ISSN: 2096-3246
Volume 58, Issue 02, 2026

FEATURE ENGINEERING AND EXPLAINABLE AI FOR PLANT DISEASE DETECTION: A COMPREHENSIVE SURVEY

Vijayalakshmi S. Abbigeri[1] **Geetha D. Devanagavi**[2]
School of Computer and Science and Engineering, REVA University, Bengaluru, India
* Corresponding author's Email: vijayalaxmisa@gmail.com

Abstract: Plant diseases represent a significant threat to global food security, causing yield losses estimated between 20% and 40% annually. Accurate, timely, and scalable disease detection is therefore critical for modern agriculture. This survey provides a comprehensive review of feature engineering approaches and deep learning techniques applied to plant disease detection, with special focus on Gradient-weighted Class Activation Mapping (GradCAM) and Explainable Artificial Intelligence (XAI) methods that bridge the gap between black-box model predictions and human interpretability. We examine traditional handcrafted feature engineering methods including color histograms, texture descriptors (GLCM, LBP, Gabor), and shape features, followed by deep feature extraction using Convolutional Neural Networks (CNNs) and transfer learning architectures such as ResNet, EfficientNet, VGG, DenseNet, and MobileNet. The survey covers key datasets including PlantVillage and PlantDoc, discusses data augmentation strategies, and critically evaluates explainability tools including GradCAM, GradCAM++, LIME, and SHAP. Emerging trends such as MLOps integration, lightweight models for edge deployment, and transformer-based architectures are also discussed. The paper highlights challenges including limited real-world data, class imbalance, and the need for domain-trustworthy interpretability, and outlines future research directions toward robust, explainable, and deployable plant disease detection systems.

Keywords: Feature engineering, plant disease detection, GradCAM, explainable AI, deep learning, transfer learning, convolutional neural networks, precision agriculture

1. Introduction

Agriculture forms the backbone of economies worldwide, and crop diseases remain one of the most critical threats to food security. According to the Food and Agriculture Organization (FAO), plant diseases are responsible for yield losses of 20% to 40% globally each year, with the economic burden falling disproportionately on smallholder farmers in developing nations [1]. Traditional disease diagnosis methods rely heavily on visual inspection by trained agronomists, a process that is slow, expensive, and not scalable to the millions of farms that require monitoring.

The convergence of computer vision, deep learning, and mobile technology has opened transformative opportunities for automated plant disease detection. By analyzing leaf images captured on smartphones or drones, machine learning models can classify diseases rapidly and at scale. A critical factor in the success of these models lies in the quality of features extracted from raw images. Feature engineering—the art and science of transforming raw data into informative representations—remains foundational even in the era of deep learning [2].

While deep convolutional neural networks (CNNs) have dramatically reduced the need for manual feature design, understanding what features these networks learn is increasingly important for building trust in agricultural AI systems. Gradient-weighted Class Activation Mapping (GradCAM), introduced by Selvaraju et al. [3], provides a powerful mechanism to visualize which regions of a leaf image contributed most to a model's disease classification decision. This explainability is not merely academic—it is essential for farmers, agronomists, and policymakers who must trust AI recommendations before acting on them.

This survey bridges the domains of feature engineering and explainable AI in the specific context of plant disease detection. We trace the evolution from handcrafted features to deep automatic feature extraction, examine how GradCAM and related XAI tools reveal learned feature representations, and discuss the latest architectures, datasets, and deployment strategies. Our review synthesizes literature from 2020 to 2025, reflecting the rapid recent advances in this field.

The remainder of this paper is structured as follows: Section 2 reviews traditional feature engineering approaches for plant disease images. Section 3 covers deep learning-based feature extraction. Section 4 examines GradCAM and explainability methods. Section 5 discusses benchmark datasets. Section 6 reviews the latest architectures and results. Section 7 covers data challenges and augmentation. Section 8 addresses deployment and MLOps. Section 9 identifies emerging trends, and Section 10 concludes the survey.

2. Traditional Feature Engineering for Plant Disease Detection

Before the advent of deep learning, plant disease detection relied on carefully crafted feature pipelines that encoded domain knowledge about how diseased leaves appear visually. These handcrafted features can be organized into three principal categories: color, texture, and shape features [4].

2.1 Color Features

Color is often the first visible indicator of plant disease. Yellowing (chlorosis), browning (necrosis), and darkening of leaf tissue are hallmarks of infection. Feature engineering approaches exploit multiple color spaces to capture these cues:

RGB Histograms: Statistical summaries of red, green, and blue channel distributions provide a global description of color composition in diseased versus healthy leaves.

HSV Color Space: The Hue-Saturation-Value (HSV) model separates chromatic content from brightness, making it more robust to illumination variation—a key challenge in field photography.

LAB Color Space: The perceptually uniform LAB space enables more meaningful color distance metrics, useful for segmenting lesion regions from healthy tissue.

Color Moments: Mean, standard deviation, and skewness computed per channel provide compact color descriptors that have been widely used in plant disease classification [5].

A major limitation of color-based features is their sensitivity to environmental conditions such as lighting, shadows, and soil contamination on leaves, which can confound disease signatures.

2.2 Texture Features

Texture features capture the spatial arrangement of pixel intensities, encoding patterns of lesion roughness, spotting, and surface variation that are characteristic of specific diseases:

Gray Level Co-occurrence Matrix (GLCM): One of the most widely used texture descriptors, GLCM captures second-order statistics (contrast, energy, homogeneity, correlation) from spatial relationships between pixel pairs. It has been successfully applied to detect fungal and bacterial diseases [6].

Local Binary Pattern (LBP): LBP encodes local texture by comparing each pixel to its circular neighborhood, producing a rotation-invariant texture histogram. Its computational efficiency makes it suitable for real-time mobile applications.

Gabor Wavelet Transform: Multi-scale, multi-orientation Gabor filters simulate biological visual processing and are effective at capturing both fine-grained lesion textures and broader diseased area patterns [7].

Histogram of Oriented Gradients (HOG): Originally developed for pedestrian detection, HOG captures local gradient orientations and has been adapted for detecting edge patterns around disease lesions.

2.3 Shape Features

Shape-based features describe the morphological characteristics of lesions and leaf boundaries:

Lesion Area and Perimeter: Simple geometric measures that quantify the extent of visible disease symptoms.

Aspect Ratio and Circularity: Shape descriptors that differentiate between round fungal spots, elongated bacterial lesions, and irregular viral damage patterns.

Hu Moments: Seven rotation, scale, and translation-invariant moment descriptors that provide compact shape signatures for leaf and lesion contours.

Feature Type	Method	Advantages	Limitations	Typical Accuracy
Color	RGB, HSV, LAB Histograms	Simple to compute, interpretable, lighting-sensitive	Poor under varied lighting	75-85%
Texture	GLCM, LBP, Gabor	Captures lesion patterns, rotation-invariant variants available	High dimensionality, computationally intensive	80-90%
Shape	HOG, Hu Moments	Compact descriptors, invariant to transformations	Requires accurate segmentation	78-88%
Combined	Feature fusion + SVM/RF	Complementary information, higher accuracy	Feature selection complexity	88-93%

Table 1: Comparison of traditional feature engineering techniques for plant disease detection.

2.4 Comparison of Traditional Feature Engineering Techniques

Table 1 presents a comparison of traditional feature engineering techniques used for plant disease detection. **Color features**, such as RGB, HSV, and LAB histograms, are simple and interpretable but are sensitive to variations in lighting conditions. **Texture-based methods**, including GLCM, LBP, and Gabor filters, effectively capture lesion patterns and achieve relatively high accuracy, although they are computationally intensive. **Shape descriptors**, such as HOG and Hu Moments, provide transformation-invariant representations but require accurate leaf segmentation. The **combined approach**, which integrates multiple feature types with classifiers such as SVM or Random Forest, demonstrates the highest classification accuracy by exploiting complementary information, despite increased feature selection complexity. Overall, combining diverse handcrafted features generally yields better performance than relying on a single feature type.

3. Deep Learning-Based Feature Extraction

The limitations of handcrafted features—their sensitivity to environmental variation, the need for expert domain knowledge in their design, and their inability to capture complex hierarchical patterns—have been largely overcome by convolutional neural networks (CNNs). CNNs automatically learn feature representations directly from raw pixel data through hierarchical layers of convolution, activation, and pooling operations [8].

3.1 CNN Architectures for Plant Disease Detection

Several CNN architectures have been adapted or purpose-built for plant disease classification:

VGGNet (VGG16/VGG19): These architectures, with their uniform 3x3 convolution stacks, were among the first deep networks applied to plant disease detection. Despite being relatively large models, they remain popular baselines due to their simplicity and strong feature extraction capabilities [9].

ResNet (ResNet50, ResNet152): Residual connections in ResNet enable training of very deep networks by mitigating the vanishing gradient problem. ResNet152 has demonstrated high performance in corn leaf disease diagnosis, with GradCAM visualization confirming that learned features align with agronomically meaningful lesion regions [10].

EfficientNet (B0-B7): EfficientNet achieves state-of-the-art accuracy with fewer parameters through compound scaling of depth, width, and resolution. EfficientNetB5, in particular, has achieved 99.07% accuracy on tomato leaf disease classification from the PlantVillage dataset [11].

DenseNet: Dense connectivity between layers promotes feature reuse and gradient flow, making DenseNet particularly effective for distinguishing visually similar disease symptoms. DenseNet121 has been combined with attention mechanisms for cassava and groundnut disease detection [12].

MobileNet (V2, V3): Depthwise separable convolutions make MobileNet significantly lighter than standard CNNs, enabling deployment on smartphones and edge devices with constrained computational resources. MobileNetV3 combined with logistic regression for binary healthy/diseased classification reduced inference time by 77.6% compared to standard deep models [13].

3.2 Transfer Learning

A major practical challenge in plant disease detection is the scarcity of labeled image data for specific crops and diseases. Transfer learning addresses this by initializing model weights from large-scale pretraining on ImageNet and fine-tuning on plant disease datasets. This approach consistently outperforms training from scratch, particularly for small datasets [14].

Recent work has shown that EfficientNet-based transfer learning outperformed VGG16, VGG19, ResNet, InceptionV3, and MobileNet on tomato leaf disease classification, demonstrating both the effectiveness of the architecture and the transfer learning paradigm [15]. Hybrid models combining transfer learning with ensemble strategies have achieved accuracies exceeding 99% on controlled datasets, though real-world performance remains more challenging.

3.3 Attention Mechanisms

Attention mechanisms extend CNNs by enabling models to focus selectively on the most informative spatial regions. The Convolutional Block Attention Module (CBAM), integrated into VGG16 (CBAM-VGG16), enhances disease localization accuracy and interpretability by computing both channel and spatial attention maps. This architecture achieved up to 98.87% accuracy across five plant disease datasets [16]. The Squeeze-and-Excitation (SE) block, another popular attention mechanism, recalibrates channel-wise feature responses to emphasize disease-relevant feature maps.

4. GradCAM and Explainable AI for Plant Disease Detection

The opacity of deep learning models—often characterized as “black boxes”—poses a fundamental challenge for agricultural AI deployment. Farmers and agronomists require not only accurate predictions but also visual confirmation that the model is detecting genuine disease symptoms rather than spurious artifacts. Explainable AI (XAI) techniques address this by providing human-interpretable insights into model decisions [17].

4.1 GradCAM: Mechanism and Application

Gradient-weighted Class Activation Mapping (GradCAM), proposed by Selvaraju et al. in 2017, generates a coarse localization heatmap highlighting image regions important for predicting a concept. The technique computes the gradient of the class score with respect to the final convolutional feature maps, then performs a weighted combination of those maps to produce the saliency visualization. The result is an overlay on the input image that highlights which leaf regions drove the classification decision.

In plant disease applications, GradCAM has consistently demonstrated that well-trained models focus on actual lesions, discolored regions, and characteristic disease patterns—validating that models learn clinically meaningful features rather than image artifacts. A study using ResNet and EfficientNet on apple disease detection showed that GradCAM significantly outperformed standard Class Activation Mapping (CAM) in localizing disease regions, particularly for multiple co-occurring lesion instances [18].

4.2 GradCAM++ and Advanced Variants

GradCAM++ extends GradCAM by using higher-order gradients to weight the feature maps, providing better localization for multiple instances of the same class in a single image. In a comparative study of plant disease XAI methods, GradCAM++ produced more precise explanations than GradCAM, particularly for images containing multiple lesions or partially diseased leaves [19]. This is especially relevant for real-world field images where diseases may manifest at different severities across the same leaf.

4.3 LIME and SHAP as Complementary XAI Methods

Local Interpretable Model-Agnostic Explanations (LIME) generates explanations by approximating the local decision boundary with an interpretable surrogate model around a specific prediction. In plant disease detection, LIME produces segmented image explanations showing which leaf regions contributed positively or negatively to the disease classification [20].

SHapley Additive exPlanations (SHAP), grounded in cooperative game theory, provides theoretically consistent feature attributions. SHAP has been applied to mulberry leaf disease classification using a CNN-ViT (Vision Transformer) hybrid model, explaining model decisions in terms of both local pixel contributions **Method**	**Type**	**Mechanism**	**Strengths**	**Limitations**
GradCAM	Gradient-based	Weighted sum of gradient-activated feature maps	Fast, spatially precise for CNNs	Limited to single instance per image
GradCAM++	Gradient-based	Higher-order gradient weighting	Better multi-instance localization	Slightly higher computational cost
LIME	Model-agnostic	Local surrogate model approximation	Works for any model type	Larger segments, less precise localization
SHAP	Model-agnostic	Shapley values from game theory	Theoretically consistent global+local	Computationally expensive for deep models

SHapley Additive exPlanations (SHAP), grounded in cooperative game theory, provides theoretically consistent feature attributions. SHAP has been applied to mulberry leaf disease classification using a CNN-ViT (Vision Transformer) hybrid model, explaining model decisions in terms of both local pixel contributions **Method**	**Type**	**Mechanism**	**Strengths**	**Limitations**
Saliency Maps	Gradient-based	First-order input gradients	Simple, fast	Noisy, less interpretable

Table 2: Comparison of Explainable AI methods applied to plant disease detection.

Dataset	**Images**	**Classes**	**Conditions**	**Key Characteristics**
PlantVillage	54,305	38	Controlled	Most widely used; uniform background; high accuracy baseline
PlantDoc	2,598	17	Real-world	Field conditions; variable lighting; harder generalization benchmark

Dataset	Images	Classes	Conditions	Key Characteristics
FieldPlant	~5,000+	Multiple	Field	Real agricultural environments; high intra-class variation
Plant Disease Expert	199,644	58	Controlled	Large-scale; evaluates model capacity

Table 3: Summary of major benchmark datasets for plant disease detection.

and global feature importance across the dataset [21].

LIME, GradCAM, and GradCAM++ have been directly compared for plant disease detection. GradCAM and GradCAM++ provide more spatially precise explanations for CNN-based models, while LIME, though more general, tends to produce larger, less localized segments [22].

4.4 Comparison of XAI Techniques

Table 2 compares various Explainable AI (XAI) methods used in plant disease detection based on their mechanisms, strengths, and limitations. Gradient-based techniques such as **Grad-CAM** and **Grad-CAM++** provide effective visual explanations for CNN models, with Grad-CAM++ offering improved localization. **LIME** and **SHAP** are model-agnostic approaches that can explain predictions from different models but require higher computational resources. **Saliency Maps** are simple and computationally efficient but may generate noisy and less interpretable explanations. The comparison indicates that each method involves a trade-off between interpretability, computational complexity, and localization performance.

5. Benchmark Datasets for Plant Disease Detection

The availability of large, labeled image datasets has been instrumental in the rapid progress of plant disease detection models. However, the characteristics of available datasets significantly influence model performance, generalization, and the realism of evaluation.

5.1 PlantVillage Dataset

The PlantVillage dataset, created at Penn State University, contains over 54,305 leaf images across 38 disease classes spanning 14 crop species. It remains the most widely used benchmark for plant disease classification research. Images were captured under controlled laboratory conditions with uniform backgrounds, which enables very high classification accuracies (often exceeding 99%) but limits direct applicability to real field conditions [23].

5.2 PlantDoc Dataset

PlantDoc was specifically designed to address the real-world limitations of PlantVillage. It contains 2,598 images across 13 plant species and 17 disease categories, collected under natural field conditions with variable lighting, backgrounds, and disease severity levels. Models tested on PlantDoc consistently show lower accuracy than on PlantVillage, highlighting the generalization gap between controlled and field conditions [24].

5.3 Other Notable Datasets

Several other datasets have emerged to address specific crops and conditions:

FieldPlant: A dataset of field plant images collected under real agricultural conditions, covering multiple disease categories with high intra-class variation [25].

Plant Disease Expert: A large-scale dataset with 199,644 images across 58 classes, enabling evaluation of model capacity and generalization at scale.

Crop-specific Datasets: Rice, wheat, maize, potato, grape, and mango disease datasets from various regional agricultural institutions provide crop-specific benchmarks relevant to local food security priorities.

Table 3 summarizes the major benchmark datasets used for plant disease detection. **PlantVillage** and **Plant Disease Expert** are large, controlled datasets widely used for model development and evaluation, while **PlantDoc** and **FieldPlant** provide real-world field images with varying environmental conditions, making them more challenging and suitable for assessing model generalization and robustness.

6. Recent Deep Learning Architectures and Results

The period 2020-2025 has seen rapid architectural innovation in plant disease detection. Beyond classical CNNs, researchers have explored hybrid architectures, transformer-based models, and multi-scale networks that push both accuracy and interpretability.

6.1 Hybrid CNN-Attention Architectures

A depthwise CNN with Squeeze-and-Excitation (SE) integration and residual skip connections was proposed for enhanced plant disease detection, achieving superior accuracy through multi-scale feature aggregation while maintaining low parameter counts suitable for resource-constrained deployment [26]. The CBAM-VGG16 architecture, which integrates attention modules into each convolutional stage, improved disease localization with GradCAM visualization confirming attention focused on lesion boundaries [27].

6.2 Transformer-Based Models

Vision Transformers (ViT) and their hybrid CNN variants have begun to challenge CNN dominance in plant disease detection. The TCLeaf-Net model, a transformer-convolution framework with global-local attention, demonstrated robust in-field lesion-level detection by combining CNN's local feature extraction with transformer's long-range dependency modeling [28]. A CNN-ViT hybrid for mulberry leaf disease classification achieved 95.05% accuracy and employed SHAP for XAI explanation, demonstrating that transformers can achieve comparable performance to CNNs with superior interpretability support [29].

6.3 Multi-Scale Feature Extraction

The Multi-Scale FeatureNet with Coordinate Attention Module (MSFNet-CAM) was proposed for cassava and groundnut leaf disease detection. By combining multi-scale features with coordinate attention and ROI-based segmentation, this architecture addressed the challenge of detecting diseases at varying spatial scales—from fine-grained early lesions to extensive late-stage infections [30]. The InsightNet framework similarly employed multi-scale convolutional features with GradCAM visualization, providing both high accuracy and transparent reasoning for sustainable agriculture applications [31].

6.4 Lightweight Models for Edge Deployment

Practical deployment on smartphones and agricultural IoT devices demands computationally efficient models. The Mob-Res architecture, combining MobileNetV2 feature extraction with residual learning, achieved 99.47% accuracy on PlantVillage with only 3.51 million parameters—demonstrating that lightweight models can match

Model	Dataset	XAI Method	Accuracy	Key Contribution
EfficientNetB5 (BotanicX-AI)	PlantVillage (Tomato)	GradCAM + LIME	99.07%	9-class tomato disease XAI
ResNet152 +	Corn	GradCAM	98.5%+	Corn-specific

Model	Dataset	XAI Method	Accuracy	Key Contribution
GradCAM	PlantVillage			interpretable diagnosis
Transfer Learning (Potato)	Potato Leaf Dataset	GradCAM++	99.55%	GradCAM++ for multi-lesion localization
CNN-ViT + SHAP	Mulberry Leaf	SHAP	95.05%	Transformer with game-theory XAI
Mob-Res (Lightweight)	PlantVillage (38 classes)	Grad-CAM	99.47%	3.51M params, mobile-ready
CBAM-VGG16	5 Disease Datasets	GradCAM	98.87%	Attention-enhanced localization

Table 4: Summary of recent deep learning models with XAI applied to plant disease detection (2022-2025).heavier architectures when carefully designed [32]. The LeafAI framework employs a two-stage hybrid approach: a logistic regression binary classifier first filters healthy leaves, followed by a MobileNetV3 disease classifier, reducing total inference time by 77.6% with only 3% accuracy loss compared to end-to-end deep learning [33].

Table 4 summarizes recent deep learning–based plant disease detection models integrated with Explainable AI (XAI) techniques from 2022–2025. The reviewed studies employ methods such as Grad-CAM, LIME, and SHAP to improve model interpretability while maintaining high classification accuracy (95–99.5%). These approaches enable visualization of disease-affected regions, support reliable diagnosis across different crops, and enhance the practical deployment of AI systems in precision agriculture.

7. Data Challenges and Augmentation Strategies

Data quality and quantity fundamentally constrain the performance and generalizability of plant disease detection models. Several interconnected challenges must be addressed in any practical deployment.

7.1 The Controlled-to-Field Gap

The most significant challenge in plant disease detection is the discrepancy between model performance on controlled benchmarks (e.g., PlantVillage) and real field conditions. In laboratory settings, leaves are photographed against uniform backgrounds with consistent lighting. In the field, images contain soil, multiple leaves, varying illumination, occlusion, and mixed disease stages. Models achieving 99%+ accuracy on PlantVillage often fall to 70-80% accuracy on real-world collections, underscoring the need for field-specific training data and robust feature engineering [34].

7.2 Class Imbalance

Plant disease datasets commonly suffer from severe class imbalance, where common diseases are heavily overrepresented relative to rare or emerging ones. Standard accuracy metrics can be misleading in such settings. Recent work has addressed this through weighted loss functions (focal loss, unified focal loss), oversampling of minority disease classes, and ensemble methods combining MobileNetV3-Small and EfficientNetV2B3 with weighted voting, achieving over 94% accuracy on imbalanced datasets [35].

7.3 Data Augmentation Techniques

Data augmentation artificially expands training datasets by applying label-preserving transformations. Standard augmentation in plant disease research includes:

Geometric Transformations: Random flipping, rotation, scaling, and cropping to improve spatial invariance.

Photometric Augmentation: Brightness, contrast, saturation, and hue perturbations that simulate diverse field lighting conditions.

Advanced Augmentations: MixUp, CutMix, and Mosaic augmentation combine multiple training samples to improve model robustness and reduce overfitting.

Generative Adversarial Networks (GAN): GANs have been used to synthesize realistic disease images for minority classes, boosting diversity beyond traditional augmentation. DCGAN-based synthetic sample generation for class-imbalanced plant disease datasets has shown measurable accuracy improvements [36].

7.4 Federated Learning for Privacy-Preserving Data Sharing

An emerging approach to data scarcity is federated learning, which enables collaborative model training across multiple farms or research institutions without centralizing sensitive agricultural data. Models are trained locally and only gradient updates (not raw images) are shared, preserving data privacy while benefiting from distributed data diversity. Recent work has explored federated learning specifically for crop disease detection in distributed agricultural settings [37].

8. Deployment, MLOps, and Precision Agriculture Integration

Translating research models into operational agricultural tools requires careful attention to deployment infrastructure, model lifecycle management, and integration with broader precision agriculture systems.

8.1 Mobile and Edge Deployment

Farmers in low-connectivity regions require offline-capable smartphone applications. Several frameworks have been developed for deploying plant disease models on mobile devices, leveraging TensorFlow Lite, PyTorch Mobile, and ONNX runtime for model quantization and compression. GradCAM visualizations have been integrated into mobile apps to provide farmers with heatmap overlays that visually indicate which leaf regions triggered disease alerts, enhancing trust and adoption [38].

8.2 Drone and UAV Integration

Unmanned aerial vehicles (UAVs) equipped with multispectral cameras enable large-scale field monitoring. Feature engineering for drone imagery requires adaptations for aerial perspective, lower image resolution, and spectral bands beyond visible light (near-infrared, red edge) that are diagnostic for plant stress. YOLO-based object detection models have been adapted for drone imagery to localize disease outbreaks within field maps, enabling targeted interventions [39].

8.3 MLOps Practices for Plant Disease Models

Machine Learning Operations (MLOps) provides the operational framework for maintaining model quality in production agricultural systems. Key MLOps practices for plant disease detection include: continuous model retraining as new disease strains and seasonal variations are observed; automated data pipeline management for image collection, preprocessing, and augmentation; model versioning and performance monitoring with drift detection; and A/B testing of new architectures before field deployment. These practices ensure that models remain accurate and reliable across growing seasons and geographic regions [40].

9. Emerging Trends and Future Directions

9.1 Foundation Models and Large Vision Models

Large pretrained vision models, including CLIP and vision-language models, are beginning to be explored for plant disease detection. These models, trained on billions of image-text pairs, offer exceptional zero-shot and few-shot capabilities that could address the persistent challenge of limited labeled data for rare or emerging disease strains.

9.2 Concept-Based Explainability

Beyond pixel-level saliency maps, concept-based XAI approaches such as TCAV (Testing with Concept Activation Vectors) aim to explain model decisions in terms of human-interpretable agronomic concepts (e.g., ‘yellowing’, ‘necrotic spots’, ‘water-soaked lesions’). Automated concept identification for plant disease classifiers has demonstrated that models learn representations that map closely to agronomist-defined disease descriptors [41].

9.3 Multi-Modal Feature Integration

Future plant disease detection systems will likely integrate multiple data modalities: leaf images, weather data, soil sensor readings, GPS location, and temporal crop history. Multi-modal feature engineering combining visual features from CNNs with environmental time-series features and agronomic metadata holds promise for earlier detection and more accurate disease risk prediction [42, 43].

9.4 Self-Supervised and Semi-Supervised Learning

Given the high cost of expert disease labeling, self-supervised pretraining on large collections of unlabeled plant images, followed by fine-tuning with limited labeled examples, represents a promising direction. Contrastive learning approaches like SimCLR and MoCo have shown strong transfer performance and may significantly reduce the labeled data requirements for new crop-disease combinations.

10. Conclusion

This survey has presented a comprehensive review of feature engineering and explainable AI approaches for plant disease detection, tracing the evolution from handcrafted color, texture, and shape features to automatically learned deep representations via CNNs and transformers. The integration of GradCAM, GradCAM++, LIME, and SHAP into plant disease detection pipelines represents a critical advance toward trustworthy agricultural AI—enabling not only accurate predictions but also human-verifiable reasoning grounded in visible disease symptoms.

Key findings from this review include: (1) EfficientNet and ResNet-based transfer learning consistently achieves 99%+ accuracy on controlled benchmarks, with GradCAM confirming alignment between model attention and actual lesion regions; (2) the controlled-to-field generalization gap remains the most pressing practical challenge, motivating investment in diverse real-world datasets such as PlantDoc and FieldPlant; (3) lightweight architectures such as MobileNet and Mob-Res have achieved near-parity with heavier models at dramatically reduced computational cost, enabling genuine mobile and edge deployment; and (4) transformer-based and hybrid CNN-ViT models are emerging as strong competitors to pure CNNs, particularly for complex multi-disease scenarios.

Future research should prioritize: development of large-scale, geographically diverse real-world disease datasets; deeper integration of XAI into farmer-facing applications with actionable natural language explanations; multi-modal models combining imagery with environmental and temporal data; and rigorous MLOps frameworks ensuring production model reliability across growing seasons.

Ultimately, the convergence of robust feature engineering, deep learning, and explainable AI offers a transformative path toward intelligent, trustworthy, and accessible plant disease management—contributing directly to global food security and the sustainability of agricultural systems.

References

[1] FAO, "The State of Food and Agriculture 2021," Food and Agriculture Organization of the United Nations, Rome, 2021.

[2] V. S. Abbigeri and G. D. Devanagavi, "Leaf Disease Detection and Pesticide Recommendation Using Pretrained CNN Models in Keras on the Augmented New Plant Diseases Dataset", *Eng. Technol. Appl. Sci. Res.*, vol. 16, no. 2, pp. 33705–33711, Apr. 2026.

[3] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," International Journal of Computer Vision, vol. 128, pp. 336-359, 2020. doi: 10.1007/s11263-019-01228-7.

[4] S. Ashurov et al., "Enhancing plant disease detection through deep learning: a Depthwise CNN with squeeze and excitation integration and residual skip connections," Frontiers in Plant Science, 2025. doi: 10.3389/fpls.2024.1505857.

[5] S. Natarajan, P. Chakrabarti, and M. Margala, "Robust diagnosis and meta visualizations of plant diseases through deep neural architecture with explainable AI," Scientific Reports, vol. 14, 2024. doi: 10.1038/s41598-024-64601-8.

[6] M. S. Hossain et al., "Color and texture-based plant leaf disease detection using KNN classifier," IEEE Access, 2019.

[7] L. Qiu, V. M. Chinchilli, and L. Lin, "Understanding deep representation learning for multi-view temporal data," arXiv preprint, 2020. doi: 10.48550/arXiv.2005.05485.

[8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, pp. 436-444, 2015. doi: 10.1038/nature14539.

[9] T.-H. Nguyen, T.-N. Nguyen, and B.-V. Ngo, "A VGG-19 Model with Transfer Learning and Image Segmentation for Classification of Tomato Leaf Disease," AgriEngineering, vol. 4, no. 4, pp. 871-887, 2022.

[10] K. Gopalan et al., "Corn leaf disease diagnosis: enhancing accuracy with ResNet152 and Grad-CAM for explainable AI," BMC Plant Biology, vol. 25, p. 440, 2025. doi: 10.1186/s12870-025-06386-0.

[11] M. Bhandari, T. B. Shahi, A. Neupane, and K. B. Walsh, "BotanicX-AI: Identification of Tomato Leaf Diseases Using an Explanation-Driven Deep-Learning Model," Journal of Imaging, vol. 9, no. 2, p. 53, 2023. doi: 10.3390/jimaging9020053.

[12] N. A. Samee et al., "High-performance parallel multi-scale attention network with explainable AI for intelligent diagnosis of leaf diseases," Scientific Reports, vol. 15, p. 42060, 2025. doi: 10.1038/s41598-025-26144-4.

[13] LeafAI Study, "Interpretable plant disease detection for edge computing: a hybrid AI approach with Grad-CAM," PMC, 2025.

[14] M. Saberi Anari, "A Hybrid Model for Leaf Diseases Classification Based on the Modified Deep Transfer Learning and Ensemble Approach for Agricultural AIoT-Based

Monitoring," Computational Intelligence and Neuroscience, 2022. doi: 10.1155/2022/6504616.

[15] S. A. Hamim and A. I. Jony, "Leaf-Based Plant Disease Detection and Explainable AI," arXiv preprint, 2024. arXiv:2404.16833.

[16] S. Kinger and V. Kulkarni, "Explainable AI for Deep Learning Based Disease Detection," in CBAM-VGG16 Study, 2022.

[17] A. Chatzimparmpas et al., "Building Trust in ML Models Through Visualization Techniques," Computer Graphics Forum, vol. 39, no. 3, pp. 713-756, 2020.

[18] Plant Disease Detection and Localization using GRADCAM, Academia.edu, 2022.

[19] S. A. Hamim et al., "GradCAM++ comparison for plant disease detection," arXiv:2404.16833, 2024.

[20] S. M. Alhammad et al., "Deep learning and explainable AI for classification of potato leaf diseases," Frontiers in Artificial Intelligence, vol. 7, 2024. doi: 10.3389/frai.2024.1449329.

[21] M. A. Hasan et al., "Mulberry leaf disease detection by CNN-ViT with XAI integration," PLoS One, vol. 20, no. 6, 2025.

[22] A. Shrotriya, A. K. Bairwa, and A. K. Sharma, "Automated Plant Disease Detection Using CNNs and YOLOv5: A Comprehensive Approach with Grad-CAM and VGG16," in Proc. MoSICom, IEEE, 2024.

[23] D. P. Hughes and M. Salathé, "An open access repository of images on plant health to enable the development of mobile disease diagnostics," arXiv:1511.08060, 2015.

[24] D. Singh et al., "PlantDoc: A dataset for visual plant disease detection," in Proc. ACM IKDD CoDS, 2020.

[25] E. Moupojou et al., "FieldPlant: A dataset of field plant images for plant disease detection and classification with deep learning," IEEE Access, vol. 11, pp. 35398-35410, 2023.

[26] A. Y. Ashurov et al., "Enhancing plant disease detection through deep learning: Depthwise CNN with SE integration," Frontiers in Plant Science, 2025.

[27] S. Kinger and V. Kulkarni, "Interpretable Plant Leaf Disease Detection Using Attention-Enhanced CNN," 2022.

[28] TCLeaf-Net, "A transformer-convolution framework with global-local attention for robust in-field lesion-level plant leaf disease detection," arXiv:2512.12357, 2024.

[29] M. A. Hasan et al., "Mulberry leaf disease detection by CNN-ViT with XAI (SHAP) integration," PLoS One, 2025.

[30] N. A. Samee et al., "High-performance parallel MSFNet-CAM for cassava and groundnut disease detection," Scientific Reports, 2025.

[31] Tamim et al., "InsightNet: A Deep Learning Framework for Enhanced Plant Disease Detection and Explainable Insights," Plant Direct, 2025. doi: 10.1002/pld3.70076.

[32] Mob-Res Study, "A lightweight and explainable CNN model for empowering plant disease diagnosis," Scientific Reports, 2025. doi: 10.1038/s41598-025-94083-1.

[33] LeafAI, "Interpretable plant disease detection for edge computing," PMC, 2025.

[34] J. G. Barbedo, “Deep learning applied to plant pathology: the problem of data representativeness,” Tropical Plant Pathology, vol. 47, pp. 85-94, 2022.

[35] Machine Learning and Deep Learning Image-Based Plant Disease Classification Survey, ScienceDirect, 2024.

[36] M. Saberi Anari, “A Hybrid Model for Leaf Diseases Classification Using DCGAN and Transfer Learning,” 2022.

[37] Federated learning for crop disease detection study, 2023.

[38] Y. Kaya and E. Gürsoy, “A review of deep learning architectures for plant disease detection,” Turkish Journal of Biology, 2025.

[39] Y. Alhwaiti et al., “Leveraging YOLO deep learning models to enhance plant disease identification,” Scientific Reports, 2025.

[40] D. Kreuzberger et al., “Machine Learning Operations (MLOps): Concepts and Architecture,” arXiv:2205.02302, 2022.

[41] Explainability of Deep Learning-Based Plant Disease Classifiers Through Automated Concept Identification, arXiv:2412.07408, 2024.

[42] V. S. Abbigeri and G. D. Devanagavi, "Feature Engineering in Machine Learning: Current Trends, Challenges, and Best Practices," Journal of Engineering and Technology Management, vol. 75, pp. 561–572, Jan. 2025.

[43] Y. Roh, G. Heo, and S. E. Whang, “Review on Data Collection Strategies for ML and Big Data,” arXiv:1811.03402, 2018.