

Plant Disease Detection Using CNN

Asmita Jadhav¹, Prof. Kumud Wasnik²

¹Computer Science and Technology Usha Mittal Institute of Technology, Mumbai, India

²H.O.D., Computer Science and Technology, Usha Mittal Institute of Technology Mumbai, India

Abstract—Plant diseases are a significant menace to world agriculture, which causes significant harvest loss and economic difficulties. Conventional techniques for the identification of diseases are time-consuming and prone to erroneous influence. In this project, a solution is presented using a folding seller of Neuron Networks (CNN) for automatic plant disease detection and classification from leaf photos. Due to its capability of learning and deriving important features from visual information, CNN is very efficient in image-based operations like disease classification. The system proposed utilizes CNN to scan plant leaf images to detect symptoms of disease like color variations, stains, and texture changes. The process starts with image processing and then unique extraction through the CNN layer, completing the disease classification. The system's functionality in image classification using trained models guarantees high accuracy when it comes to identifying a range of plant diseases. The project also presents a friendly interface where users upload images of leaves, are diagnosed with disease, and seek progression of disease procedures via a personalized account. The system under consideration combines extensions like user authentication and data storage to enable users to track the health of the system in the long term. The aim is to enormously enhance plant disease management, minimize human error, and push towards more sustainable farming practices.

Index Terms—plants disease, image classification, machine learning model, convolutional neural network, disease prediction

I. INTRODUCTION

The agricultural industry has encountered serious challenges brought about by the rising incidence of plant pathogens, which can cause huge losses of crops and decreased food security. Traditional methods of disease identification often rely on manual inspection, which can be time consuming and prone to human error. To address these challenges, the integration of technology in agriculture has gained momentum, with machine learning and computer vision emerging as

powerful tools for improving plant health management. A Convolutional Neural Network (CNN) is a type of artificial intelligence (AI) model that helps computers "see" and understand images, just like humans do. It is especially good at recognizing patterns, shapes, and colors in pictures. CNNs are widely used in fields like medical imaging, self-driving cars, and, in this case, plant disease detection.

How CNN Works in Plant Disease Detection:

- Collecting Images – First, a large number of images of plant leaves are collected. These images include both healthy leaves and leaves affected by various diseases.
- Preprocessing the Images – The images are cleaned, resized, and sometimes enhanced to improve clarity. This helps the CNN model learn better.
- Feature Extraction – The CNN automatically identifies important patterns in the images, such as color changes, spots, or texture differences that indicate disease. It does this through multiple layers:
 - Convolution Layer: Detects small patterns like edges and spots. Pooling Layer: Reduces the image size while keeping the important information.
 - Fully Connected Layer: Combines all detected patterns to make a final decision.

Training the Model: The CNN is trained using a dataset of labeled images (e.g., "Healthy," "Blight," "Mildew"). It learns by repeatedly analyzing these images and adjusting itself to recognize different diseases correctly.

Testing and Prediction: Once trained, the CNN can analyze new images and predict whether a plant is healthy or diseased.

II. RELATED WORK

The domain of plant disease detection has seen a growing emphasis on hybrid models, transfer

learning, explainable AI, and feature optimization to improve model robustness, interpretability, and generalizability.

- Raghuram and Borah (2025) [7] developed a *Hybrid Learning Model (HLM)* by integrating *Deep Reinforcement Learning* with *Transfer Learning* using MobileNetV2 as the base feature extractor. This approach enhanced adaptability to diverse environmental conditions and improved classification accuracy in tomato leaf disease detection.

- Kumari et al. (2025) [9] introduced a feature-enhanced, explainable framework that combines *Hough Transform (HT)* and *Discrete Wavelet Transform (DWT)* across multiple color spaces. Achieving *99% classification accuracy*, the model prioritized feature visibility and transparency, aiding agronomists in understanding model decisions.

- Aishwarya and Reddy (2023) [10] proposed an *ensemble CNN architecture* combining *DenseNet169*, *InceptionV3*, and *Xception*. Their non-linear fusion strategy of decision scores led to an accuracy of *98.46%* in groundnut leaf disease classification, demonstrating the effectiveness of model ensembling in handling intra-class variability.

EXISTING SYSTEMS

Recent advancements in deep learning have led to the development of efficient, real-time systems for plant disease detection. These systems integrate various CNN architectures with mobile and edge platforms to enhance accessibility and usability for end-users, particularly farmers.

- Vinay et al. (2025) [6] designed a hybrid deep learning model by combining *ResNet-50* and *InceptionV3*. The model, enhanced with data augmentation techniques, achieved a classification accuracy of *97%* and demonstrated high recall in detecting diseases across various crops, especially in early stages.

- Attri et al. (2025) [8] introduced an *Edge AI-based solution* optimized for low-resource environments using *MobileNetV2*. Deployed on embedded hardware, the model targeted mango and apple leaf disease detection and achieved *93.1% accuracy*, facilitating real-time, offline diagnosis for farmers in rural regions.

- Shewale and Daruwala (2023) [11] presented a CNN-driven approach specifically focused on early detection of tomato plant diseases. Their system

incorporated customized preprocessing techniques and demonstrated significant potential for low-latency classification in agricultural monitoring applications.

III. PROPOSED SYSTEM

A. System Design

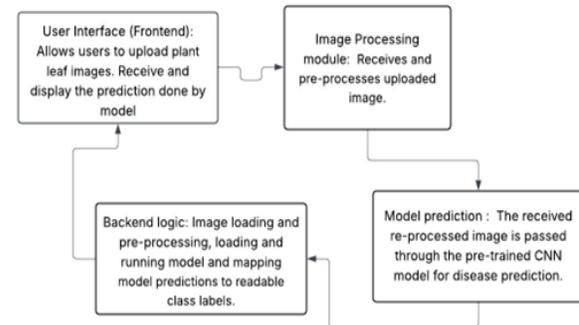


Fig. 1. System design

The proposed system automates plant disease identification using a modular image classification pipeline powered by deep learning. It integrates a Convolutional Neural Network (CNN) and an interactive web interface built with Streamlit. The architecture comprises the following components:

- User Interface: Developed using Streamlit, allowing users to:
 - Upload plant leaf images,
 - Display prediction results and uploaded images,
 - Access a user-friendly interface suitable for non-technical users.
- Image Preprocessing: Utilizes OpenCV and TensorFlow/Keras preprocessing utilities to:
 - Resize images to 128×128 pixels,
 - Normalize pixel values,
 - Convert image color format from BGR to RGB.
- CNN Model: Implemented in TensorFlow/Keras and trained on the Kaggle “New Plant Diseases Dataset” containing over 87,000 images across 38 classes. The model performs multi-class classification using:
 - Multiple convolutional and pooling layers,
 - ReLU activations and Softmax output,
 - Data augmentation techniques such as rotation, flipping, and zooming.
- Backend Logic: Written in Python to:
 - Load the trained model and run inference,
 - Map prediction indices to class names,

- Return results to the interface.
- Data Flow: The workflow follows these steps:
 - 1) User uploads a plant leaf image via Streamlit,
 - 2) Image is preprocessed,
 - 3) CNN model performs inference,
 - 4) Predicted class and confidence score are displayed,
 - 5) (Optional) Results are logged.

B. System Development

Technologies Used:

- Programming Languages and Libraries: Python, TensorFlow, Keras, NumPy, Pandas, OpenCV, PIL, Scikit-learn, Matplotlib, Seaborn.
- Development Tools: Streamlit for frontend development, Jupyter Notebook for prototyping, Git/GitHub for version control.

Implementation Details:

- Image Preprocessing: All images are resized to 128×128 pixels, normalized, and color-corrected. Class labels are encoded.
- Model Architecture: The CNN model consists of convolutional, max-pooling, and dropout layers with batch normalization. It uses ReLU for intermediate layers and Softmax for output.
- Training and Evaluation: The dataset is split into training, validation, and testing sets. Evaluation metrics include accuracy, precision, recall, and confusion matrices.
- Web Application: Streamlit interface supports real-time predictions and displays disease names and probabilities.
- Deployment: The application is packaged with a virtual environment and prepared for deployment on Streamlit.

IV. METHODOLOGY

Overview:

This study uses a Convolutional Neural Network (CNN) to classify plant leaf images as healthy or diseased. The goal is to develop an accurate, scalable model that assists in early detection of plant diseases, improving agricultural outcomes through automation and real-time support.

Research Design:

A supervised deep learning pipeline was followed:

- Data Collection: Used a publicly available dataset

with 87,848 images spanning 38 disease classes.

- Preprocessing: Images were resized, normalized, and augmented to improve model generalization.
- Model Building: Developed a CNN using TensorFlow and Keras with convolutional, pooling, dropout, and dense layers.
- Training and Validation: The model was trained using categorical cross-entropy loss and the Adam optimizer. Validation data helped monitor performance.
- Evaluation: Accuracy, precision, recall, F1-score, and confusion matrix were used to assess performance.

Sampling Strategy:

A stratified sampling approach ensured balanced class representation:

- Training Set: 70% (70,295 images)
- Validation Set: 15% (17,572 images)
- Testing Set: 15% (13,177 images)

Tools and Libraries:

The model was implemented in Python using the following tools:

- TensorFlow and Keras – Model design and training
- OpenCV – Image processing
- Matplotlib, Seaborn – Data visualization
- NumPy, Pandas, Scikit-learn – Data manipulation and performance metrics

Ethical Considerations:

The dataset is open-access and contains no personal or sensitive information. All images were used in accordance with the dataset's license, ensuring compliance with AI research ethics.

Limitations and Assumptions:

- Limitations: Potential dataset bias, variable image quality, and limited generalizability to real-field environments.
- Assumptions: Images are correctly labeled and augmented data accurately represents real-world variance.

System Requirements:

Hardware:

- Recommended: Intel i5 or higher, 4 GB+ RAM, SSD (128 GB+), Full HD monitor
- Minimum: Intel i3, 2 GB RAM, HDD (80 GB),

SVGA monitor

Software:

- Operating System: Windows 10 (64-bit), macOS, or Linux
- Development Tools: Python 3.x, Jupyter Notebook, Py- Charm or VS Code
- Libraries: TensorFlow, Keras, OpenCV, Scikit-learn, Mat-plotlib, Pandas

V. RESULTS AND DISCUSSION

In the proposed system, a Convolutional Neural Network (CNN) model was developed to classify plant diseases from leaf images, achieving impressive performance. During the training phase, the CNN model was trained over 10 epochs. The training began with a modest accuracy of around 39.7%, which quickly improved as the model learned relevant features from the image data. By the end of training, the model achieved a training accuracy of 98.03% and a validation accuracy of 94.55%, along with a significant reduction in loss values—indicating that the model had effectively learned the distinguishing features of each class. The relatively small gap between training and validation accuracy also suggests that the model did not overfit the training data and is capable of generalizing well to unseen images.

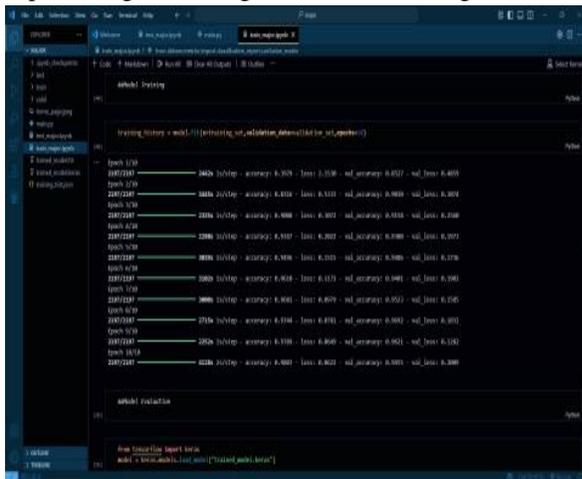


Fig. 2. Model Training for 10 epochs

The steadily decreasing loss values across epochs further confirmed that the model was learning effectively.

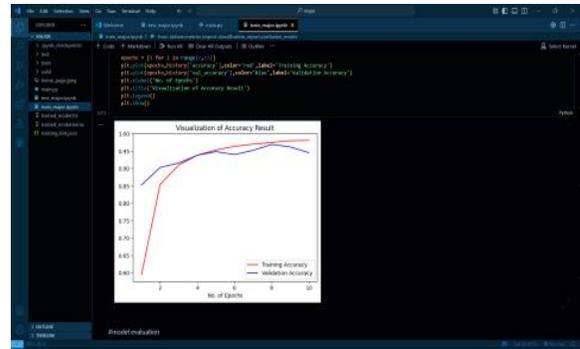


Fig. 3. Model accuracy graph

Post-training, the model was evaluated by making predictions on test samples. The output class index from the model was mapped back to human-readable class names using a predefined list derived from the dataset’s folder structure. In one such prediction, using the class labels derived from the directory structure, it correctly identified a test image as "Tomato → Tomato Yellow Leaf Curl Virus", demonstrating its ability to accurately map learned features to real-world plant disease categories.

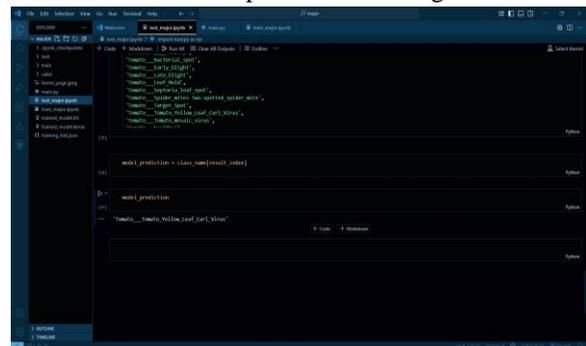


Fig. 4. Model testing

Overall, the project demonstrates the effectiveness of CNNs in plant disease detection, offering a scalable and efficient tool for modern agriculture. With high accuracy and reliable performance, this model could contribute significantly to early disease diagnosis, reduced pesticide usage, and improved crop health management.

CONCLUSION

In this research, a deep learning-based system was developed for automated plant disease detection using Convolutional Neural Networks (CNNs). By training on a large, diverse dataset of leaf images, the model demonstrated high accuracy in identifying various plant diseases, enabling early diagnosis and effective

crop management. The approach reduces reliance on manual inspection, offering a scalable and real-time solution that can be integrated into mobile or edge devices for use directly in agricultural fields. This work contributes to sustainable farming by promoting timely intervention, reducing yield loss, and supporting farmers with accessible, technology-driven plant health monitoring tools.

ACKNOWLEDGMENT

I have great pleasure in expressing my gratitude to all those who have contributed and motivated me during the project work. I am extremely thankful and pay my gratitude to my guide Prof. Kumud Wasnik for her valuable guidance and support towards me which encouraged me in the completion of this project.

REFERENCE

- [1] Neha Sharma, Vibhor Jain and Anju Mishra, "An Analysis of Convolutional Neural Networks for Image Classification", (ICCIDS 2018)
- [2] Mingyuan Xin and Yong Wang, "Research on image classification model based on deep convolution neural network", (EURASIP Journal on Image and Video Processing 2019)
- [3] Muhammad Shoaib1, Babar Shah, Shaker El-Sappagh, Akhtar Ali, Asad Ullah, Fayadh Alenezi, Tsanko Gechev, Tariq Hussain and Farman Ali, "An advanced deep learning models-based plant disease detection: A review of recent research", (Frontiers in Plant Science 2023)
- [4] Pranesh Kulkarni, Atharva Karwande, Tejas Kolhe, Soham Kamble, Akshay Joshi and Medha Wyawahare, "Plant Disease Detection Using Image Processing and Machine Learning",
- [5] Kazi Naimur Rahman, Sajal Chandra Banik, Raihan Islam, and Arafath Al Fahim, "A real time monitoring system for accurate plant leaves disease detection using deep learning", *Crop Design*, vol. 4, no. 1, pp. 100092, 2025. doi: 10.1016/j.crope.2024.100092.
- [6] Vinay K, Vempalli Surya, Thushar S, Tripty Singh, and Apurvanand Sahay, "A Deep Learning Framework for Early Detection and Diagnosis of Plant Diseases", *Procedia Computer Science*, vol. 258, pp. 1435–1445, 2025. International Conference on Machine Learning and Data Engineering. doi: 10.1016/j.procs.2025.04.376.
- [7] Kadambari Raghuram and Malaya Dutta Borah, "A Hybrid Learning Model for Tomato Plant Disease Detection using Deep Reinforcement Learning with Transfer Learning", *Procedia Computer Science*, vol. 252, pp. 341–354, 2025. 4th International Conference on Evolutionary Computing and Mobile Sustainable Networks. doi: 10.1016/j.procs.2024.12.036.
- [8] Ishana Attri, Lalit Kumar Awasthi, and T.P. Sharma, "TinyML for Plant Disease Detection: Efficient Edge AI Solutions for Apple and Mango Leaves", *Procedia Computer Science*, vol. 258, pp. 2870–2877, 2025. International Conference on Machine Learning and Data Engineering. doi: 10.1016/j.procs.2025.04.547.
- [9] Saumya Kumari, Resham Jha, Aarushi Ray, Junali Jasmine Jena, Mahendra Kumar Gourisaria, and Anjan Bandyopadhyay, "An Explainable AI Approach for Potato Plant Disease Detection using Enhanced Feature Engineering", *Procedia Computer Science*, vol. 258, pp. 3570–3579, 2025. International Conference on Machine Learning and Data Engineering. doi: 10.1016/j.procs.2025.04.612.
- [10] Aishwarya M.P. and Padmanabha Reddy, "Ensemble of CNN models for classification of groundnut plant leaf disease detection", *Smart Agricultural Technology*, vol. 6, p. 100362, 2023. doi: 10.1016/j.atech.2023.100362.
- [11] Mitali V. Shewale and Rohin D. Daruwala, "High performance deep learning architecture for early detection and classification of plant leaf disease", *Journal of Agriculture and Food Research*, vol. 14, p. 100675, 2023. doi: 10.1016/j.jafr.2023.100675.
- [12] Shah Faisal, Kashif Javed, Sara Ali, Areej Alasiry, Mehrez Marzougui, Muhammad Attique Khan and Jae-Hyuk Cha, "Deep Transfer Learning Based Detection and Classification of Citrus Plant Diseases", *Computers, Materials and Continua*, vol. 76, no. 1, pp. 895–914, 2023. doi: 10.32604/cmc.2023.039781.
- [13] Parul Sharma, Yash Paul Singh Berwal and Wiqas Ghai, "Performance analysis of deep learning CNN models for disease detection in plants using image segmentation", *Information Processing in Agriculture*, vol. 7, no. 4, pp. 566–574, 2020. doi: 10.1016/j.inpa.2019.11.001.

- [14] S. Babu, M. Maravarman, and R. Pitchai, "Detection of Rice Plant Disease Using Deep Learning Techniques," *Journal of Mobile Multimedia*, vol. 18, no. 3, pp. 757-770, 2022. doi: 10.13052/jmm1550-4646.18314.
- [15] A. Shrotriya, A. K. Sharma, A. K. Bairwa, and R. Manoj, "Hybrid Ensemble Learning With CNN and RNN for Multimodal Cotton Plant Disease Detection," *IEEE Access*, vol. 12, pp. 198028-198045, 2024. doi: 10.1109/ACCESS.2024.3515843.
- [16] C. K. Sunil, C. D. Jaidhar, and N. Patil, "Cardamom Plant Disease Detection Approach Using EfficientNetV2," *IEEE Access*, vol. 10, pp. 789-804, 2022. doi: 10.1109/ACCESS.2021.3138920.
- [17] R. Rashid, W. Aslam, R. Aziz, and G. Aldehim, "An Early and Smart Detection of Corn Plant Leaf Diseases Using IoT and Deep Learning Multi-Models," *IEEE Access*, vol. 12, pp. 23149-23162, 2024. doi: 10.1109/ACCESS.2024.3357099.
- [18] D. Luo, Y. Xue, X. Deng, B. Yang, H. Chen, and Z. Mo, "Citrus Diseases and Pests Detection Model Based on Self-Attention YOLOV8," *IEEE Access*, vol. 11, pp. 139872-139881, 2023. doi: 10.1109/ACCESS.2023.3340148.
- [19] V. K. Vishnoi, K. Kumar, B. Kumar, S. Mohan, and A. A. Khan, "Detection of Apple Plant Diseases Using Leaf Images Through Convolutional Neural Network," *IEEE Access*, vol. 11, pp. 6594-6609, 2023. doi: 10.1109/ACCESS.2022.3232917.
- [20] H. Islam, M. Nahiduzzaman, M. A. H. Pramanik, M. K. Syfullah, S. M. Shahriar, A. Sultana, M. Ahsan, J. Haider, A. Khandakar, and M. E. H. Chowdhury, "Plant Disease Classifier: Detection of Dual-Crop Diseases Using Lightweight 2D CNN Architecture," *IEEE Access*, vol. 11, pp. 110627-110643, 2023. doi: 10.1109/ACCESS.2023.3320686.
- [21] X. Zhang, Y. Mao, Q. Yang, and X. Zhang, "A Plant Leaf Disease Image Classification Method Integrating Capsule Network and Residual Network," *IEEE Access*, vol. 12, pp. 44573-44585, 2024. doi: 10.1109/ACCESS.2024.3377230.
- [22] S. Chen, Y. Liao, F. Lin, and B. Huang, "An Improved Lightweight YOLOv5 Algorithm for Detecting Strawberry Diseases," *IEEE Access*, vol. 11, pp. 54080-54092, 2023. doi: 10.1109/ACCESS.2023.3282309.
- [23] Q. Zeng, X. Ma, B. Cheng, E. Zhou, and W. Pang, "GANs-Based Data Augmentation for Citrus Disease Severity Detection Using Deep Learning," *IEEE Access*, vol. 8, pp. 172882-172891, 2020. doi: 10.1109/ACCESS.2020.3025196.