IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 12, December 2025

DOI: 10.17148/IJARCCE.2025.141224

Automated Plant Disease Detection Using Convolutional Neural Networks

Karthik S G[1], Keerthan Gowda K[2], Prakyath S[3], Himanth M[4], Prof. Malashree M S[5]

Student, Department of Computer Science and Engineering, Maharaja Institute of Technology Mysore, Belawadi Mandya, Karnataka, India[1,2,3,4]

Professor, Department of Computer Science and Engineering, Maharaja Institute of Technology Mysore, Belawadi Mandya, Karnataka, India[5]

Abstract: This project presents an automated plant disease detection system developed using Convolutional Neural Networks (CNNs) to support early and accurate diagnosis of crop diseases. Plant diseases significantly impact global agricultural productivity, and traditional manual inspection methods are often slow, inconsistent, and dependent on expert knowledge. To address these challenges, the proposed system leverages deep learning to classify diseases from leaf images with improved precision and reliability. A large dataset consisting of over 87,000 healthy and diseased leaf images across 38 classes was preprocessed and used to train a custom CNN model. The model effectively extracts spatial features from input images and achieves high performance, recording approximately 99% training accuracy and 97% validation accuracy. The solution is deployed as an interactive web application built with Streamlit, enabling users—particularly farmers and agronomists—to upload leaf images and receive real-time disease predictions. By offering a fast, affordable, and scalable diagnostic tool, this work contributes to smarter agricultural practices, timely disease management, reduced dependency on expert intervention, and overall enhancement of crop health monitoring. The study also highlights the potential of CNN-based systems to transform traditional plant disease diagnosis through efficient, user-friendly, and technology-driven approaches.

Keywords: Plant Disease Detection, Convolutional Neural Networks (CNNs), Deep Learning, Image Processing, Machine Learning, Feature Extraction, Automated Diagnosis, Agriculture Technology, Leaf Image Classification, Training and Validation, Dataset Preparation, Image Preprocessing, Transfer Learning, Streamlit Web Application, TensorFlow/Keras, Real-Time Prediction, Mobile/Field Deployment, Accuracy and Performance Metrics, Sustainable Agriculture, Precision Agriculture, Early Disease Detection, Computer Vision, Data Augmentation, Plant Health Monitoring, Model Evaluation, Classification Models, Disease Recognition System, Web-Based Interface, Model Optimization, Field Images, PlantVillage Dataset, Hyperspectral Imaging, Few-Shot Learning (FSL), Generative Adversarial Networks (GANs), Image Segmentation, Object Detection, Decision Support Systems (DSS).

I. INTRODUCTION

Agriculture is a vital sector that supports global food production, yet crop losses caused by plant diseases continue to pose a major challenge to farmers and the agricultural economy. Early and accurate disease detection plays a crucial role in preventing large-scale damage, but traditional methods—such as manual inspection or expert diagnosis—are often slow, inconsistent, and inaccessible to farmers in remote regions. With the rapid advancement of artificial intelligence, deep learning, and image processing technologies, automated solutions have emerged as reliable tools for plant health assessment.

Convolutional Neural Networks (CNNs) have proven to be highly effective for visual recognition tasks due to their ability to learn complex patterns directly from images. By using CNNs to analyse leaf photographs, plant diseases can be identified with greater speed, precision, and consistency when compared to conventional techniques. This technology helps bridge the gap between expert knowledge and field-level decision-making, offering farmers accurate and timely insights into crop conditions.

The aim of this project is to develop an automated plant disease detection system that leverages CNNs to classify healthy and diseased leaves. The system uses a large image dataset to train a deep learning model capable of distinguishing between multiple disease categories. The proposed approach includes image preprocessing, model training, performance evaluation, and deployment through a user-friendly web interface, allowing real-time disease prediction from uploaded leaf images. By integrating AI-driven diagnosis with accessible digital tools, this project supports sustainable agriculture, reduces dependency on expert intervention, minimizes crop losses, and promotes smarter resource management.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 12, December 2025
DOI: 10.17148/IJARCCE.2025.141224

II. LITERATURE REVIEW

Research in automated plant disease detection has progressed rapidly with the adoption of machine learning and deep learning techniques, especially Convolutional Neural Networks (CNNs). Recent studies highlight the transition from traditional feature-engineering methods to more advanced, data-driven approaches capable of performing effectively in real agricultural environments.

Moupojou et al. introduce the *FieldPlant* dataset, addressing a major limitation in plant disease research—the scarcity of real-world field images. Their dataset contains over 5,170 images captured under natural farm conditions, including complex backgrounds, lighting variations, and occlusions. The study demonstrates that deep learning models such as MobileNet, VGG16, InceptionV3, and InceptionResNetV2 achieve significantly better generalization when trained on realistic field datasets compared to controlled laboratory datasets. This work highlights the importance of diverse, well-annotated datasets for practical disease detection systems **[1]**.

In contrast, Maniyath et al. explore classical machine learning approaches for plant disease classification. Their system relies on feature extraction techniques such as Histogram of Oriented Gradients (HOG), Hu Moments, Haralick textures, and Color Histograms to classify papaya leaf diseases. Despite using a relatively small dataset of only 160 images, the Random Forest classifier outperforms other models like SVM, KNN, and Logistic Regression. The study shows that traditional machine learning methods can still be useful in low-resource agricultural settings, though their accuracy and scalability remain limited **[2]**.

Li, Zhang, and Wang present a comprehensive review of deep learning methods for plant disease detection using leaf images. They highlight the strengths of CNN architectures such as AlexNet, ResNet, VGG16, GoogLeNet, and Inception, which automatically learn visual features from images without manual feature engineering. The review also emphasizes the importance of transfer learning, data augmentation, and interpretability techniques such as saliency maps. Furthermore, emerging approaches like Few-Shot Learning (FSL) and Generative Adversarial Networks (GANs) are identified as promising solutions to dataset scarcity and early-stage disease recognition **[3]**.

Kumar examines the broader role of artificial intelligence in agriculture, particularly in large-scale monitoring and precision farming. The study outlines how AI technologies—including machine learning, deep learning, remote sensing, and drone-based imaging—are being integrated to improve early detection, surveillance, and decision support. Despite these advancements, challenges remain, such as limited access to AI tools in rural regions, the black-box nature of deep learning models, and the lack of datasets for rare diseases. The paper calls for more interpretable, lightweight, and accessible AI solutions for agricultural communities **[4]**.

Bhargava et al. provide an extensive overview of computer vision and artificial intelligence applications across the entire plant disease detection pipeline. Their review discusses traditional machine learning, deep learning, hyperspectral imaging, molecular diagnostic techniques (such as ELISA, PCR, LAMP), and emerging few-shot learning approaches. The study highlights ongoing challenges such as real-time image variability, difficulty detecting early-stage infections, and the need for scalable models that can adapt to different crops and disease symptoms. The authors argue for more integrated systems combining deep learning with advanced imaging and mobile deployment for practical field use **[5]**.

Overall, the literature shows a clear progression from handcrafted feature-based models to advanced deep learning systems that can support real-time, accurate, and scalable plant disease detection. These developments underline the importance of dataset quality, model robustness, interpretability, and accessible deployment technologies for real-world agricultural applications.

III. SYSTEM ARCHITECTURE & WORKFLOW

3.1 System Architecture

The proposed system is designed as a web-based plant disease detection platform that integrates a deep learning model with an intuitive user interface. At a high level, the architecture consists of four main components: the dataset storage, the model training environment, the trained CNN model, and the Streamlit-based web application.

The dataset component holds more than 87,000 RGB images of healthy and diseased plant leaves belonging to 38 different classes. These images are organized into training, validation, and test sets and act as the primary input for model development. The training environment, built using Python and TensorFlow/Keras, is responsible for preprocessing the images, defining the CNN architecture, and learning discriminative features for disease classification. Once training is

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 12, December 2025

DOI: 10.17148/IJARCCE.2025.141224

completed, the best-performing model is stored in a serialized format (e.g., trained_plant_disease_model.keras) for later use during inference.
On the deployment side, the Streamlit web application serves as the frontend interface. It provides different pages such as “Home,” “About,” and “Disease Recognition” to guide the user.
The backend of the web app loads the saved CNN model and exposes a prediction function that accepts user-uploaded leaf images. This setup allows real-time communication between the web interface and the deep learning model, enabling users like farmers and agronomists to obtain disease predictions directly from a browser without interacting with the underlying code.

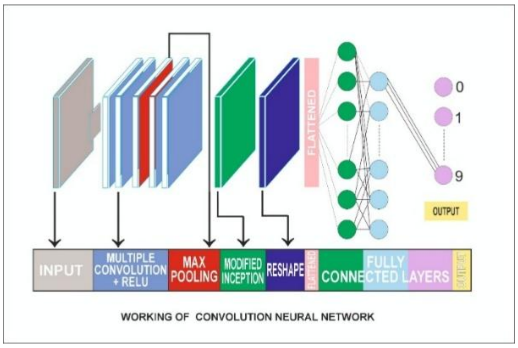

Figure 1: Working of CNN

3.2 Workflow

The overall workflow of the system can be divided into two phases: **model development** and **runtime prediction**.

3.2.1 Model Development Phase

1. **Dataset Preparation**
   The process begins with collecting and organizing the leaf image dataset. Images are grouped into 38 classes representing various plant species and disease types, including healthy leaves. The dataset is then split into training and validation subsets, while a small set of unseen test images is reserved for final evaluation.
2. **Image Preprocessing**
   All images are resized to a fixed resolution (128×128 pixels) to ensure a uniform input size for the CNN. Pixel values are normalized to a [0, 1] range to stabilize and speed up training. Data augmentation techniques (such as rotation or flipping) may be applied offline to increase variability and improve the model’s ability to generalize.
3. **CNN Model Design and Training**
   A custom CNN architecture is constructed using TensorFlow and Keras. The network typically includes convolutional layers for feature extraction, pooling layers for dimensionality reduction, and dense layers for final classification into 38 disease categories. The model is trained using categorical cross-entropy loss and an optimizer such as Adam. During training, performance metrics like accuracy and loss are monitored on both training and validation sets, and the model parameters are updated iteratively over multiple epochs. The final model achieves around 99% training accuracy and 97% validation accuracy, indicating strong generalization on the given dataset.
4. **Model Saving**
   After satisfactory performance is obtained on the given dataset, the trained CNN is saved in a file (trained_plant_disease_model.keras). This file encapsulates the network architecture and learned weights, allowing the model to be reloaded later without retraining.

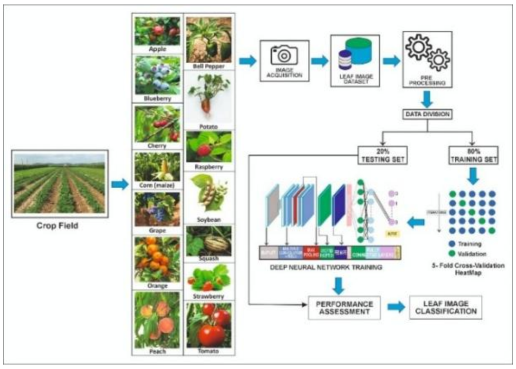

Figure 2: Workflow Diagram

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 12, December 2025
DOI: 10.17148/IJARCCE.2025.141224

3.2.2 Runtime Prediction Phase

1. **User Image Upload**
   In the deployed Streamlit application, the user navigates to the “Disease Recognition” page and uploads a plant leaf image from a local device such as a smartphone or computer.
2. **Preprocessing for Inference**
   The uploaded image is read by the backend, resized to 128×128 pixels, converted to an array, and normalized in the same way as during training. This ensures compatibility with the CNN’s expected input format.
3. **Model Inference**
   The preprocessed image is passed to the loaded CNN model. The network computes a probability distribution over the 38 disease classes using a softmax and activation RELU output layer. The class with the highest probability is selected as the predicted disease label.
4. **Result Display**
   The predicted class name is mapped from its index to a human-readable disease label. (e.g., “Tomato Late Blight,” “Apple Scab,” etc.) and displayed to the user along with the uploaded image. This immediate feedback allows users to quickly understand the likely disease affecting the plant and take appropriate action.

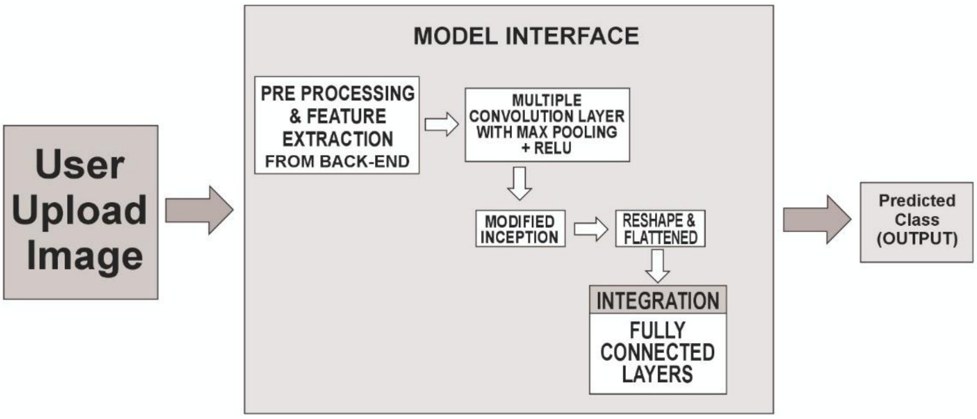

Figure 3: Class Prediction Diagram

IV. RESULTS AND DISCUSSION

Results

The system was trained using a large dataset of more than 87,000 leaf images categorized into 38 classes, representing both healthy and diseased conditions. After preprocessing and model training, the Convolutional Neural Network (CNN) achieved strong performance, recording approximately **99% training accuracy** and **97% validation accuracy**, indicating that the model was successful in learning discriminative features while maintaining good generalization to unseen data. These performance metrics were monitored throughout several epochs, with consistent reductions in training and validation loss, confirming stable learning behavior. The final trained model was saved and deployed through a Streamlit web interface, where real-time testing using unseen images further demonstrated the model’s ability to correctly classify plant diseases based on visual characteristics.

During deployment, the model was evaluated using 33 test images that were not part of training or validation. The system processed each uploaded leaf image, resized it to 128×128 pixels, normalized pixel values, and generated predictions through the softmax classification layer. The model successfully identified major diseases such as tomato blight, apple scab, corn rust, and potato late blight, demonstrating practical functionality in real-world scenarios. The prediction output was displayed instantly along with the uploaded image, making the system efficient for real-time diagnosis in field conditions.

Discussion

The strong accuracy values indicate that the CNN architecture used in this project is well-suited for plant disease classification. The model benefits from multiple convolutional layers that capture essential spatial patterns related to color changes, lesion shapes, and texture variations typical of infected leaves. The high validation accuracy suggests that the model is not only learning these features but also generalizing effectively to new data. However, the slight gap between training and validation accuracy indicates that while the model is robust, performance could improve further with additional augmentation or a deeper network design to reduce mild overfitting. One of the strengths of the proposed system is its integration with a user-friendly Streamlit interface.

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 12, December 2025
DOI: 10.17148/IJARCCE.2025.141224

This interface transforms a complex deep learning model into a tool accessible to farmers, students, and agricultural professionals. Users can upload images directly from their devices and receive immediate results, eliminating the need for expert intervention or lab-based diagnostics. This demonstrates the practical value of combining AI-powered models with lightweight, interactive web technologies.

However, certain challenges remain. The performance of the system depends heavily on the diversity of the dataset. Although the dataset used in this project is large, it mainly includes images with clear visibility of the leaf and relatively consistent lighting conditions. Real-world field environments often involve factors such as background clutter, shadows, overlapping leaves, and varying image quality. If the model encounters conditions significantly different from those in the training dataset, prediction accuracy may decrease. This highlights the need for additional field-based training data and more advanced augmentation strategies to improve real-world performance.

Overall, the results demonstrate that the proposed CNN-based plant disease detection system is reliable, efficient, and practical for real-time agricultural disease monitoring. With further refinement—such as incorporating field images, implementing transfer learning with deeper architectures, and enhancing interpretability—the system has the potential to become a powerful decision-support tool in sustainable agriculture.

V. CONCLUSION

This work demonstrates the effectiveness of Convolutional Neural Networks (CNNs) for automated plant disease detection using leaf images. The proposed system, trained on a large dataset encompassing multiple crop species and disease categories, achieved high accuracy and proved capable of reliably distinguishing between healthy and infected leaves. By integrating the trained model into a Streamlit-based web application, the system offers an accessible and user-friendly platform that enables real-time disease diagnosis without requiring expert knowledge. While performance may vary with field conditions such as lighting and background complexity, the results indicate strong potential for applying deep learning to support timely decision-making and reduce crop losses. Overall, the system provides a practical and scalable solution that contributes to advancing precision agriculture through efficient, technology-driven plant health monitoring.

REFERENCES

[1] E. Moupojou *et al.*, "FieldPlant: A Dataset of Field Plant Images for Plant Disease Detection and Classification with Deep Learning," *IEEE Access*, 2023.
[2] R. Maniyath, S. P. V. Vinod, M. Niveditha, R. Pooja, N. Prasad, N. Shashank, and R. Hebbar, "Plant Disease Detection Using Machine Learning," 2018.
[3] L. Li, S. Zhang, and B. Wang, "Plant Disease Detection and Classification by Deep Learning – A Review," *IEEE Access*, 2021.
[4] S. Kumar, "Transforming Agriculture through Artificial Intelligence: Advancements in Plant Disease Detection, Applications, and Challenges," 2024.
[5] A. Bhargava, A. Shukla, O. P. Goswami, M. H. Alsharif, P. Uthansakul, and M. Uthansakul, "Plant Leaf Disease Detection, Classification, and Diagnosis Using Computer Vision and Artificial Intelligence: A Review," *IEEE Access*, 2024.
[6] Makerere AI Lab & NaCRRI. (2020). *iBean Dataset (GitHub Repository)*.
[7] M. Sharif et al., "Detection and classification of citrus diseases…," *Computers and Electronics in Agriculture*, vol. 150, pp. 220–234, 2018.
[8] P. K. Sethy, "Rice leaf disease image samples," *Mendeley Data*, 2020.
[9] D. O. Oyewola et al., "Detecting cassava mosaic disease using deep residual CNN," *PeerJ Computer Science*, vol. 7, p. e352, 2021
[10] AI Challenger (2018). *AI Challenger 2018 Dataset*.
[11] L. C. Ngugi et al., "Tomato leaf segmentation algorithms…," *Computers and Electronics in Agriculture*, vol. 178, 2020.
[12] M. Ahmad et al., "Plant disease detection in imbalanced datasets…," *IEEE Access*, vol. 9, pp. 140565–140580, 2021
[13] D. Wang et al., "DHBP: A dual-stream hierarchical bilinear pooling model…," *Computers and Electronics in Agriculture*, vol. 195, 2022.
[14] L. Li et al., "Plant disease detection and classification by deep learning—A review," *IEEE Access*, vol. 9, pp. 56683–56698, 2021.

[15] M. Nagaraju and P. Chawla, “Systematic review of deep learning techniques…,” *Int. J. Syst. Assurance Engineering & Management*, vol. 11, no. 3, 2020.
[16] M. Sandler et al., “MobileNetV2: Inverted residuals and linear bottlenecks,” *CVPR*, 2018, pp. 4510–4520.
[17] E. Elfatimi et al., “Beans leaf diseases classification using MobileNet models,” IEEE Access, vol. 10, pp. 9471–9482, 2022.
[18] K. Kiratiratanapruk et al., “Using deep learning…detect rice diseases,” *Int. Conf. Industrial Engineering and Applications*, 2020.
[19] A. Khattak et al., “Automatic detection of citrus leaf diseases…,” *IEEE Access*, vol. 9, 2021.
[20] H. Amin et al., “End-to-end deep learning model for corn leaf disease classification,” *IEEE Access*, vol. 10, 2022.
[21] S. M. Hassan and A. K. Maji, “Plant disease identification using a novel CNN,” *IEEE Access*, vol. 10, 2022.
[22] F. Alvaro et al., “A robust deep-learning-based detector for real-time tomato plant diseases,” *Sensors*, vol. 17, 2017.
[23] *List of Countries by Tomato Production – Wikipedia*. Accessed 2022
[24] Y. Ma et al., “Multiple object detection and tracking…,” *Int. J. Pattern Recognition and Artificial Intelligence*, vol. 31, 2017