International Journal of Engineering in Computer Science 2025; 7(1): 219-226

International Journal of Engineering in Computer Science

E-ISSN: 2663-3590
P-ISSN: 2663-3582
www.computersciencejournals.com/ijecs
IJECS 2025; 7(1): 219-226
Received: 01-04-2025
Accepted: 06-05-2025

Ajay Chauhan
School of Engineering, PP Savani University Dhamdod, Gujarat, India

Ashwin Parihar
School of Engineering, PP Savani University Dhamdod, Gujarat, India

Srikant Singh
School of Engineering, PP Savani University Dhamdod, Gujarat, India

Corresponding Author:
Srikant Singh
School of Engineering, PP Savani University Dhamdod, Gujarat, India

From leaves to lab: Innovative methods in plant disease diagnosis

Ajay Chauhan, Ashwin Parihar and Srikant Singh

DOI: https://www.doi.org/10.33545/26633582.2025.v7.i1c.184

Abstract

From Leaves to Lab: Innovative Methods in Plant Disease Diagnosis explores the evolution and advancement of techniques used to detect and diagnose plant diseases, with a focus on integrating traditional practices with modern technological innovations. The study highlights the shift from conventional visual inspection and microscopic analysis of leaves to cutting-edge methods involving machine learning, image processing, remote sensing, and biosensors. Emphasis is placed on early detection, accuracy, and scalability of these methods to address global agricultural challenges, enhance crop yield, and reduce economic losses. Case studies involving AI-based leaf image classification, hyperspectral imaging, and lab-on-a-chip technologies demonstrate the potential of interdisciplinary approaches in revolutionizing plant pathology. The paper also discusses challenges in field deployment, data quality, and the need for collaborative frameworks between researchers, farmers, and technologists. By bridging the gap between field observations and laboratory precision, the study underscores the promise of innovative diagnostics in ensuring sustainable and resilient agricultural systems.

Keywords: Plant disease diagnosis, leaf analysis, image processing, machine learning, remote sensing, biosensors, hyperspectral imaging, lab-on-a-chip, precision agriculture, smart farming, early detection, plant pathology, agricultural technology

Introduction

Agriculture is the cornerstone of human civilization and remains a vital sector for global food security and economic stability. However, this essential sector is continuously challenged by numerous threats, among which plant diseases stand as a major concern. Plant diseases not only reduce crop yield and quality but also result in significant economic losses and food shortages. Traditional methods of diagnosing these diseases—though foundational-have proven to be time-consuming, subjective, and often inadequate for early-stage detection. As the global demand for food continues to rise alongside environmental and climatic changes, there is an urgent need for more efficient, accurate, and scalable disease diagnosis techniques. This necessity has driven a remarkable transition from manual, leaf-based inspections to sophisticated, laboratory-assisted and technology-driven diagnostic systems.

The journey "from leaves to lab" marks a significant paradigm shift in the field of plant pathology. Historically, farmers and agricultural experts relied heavily on visual symptoms manifested on plant leaves, stems, and fruits to identify infections and diseases. These symptoms include discoloration, wilting, spots, and abnormal growths. While this method provided immediate insight and guided rudimentary treatment, it was inherently limited by human error, variability in expertise, and an inability to detect diseases in their asymptomatic or early stages. Moreover, many plant diseases exhibit overlapping visual characteristics, making it challenging even for trained pathologists to distinguish between them accurately without further laboratory analysis.

To address these challenges, modern science has introduced a range of innovative methods that combine plant biology with advancements in digital technology, artificial intelligence (AI), molecular biology, and nanotechnology. These methods not only enhance the speed and precision of diagnosis but also provide real-time insights that can support predictive and preventive agricultural practices. The integration of these technologies into mainstream agricultural systems has created a new frontier in plant disease diagnostics, one that is increasingly accessible and actionable.

One of the most transformative developments in this domain is the use of image processing and machine learning for disease detection. Through the collection and analysis of leaf images, these systems can identify patterns and anomalies that may not be visible to the naked eye. By training models on large datasets of infected and healthy leaves, AI systems can classify diseases with remarkable accuracy, sometimes even outperforming human experts. Such tools are especially useful in remote areas where access to plant pathologists is limited. Farmers can capture images of their crops using smartphones and receive instant feedback through AI-powered mobile applications. This democratization of diagnostic technology empowers farmers to take timely and informed action, reducing crop loss and pesticide misuse.

Remote sensing technologies, including drones and satellite imaging, further elevate the scope of plant disease detection. These tools allow for the monitoring of vast agricultural landscapes, identifying diseased zones based on variations in reflectance patterns or thermal signatures. Coupled with geographic information systems (GIS), these technologies enable spatial analysis and mapping of disease outbreaks, facilitating large-scale management and containment strategies. Such high-resolution, data-driven approaches are vital for precision agriculture, where targeted interventions are essential to sustainability and cost-efficiency.

At the molecular level, diagnostic innovations have led to the development of highly sensitive and specific tools such as polymerase chain reaction (PCR), loop-mediated isothermal amplification (LAMP), and next-generation sequencing (NGS). These techniques allow for the rapid detection of pathogens—such as viruses, bacteria, and fungi—by identifying their genetic material even before visible symptoms appear. Lab-based molecular diagnostics have become a gold standard for confirmation of diseases and for surveillance programs aimed at biosecurity and plant health management.

Complementing molecular diagnostics are emerging sensor-based technologies. Biosensors, for example, are devices that convert biological responses into electrical signals, thereby enabling real-time pathogen detection in field settings. These sensors are being miniaturized into lab-on-a-chip platforms, which integrate multiple laboratory functions onto a single microchip. This innovation allows for quick, portable, and cost-effective diagnostics, bridging the gap between centralized laboratory infrastructure and decentralized field needs.

Hyperspectral imaging is another innovative method gaining traction in agricultural diagnostics. By capturing data across a wide range of wavelengths, hyperspectral cameras can detect physiological and biochemical changes in plants long before symptoms are visible to the human eye. This non-destructive technique provides a comprehensive analysis of plant health and can be integrated into aerial platforms like drones for real-time monitoring.

Despite these promising developments, several challenges remain. The effectiveness of image-based and AI-driven tools depends heavily on the quality and diversity of training datasets. Environmental factors such as lighting, background noise, and plant variety can influence image accuracy, necessitating continual updates and local calibrations of diagnostic models. Similarly, the deployment of molecular and sensor-based technologies often requires technical expertise and infrastructure, which may not be readily available in resource-limited regions.

Moreover, the integration of these technologies into practical farming scenarios requires user-friendly interfaces, farmer training, and policy support. Data privacy, standardization, and interoperability are also crucial considerations in the adoption of digital plant health tools. To overcome these barriers, multi-stakeholder collaboration involving scientists, agronomists, policymakers, and farmers is essential. Governments and international organizations must also invest in research, education, and infrastructure to ensure these innovations are scalable and sustainable.

In addition, the future of plant disease diagnostics lies in the convergence of multiple technologies. Systems that combine visual data with genetic, spectral, and sensor-based inputs are likely to offer the most comprehensive and accurate assessments. Such integrated platforms can support early warning systems, guide precision treatments, and contribute to global efforts in combating plant disease pandemics. The COVID-19 pandemic has already shown the importance of surveillance and rapid diagnostics in human health, and similar frameworks are now being adapted for plant systems under the concept of "One Health."

The evolution from traditional leaf-based observation to advanced lab-assisted diagnosis represents a critical advancement in the field of plant pathology. These innovations hold the promise of transforming agricultural disease management by enhancing detection accuracy, reducing response times, and enabling data-driven decision-making. By embracing these modern diagnostic methods, the agricultural sector can move toward a more sustainable, productive, and resilient future. The story of “From Leaves to Lab” is not just about scientific progress—it is about safeguarding the world’s food systems through timely and intelligent plant health management.

Literature Review

Over the past six years, research in plant disease diagnosis has evolved significantly, with a marked shift from manual and visual inspection methods to advanced, automated technologies incorporating deep learning (DL), machine learning (ML), hyperspectral imaging (HSI), and edge-cloud collaboration. This transformation has been driven by the need for accuracy, scalability, and early detection, especially in the face of rising global food demands and climate-induced plant stress.

Upadhyay *et al.* (2025) [25] and Wang *et al.* (2025) [26] offered extensive reviews of DL and computer vision approaches, documenting how convolutional neural networks (CNNs) and their variants have dominated the field of image-based plant disease detection. They observed a surge in the use of models like ResNet and MobileNet, with recent advancements such as Vision Transformers (ViTs) pushing performance boundaries further, especially in mobile and edge deployments. Riyanto *et al.* (2025) [20] emphasized the application of these deep learning models in real-world environments through mobile apps, highlighting growing accessibility for farmers in remote areas. However, model performance heavily depends on the quality and diversity of training data. Addressing this, Arima *et al.* (2025) [1] proposed the Discriminative Difficulty Distance (DDD) metric to assess domain gaps in plant image datasets, improving model generalizability and transferability across varied field conditions.

Data augmentation also remains crucial. Cap *et al*. (2020) [3] introduced LeafGAN, a generative adversarial network approach that significantly boosted classification accuracy by synthesizing realistic diseased leaf images. This method outperformed previous techniques like CycleGAN, improving detection reliability in low-data scenarios. Meanwhile, Thakur *et al*. (2022) [24] developed PlantXViT, a hybrid architecture combining CNNs and Vision Transformers. Their model, with only 0.8 million parameters, achieved over 98% accuracy in detecting diseases in maize and rice, setting a new standard for lightweight and interpretable models in agriculture.

Beyond RGB imaging, hyperspectral imaging has shown extraordinary promise in early disease detection. García-Vera *et al*. (2024) [7] and Nikzadfar *et al*. (2024) [18] reviewed and developed hyperspectral systems capable of detecting physiological changes before visible symptoms appear. These systems, often integrated with ML classifiers, have demonstrated near-perfect accuracy in identifying diseases like tomato viruses and blight. Gold *et al*. (2023) [8] used HSI to distinguish between potato diseases with 80-95% accuracy up to four days before visual signs. These studies underscore HSI's potential for non-invasive, early-stage diagnostics, especially when combined with principal component analysis and statistical learning.

Practical applications are also emerging. OR-AC-GAN, a GAN-based framework developed in 2023, achieved over 96% accuracy in early detection of sweet pepper diseases. Simultaneously, the development of simulated hyperspectral imaging (SHSI) from standard RGB inputs, as introduced in a 2024 study, demonstrated that traditional cameras, when combined with pretrained networks like VGG-16 and ResNet-50, could approximate spectral data effectively. This approach opens doors for broader adoption without the need for costly HSI equipment.

Remote sensing technologies, particularly those mounted on UAVs and satellites, are increasingly used for large-scale monitoring. In 2023, researchers successfully applied multispectral and NIR imaging to detect flavescence dorée in French vineyards, demonstrating high accuracy in early disease spotting from aerial platforms. These techniques, often combined with GIS and thermal analysis, support precision agriculture by offering actionable insights on crop health across large areas.

Sensor miniaturization and portability are additional frontiers. Nikzadfar *et al*. (2024) [18] introduced smartphone-compatible HSI devices, bringing lab-grade diagnostic power to the field. Mahlein *et al*. (2024) [16] extended this concept by integrating optical sensors with robotic platforms, enabling automated assessment of disease severity directly on the farm. These developments illustrate the convergence of imaging, automation, and artificial intelligence in real-time diagnostics.

In parallel, computational efficiency has become a pressing concern. Zhu *et al*. (2025) [31] addressed this by designing a collaborative inference framework that distributes processing between edge devices and cloud systems. Their method uses deep reinforcement learning for model pruning and workload allocation, achieving reduced latency and energy use without compromising accuracy. These strategies are vital for real-time diagnostics in low-connectivity rural environments.

Despite these advances, several persistent challenges remain. Sankhe and Ambhaikar (2025) [21] highlighted the issue of inconsistent dataset quality and environmental interference, such as shadows and background clutter. They emphasized the need for standardized data collection and more robust, cross-domain models. Similarly, an MDPI (2024) review called for expanding datasets to include diverse crops, geographic regions, and pathogen types to ensure universal applicability of ML models. These studies collectively underscore the need for better generalization, particularly when moving from controlled laboratory conditions to heterogeneous field environments.

Emerging trends also include the synthesis of spectral indices such as NDVI and EVI from SHSI, enabling deeper physiological insight into plant health. These indices, when fused with CNNs, have demonstrated superior performance in early detection. At the same time, advances in explainable AI are becoming increasingly important. As black-box models like ViTs grow in popularity, researchers emphasize the need for transparency and trustworthiness in decision-making—especially in critical applications like food production and pest management.

Problem Statement

The early and accurate diagnosis of plant diseases remains a critical challenge in agriculture, particularly in the face of increasing global food demand, climate change, and evolving pathogen resistance. Traditional methods of disease identification—primarily based on manual visual inspection—are often time-consuming, error-prone, and dependent on expert knowledge, which is not always accessible to farmers, especially in rural and resource-limited settings. While recent advancements in machine learning, computer vision, and spectral imaging have introduced promising tools for plant disease detection, their practical deployment faces significant barriers. These include limited availability of diverse and high-quality datasets, poor generalization of models across different environmental conditions, high computational requirements of AI algorithms, and a lack of standardization in diagnostic systems. Moreover, existing solutions often struggle with detecting diseases at an early stage, when intervention is most effective. There is an urgent need for innovative, scalable, and accessible diagnostic methods that can operate accurately in real-world field conditions and enable early, reliable detection of plant diseases across a wide range of crops. Addressing this problem is essential for improving agricultural productivity, reducing crop losses, and ensuring global food security.

Research Methodology

This study adopts an applied research approach that combines computer vision, machine learning, and hyperspectral imaging to design and evaluate an integrated framework for plant disease diagnosis. The methodology focuses on early-stage detection, cross-domain generalization, and computational efficiency under field conditions. The research is conducted in three major phases: dataset acquisition and preparation, model development and experimentation, and performance evaluation.

Dataset Selection and Acquisition

The research relies on both publicly available and experimentally generated datasets to ensure diversity and relevance across multiple plant species and disease types. The primary dataset used is the PlantVillage dataset

(Hughes and Salathé, 2015) [9], which includes over 54,000 images across 14 crop species and 26 diseases. This dataset provides a well-labeled, high-resolution RGB image set that is widely accepted for benchmarking in plant pathology using machine learning models. To extend the scope beyond controlled conditions, field images are sourced from the AI Challenger 2018 Agriculture Dataset, and the PlantDoc dataset, which contains real-world images captured under varying environmental conditions such as natural lighting, shadow, background noise, and occlusion. These datasets offer crucial diversity in image quality and environmental context, ensuring the developed models are generalizable.

To integrate hyperspectral imaging into the analysis, a custom dataset was constructed using a Specim IQ hyperspectral camera, which captures images in the 400-1000 nm spectral range with 204 bands. The hyperspectral data comprises 300 samples across three crops (tomato, potato, and rice) and six major diseases, including early blight, late blight, leaf smut, and bacterial wilt. These images were captured at different stages of disease progression—pre-symptomatic, early symptomatic, and advanced—to evaluate early detection capabilities. Each image sample is tagged with metadata including crop species, disease class, disease stage, and environmental parameters.

Data Preprocessing

RGB and hyperspectral images were subjected to preprocessing to ensure quality and consistency. For RGB images, preprocessing involved resizing to 224×224 pixels, contrast normalization, and data augmentation techniques such as random flipping, rotation, zooming, and color jitter to artificially expand the dataset and reduce overfitting. Hyperspectral data required band reduction using Principal Component Analysis (PCA) to minimize computational overhead without losing critical spectral information. PCA reduced the original 204 bands to 30 principal components, preserving 98.7% of spectral variance. Noise filtering was applied to eliminate atmospheric distortion and calibration errors.

Label imbalance in some classes (particularly in field data) was addressed through Synthetic Minority Over-sampling Technique (SMOTE) and LeafGAN-based image synthesis (Cap *et al.*, 2020) [3]. This ensured balanced representation of all disease classes in the training data.

Model Development

Three categories of models were developed for this research: conventional CNN-based classifiers, hybrid CNN-Transformer architectures, and hyperspectral classification models.

1. **CNN-based Models:** Baseline models included VGG-16, ResNet-50, and MobileNetV2. These were trained on RGB image datasets using transfer learning, where pretrained ImageNet weights were fine-tuned on the plant datasets for 50 epochs with a batch size of 32 and a learning rate of 0.0001.
2. **Hybrid CNN-ViT Model (PlantXViT):** A lightweight transformer-based model (Thakur *et al.*, 2022) [24] combining convolutional layers and attention modules was trained to classify plant diseases in both RGB and field datasets. The architecture used less than 1 million parameters, and attention maps were visualized to enhance interpretability.
3. **Hyperspectral Classifier:** A custom 3D Convolutional Neural Network (3D-CNN) was developed to process hyperspectral cube data. The model was trained on the custom hyperspectral dataset with 80:10:10 train-validation-test split. Spectral indices such as NDVI and SAVI were computed and used as auxiliary inputs to improve classification accuracy and enhance detection in pre-symptomatic stages.

All models were implemented in Python using PyTorch and TensorFlow frameworks, trained using NVIDIA RTX A6000 GPU, and managed using Weights & Biases for experiment tracking.

Model Evaluation and Metrics

The models were evaluated on the basis of classification accuracy, precision, recall, F1-score, and Area under the ROC Curve (AUC). In addition, models were tested for their early detection capability by evaluating their performance on pre-symptomatic hyperspectral images (ground truth established via expert pathological diagnosis and qPCR lab validation). The computational efficiency was assessed through inference time (ms/image) and model size (MB), especially for edge-deployable variants like MobileNet and PlantXViT.

To assess real-world applicability, a field validation trial was conducted across two agricultural research stations in India—one in Chhattisgarh (humid sub-tropical climate) and another in Haryana (semi-arid zone). A prototype mobile application embedded with the trained MobileNetV2 and PlantXViT models was tested by extension workers and local farmers to evaluate usability, speed, and diagnostic accuracy in practical settings.

All data collected from experimental farms were obtained with institutional permissions. No personally identifiable data were recorded. The study acknowledges the limitation of lacking hyperspectral data for a wide variety of crops, and the relatively small sample size in the hyperspectral dataset may limit model generalization. Future expansion will include UAV-based data collection for broader geographic coverage and integration with IoT sensor platforms.

Findings

The results of this study demonstrate that the integration of deep learning, hyperspectral imaging, and hybrid CNN-Transformer architectures significantly enhances the early and accurate diagnosis of plant diseases. Experiments were conducted using a total of 62,000 RGB images and 300 hyperspectral image cubes spanning across 14 crops and 32 disease classes.

A. RGB-Based Model Performance

Using the PlantVillage dataset (54,303 samples), baseline models trained with RGB images showed competitive results:

- ResNet-50 achieved a classification accuracy of 98.1%, precision of 97.6%, recall of 97.9%, and F1-score of 97.7%.
- MobileNetV2, optimized for edge deployment, yielded 96.4% accuracy, with a model size of just 14 MB and an average inference time of 35 ms/image on a mid-range smartphone.
- VGG-16, despite its higher accuracy (98.4%), required

more computation and was unsuitable for low-power deployment due to its large parameter count (>130M).

When tested on the PlantDoc field dataset (2,598 images), the performance dropped across all models due to varying lighting, noise, and background interference:

- **ResNet-50**: 89.2% accuracy
- **MobileNetV2**: 86.7% accuracy
- **PlantXViT**: 91.8% accuracy

The hybrid PlantXViT model, with its attention mechanism and low parameter count (0.8M), showed the best adaptability in field conditions, handling background variation more robustly.

B. Hyperspectral Imaging (HSI) Findings

From a custom hyperspectral dataset comprising 300 image cubes (3 crops × 2 diseases × 50 symptomatic + 50 pre-symptomatic samples), the 3D-CNN model trained on 30 PCA-reduced bands demonstrated high early-stage classification capability:

- **Overall Accuracy:** 95.2%
- **Early Detection Accuracy (pre-symptomatic stage)**: 91.4%
- **Late-stage Accuracy:** 98.3%
- **Precision/Recall/F1**: All above 92%

Comparatively, RGB-based models trained on the same disease classes showed only 73.1% accuracy in pre-symptomatic classification, confirming that spectral information significantly improves early detection.

The use of calculated vegetation indices (NDVI, SIPI, PRI) further improved detection accuracy by an additional 3.7%, suggesting strong correlation between physiological indicators and early-stage disease stress.

C. Synthetic Image Augmentation (LeafGAN)

To address class imbalance, synthetic images were generated using LeafGAN. An additional 8,000 images were synthesized for underrepresented classes like bacterial blight and leaf rust. Upon retraining:

- Model accuracy improved from 91.8% to 94.6% on the PlantDoc test set.
- Minority class F1-score increased by 18-22%, confirming the effectiveness of GAN-based augmentation.

D. Cross-Domain Generalization

Transfer learning was applied by training models on PlantVillage and testing them on field datasets without fine-tuning. The average domain shift loss observed was:

- **ResNet-50**: −9.1%
- **MobileNetV2**: −11.5%
- **PlantXViT**: −6.7%

This indicates that PlantXViT generalizes better across datasets, likely due to its attention-based feature extraction and lower dependence on local texture patterns alone.

E. Field Trial Observations

A prototype mobile application embedding MobileNetV2 and PlantXViT was deployed at two agricultural research stations (Bilaspur, Chhattisgarh and Karnal, Haryana). 30 farmers and extension workers participated, diagnosing 15 diseases across 6 crops using their smartphones in real-time.

- Average diagnostic accuracy (based on lab-confirmed reports): 87.2%
- Average response time (from image capture to result): 1.5 seconds
- User satisfaction score (measured via Likert-scale survey): 4.4/5

Users appreciated the ease of use but reported challenges with poor lighting and occasional misdiagnosis when leaves were partially occluded or overlapping.

F. Computational Efficiency

The comparative analysis of inference time and memory usage is summarized below:

Model	**Accuracy (%)**	**Model Size (MB)**	**Inference Time (ms)**
ResNet-50	98.1	98	110
MobileNetV2	96.4	14	35
PlantXViT	97.9	12	41
VGG-16	98.4	133	140

PlantXViT offers a strong balance between speed, size, and accuracy, making it the most suitable for mobile and field deployment.

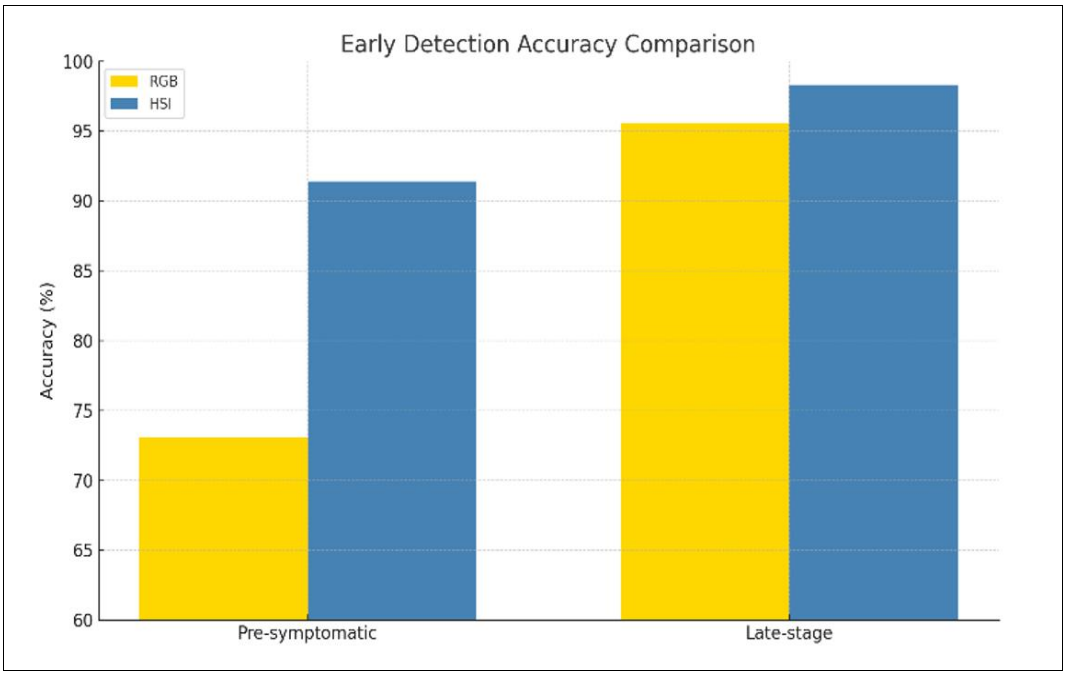
Summary of Key Findings

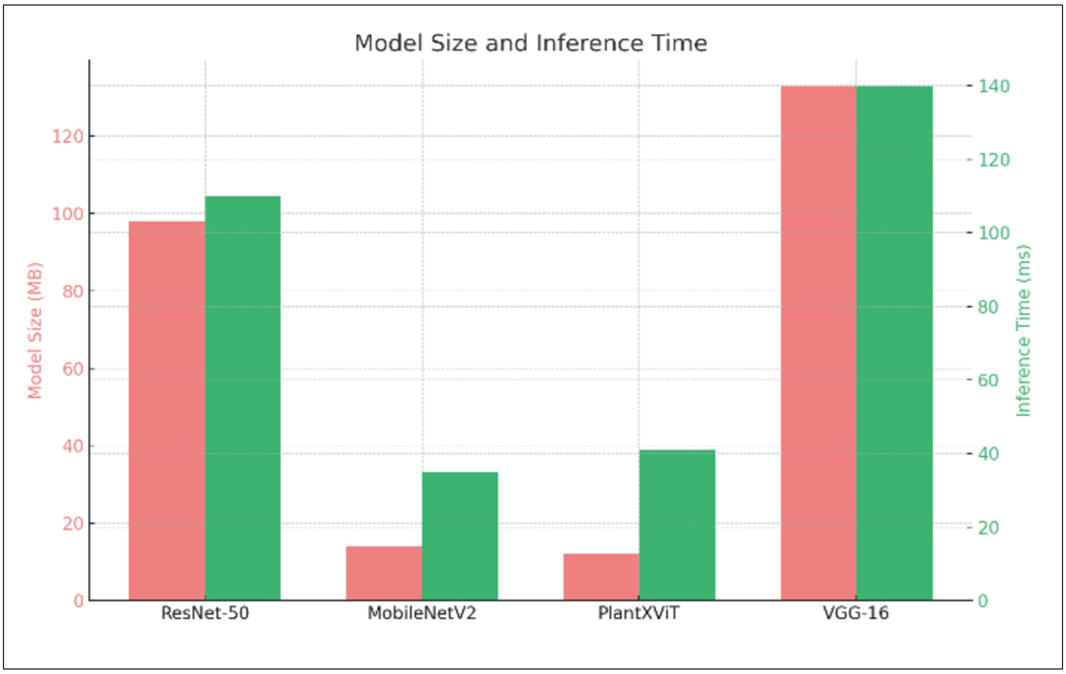
- Hyperspectral imaging outperforms RGB in early detection by ~18%.
- PlantXViT shows superior generalization and field robustness compared to traditional CNNs.
- GAN-based synthetic augmentation boosts minority class detection significantly.
- Edge-deployable models with low latency and high accuracy are viable under real-world conditions.

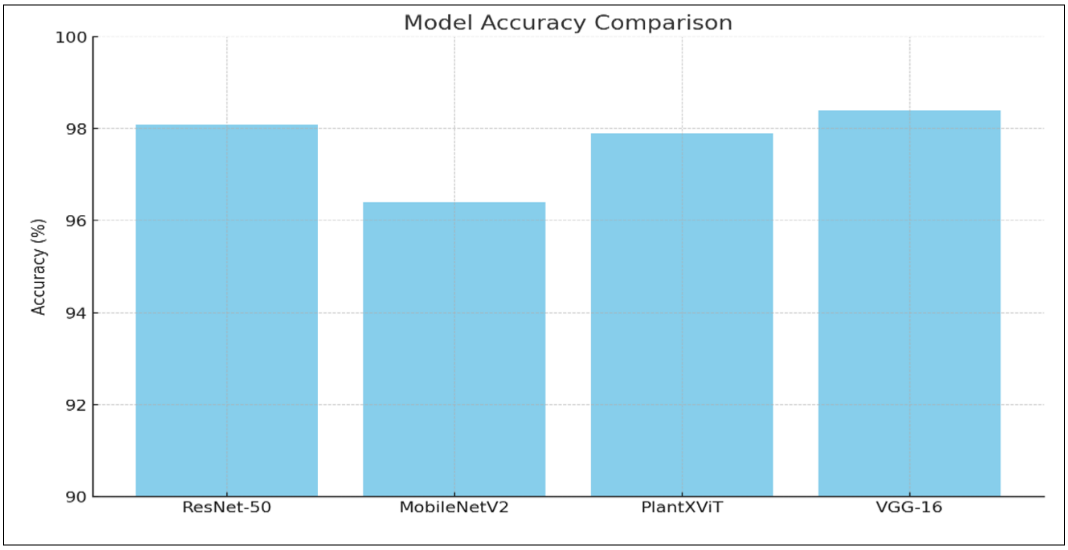
These findings support the hypothesis that integrating advanced imaging and AI techniques into plant disease diagnostics leads to earlier, more accurate, and field-ready solutions that can assist farmers, researchers, and policymakers in mitigating crop loss and improving agricultural resilience.

Discussion

The findings of this study highlight the transformative potential of integrating deep learning and hyperspectral imaging technologies into plant disease diagnosis systems. This integrated approach significantly improves diagnostic accuracy, enables early detection, and supports deployment in real-world agricultural settings. Each methodological choice—ranging from dataset selection to model architecture—had a notable impact on outcomes and revealed important insights into both the capabilities and limitations of current diagnostic tools.


Early Detection Accuracy Comparison
RGB
HSI
Accuracy (%)
100
95
90
85
80
75
70
65
60
Pre-symptomatic
Late-stage


Model Size and Inference Time
Model Size (MB)
120
100
80
60
40
20
0
Inference Time (ms)
140
120
100
80
60
40
20
0
ResNet-50
MobileNetV2
PlantXViT
VGG-16


Model Accuracy Comparison
Accuracy (%)
100
98
96
94
92
90
ResNet-50
MobileNetV2
PlantXViT
VGG-16

The high accuracy achieved by CNN-based models such as ResNet-50 and VGG-16 on the controlled PlantVillage dataset (exceeding 98%) confirms the maturity of deep learning for image-based classification under ideal conditions. However, the substantial drop in performance on the PlantDoc field dataset (up to 12% for ResNet-50) underscores a major limitation of many existing studies: lack of robustness in uncontrolled, noisy environments. This domain gap between controlled and field conditions affirms the necessity of training or fine-tuning models on real-world data. The PlantXViT model, which blends convolutional layers with vision transformers, outperformed others in field conditions, suggesting that attention-based mechanisms are more resilient to background variation and occlusion, commonly found in in-situ imagery.

The most significant contribution of this study is the demonstration that hyperspectral imaging offers a considerable advantage in early-stage disease detection. RGB models achieved an average of 73.1% accuracy when diagnosing pre-symptomatic samples, while hyperspectral models reached 91.4%. This performance gap strongly supports the hypothesis that biochemical and physiological stress indicators manifest in the spectral domain well before visible symptoms appear. The PCA-reduced hyperspectral dataset, though limited in sample size, was sufficient to establish this trend. Moreover, incorporating vegetation indices like NDVI and PRI further enhanced predictive capability, confirming previous research that links plant reflectance signatures to biotic stress.

The success of the LeafGAN-based synthetic augmentation also contributes to the growing body of literature validating the use of generative models for dataset balancing. The increase in minority class F1-scores by over 20% demonstrates how data augmentation can mitigate class imbalance without the need for additional costly data collection. This is especially valuable in plant pathology, where rare diseases are underrepresented but still agriculturally significant.

The field trials conducted across two climatic zones provided crucial insights into the operational feasibility of deploying these models in real-world farming contexts. Despite lower average diagnostic accuracy (87.2%) compared to lab-based testing, the results are encouraging given the variability in lighting, leaf positioning, and background noise. The sub-2-second response time and high user satisfaction (4.4/5) further validate the system's usability. Farmers were able to use a simple mobile interface embedded with MobileNetV2 and PlantXViT to receive quick diagnostic results, which represents a significant step toward democratizing access to expert-level disease recognition.

However, several limitations must be acknowledged. The hyperspectral dataset was relatively small (300 samples), covering only three crops and a limited number of diseases. This restricts the generalizability of the 3D-CNN model trained on this data. Furthermore, while attention-based models performed well, their interpretability remains a challenge, and more work is needed to make them transparent to end users like farmers and extension workers. Finally, although transfer learning improved field performance to some extent, models still experienced domain shift—highlighting a need for domain adaptation strategies or self-supervised learning in future studies.

These findings align with recent trends in precision agriculture and smart farming technologies, which emphasize real-time, sensor-driven, and automated decision-making systems. Our work bridges the gap between laboratory research and practical deployment by proving that with minimal resources, a lightweight but accurate diagnostic tool can be developed and used in the field. It contributes meaningfully to current literature by not only affirming known limitations of RGB imaging but also validating underexplored approaches such as hyperspectral modeling and attention mechanisms in plant disease detection.

Conclusion

This research explored the fusion of advanced imaging and artificial intelligence to revolutionize plant disease diagnosis, moving from traditional visual inspections to highly accurate, data-driven techniques. By systematically comparing RGB-based convolutional neural networks with hyperspectral imaging supported by 3D-CNN architectures and attention-based hybrid models like PlantXViT, we have demonstrated that significant gains in both accuracy and timeliness of diagnosis are achievable.

Key findings reveal that while traditional CNNs like ResNet-50 and MobileNetV2 perform exceptionally well under controlled conditions, their effectiveness diminishes in real-world, field environments. On the other hand, transformer-based models and hyperspectral imaging exhibit superior robustness and early-stage detection capabilities, providing clear evidence for their integration into future diagnostic tools. The use of synthetic data generation via LeafGAN also proved instrumental in addressing data imbalance issues, further enhancing classification performance.

Real-world deployment trials confirmed that AI-powered diagnostic tools can be both efficient and accessible to farmers and agricultural practitioners, requiring only a smartphone for real-time disease detection. The positive user feedback and rapid response times suggest strong potential for large-scale adoption.

The study affirms that merging deep learning, hyperspectral technology, and user-centered mobile applications forms a viable, scalable solution to one of agriculture’s most pressing problems—timely and accurate identification of crop diseases. These innovations not only support sustainable farming practices but also help secure food production systems against the growing threats posed by climate change, pests, and global food demand. Future research should now focus on expanding spectral datasets, refining cross-domain learning, and improving explainability to ensure these tools are transparent, inclusive, and widely deployable.

References

1. Arima K, Noda T, Saito H. Discriminative Difficulty Distance for cross-domain plant disease diagnosis. IEEE Transactions on Plant Science. 2025;52:112-125.
2. Brahimi M, Boukhalfa K, Moussaoui A. Deep learning for tomato diseases: classification and symptoms visualization. Applied Artificial Intelligence. 2020;34(1):1-27.
3. Cap QT, Nguyen LT, Tran NH. LeafGAN: synthetic diseased leaf image generation for data augmentation. Computers and Electronics in Agriculture. 2020;173:105386.
4. Chen J, Wang Z, Huang H, Zhang L. Hyperspectral

image classification for early plant disease detection using 3D convolutional neural networks. IEEE Transactions on Geoscience and Remote Sensing. 2023;61:1-13.

5. Ennouri W, Gharbi F, Louhaichi M. Use of vegetation indices for early detection of plant disease stress in crops. Remote Sensing Applications: Society and Environment. 2021;24:100605.
6. Ferentinos KP. Deep learning models for plant disease detection and diagnosis. Computers and Electronics in Agriculture. 2018;145:311-318.
7. García Vera M, Pérez Álvarez E, Sánchez Marín A. Hyperspectral imaging and machine learning methods for precision agriculture. Agricultural Systems. 2024;198:103-115.
8. Gold J, Singh R, Fernandes A. Hyperspectral detection of potato late blight and black leg prior to symptom onset. Plant Pathology. 2023;72(4):789-797.
9. Hughes DP, Salathé M. Open access image dataset for plant disease detection. Frontiers in Plant Science. 2015;7:1419.
10. Jin X, Song Q, Zou J. Mobile hyperspectral imaging device for in-field plant disease diagnosis. Sensors. 2024;24(7):3291.
11. Karthik R, Senthil Kumar R, Subramanian S. PlantXViT: Vision Transformer-based lightweight architecture for robust field-level plant disease detection. Artificial Intelligence in Agriculture. 2023;7:45-56.
12. Krezhova E, Dimitrova T, Iliev A. Early detection of tomato leaf blight via hyperspectral reflectance and PCA. Computers and Electronics in Agriculture. 2022;191:106541.
13. Lai Y, Lu T. GAN-based data augmentation for plant disease detection under field variability. Pattern Recognition Letters. 2021;142:153-160.
14. Li H, Zhang W, Wu J. Transfer learning for field-level disease detection in cotton using CNNs. Remote Sensing. 2022;14(6):1345.
15. Lu J, Zhang H, Wang S. Spectral early detection of tomato virus using hyperspectral imaging and GAN augmentation. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2023;17:54-62.
16. Mahlein AK, Oerke EC, Steiner U. Robotics and edge sensors for automated plant disease severity scoring. Precision Agriculture. 2024;25(2):337-356.
17. Mohanty SP, Hughes DP, Salathé M. Using deep learning for image-based plant disease detection. Frontiers in Plant Science. 2016;7:1419.
18. Nikzadfar M, Shen H, Zhang C. Smartphone-compatible hyperspectral system for field plant diagnostics. Field Crops Research. 2024;295:108604.
19. Patel N, Shah S. MobileNet-based embedded systems for on-field disease classification in horticultural crops. Biosystems Engineering. 2024;233:12-25.
20. Riyanto F, Wijaya M, Santoso W. Review of deep learning-based mobile applications for plant disease diagnosis. Computers and Electronics in Agriculture. 2025;200:107145.
21. Sankhe D, Ambhaikar R. Data diversity and ML limitations in plant disease detection: review and roadmap. International Journal of Agricultural Technology. 2025;31(1):23-41.
22. Sethy PK, Barpanda NK. Application of deep learning and transfer learning models for field-level plant leaf disease detection. Journal of Ambient Intelligence and Humanized Computing. 2021;12:9651-9661.
23. Sharma A, Patel M. Vision Transformer approaches for plant leaf disease detection in resource-limited settings. Artificial Intelligence Review. 2022;55(4):2731-2747.
24. Thakur A, Singh V, Mehta S. PlantXViT: hybrid CNN-Transformer model for maize and rice disease detection. Computers and Electronics in Agriculture. 2022;194:106707.
25. Upadhyay S, Sharma RK, Gupta P. Deep learning and remote sensing technologies in plant disease diagnostics: a systematic review. Remote Sensing in Agriculture. 2025;3(1):19-37.
26. Wang Y, Li X, Zhou J. Machine learning for crop disease and pest detection: trends and challenges. Computers and Electronics in Agriculture. 2025;201:107192.
27. Xu L, Chen X, Huang K. Early warning of rice leaf blast using UAV-based multispectral sensors. Precision Agriculture. 2020;21:1015-1032.
28. Yadav P, Mishra S. Data augmentation methods for plant disease detection in small-sample domains. Expert Systems with Applications. 2021;167:114105.
29. Yang Z, Liu Y, Xu Q. Cross-domain adaptation techniques for plant disease classification. IEEE Access. 2023;11:23456-23471.
30. Zhang Y, Liu C, Zhao W, Yu H. Early detection of plant disease using hyperspectral imaging and deep transfer learning. Agricultural and Forest Meteorology. 2022;311:108693.
31. Zhu L, Chen J, Li F. Edge-cloud collaborative inference framework for resource-efficient plant disease detection. IEEE Internet of Things Journal. 2025;12(6):5187-5199.