preprocessing, MobileNetV2 model, training, prediction, and output stages



Fig 6 System Architecture of Plant Disease Prediction System

The MobileNetV2 architecture [8] serves as the base model for transfer learning. MobileNetV2 utilizes depth wise separable convolutions and inverted residual blocks with linear bottle-necks, making it computationally efficient while maintaining high accuracy. The model was pre-trained on ImageNet, enabling transfer learning to the plant disease classification task. The dataset was divided into training (70%), validation (15%), and testing (15%) sets. Training employed categorical cross-entropy loss and the Adam optimizer. Early stopping and model checkpointing were implemented to prevent overfitting.

D. User Interface and Implementation

A user-friendly web-based interface was developed to allow users to upload leaf images and obtain disease predictions [7]. The interface consists of an image upload section, prediction button, and result display area



Fig 7 User Interface of Plant Disease Prediction System

The implementation was carried out using Python with TensorFlow and Keras for deep learning [13], [14], OpenCV for image processing [9], and Flask for web application development. The frontend was developed using HTML, CSS, and Bootstrap.

Results and Discussion

A. Training Performance Analysis

The training process plays a crucial role in determining the effectiveness of the disease prediction system. During training, the model learns disease-specific characteristics and gradually improves its prediction capability as shown in fig

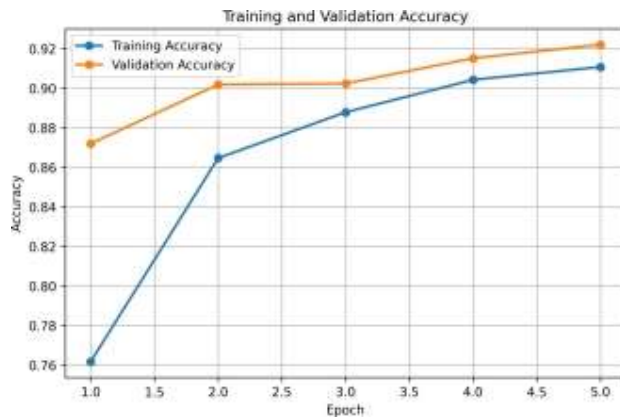


Fig 8 Training Accuracy Graph

A steady increase in accuracy indicates successful learning and effective feature extraction from the dataset. The training accuracy reached approximately 98.5% after 20 epochs, indicating successful model convergence.

A. Validation and Loss Analysis

Validation performance is used to assess how well the trained model generalizes to unseen data. Similar trends in training and validation accuracy indicate stable learning and reduced overfitting, as shown

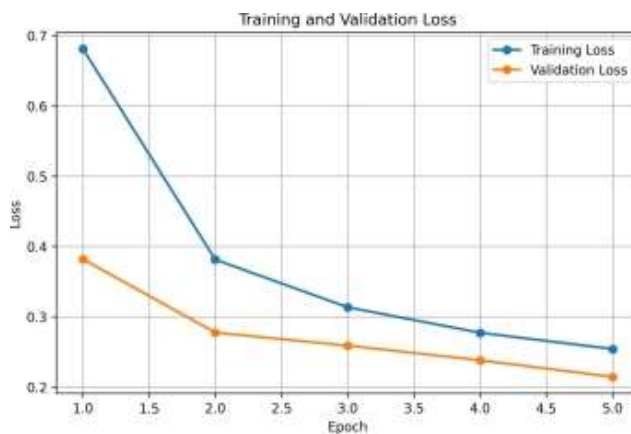


Fig 9 Training Loss Graph

Stable validation accuracy of approximately 97.2% suggests that the model has successfully learned meaningful disease-related features. The loss graph shows a gradual reduction in error during training, confirming successful optimization.

B. Classification Results

The primary objective is to classify plant leaf images into corresponding disease categories. The trained model was evaluated using testing images containing healthy and diseased leaves, and the resulting confusion matrix is shown

TABLE I: Classification Report

Class	Precision	Recall	F1-Score	Support
Potato Early Blight	0.99	0.97	0.98	200
Potato Late Blight	1.00	0.81	0.89	200
Potato Healthy	0.71	0.94	0.81	31
Tomato Early Blight	0.91	0.64	0.75	200
Tomato Late Blight	0.90	0.92	0.91	382
Tomato Leaf Mold	0.77	0.97	0.86	191
Tomato Septoria Spot	0.87	0.93	0.90	355
Tomato Healthy	0.97	0.99	0.98	319
Accuracy	90.47%			
Macro Avg	0.89	0.90	0.89	1878
Weighted Avg	0.91	0.90	0.90	1878

TABLE II: Overall Performance Metrics

Metric	Value
Accuracy	90.47%
Precision	91.18%
Recall	90.47%
F1 Score	90.29%



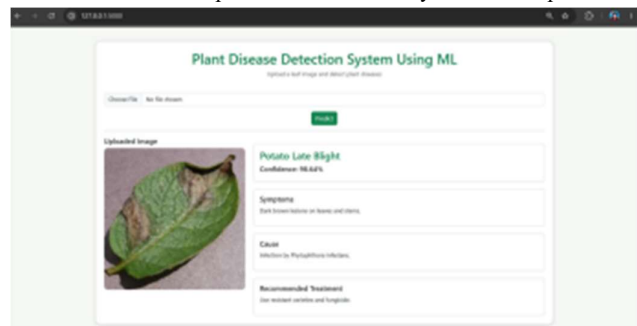
Fig 10 Confusion Matrix for Disease Classification

The confusion matrix provides a detailed representation of classification performance. The overall test accuracy achieved was 90.47%, with precision of 91.18% and F1-score of 90.29%, as summarized in Table I and Table II.

Analysis reveals that most disease categories are accurately classified, with some confusion observed between diseases exhibiting similar symptoms, such as Early Blight and Late Blight. These diseases share similar visual characteristics including dark spots and lesions, leading to occasional misclassification.

D. Prediction Output Analysis

The developed web interface allows users to upload leaf im-ages and receive disease predictions instantly. The prediction



output includes the identified disease category and confidence score

Fig 11 Disease Prediction Result

The output screen displays the uploaded image along with the predicted disease name and confidence score. Users can obtain disease predictions within a short processing time, enabling rapid decision-making and timely intervention.

E. Discussion

The results demonstrate the effectiveness of image-based disease identification techniques in agriculture. The system successfully classified healthy and diseased plant leaves and provided reliable predictions through an automated workflow. The training and validation results indicate that the model learned relevant disease characteristics and achieved stable performance during testing.

The developed web interface enhances usability by providing an accessible platform for disease diagnosis. Users can up-load images and obtain results quickly without requiring expert knowledge or specialized equipment. The average prediction time was approximately 2–3 seconds per image, demonstrating practical applicability.

Although the system demonstrates promising performance, certain limitations remain. Prediction accuracy may be influenced by image quality, lighting conditions, background complexity, and disease severity. Future improvements may focus on expanding the dataset and incorporating additional crop varieties to further enhance system robustness.

Conclusion and Future Scope

A. Conclusion

The Plant Disease Prediction System developed in this project provides an efficient and automated solution for identifying diseases in tomato and potato plants using leaf im-ages. Plant diseases are one of the major factors affecting agricultural productivity and crop quality. Early detection of these diseases is essential for minimizing crop losses and ensuring healthy plant growth. The proposed system addresses this challenge by providing a simple and reliable method for disease identification.

The project utilizes image analysis and intelligent classification techniques to examine visual symptoms present on plant leaves. The system is capable of distinguishing between

healthy and diseased leaves and predicting the corresponding disease category. Through dataset preparation, image preprocessing, model training, and prediction analysis, a complete disease detection workflow was successfully implemented.

The developed web-based interface allows users to upload leaf images and obtain disease predictions in a user-friendly manner, improving accessibility and reducing dependence on manual disease inspection. The obtained results demonstrate that the system can effectively identify plant diseases with 90.47% accuracy.

The project highlights the growing importance of technology in modern agriculture. Automated disease detection systems can assist farmers in monitoring crop health, taking timely preventive actions, and improving overall productivity. Such systems contribute to sustainable farming practices by supporting efficient crop management and reducing unnecessary resource usage.

B. Future Scope

The field of agricultural technology is continuously evolving, creating numerous opportunities for enhancing plant disease prediction systems. Several improvements can be incorporated in the future.

1) *Dataset Expansion:* The system can be enhanced by expanding the dataset to include additional crop varieties and disease categories [4]. A larger and more diverse dataset would improve recognition of a wider range of diseases.

2) *Mobile Application:* The system can be extended to support real-time disease detection through mobile applications. Farmers could capture leaf images directly using smartphone cameras and obtain instant predictions [1].

3) *Cloud Integration:* Integration with cloud-based platforms would allow disease predictions to be performed remotely while maintaining centralized storage of agricultural data [7].

4) *IoT Integration:* Sensors deployed in agricultural fields can monitor environmental parameters such as temperature, humidity, soil moisture, and light intensity. Combining sensor

data with disease prediction systems can provide comprehensive crop health monitoring [3].

5) *Treatment Recommendations:* Future versions may include treatment recommendations based on the identified disease, suggesting preventive measures, pesticides, or agricultural practices to control infection.

6) *Drone Integration:* The system can be integrated with drone-based imaging technologies for large-scale agricultural monitoring, enabling rapid disease detection across extensive farmland areas.

7) *Multilingual Support:* Providing the interface in regional languages would increase usability among farmers from different regions and improve technology adoption.

In conclusion, the future scope extends beyond disease prediction and includes the development of intelligent agricultural decision-support systems. Continued advancements in image analysis, artificial intelligence, cloud computing, and IoT technologies will further enhance the effectiveness of plant disease management solutions.

Acknowledgment

The authors express their sincere gratitude to the Department of Electronics & Telecommunication Engineering, MGM's College of Engineering, Nanded, for providing the necessary facilities and support for this research work.

References

- [1] S. Mohanty, D. Hughes, and M. Salathe, "Using Deep Learning for Image-Based Plant Disease Detection," *Frontiers in Plant Science*, vol. 7, pp. 1–10, 2016.
- [2] D. P. Hughes and M. Salathe, "An Open Access Repository of Images on Plant Health to Enable the Development of Mobile Disease Diagnostics," *PlantVillage Dataset*, 2015. [Online]. Available: <https://plantvillage.psu.edu>
- [3] M. Too, S. Yujian, N. Njuki, and L. Yingchun, "A Comparative Study of Fine-Tuning Deep Learning Models for Plant Disease Identification," *Computers and Electronics in Agriculture*, vol. 161, pp. 272–279, 2019.

- [4] J. G. A. Barbedo, "Impact of Dataset Size and Variety on the Effectiveness of Deep Learning and Transfer Learning for Plant Disease Classification," *Computers and Electronics in Agriculture*, vol. 153, pp. 46–53, 2018.
- [5] K. P. Ferentinos, "Deep Learning Models for Plant Disease Detection and Diagnosis," *Computers and Electronics in Agriculture*, vol. 145, pp. 311–318, 2018.
- [6] K. P. Singh and S. Gupta, "Plant Disease Detection Using Image Processing Techniques: A Review," *International Journal of Computer Applications*, vol. 124, no. 16, pp. 6–10, 2015.
- [7] R. Patel and M. Patel, "A Survey on Plant Disease Detection and Classification Using Machine Learning," *International Journal of Engineering Research and Technology*, vol. 8, no. 5, pp. 102–107, 2019.
- [8] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. Chen, "Mo-bileNetV2: Inverted Residuals and Linear Bottlenecks," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 4510–4520.
- [9] G. Bradski, "The OpenCV Library," *Dr. Dobbs's Journal of Software Tools*, vol. 25, no. 11, pp. 120–126, 2000.
- [10] F. Chollet, *Deep Learning with Python*, 2nd ed. Shelter Island, NY, USA: Manning Publications, 2021.
- [11] A. Geron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 3rd ed. Sebastopol, CA, USA: O'Reilly Media, 2022.
- [12] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed. New York, NY, USA: Pearson Education, 2018.
- [13] TensorFlow Development Team, "TensorFlow: An End-to-End Open Source Machine Learning Platform." [Online]. Available: <https://www.tensorflow.org>
- [14] Keras Development Team, "Keras API Documentation." [Online]. Available: <https://keras.io>