International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 11 | Nov - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

Recent Trends in Plant Leaf Disease Classification: From Image Processing to Explainable Artificial Intelligence

K. Pavithra Devi
Research Scholar,
Department of Computer science (Artificial intelligence &Data Science),
Dr.SNS Rajalakshmi College of Arts and Science, Coimbatore
pavithradevisps@gmail.com.

Dr.P.Shivaranjani
Head, Department of Computer Science with Artificial Intelligence,
Machine Learning and Data Science,
Dr.SNS Rajalakshmi College of Arts and Science, Coimbatore.
hod.gcd@drsnsrcas.ac.in

Abstract
Plant leaf diseases significantly affect agricultural productivity and crop quality, making timely and accurate disease diagnosis essential for sustainable farming. Recent advances in artificial intelligence have transformed plant disease classification from conventional image processing techniques to sophisticated deep learning and explainable artificial intelligence (XAI) frameworks. This survey presents a comprehensive review of state-of-the-art methods for plant leaf disease classification published between 2023 and 2026. The study examines publicly available benchmark datasets, image preprocessing, leaf segmentation, data augmentation, feature extraction, and feature selection techniques, followed by an extensive analysis of machine learning and deep learning approaches for disease classification. Furthermore, recent developments in XAI are reviewed to highlight their role in improving model transparency, interpretability, and user trust. The strengths, limitations, and research trends of existing methods are systematically discussed, with particular emphasis on lightweight models, transformer-based architectures, ensemble learning, few-shot learning, and real-time deployment. Finally, the survey identifies key challenges, including limited dataset diversity, poor generalization under field conditions, computational complexity, and insufficient explainability, and outlines future research directions toward robust, efficient, and interpretable plant disease diagnosis systems for smart agriculture.

Keywords: Plant Leaf Disease Classification, Deep Learning, Image Processing, Feature Extraction, Precision Agriculture.

1. Introduction

Agriculture is fundamental to global food security and economic development, with healthy crops being essential for sustainable food production. However, plant diseases significantly reduce crop yield and quality, causing substantial economic losses worldwide. Early and accurate detection of plant diseases is therefore crucial for minimizing crop damage, improving agricultural productivity, and supporting precision farming. Conventional disease diagnosis primarily relies on visual inspection by experienced farmers or plant pathologists, which is often time-consuming, subjective, and unsuitable for large-scale agricultural monitoring.

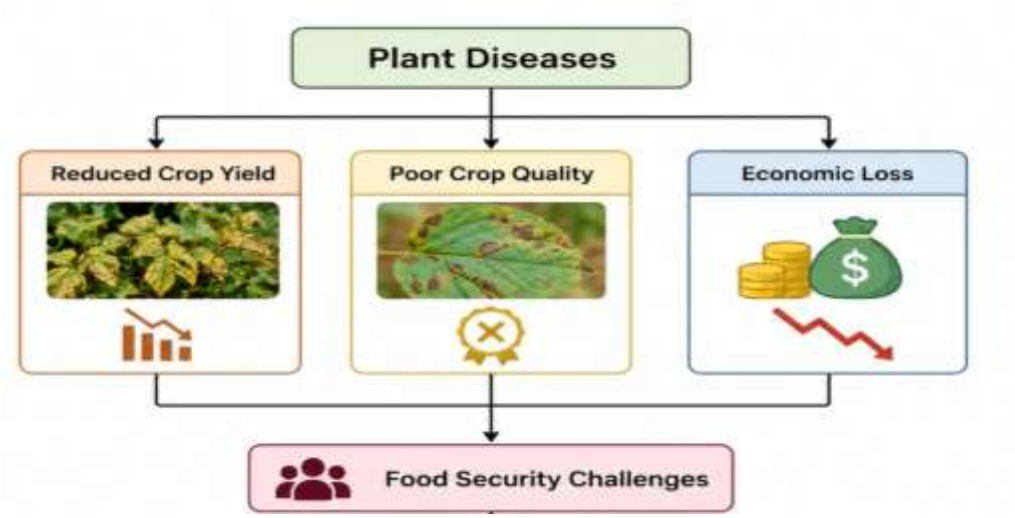


Fig.1. Impact of Plant Diseases on Agriculture

Recent advances in Artificial Intelligence (AI) have enabled automated plant leaf disease diagnosis using image-based analysis. Initially, machine learning techniques were employed by combining handcrafted image features with classification algorithms to identify disease symptoms. More recently, deep learning approaches have demonstrated superior performance by automatically learning discriminative features from leaf images, thereby reducing the need for manual feature engineering. In addition to classification methods, image preprocessing, leaf segmentation, data augmentation, feature extraction, feature selection, and Explainable Artificial Intelligence (XAI) have become integral components for improving classification accuracy, robustness, and model interpretability. [1]

Despite considerable progress, several challenges remain, including limited real-field datasets, variations in illumination and background conditions, class imbalance, visually similar disease symptoms, and insufficient model interpretability. Furthermore, many existing studies focus primarily on classification performance while providing limited discussion of the complete disease diagnosis pipeline. Therefore, a comprehensive survey covering the entire workflow is needed to provide a holistic understanding of current methodologies and identify future research opportunities. [2]

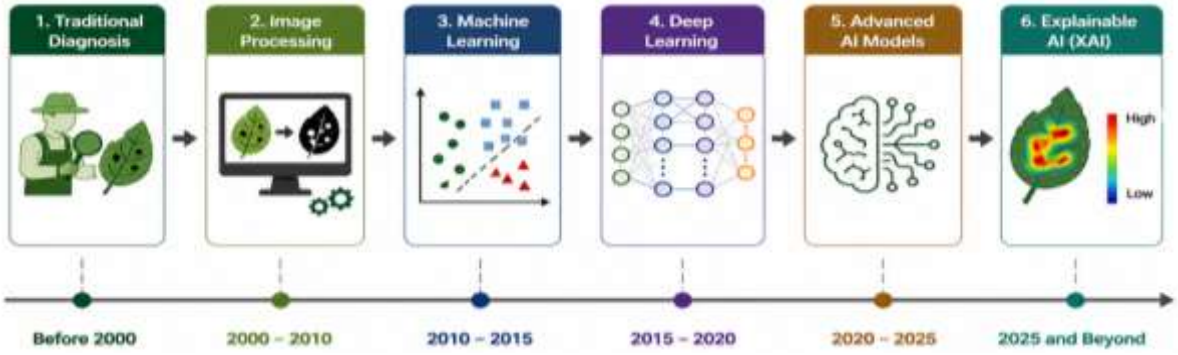


Fig.2. Evolution of Plant Leaf Disease Detection

This survey presents a comprehensive review of plant leaf disease classification using machine learning and deep learning techniques. It systematically discusses publicly available datasets, image preprocessing, segmentation, data augmentation, feature extraction, feature selection, classification approaches, explainable artificial intelligence, evaluation metrics, research challenges, and future research directions. The survey aims to provide researchers and practitioners with a structured reference for developing accurate, reliable, and interpretable intelligent plant disease diagnosis systems.

2. Plant Leaf Diseases

Plant diseases are among the primary factors limiting agricultural productivity and crop quality worldwide. They are caused by a variety of pathogenic organisms, including fungi, bacteria, viruses, nematodes, and phytoplasmas, as well as unfavorable environmental conditions. These diseases affect different parts of plants, particularly leaves, stems, fruits, and roots, resulting in reduced photosynthetic activity, stunted growth, poor-quality produce, and significant yield losses. Early identification of plant diseases is therefore essential for implementing timely control measures, minimizing economic losses, and promoting sustainable agricultural practices.

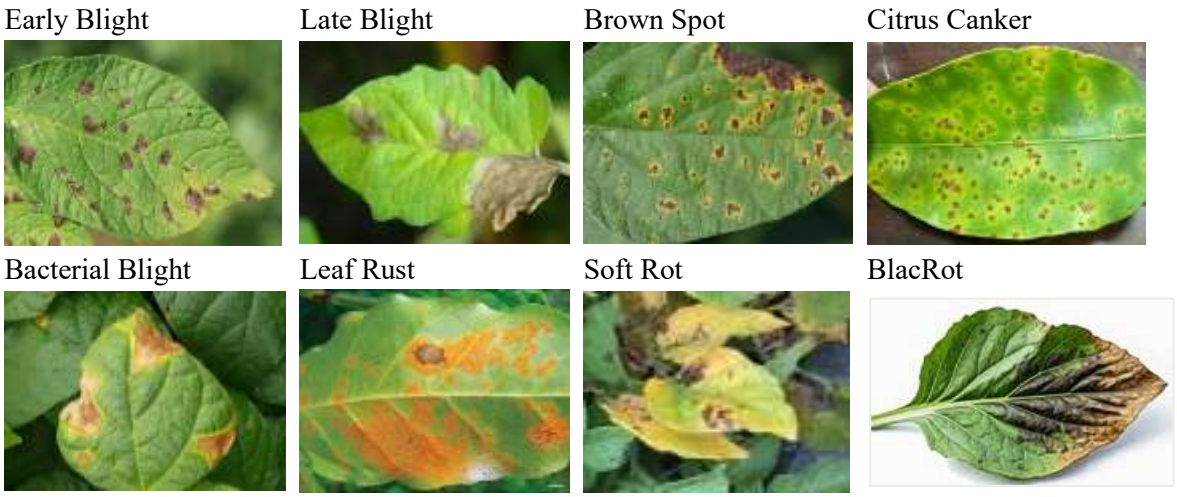


Fig.3. Common Leaf Disease Images

Figure 3 illustrates several representative plant leaf diseases commonly encountered in agricultural applications. Among various plant organs, leaves are the most commonly analyzed for disease diagnosis because they exhibit visible symptoms during the early stages of infection. Typical symptoms include leaf spots, chlorosis, necrosis, blight, mildew, rust, curling, and discoloration, which can be effectively captured using digital imaging techniques. Advances in computer vision and artificial intelligence have facilitated automated analysis of these visual symptoms, enabling rapid and non-destructive disease diagnosis. [3]

Plant diseases can be broadly categorized according to their causal agents. **Fungal diseases** are the most prevalent and include conditions such as early blight, late blight, powdery mildew, leaf mold, and rust. **Bacterial diseases** often manifest as leaf spots, wilting, or cankers, while **viral diseases** commonly produce mosaic patterns, yellowing, leaf curling, and deformation. In addition, nutrient deficiencies and abiotic stresses may produce symptoms similar to infectious diseases, making accurate classification a challenging task.

Table.1. Common Plant Leaf Diseases

Crop	Disease	Causal Agent	Common Symptoms
Tomato	Early Blight	Fungus	Brown concentric spots
Tomato	Late Blight	Oomycete	Dark lesions with white mold
Potato	Late Blight	Oomycete	Water-soaked lesions
Potato	Soft Rot	Bacteria	Soft mushy tissue, tissue decay
Rice	Brown Spot	Fungus	Circular brown spots
Apple	Apple Scab	Fungus	Olive-green lesions
Corn	Common Rust	Fungus	Orange pustules
Grapes	Black Rot	Fungus	Dark circular lesions
Citrus	Citrus Canker	Bacteria	Raised corky lesions
Cotton	Bacterial Blight	Bacteria	Angular water-soaked spots
Wheat	Leaf Rust	Fungus	Orange-brown pustules

Table 1 summarizes common diseases, their causal agents, and typical visual symptoms. Numerous crops have been extensively investigated for automated disease detection owing to their agricultural importance and the availability of publicly accessible datasets. Tomato, potato, rice, maize (corn), apple, grape, citrus, cotton, soybean, and wheat are among the most frequently studied crops in machine learning and deep learning research.

3. Recent Advances in Plant Leaf Disease Classification

Plant leaf disease classification has attracted significant research interest due to advancements in computer vision and artificial intelligence. Early studies primarily relied on traditional image processing and machine learning techniques using handcrafted features for disease recognition. Recent research has increasingly focused on deep learning architectures, transfer learning, semantic segmentation, attention mechanisms, transformer-based models, and explainable artificial intelligence (XAI) to improve classification accuracy, robustness, and model interpretability. These

advancements have enabled automated disease diagnosis under diverse agricultural conditions and accelerated the development of intelligent precision agriculture systems. Table 2 summarizes representative studies on plant leaf disease classification, including the datasets, models, and key findings.

Table.2. Comparative Analysis of Existing Plant Leaf Disease Classification Studies

Authors	Dataset	Models Used	Key Findings
Srinivasan et al. (2026) [4]	Tomato, Corn, Potato datasets	Hybrid Deep Neural Transformer + XAI	Improved classification accuracy and model interpretability.
Krishna et al. (2025) [5]	PlantDoc + Web images	EfficientNet-B0, EfficientNet-B3, ResNet50, DenseNet201	Multi-dataset learning enhanced disease classification robustness.
Abd Algani et al. (2023) [6]	PlantVillage	Optimized CNN	Optimized deep learning improved disease recognition performance.
Leelavathi & Kalamani (2025) [7]	Medicinal Plant Dataset	CLAHE, Adaptive Threshold, VGG16, HZKHO, Attention-DenseNet	Integrated framework significantly improved classification accuracy.
Polly & Devi (2024) [8]	Apple, Tomato, Corn datasets	YOLOv8, DeepLabV3+, U-Net, CNN	Semantic segmentation improved disease localization and classification.
Hukkeri et al. (2024) [9]	PlantVillage (38 Classes)	AlexNet, VGG16, VGG19, ResNet50, DenseNet121, InceptionV3, EfficientNetB0	EfficientNetB0 achieved the highest classification accuracy.

Despite substantial progress, several challenges remain unresolved. Most existing studies focus primarily on improving classification accuracy while giving comparatively less attention to the complete disease diagnosis pipeline, including preprocessing, segmentation, feature engineering, and explainability. In addition, many models are evaluated on laboratory-controlled datasets with limited validation under real-field conditions, restricting their practical applicability. Therefore, a comprehensive survey that systematically reviews datasets, image processing techniques, feature engineering methods, machine learning, deep learning, and explainable AI is essential to provide a holistic understanding of current developments and identify future research directions.

4. Methodology

The overall plant leaf disease classification workflow consists of image acquisition, image preprocessing, leaf segmentation, data augmentation, feature extraction, feature selection, disease classification, and explainable disease diagnosis. Each stage contributes to improving the robustness, accuracy, and interpretability of automated disease recognition systems.

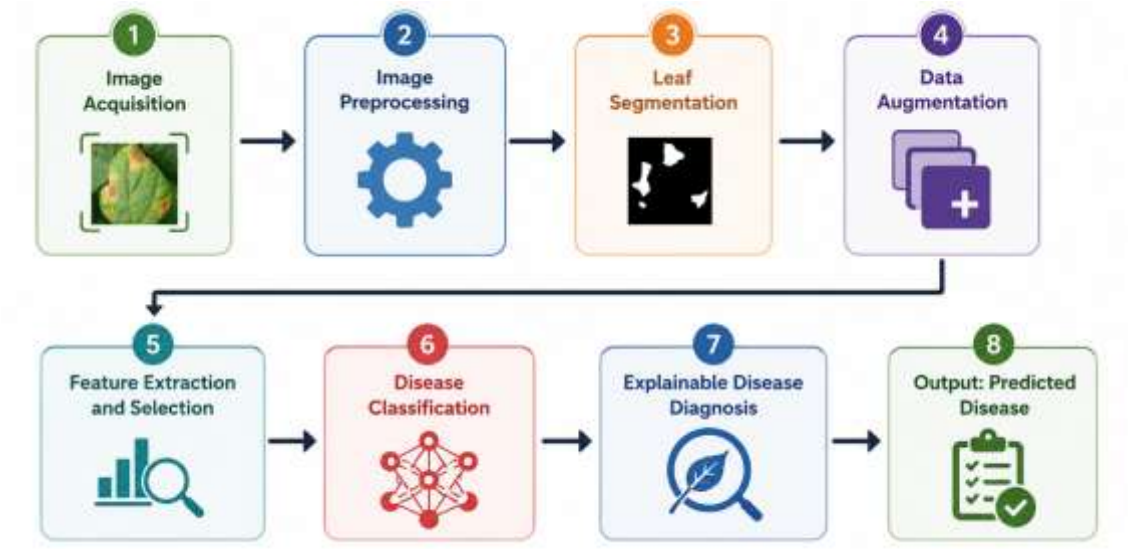


Fig.4. Plant Leaf Disease Classification Pipeline

Figure 4 presents the overall methodology for plant leaf disease classification, encompassing image acquisition, preprocessing, segmentation, data augmentation, feature extraction and selection, disease classification, and explainable AI-based diagnosis for accurate and interpretable disease prediction.

4.1. Datasets

Publicly available datasets play a crucial role in developing and evaluating automated plant leaf disease classification systems. These datasets provide large collections of healthy and diseased leaf images captured under controlled or real-field conditions, enabling researchers to train, validate, and benchmark machine learning and deep learning models.

i. **PlantVillage -** The PlantVillage dataset is one of the most widely used benchmark datasets for plant leaf disease classification. It contains **54,306** images of healthy and diseased plant leaves belonging to **38 classes** across **14 crop species**. The images were captured under controlled laboratory conditions with homogeneous backgrounds, making the dataset suitable for benchmarking machine learning and deep learning models. [10]

ii. **PlantDoc Dataset** - The PlantDoc dataset is a real-world benchmark dataset collected under natural environmental conditions. It contains **2,598** annotated images from **13 crop species** covering **27 disease classes**, with varying illumination, occlusions, and complex backgrounds. The dataset is widely used to evaluate the robustness and generalization capability of plant disease classification models. [11]

iii. **Plant Pathology 2020 Dataset** - The Plant Pathology 2020 dataset was introduced as part of the FGVC7 Kaggle Challenge for apple leaf disease classification. It contains **3,651** high-resolution images belonging to **four classes**: healthy, apple scab, cedar apple rust, and multiple diseases. The images were collected under real orchard conditions with varying lighting, viewing angles, and backgrounds, providing a realistic benchmark for disease recognition. [12]

iv. **Plant Pathology 2021 Dataset -** The Plant Pathology 2021 dataset is an extended version of the Plant Pathology 2020 dataset, released for the FGVC8 Kaggle Challenge. It contains **18,632** high-quality RGB images of apple leaves with **12** disease classes, including additional disease categories and multi-label annotations. The dataset represents real-field scenarios with non-homogeneous backgrounds, varying leaf maturity, and diverse imaging conditions, making it a challenging benchmark for advanced deep learning models. [13]

v. **Cassava Leaf Disease Dataset** - The Cassava Leaf Disease dataset was developed to facilitate automated diagnosis of cassava diseases that threaten food security, particularly in Africa. It contains **21,367** field images categorized into **five classes**, including four major diseases and one healthy class. The dataset is extensively used for evaluating deep learning models under real agricultural field conditions. [14]

Table.3. Publicly Available Datasets for Plant Disease Classification

Dataset	Crop Types	Images	Classes	Characteristics
PlantVillage	14 crops	54,306	38	Laboratory-controlled images; most widely used benchmark dataset
PlantDoc	13 crops	2,598	27	Real-field images with complex backgrounds
Plant Pathology 2020	Apple	3,651	4	Field images collected under natural conditions
Plant Pathology 2021	Apple	18,632	12	Multi-label disease classification dataset
Cassava Leaf Disease	Cassava	21,367	5	Real-world disease images for food security applications

The availability of standardized datasets has significantly accelerated the development of robust disease diagnosis systems and facilitated fair comparison among different classification approaches.

4.2. Image Preprocessing

Image preprocessing improves the quality and consistency of leaf images before feature extraction and classification. It minimizes the effects of noise, illumination variations, and background interference, thereby enhancing the performance of disease classification models.

Table.4. Common Image Preprocessing Techniques

Technique	Description
Image Resizing	Rescales images to a fixed size for uniform model input.
Noise Removal	Eliminates unwanted noise using smoothing filters.
Contrast Enhancement	Improves visibility of disease symptoms.
Histogram Equalization	Enhances global image contrast.
CLAHE	Improves local contrast while preventing over-enhancement.
Median Filter	Removes salt-and-pepper noise while preserving edges.
Gaussian Filter	Reduces image noise through Gaussian smoothing.
Background Removal	Removes unnecessary background to isolate the leaf.
Color Space Conversion	Converts RGB images to HSV, LAB, or grayscale for better analysis.

4.3. Leaf Segmentation

Leaf segmentation separates the leaf or infected region from the background, allowing the classification model to focus on disease-affected areas. Effective segmentation reduces background interference and improves disease localization.

Table.5. Common Leaf Segmentation Techniques

Technique	Description
Thresholding	Separates foreground and background using intensity values.
Otsu Thresholding	Automatically determines the optimal threshold value.
K-Means Clustering	Groups pixels into clusters based on colour similarity.
Watershed Algorithm	Segments touching or overlapping regions.
U-Net	CNN-based semantic segmentation for precise leaf extraction.
U-Net++	Improved U-Net with nested skip connections.

DeepLabV3+	Multi-scale semantic segmentation using atrous convolution.
Mask R-CNN	Instance segmentation for detecting individual leaves.
YOLO Segmentation	Performs simultaneous object detection and segmentation.

4.4. Data Augmentation

Data augmentation artificially increases dataset diversity by generating modified versions of training images. It helps reduce overfitting and improves the generalization capability of machine learning and deep learning models.

Table.6. Common Data Augmentation Techniques

Technique	**Description**
Rotation	Rotates images at different angles.
Flipping	Horizontally or vertically flips images.
Scaling	Changes image size while preserving content.
Translation	Shifts images horizontally or vertically.
Cropping	Extracts different regions of an image.
Brightness Adjustment	Varies image brightness to simulate lighting changes.
MixUp	Combines two images and their labels for training.
CutMix	Replaces image regions with patches from another image.
Mosaic	Combines multiple images into a single training sample.
GAN-based Augmentation	Generates synthetic leaf images using generative models.

Image preprocessing, leaf segmentation, and data augmentation collectively enhance image quality, isolate disease-affected regions, and improve dataset diversity. These techniques provide high-quality inputs for subsequent feature extraction and classification, resulting in more accurate and robust plant leaf disease diagnosis.

Table.7. Representative Studies on Image Preprocessing and Leaf Segmentation

Authors	**Dataset**	**Pre-processing**	**Segmentation**
AlArfaj et al. (2023) [15]	Pepper Bell Leaf Disease Dataset	Multi-step pre-processing, image enhancement	U-Net
Leelavathi & Kalamani (2025) [7]	Mendeley Medicinal Plant Leaf Dataset	Median Filter, CLAHE	Adaptive Threshold Segmentation
Polly & Devi (2024) [8]	Apple, Tomato, and Corn Leaf Images	ROI extraction, background removal	YOLOv8, DeepLabV3+, U-Net
Yang et al. (2023) [16]	Plant Leaves Dataset	Leaf detection, ROI extraction	YOLOv8 + Improved DeepLabV3+
Hu et al. (2024) [17]	Plant Leaf Disease Dataset	—	U-Net3+ with SK-Attention
Kumar et al. (2023) [18]	PlantVillage Dataset	Image enhancement	Image segmentation
Singh et al. (2025) [19]	Leaf Disease Segmentation Dataset	Image preprocessing	U-Net

Although significant progress has been made in image preprocessing and leaf segmentation, several challenges remain. Many studies rely on controlled datasets and crop-specific segmentation methods, limiting their generalization to diverse field environments with complex backgrounds, varying illumination, and occlusions. Furthermore, preprocessing and segmentation techniques are often evaluated independently rather than as adaptive components of a unified disease

classification pipeline. Developing robust, lightweight, and generalized preprocessing and segmentation frameworks capable of handling multiple crops and real-world imaging conditions remains an important research direction.

4.5. Feature Extraction

Feature extraction is the process of transforming preprocessed leaf images into informative representations that capture disease-specific characteristics. Effective feature extraction enhances the discriminative ability of classification models by identifying relevant visual patterns related to color, texture, shape, and lesion morphology. Depending on the classification approach, features may be extracted manually using traditional image processing techniques or automatically through deep learning models.

Table.8. Common Feature Extraction Techniques

Feature Extraction Technique	Description
Color Features	Extracts color information (RGB, HSV, LAB) to identify discoloration and lesion patterns.
GLCM	Extracts statistical texture features such as contrast, correlation, energy, and homogeneity.
LBP	Captures local texture patterns by encoding neighborhood pixel relationships.
HOG	Describes edge and gradient orientation information for disease region representation.
CNN Features	Automatically learns hierarchical features from leaf images using convolutional layers.
Hybrid Feature Fusion	Combines handcrafted features (e.g., GLCM, LBP, HOG) with deep features to improve classification performance.

Various feature extraction techniques have been employed to represent disease-specific characteristics from plant leaf images. Traditional methods primarily utilize handcrafted descriptors based on color, texture, and edge information, whereas recent studies increasingly rely on deep learning models to automatically learn discriminative features.

Table.9. Studies on Feature Extraction Techniques for Plant Leaf Disease Classification

Authors	Dataset	Feature Extraction Technique	Key Contribution
Nandhini & Bhavani (2020) [20]	Plant leaf disease dataset	Color, Shape, GLCM	Improved disease representation using handcrafted features.
Yogeshwari & Thailambal (2023) [21]	PlantVillage	GLCM + PCA	Enhanced feature representation using texture descriptors.
Chetan et al. (2024) [22]	PlantVillage	Hybrid Color, Texture, Shape Features	Improved feature representation through hybrid descriptors.
Leelavathi & Kalamani (2025) [7]	Mendeley Medicinal Plant Dataset	Multi-scale VGG16 + Color + Texture + Shape	Combined deep and handcrafted features for improved classification.
Gautam et al. (2025) [23]	Mendeley Medicinal Leaf Dataset	LBP + HOG + CNN Feature Fusion	Hybrid feature fusion enhanced classification performance.
Babu & Prasad (2025) [24]	Apple, Grape, Corn Datasets	HOG + ORB + SIFT + VGG + ResNet	Integrated handcrafted and deep features for robust detection.

Existing feature extraction methods either rely on handcrafted descriptors with limited discriminative capability or computationally intensive deep learning models. Developing efficient hybrid feature extraction techniques that provide robust, generalized, and lightweight feature representations across different crops and environmental conditions remains a significant research challenge.

4.6. Feature Selection

Feature selection is the process of identifying the most relevant features from the extracted feature set while eliminating redundant and irrelevant information. It reduces feature dimensionality, improves computational efficiency, minimizes overfitting, and enhances the performance of machine learning and deep learning models. Effective feature selection enables classifiers to focus on the most discriminative disease characteristics, resulting in improved classification accuracy and model generalization.

Table.10. Common Feature Selection Techniques

Technique	Description
Principal Component Analysis (PCA)	Reduces feature dimensionality by transforming correlated features into principal components.
ReliefF	Ranks features based on their ability to distinguish between different classes.
Genetic Algorithm (GA)	Selects optimal feature subsets using evolutionary optimization.
Particle Swarm Optimization (PSO)	Optimizes feature subsets by simulating swarm intelligence.
Neighborhood Component Analysis (NCA)	Learns feature importance by maximizing classification accuracy.
Recursive Feature Elimination (RFE)	Iteratively removes less important features to obtain an optimal feature subset.

Feature selection techniques identify the most informative features while removing redundant and irrelevant data. Table 10 presents the commonly used feature selection methods in plant leaf disease classification.

Table.11. Studies on Feature Selection Techniques for Plant Leaf Disease Classification

Author	Dataset	Feature Selection Technique	Key Contribution
Yogeshwari & Thailambal (2023) [21]	PlantVillage	Principal Component Analysis (PCA)	Reduced feature dimensionality for efficient classification.
Aggarwal et al. (2023) [25]	Rice Leaf Disease Dataset	Deep Feature Selection	Selected discriminative deep features for ML classification.
Ji et al. (2023) [26]	PlantVillage	Salp Swarm Algorithm (SSAFS)	Optimized feature subsets using swarm intelligence.
Lakineni et al. (2023) [27]	PlantVillage	Optimal Feature Selection	Improved classification using optimal CNN feature selection.
Dubey & Choubey (2024) [28]	Paddy Leaf Disease Dataset	Adaptive Feature Selection (AFS)	Adaptively selected informative features for disease classification.
Leelavathi & Kalamani (2025) [7]	Mendeley Medicinal Plant Dataset	HZKHO Optimization	Optimized weighted feature selection for improved classification.
Gautam et al. (2025) [23]	Mendeley Medicinal Leaf Dataset	Neighborhood Component Analysis (NCA)	Enhanced feature discrimination using NCA.

Anand et al. (2025) [29]	Plant & Medicinal Plant Datasets	Quantum-Inspired Evolutionary Feature Selection (QEFS)	Improved feature optimization using quantum-inspired evolution.

Despite significant advancements in feature selection techniques, several challenges remain. Many existing methods are computationally expensive, dataset-specific, and exhibit limited generalization across different crop species and disease conditions. Furthermore, optimization-based feature selection algorithms often prioritize classification accuracy while overlooking computational efficiency and scalability. Therefore, there is a need for adaptive, lightweight, and generalized feature selection frameworks that can effectively identify discriminative features while reducing redundancy and supporting real-time plant leaf disease classification.

4.7. Classification

Classification is the final stage of the plant leaf disease diagnosis pipeline, where the extracted and selected features are assigned to their corresponding disease categories. Depending on the learning paradigm, classification approaches are broadly categorized into machine learning (ML) and deep learning (DL) methods. Machine learning algorithms generally rely on handcrafted or selected features, whereas deep learning models automatically learn hierarchical feature representations from raw images. Recent studies have increasingly adopted deep learning architectures due to their superior accuracy and ability to handle complex disease patterns.

4.7.1. Machine Learning-Based Classification

Machine learning classifiers utilize manually extracted or optimized features to identify plant diseases. These methods are computationally efficient and perform well on small- and medium-sized datasets but often depend on effective feature engineering.

Table.12. Common Machine Learning Classifiers

Classifier	**Description**
Support Vector Machine (SVM)	Separates disease classes using an optimal decision boundary.
Random Forest (RF)	Ensemble of decision trees for robust classification.
K-Nearest Neighbors (KNN)	Classifies samples based on the nearest neighbors.
Decision Tree (DT)	Uses hierarchical decision rules for classification.
Naïve Bayes (NB)	Probabilistic classifier based on Bayes' theorem.
Extreme Gradient Boosting (XGBoost)	Gradient boosting algorithm for accurate disease prediction.

Although machine learning classifiers provide satisfactory performance on handcrafted features, their effectiveness depends heavily on feature quality and extensive feature engineering. They also struggle to capture complex visual patterns in large-scale datasets, limiting their applicability to real-world plant disease classification.

Table.13. Recent Studies on Machine Learning Based Plant Disease Classification

Author(s)	**Dataset**	**Machine Learning Model(s)**	**Key Contribution**
Aggarwal et al. (2023) [25]	Rice Leaf Disease Dataset	Extra Trees, Histogram Gradient Boosting	Improved classification using selected deep features.
Anand et al. (2025) [29]	Plant & Medicinal Plant Datasets	RF, SVM, DT, KNN, NB	Enhanced classification using QEFS-selected features.
Singla et al. (2023) [30]	PlantVillage	SVM, RF, DT, KNN, LR	Compared classical ML classifiers.
Sridhar & Angamuthu (2025) [31]	Multi-crop Dataset	Multiclass SVM (Linear, RBF, Quadratic, Cubic)	Optimized SVM kernels for disease classification.

Sujatha et al. (2025) [32]	PlantVillage	SVM, RF, DT, KNN, NB	Compared ML and DL classification methods.
Sankhe (2026) [33]	Plant Leaf Disease Dataset	SVM, RF, KNN, CNN	Developed an efficient ML–DL classification framework.

Despite the effectiveness of machine learning classifiers, their performance largely depends on high-quality handcrafted features and extensive feature engineering. Although recent studies have integrated machine learning with deep learning and optimization techniques, many models remain dataset-specific and exhibit limited robustness under real-world agricultural conditions. Therefore, developing generalized, lightweight, and scalable classification frameworks capable of maintaining high accuracy across diverse crops and environmental conditions remains a significant research challenge.

4.7.2. Deep Learning-Based Classification

Deep learning models automatically learn discriminative features directly from plant leaf images, eliminating the need for manual feature engineering. Their ability to model complex visual characteristics has made them the dominant approach for plant leaf disease classification, particularly with the availability of large annotated datasets.

Table.14. Common Deep Learning Models

Model	**Description**
Convolutional Neural Network	Automatically learns hierarchical image features.
ResNet	Employs residual learning to train deeper networks effectively.
DenseNet	Reuses feature maps through dense connections for efficient learning.
EfficientNet	Balances network depth, width, and resolution for improved performance.
Vision Transformer (ViT)	Captures global image relationships using self-attention.
Swin Transformer	Uses shifted window attention for efficient vision tasks.
Hybrid CNN–Transformer	Combines local CNN features with global transformer representations.

Despite achieving state-of-the-art performance, deep learning models often require large annotated datasets, substantial computational resources, and extensive training time. Their limited interpretability and reduced generalization under varying field conditions remain significant challenges, highlighting the need for lightweight, explainable, and robust classification frameworks.

Table.15. Recent Studies on Deep Learning Based Plant Disease Classification

Authors	**Dataset**	**Deep Learning Model**	**Key Contribution**
Srinivasan et al. (2026) [4]	PlantVillage + Multiple Public Datasets	Hybrid Deep Neural Transformer	Combined transformer learning with XAI.
Ali et al. (2024) [34]	PlantVillage	Ensemble of DenseNet121, MobileNetV2, InceptionV3, Xception	Improved classification using ensemble learning.
Saad & Salman (2024) [35]	Field Images (5 crops, 25 classes)	Siamese Neural Network (One-Shot Learning)	Enabled one-shot disease classification.

Kohut et al. (2025) [36]	PlantVillage, TomatoVillage, Kaggle	YOLOv8, EfficientDetV2, Faster R-CNN	Evaluated lightweight detection models.
Reddy et al. (2025) [37]	PlantVillage	Lightweight CNN	Developed a mobile disease diagnosis system.
Arshad et al. (2026) [38]	PlantVillage	ConvGeM-Next (ConvNeXt + GeM Pooling)	Enhanced classification using ConvNeXt architecture.
Duhan et al. (2026) [39]	PlantDoc, Mango, Soybean Datasets	MobileNetV2, EfficientNetB0, ShuffleNetV2, NASNet-Mobile	Compared lightweight deep learning models.
Devarajan et al. (2026) [40]	Plant Disease Dataset	CNN-based PDD-DL Framework	Enabled real-time disease diagnosis.

Despite the remarkable success of deep learning models in plant leaf disease classification, several challenges remain. Most existing models require large annotated datasets, high computational resources, and extensive training time, limiting their deployment on resource-constrained devices. Although recent studies have explored lightweight architectures, transformers, ensemble learning, and one-shot learning, many approaches still exhibit limited generalization across different crops, disease stages, and real-world field conditions. Therefore, developing efficient, robust, and generalized deep learning frameworks that provide high accuracy while ensuring computational efficiency and practical field deployment remains an important research direction.

4.7.3. Explainable Artificial Intelligence (XAI)

Explainable Artificial Intelligence (XAI) enhances the transparency and interpretability of deep learning models by explaining how predictions are generated. In plant leaf disease classification, XAI techniques help identify disease-affected regions and highlight the visual features that influence model decisions. This improves model reliability, facilitates expert validation, and increases user trust, making AI-based disease diagnosis more suitable for practical agricultural applications.

Table.16. Common XAI Techniques

XAI Technique	Description
Grad-CAM	Produces heatmaps highlighting image regions that influence model predictions.
Grad-CAM++	Improved Grad-CAM for more accurate localization of multiple disease regions.
LIME	Explains individual predictions by approximating the model locally with an interpretable model.
SHAP	Quantifies the contribution of each feature to the prediction using Shapley values.
Integrated Gradients	Measures feature importance by integrating gradients from a baseline to the input image.
Attention Visualization	Visualizes attention maps to show regions emphasized by attention-based models.

Although XAI techniques have improved the interpretability of plant disease classification models, most existing studies focus on visual explanations without quantitatively assessing their reliability or consistency. Furthermore, many explanation methods are computationally expensive and model-specific, limiting their applicability in real-time

agricultural systems. Developing efficient, model-agnostic, and trustworthy XAI frameworks that provide accurate and clinically meaningful explanations across diverse crops and disease conditions remains an important research challenge.

Table.17. Recent Studies on Explainable AI for Plant Leaf Disease Classification

Author(s)	Dataset	XAI Technique	Key Contribution
Srinivasan et al. (2026) [4]	PlantVillage + Public Datasets	XAI with Hybrid Transformer	Integrated transformer learning with XAI.
Natarajan et al. (2024) [41]	PlantVillage	Grad-CAM	Visualized disease-affected regions.
Oad et al. (2024) [42]	PlantVillage	LIME	Improved prediction interpretability.
Alhammad et al. (2025) [43]	Potato Leaf Dataset	Grad-CAM	Combined transfer learning with explainability.
El-Behery et al. (2026) [44]	Rice & Sugarcane Datasets	Grad-CAM	Enabled explainable few-shot disease detection.

Despite recent progress in explainable AI for plant leaf disease classification, several limitations remain. Most studies primarily rely on Grad-CAM, with relatively few exploring complementary explanation methods such as SHAP, Integrated Gradients, or model-agnostic explainability. Moreover, existing approaches mainly provide qualitative visualizations without quantitatively evaluating explanation fidelity, consistency, or clinical relevance. Developing robust, model-agnostic, and computationally efficient XAI frameworks capable of delivering trustworthy explanations across diverse crops, disease conditions, and real-world agricultural environments remains an important research challenge.

5. Challenges and Future Research Directions

This section summarizes the current limitations in plant leaf disease classification and outlines promising research opportunities.

Table.18. Challenges and Future Research Directions

Challenge	Future Research Direction
Limited annotated datasets	Develop large-scale and diverse benchmark datasets
Poor generalization across crops	Design crop-independent and domain-adaptive models
Complex field backgrounds	Improve robust segmentation and background removal techniques
Class imbalance	Employ advanced augmentation and synthetic data generation
High computational cost	Develop lightweight and edge-deployable models
Limited model interpretability	Integrate reliable and quantitative XAI techniques
Real-time deployment	Optimize models for mobile and IoT-based smart agriculture
Multimodal data integration	Combine RGB, hyperspectral, thermal, and UAV imagery
Early disease detection	Detect diseases before visible symptoms appear
Standardized evaluation	Establish common datasets and benchmarking protocols

Future research should focus on developing generalized and computationally efficient models capable of operating under diverse environmental conditions. Integrating multimodal sensing, vision transformers, federated learning, foundation models, and explainable AI can significantly improve disease diagnosis while enhancing transparency and scalability. Furthermore, standardized benchmarking datasets and edge-based deployment strategies are essential for translating laboratory models into practical smart agriculture applications.

6. Conclusion

Plant leaf disease classification has progressed from conventional image processing and machine learning techniques to advanced deep learning and explainable AI frameworks. Recent developments in transformers, lightweight architectures, ensemble learning, and few-shot learning have substantially improved classification performance and deployment capability. However, challenges related to dataset diversity, computational complexity, model generalization, and

interpretability continue to limit real-world adoption. Future research should prioritize robust, explainable, and resource-efficient models capable of supporting reliable disease diagnosis across diverse crops and agricultural environments.

Reference

1. Shafay, M., Hassan, T., Owais, M., Hussain, I., Khawaja, S. G., Seneviratne, L., & Werghi, N. (2025). Recent advances in plant disease detection: challenges and opportunities. Plant methods, 21(1), 140. https://doi.org/10.1186/s13007-025-01450-0
2. Shi, T., Liu, Y., Zheng, X. et al. Recent advances in plant disease severity assessment using convolutional neural networks. Sci Rep 13, 2336 (2023). https://doi.org/10.1038/s41598-023-29230-7
3. Zhao J, Xu L, Ma Z, Li J, Wang X, Liu Y and Du X (2025) A review of plant leaf disease identification by deep learning algorithms. Front. Plant Sci. 16:1637241. doi: 10.3389/fpls.2025.1637241
4. Srinivasan, S., A, R., B. A, N. et al. Multi-class classification of plant leaf diseases using a hybrid deep neural transformer system and explainable AI techniques. Sci Rep 16, 18161 (2026). https://doi.org/10.1038/s41598-026-48103-3
5. Krishna, M. S., Machado, P., Otuka, R. I., Yahaya, S. W., Neves dos Santos, F., & Ihianle, I. K. (2025). Plant Leaf Disease Detection Using Deep Learning: A Multi-Dataset Approach. J, 8(1), 4. https://doi.org/10.3390/j8010004
6. Yousef Methkal Abd Algani, Orlando Juan Marquez Caro, Liz Maribel Robladillo Bravo, Chamandeep Kaur, Mohammed Saleh Al Ansari, B. Kiran Bala, Leaf disease identification and classification using optimized deep learning, Measurement: Sensors, 2023, https://doi.org/10.1016/j.measen.2022.100643.
7. Leelavathi, R., Kalamani, M. Medicinal plant leaf disease classification using optimal weighted features with dilated adaptive DenseNet and attention mechanism. Sci Rep 15(2025). https://doi.org/10.1038/s41598-025-20629-y
8. Roshni Polly, E. Anna Devi, Semantic segmentation for plant leaf disease classification and damage detection: A deep learning approach, Smart Agricultural Technology, 2024 https://doi.org/10.1016/j.atech.2024.100526.
9. Hukkeri G, Soundarya B, Gururaj H, Ravi V. Classification of Various Plant Leaf Disease Using Pretrained Convolutional Neural Network On Imagenet . Open Agric J, 2024, http://dx.doi.org/10.2174/0118743315305194240408034912.
10. https://www.kaggle.com/datasets/emmarex/plantdisease
11. https://www.kaggle.com/datasets/abdulhasibuddin/plant-doc-dataset
12. https://www.kaggle.com/c/plant-pathology-2020-fgvc7
13. https://www.kaggle.com/competitions/plant-pathology-2021-fgvc8
14. https://www.kaggle.com/c/cassava-leaf-disease-classification
15. A. A. AlArfaj et al., "Multi-Step Preprocessing With UNet Segmentation and Transfer Learning Model for Pepper Bell Leaf Disease Detection," in IEEE Access, vol. 11, pp. 132254-132267, 2023, doi: 10.1109/ACCESS.2023.3334428
16. Yang, T., Zhou, S., Xu, A., Ye, J., & Yin, J. (2023). An Approach for Plant Leaf Image Segmentation Based on YOLOV8 and the Improved DEEPLABV3+. Plants, 12(19), 3438. https://doi.org/10.3390/plants12193438
17. Huili Hu, Xi Ye, Zhihong Yang, and Quanxin Zhang. 2024. Research on Image Segmentation Method of Plant Leaf Diseases Based on SK-Attention Attention Mechanism. In Proceedings of the 2024 4th International Conference on Internet of Things and Machine Learning (IoTML '24). Association for Computing Machinery, New York, NY, USA, 156–160. https://doi.org/10.1145/3697467.3697628
18. A. S. Kumar, M. Ramesh, M. Arpana, A. Sudarshanam, N. T. Velusudha and T. Suthar, "Image Processing Techniques for Leaf Disease Detection based on ELM-SSA Approach," 2023 3rd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), Bengaluru, India, 2023, pp. 311-316, doi: 10.1109/ICIMIA60377.2023.10426262.
19. Singh, G., et al. (2025). Enhanced Leaf Disease Segmentation Using U-Net Architecture for Precision Agriculture: A Deep Learning Approach. Food Science & Nutrition. DOI: 10.1002/fsn3.70594.
20. N. Nandhini and R. Bhavani, "Feature Extraction for Diseased Leaf Image Classification using Machine Learning," 2020 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2020, pp. 1-4, doi: 10.1109/ICCCI48352.2020.9104203.

21. Yogeshwari, M., & Thailambal, G. (2023). Automatic feature extraction and detection of plant leaf disease using GLCM features and convolutional neural networks. Materials Today: Proceedings, 81(2), 530–536. DOI: 10.1016/j.matpr.2021.03.700.
22. Chetan, H. R., Rajanna, G. S., Nandini, G., Santosh, K. C., Puneeth, B. H., Gururaj, K. S., Abdul Razak, M. S. (2024). Plant leaf disease detection using hybrid feature extraction techniques. Archives of Phytopathology and Plant Protection, 57(3), 200–218. https://doi.org/10.1080/03235408.2024.2347985
23. Gautam, V., Kaur, G., Ghantasala, G.S.P. et al. A lightweight deep learning method for medicinal leaf image classification using feature fusion. Sci Rep 15, 34417 (2025). https://doi.org/10.1038/s41598-025-17436-w
24. N. R. Babu and J. L. Prasad, "A Novel Multi-Feature Extraction Approach for Plant Disease Detection Using Transfer Learning and Handcrafted Methods," 2025 Second International Conference on Networks and Soft Computing (ICNSoC), Vadlamudi, India, 2025, pp. 354-359, doi: 10.1109/ICNSoC66817.2025.00073.
25. Aggarwal, M., Khullar, V., Goyal, N., Singh, A., Tolba, A., Thompson, E. B., & Kumar, S. (2023). Pre-Trained Deep Neural Network-Based Features Selection Supported Machine Learning for Rice Leaf Disease Classification. Agriculture, 13(5), 936. https://doi.org/10.3390/agriculture13050936
26. Ji, Z., Xie, X., Wu, Y., et al. (2023). A Novel Feature Selection Strategy Based on Salp Swarm Algorithm for Plant Disease Detection. Plant Phenomics. DOI: 10.34133/plantphenomics.0039.
27. P. K. Lakineni, D. J. Reddy, M. Chitra, R. Umapriya, L. V. Kannan and S. R. Barkunan, "Optimal Feature Selection and Classification Using Convolutional Neural Network-Based Plant Disease Prediction," 2023 IEEE International Conference on Integrated Circuits and Communication Systems (ICICACS), Raichur, India, 2023, pp. 1-6, doi: 10.1109/ICICACS57338.2023.10099630.
28. Dubey, R.K., Choubey, D.K. Adaptive feature selection with deep learning MBi-LSTM model based paddy plant leaf disease classification. Multimed Tools Appl 83, 25543–25571 (2024). https://doi.org/10.1007/s11042-023-16475-7
29. Anand, K., Jain, B., Mittal, H. et al. QEFS: A novel plant disease prediction approach using quantum-inspired evolutionary feature selection. Appl Intell 55, 101 (2025). https://doi.org/10.1007/s10489-024-05990-1
30. Singla, R. S., Gupta, A., Gupta, R., Tripathi, V., Naruka, M. S., & Awasthi, S. (2023). Plant Disease Classification Using Machine Learning. 2023 International Conference on Disruptive Technologies (ICDT). DOI: 10.1109/ICDT57929.2023.10151118
31. Sridhar, P., & Angamuthu, P. (2025). Enhancing image based classification for crop disease detection using a multiclass SVM approach with kernel comparison. Scientific Reports, 15, 40055. DOI: 10.1038/s41598-025-23568-w.
32. Sujatha, R., Krishnan, S., Chatterjee, J.M. et al. Advancing plant leaf disease detection integrating machine learning and deep learning. Sci Rep 15, 11552 (2025). https://doi.org/10.1038/s41598-024-72197-2
33. Sankhe, S.R. An efficient plant leaf disease classification using machine learning and deep learning models. Environ Dev Sustain (2026). https://doi.org/10.1007/s10668-026-07628-3
34. Ali, A. H., Youssef, A., Abdeldl, M., & Raja, M. A. (2024). An ensemble of deep learning architectures for accurate plant disease classification. Ecological Informatics, 81, 102618. DOI: 10.1016/j.ecoinf.2024.102618.
35. Saad, M.H., Salman, A.E. A plant disease classification using one-shot learning technique with field images. Multimed Tools Appl 83, 58935–58960 (2024). https://doi.org/10.1007/s11042-023-17830-4
36. Kohut, N., Basystiuk, O., Shakhovska, N., & Melnykova, N. (2025). Detecting Plant Diseases Using Machine Learning Models. Sustainability, 17(1), 132. DOI: 10.3390/su17010132.
37. B. Ramana Reddy, G. Kalnoor, M. Devashish and P. Sai Karthik Reddy, "Deep Learning Based Mobile Application for Automated Plant Disease Detection," in IEEE Access, vol. 13, pp. 107917-107925, 2025, doi: 10.1109/ACCESS.2025.3581099.
38. Arshad Z, Javed A and Saudagar AKJ (2026) ConvGeM-next: a deep learning framework for plant disease detection. Front. Plant Sci. 17:1763739. doi: 10.3389/fpls.2026.1763739
39. Duhan, S., Gulia, P., Gill, N.S. et al. Evaluation of lightweight and efficient deep learning models for plant disease classification. Discov Internet Things 6, 43 (2026). https://doi.org/10.1007/s43926-026-00310-0
40. Devarajan, D., Allafi, R., Obayya, M., & Nemri, N. (2026). AI based real time disease diagnosis in plants using deep learning driven CNNs. Scientific reports, 16(1), 4587. https://doi.org/10.1038/s41598-025-34681-1

41. Natarajan, S., Chakrabarti, P. & Margala, M. Robust diagnosis and meta visualizations of plant diseases through deep neural architecture with explainable AI. Sci Rep 14, 13695 (2024). https://doi.org/10.1038/s41598-024-64601-8
42. A. Oad et al., "Plant Leaf Disease Detection Using Ensemble Learning and Explainable AI," in IEEE Access, vol. 12, pp. 156038-156049, 2024, doi: 10.1109/ACCESS.2024.3484574.
43. Alhammad SM, Khafaga DS, El-hady WM, Samy FM and Hosny KM (2025) Deep learning and explainable AI for classification of potato leaf diseases. Front. Artif. Intell. 7:1449329. doi: 10.3389/frai.2024.1449329
44. El-Behery, H., Attia, AF. & Rezk, N.G. An explainable deep learning framework for few shot crop disease detection in rice and sugarcane using CNN based feature extraction. Sci Rep 16, 8272 (2026). https://doi.org/10.1038/s41598-026-37501-2