International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 14, Issue 12, December 2025)

Plant Disease Prediction (2022–2025)

Brahmbhatt Jaimish[1], Ruchika Dungarani[2]

[1]PG Student, Artificial Intelligence and Machine Learning
[2]Head of Department, Artificial Intelligence and Machine Learning

***Abstract*— Plant diseases present a major threat to worldwide agriculture, causing significant losses in crop yield and quality, alongside negative economic consequences. Accurate and prompt disease identification is crucial for successful crop management. However, traditional diagnosis depends on manual expert assessment, which is often subjective, time consuming, and difficult to implement across large areas. The development of Machine Learning (ML) and Deep Learning (DL) technologies has paved the way for automated plant disease prediction systems, offering rapid, dependable, and scalable analysis. Early approaches utilized classical ML methods like Support Vector Machines (SVM), Random Forests (RF), and k-Nearest Neighbours (k-NN). These methods typically depended on manually engineered features, such as descriptors of texture, color histograms, and structural characteristics. More recently, the rise of Deep Learning (DL) has established Convolutional Neural Networks (CNNs) as the leading technique. CNNs excel because they can automatically learn complex, hierarchical feature representations directly from raw image data. State of-the-art results on common datasets, such as Plant Village, have been achieved using prominent CNN architectures—including VGG, ResNet, Inception, and EfficientNet—often leveraging transfer learning. Current research is focused on enhancing the stability of these systems under diverse real world conditions by incorporating advanced techniques like attention mechanisms, image segmentation models, and ensemble learning.**

***Keywords*— Plant Disease Prediction, Deep Learning, Convolutional Neural Networks (CNN), IoT Based Crop Monitoring, Hyperspectral Imaging, Environmental Feature Integration**

I. Introduction

Plant diseases represent a significant hurdle to global agriculture, leading to major reductions in yield, poorer crop quality, and negative economic effects. Timely and precise identification of diseases is crucial for effective crop management. Traditionally, diagnostic methods rely on manual visual assessment by specialists, which is often inherently subjective, requires intensive labor, and is unsuitable for use across large-scale operations. The evolution of Machine development of automated prediction systems for plant diseases, capable of providing analysis that is rapid, reliable, and scalable.

Classical ML approaches, such as Support Vector Machines (SVM), Random Forests (RF), and k-Nearest Neighbors (k-NN), were widely explored. These typically depended on handcrafted features like color histograms, texture descriptors, and morphological properties. With the rise of DL, Convolutional Neural Networks (CNNs) have emerged as the dominant method because they can automatically learn hierarchical feature representations directly from raw images. Architectures like VGG, ResNet, Inception, and EfficientNet, frequently combined with transfer learning, have achieved state-of-the-art results on benchmark datasets like Plant Village.

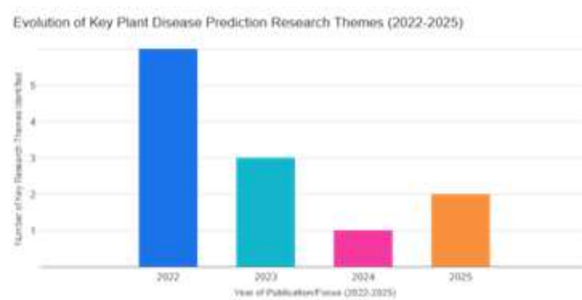

Fig. 1: Evolution Timeline (2022–2025)

Current research is advancing further by integrating attention mechanisms, image segmentation models, and ensemble learning to enhance model robustness against real-world variability. In recent years, modern deep learning frameworks and Large Language Models have transformed pattern analysis and image recognition, allowing for accurate classification of complex visual data in various fields. However, traditional agricultural ML models often operate passively, basing predictions only on data provided at a single point in time, such as an input image. These models do not automatically incorporate environmental factors, track changes over time, or adapt to evolving disease symptoms without repeated human input.

II. Literature Survey

Contemporary Early research in plant pathology utilizing computational methods heavily relied on classical machine learning (ML) techniques.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 14, Issue 12, December 2025)

These pioneering works required a multi-step process: digital image capture, followed by manual feature engineering to extract relevant visual information (such as color, texture, and shape descriptors) from the infected leaf areas. These hand-designed features were then fed into traditional classifiers like Support Vector Machines (SVM) or k-Nearest Neighbors (k-NN). For instance, initial studies demonstrated the feasibility of using color and texture analysis combined with SVM for rice disease classification, or neural networks coupled with color-based segmentation for citrus disease identification. While proving the concept of digital diagnostics, these approaches were often limited by their sensitivity to environmental noise (e.g., lighting variations and complex field backgrounds) and the significant effort required for feature selection.

TABLE I.
LITERATURE SURVEY

Year	Method / Model	Advantages	Limitations
2015 – 2016	CNN-based Classification	High accuracy compared to traditional ML	Limited to lab-quality images; poor field generalization
2018 – 2020	Transfer Learning (VGG, ResNet, Inception)	Less data required, high performance	Dependent on pretrained dataset domain
2021 – 2023	MobileNet / Lightweight CNNs	Real-time prediction, low computation	Slight drop in accuracy vs. larger models
2023 – 2025	YOLO-based Object Detection	Works in field environments; supports bounding boxes	More computational cost than pure classifiers
2024 – 2025	Attention-based Networks (CBAM, ViT)	Better feature extraction, improved accuracy	Research still evolving, training complexity

The landscape changed dramatically with the rise of Deep Learning (DL). Convolutional Neural Networks (CNNs) emerged as the dominant paradigm due to their inherent ability to autonomously learn hierarchical features directly from raw image pixels, eliminating the need for laborious manual feature extraction. A pivotal moment was the successful deployment of deep CNNs on the Plant Village dataset, which established a common benchmark and led to models achieving high accuracy across numerous crop species. Subsequent work leveraged transfer learning, demonstrating that fine-tuning pretrained architectures (like VGG16, ResNet50, and InceptionV3) significantly boosted performance and reduced training time, especially when dealing with limited agricultural datasets. Current Frontiers (2022–2025):

Recent literature, particularly from 2022 onwards, highlights a concerted effort to move AI models out of controlled labsettings and into practical, real-world farm environments by addressing issues of robustness, efficiency, and predictive capability:

Advanced Architectures: The focus has expanded beyond standard CNNs to include lightweight models like EfficientNet (for speed and reduced computational load) and the deployment of Vision Transformers (ViT), which offer powerful global feature extraction capabilities for complex disease patterns.

Real-World Integration: Research is heavily invested in making models deployable and actionable:

Edge Computing and Mobile Deployment: Developing optimized, lightweight models capable of running directly on smartphones and IoT devices to provide offline, realtime disease identification for farmers.

Multimodal Fusion: Integrating visual data with non-image data collected by IoT sensors (e.g., weather conditions, humidity, soil moisture) to create more robust disease prediction models.

Forecasting and Prediction: Utilizing techniques like Spatiotemporal Models (LSTM) to analyze historical and environmental data, enabling the forecasting of disease outbreaks before symptoms are visually apparent on the crop, thus enabling proactive management.

Severity Assessment: Advanced methods, such as UNetbased image segmentation, are used to accurately quantify the severity of lesions on leaves, moving beyond simple classification to provide actionable insights into the extent of the damage.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 14, Issue 12, December 2025)

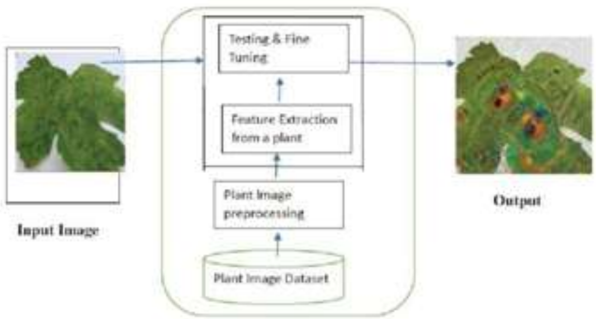

Fig. 2: Plant Disease Prediction Architecture

III. Methodology

The methodology for plant disease prediction using machine learning typically follows a structured workflow composed of data acquisition, preprocessing, segmentation, feature extraction, model training, and evaluation. This section outlines the commonly adopted approaches and techniques across literature.

A. Data Acquisition

The prediction process begins with the collection of plant leaf images or sensor-based measurements. Image datasets such as Plant Village, PlantDoc, and field captured imagery from smartphones, drones, or greenhouse cameras are widely used. Some studies also incorporate environmental data, including temperature, humidity, and soil conditions, to support multimodal prediction. Ensuring dataset diversity is essential for improving model robustness across varying environments and plant varieties.

B. Data Preprocessing

Preprocessing aims to enhance image quality and reduce noise caused by inconsistent lighting, shadows, or background clutter. Typical operations include image resizing, normalization, histogram equalization, and color space transformations (e.g., RGB to HSV or LAB). Data augmentation techniques—such as rotation, flipping, zooming, and color jittering—are employed to artificially expand the dataset and prevent overfitting, especially when training deep learning models.

C. Segmentation

Segmentation is used to isolate the leaf or diseased region from the background, thereby improving feature extraction and classification accuracy. Traditional segmentation techniques include thresholding, region growing, and clustering-based methods (e.g., K-means).

Recent works utilize deep learning-based segmentation models such as UNet, SegNet, and Mask RCNN, which offer higher precision and are more effective in complex field conditions.

D. Feature Extraction

Feature extraction differs between classical machine learning and deep learning approaches. Classical models rely on handcrafted features like texture descriptors (GLCM, LBP), color histograms, and shape-based metrics. Deep learning, particularly Convolutional Neural Networks (CNNs), automatically extract hierarchical features without manual intervention. Transfer learning approaches using pretrained networks such as VGG16, ResNet50, DenseNet, or EfficientNet have become prevalent, significantly reducing training time while improving accuracy.

E. Model Training and Classification

Classical algorithms—such as Support Vector Machines (SVM), Random Forests, Decision Trees, and K-Nearest Neighbors (k-NN)—are trained using the extracted handcrafted features. Deep learning models, on the other hand, employ CNNs or hybrid frameworks such as CNN + SVM, attention-based networks, and ensemble learning models. Training is performed using backpropagation with optimizers like Adam or SGD, combined with regularization techniques such as dropout and batch normalization to enhance generalization.

G. Performance Evaluation

To assess model effectiveness, standard evaluation metrics are used, including accuracy, precision, recall, F1-score, and confusion matrices. For segmentation tasks, Intersection over Union (IoU) is commonly employed. Studies also evaluate inference time, computational cost, and model size to determine suitability for real-time deployment on mobile or edge devices. Cross-dataset validation and field testing are recommended to ensure robustness in real-world agricultural environments.

IV. Problem Statement

The fastest Accurate and timely detection of plant diseases is essential for mitigating crop losses and ensuring sustainable agricultural productivity. However, conventional disease identification methods rely heavily on manual visual inspection by farmers or agricultural experts, which is inherently subjective, labor-intensive, and inefficient for large-scale farming operations.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 14, Issue 12, December 2025)

These traditional approaches are further constrained by limited expert availability, variability in disease symptoms across growth stages, and environmental influences on visual characteristics.

Although machine learning and deep learning techniques have shown promising results in plant disease prediction, several critical challenges still hinder their widespread adoption in real-world agricultural settings. First, most existing models are trained on controlled laboratory datasets such as Plant Village, which lack the environmental complexities present in field conditions, including inconsistent lighting, background clutter, occlusions, and overlapping leaves.

V. PURPOSED ALGORITHM

The proposed algorithm aims to improve accuracy, robustness, and generalization capability of plant disease prediction under real-world agricultural conditions. To address limitations in existing methods—such as poor performance on field images, sensitivity to background noise, and limited interpretability, a hybrid deep learning framework combining segmentation, attention-based feature extraction, and machine learning classification is proposed. The overall architecture consists of four major modules: image preprocessing, leaf segmentation, feature extraction with attention, and final classification.

A. Preprocessing Module

Input images captured from mobile devices or field cameras are first resized, normalized, and enhanced using contrast adjustment and noise reduction filters. Data augmentation operations such as rotation, zooming, horizontal flipping, and color jittering are applied to increase dataset variability and reduce the risk of overfitting.

B. Leaf Segmentation Using U-Net

To isolate relevant regions and minimize background interference, a U-Net–based segmentation network is employed.

U-Net generates precise masks that extract the leaf area while removing soil, sky, shadows, and surrounding vegetation. This improves subsequent feature extraction and reduces the impact of environmental noise frequently present in field images.

C. Feature Extraction Using CNN With Attention Mechanism

A modified Convolutional Neural Network (CNN), integrated with a channel-spatial attention mechanism, is utilized to capture both global and local disease patterns. The attention blocks enhance discriminative feature regions—such as lesion color, texture, and shape—while suppressing irrelevant background features. Transfer learning from pretrained models (e.g., ResNet50 or EfficientNet-B0) is applied to accelerate convergence and improve accuracy on limited datasets.

D. Classification Using Support Vector Machine (SVM)

Instead of using the CNN's fully connected layer for final classification, features from the attention module are passed to a Support Vector Machine (SVM). SVM provides better class separation for small and imbalanced datasets, enhancing prediction reliability in multidisease scenarios.

E. Algorithm Workflow

- The proposed algorithm follows the sequence below:
- Input: Plant leaf image
- Preprocessing: Resize, normalize, enhance, and augment
- Segmentation: Apply U-Net to extract leaf region
- Feature Extraction: Use CNN + attention mechanism to extract multi-level features
- Classification Apply SVM for final disease prediction
- Output Predicted disease label with confidence score

Fig 3. Purposed Algorithm Plant Disease Prediction

VI. CONCLUSION

Plant disease prediction has emerged as a critical application domain where machine learning and computer vision significantly enhance agricultural productivity and sustainability. This survey reviewed key methodologies, including traditional machine learning classifiers, deep convolutional neural networks, transfer learning models, and hybrid architectures, all of which have demonstrated effective performance in identifying diseases across diverse crop species.

Although current techniques achieve high accuracy in controlled environments, their real-world applicability is still constrained by challenges such as varying illumination, occlusions, background noise, and limited availability of annotated field images.

Furthermore, the need for scalable, low-cost, and hardware efficient solutions remains unmet, particularly for deployment in resource constrained agricultural regions. Recent advancements in lightweight deep learning models, multimodal sensing, and mobile-based inference provide promising pathways toward practical adoption. However, issues related to dataset imbalance, domain adaptation, generalization across climatic regions, and the interpretability of model decisions continue to require substantial research attention.

REFERENCES

[1] Z. Wang, X. Zhang, and L. Wang, “A comprehensive survey on deep learning for smart agriculture and plant disease prediction,” Information Fusion, vol. 91, pp. 53–82, 2024.

[2] G. S. Dhiman and V. Kumar, “A survey of deep learning techniques for plant disease prediction,” Journal of King Saud University – Computer and Information Sciences, 2023.

[3] H. Kaur, D. Prashar, and V. Kumar, “A review on the application of deep learning models in plant disease detection,” Journal of Ambient Intelligence and Humanized Computing, 2023.

[4] M. Shoaib, et al. "Recent advancements and challenges in deep learning-based plant disease identification." Frontiers in Plant Science, 2023.

[5] N. Shelar, & V. Gopi. "Deep learning mechanisms for efficient plant disease-identifying techniques." International Journal of Research in Engineering and Science, 2023.

[6] Anon. "Deep Learning in Multimodal Fusion for Sustainable Plant Care: A Comprehensive Review." MDPI, 2025.

[7] Anon. “Multimodal Fusion Framework for RealTime Crop Disease Forecasting.” ResearchGate Article, 2025.

[8] Anon. “An intelligent identification for pest and disease detection in wheat leaf based on environmental data using multimodal data fusion.” Frontiers in Plant Science, 2025.

[9] Anon. “Cross-Modal Data Fusion via Vision Language Model for Crop Disease Recognition” PMC - NIH, 2025.

[10] Anon. “A Multimodal Data Fusion and Embedding Attention Mechanism-Based Method for Eggplant Disease Detection.” Agronomy, 2024.

[11] P. Tripathi, & R. Nandi, “AI-based crop disease prediction using environmental parameters.” Environmental Monitoring and Assessment, vol. 195, pp. 1–17, 2023.

[12] Anon. “Automatic fused multimodal deep learning for plant identification.” Frontiers in Plant Science, 2025.

[13] B. Wang, “Zero-exemplar deep continual learning for crop disease recognition: A study of total variation attention regularization in vision transformers.” Frontiers in Plant Science, vol. 14, 2024.

[14] U. Barmanetal."ViT-SmartAgri:Vision Transformer and Smartphone-Based Plant Disease Detection for Smart Agriculture." Agronomy, vol. 14, 2024.

[15] A. Ait Nasser, & M.A. Akhloufi. "A Hybrid Deep Learning Architecture for Apple Foliar Disease Detection." Computers, vol. 13, 116, 2024. Y. Liu, et al. "High-Precision Tomato Disease

[16] Detection Using NanoSegmenter Based on Transformer and Lightweighting." Plants, vol. 12, 2559, 2023.

[17] Anon. "The Power of Vision Transformers and Acoustic Sensors for Cotton Pest Detection." IEEE Open J. Comput. Soc., 2024.

[18] S. Varghese. "Crop Disease Detection and Prediction using YOLOv8s." International Journal of Novel Research and Development, 2024.

[19] A. Verma, et al. "A unified light-weight CNN model for corn, rice, and wheat disease detection." Applied Engineering Research, 2024.

[20] K. Prathiksha, et al. "Detection of grape leaf disease using a VGG-19 based CNN model." Journal of Advanced Research in Computer Engineering, 2024.

[21] P. Thakur, et al. "VGG-ICNN model: A lightweight CNN for efficient crop disease detection." Expert Systems with Applications, 2023.

[22] S. Archanaa, et al. "Convolutional Neural Networks (CNNs) in the diagnosis of diseases in tomatoes." Procedia Computer Science, 2023.

[23] T. Sabbah, T., & M. A. Khan. "A lightweight modified CNN model detects wheat diseases with high accuracy." IEEE Access, 2023.

[24] A. T. Khan, S. M. Jensen, et al. "Plant disease detection model for edge computing devices." Frontiers in Plant Science, vol. 14, 2023.

[25] Anon. "Edge AI and CNN for Real-Time Plant Disease Detection using IOT." International Journal of Innovative Research in Science, 2025.

[26] Anon. "An Edge Computing-Based Solution for Real-Time Leaf Disease Classification using Thermal Imaging." arXiv preprint, 2024.

[27] Anon. "Deep Learning Based Automated Detection and Classification of Plant Leaf Diseases using Convolutional Neural Networks." International Journal of Innovative Research in Computer and Communication Engineering, 2025.

[28] Anon. "Novel Edge Device System for Plant Disease Detection with Deep Learning Approach." International Journal of Intelligent Systems and Applications in Engineering, 2024.

[29] M. Jung, J. S. Song, et al. "Construction of deep learning-based disease detection model in plants." Scientific Reports, vol. 13, 7331, 2023.

[30] M. S. M. Asaari, et al. "Detection of Plant Stress Condition with Deep Learning Based Detection Models." 2023 International Conference on Energy, Power, Environment, Control, and Computing (ICEPECC), 2023.