AI Powered System for Early Plant Disease Detection

Ranveer Singh1 , Ritesh Mishra2 , Shubham Kumar Tripathi3 , Saroj Singh4

[1]UG Scholar, Dept. of CSE, Babu Banarasi Das Institute of Tech. & Manag., Lucknow, UP, India
[2]UG Scholar, Dept. of CSE, Babu Banarasi Das Institute of Tech. & Manag., Lucknow, UP, India
[3]UG Scholar, Dept. of CSE, Babu Banarasi Das Institute of Tech. & Manag., Lucknow, UP, India
[4]Assistant professor, Dept. of CSE, Babu Banarasi Das Institute of Tech. & Manag., Lucknow, UP, India

Emails: ranveersinghayush@gmail.com[1] , mishraritik287@gmail.com[2] , skt925243@gmail.com[3] , saroj.cse10@bbdnitm.ac.in[4]

Abstract

Early detection of plant diseases is critical for improving crop yield, reducing economic losses, and ensuring sustainable agriculture. Recent advancements in artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), have significantly enhanced the accuracy and efficiency of plant disease identification systems. Studies such as Sujatha et al.
[1] demonstrate the effectiveness of DL models like VGG-16 and Efficient DenseNet, achieving accuracies up to 97.2%, while hybrid approaches combining Inception v3 with SVM further improve classification performance. Vision Transformer-based frameworks such as PLA-ViT proposed by Murugavalli et al. [2] address limitations of traditional convolutional neural networks (CNNs) by capturing global and local dependencies, enabling superior disease localization and classification. Additionally, lightweight architectures like MobileNetV2 [3] and attention-based CNN models [5] facilitate deployment on resource-constrained devices, making AI-driven solutions accessible to farmers in developing regions.

Furthermore, emerging technologies such as hyperspectral imaging [6], UAV-based remote sensing [13], and ensemble learning methods [9], [14] contribute to earlier and more precise disease detection under real-world conditions. While datasets like PlantVillage have enabled high model accuracy, studies highlight challenges in generalization to field environments [4]. Techniques including transfer learning [12], federated learning [17], and data augmentation using GANs [25] have been proposed to address issues of data scarcity, privacy, and variability. Overall, the integration of AI with IoT and edge computing technologies provides a scalable and efficient framework for early plant disease detection, supporting precision agriculture and enabling timely intervention strategies.

Keywords— Plant Disease Detection, Artificial Intelligence, Machine Learning, Deep Learning, Convolutional Neural Networks, Vision Transformers, Hyperspectral Imaging, Precision Agriculture, Transfer Learning, IoT, Edge Computing.

1.INTRODUCTION

The agricultural sector plays a vital role in global food security, and plant diseases pose a significant threat to crop productivity and quality. Early detection of plant diseases is essential to minimize yield loss, reduce excessive pesticide use, and ensure sustainable farming practices. Traditional disease detection methods rely heavily on manual inspection by experts, which is time-consuming, subjective, and often inaccurate in large-scale farming environments. With the rapid advancement of Artificial Intelligence (AI), automated plant disease detection systems have emerged as efficient alternatives, leveraging Machine Learning (ML) and Deep Learning (DL) techniques for accurate and timely diagnosis [7].

Recent studies demonstrate that AI-driven models significantly improve disease detection accuracy. Deep learning architectures such as VGG, ResNet, DenseNet, and EfficientNet have shown remarkable performance in classifying plant diseases from leaf images [1], [12], [21]. These models automatically extract complex features such as texture, color, and lesion patterns, eliminating the need for the manual feature engineering .Hybrid approaches

integrating ML and DL techniques enhance classification performance and improve early detection capabilities under diverse environmental conditions [1].

Another significant advancement in early plant disease detection is the integration of data augmentation and synthetic data generation techniques to overcome the challenge of limited datasets. Generative Adversarial Networks (GANs) have been effectively used to create realistic synthetic plant leaf images, thereby improving model training and reducing overfitting [25]. Additionally, modern approaches such as cross-domain learning and domain adaptation aim to bridge the gap between laboratory and real-world conditions, ensuring better generalization of models across diverse agricultural environments [15]. These advancements highlight the growing emphasis on developing robust, scalable, and adaptable AI-based systems capable of supporting real-time disease detection and decision-making in precision agriculture.

A. Role of Artificial Intelligence in Plant Disease Detection

Artificial Intelligence has transformed agricultural practices by enabling automated disease identification using image-based analysis and predictive modeling. ML and DL models typically follow a structured pipeline that includes image acquisition, preprocessing, segmentation, feature extraction, and classification [7]. Convolutional Neural Networks (CNNs) are particularly effective due to their ability to learn hierarchical feature representations directly from raw images. Studies have shown that CNN- based models outperform traditional ML techniques such as Support Vector Machines (SVM) and Random Forests in terms of accuracy and scalability [1], [8].

Moreover, transfer learning has further accelerated the adoption of AI in agriculture. Pre-trained models such as MobileNetV2, EfficientNet, and ResNet can be fine-tuned on domain-specific datasets, reducing training time and computational requirements while maintaining high accuracy [3], [12]. This approach is especially beneficial for developing real-time applications on smartphones and edge devices, making disease detection accessible to farmers in rural and resource-limited regions.

B. Advancements in Deep Learning and Vision-Based Models

Recent advancements in deep learning have introduced innovative architectures that enhance the performance of plant disease detection systems. Vision Transformers (ViTs) have emerged as a powerful alternative to CNNs by utilizing self- attention mechanisms to capture both global and local dependencies within images. The PLA-ViT framework proposed by Murugavalli et al. [2] demonstrates superior performance in disease classification and localization, overcoming the limitations of CNNs that focus primarily on local features.

Object detection models such as YOLOv4 and Faster R-CNN enable real-time detection and localization of multiple diseases within a single image, making them suitable for field applications [8], [16]. Attention mechanisms integrated into CNN architectures further improve feature extraction by focusing on the most relevant regions of the image, leading to higher classification accuracy and efficiency [5]. Ensemble learning techniques also combine multiple models to achieve improved robustness and generalization across diverse datasets [9], [14].

C. Emerging Technologies for Early Disease Detection

Beyond traditional image-based methods, emerging technologies are playing a crucial role in advancing early plant disease detection. Hyperspectral imaging (HSI) allows the capture of detailed spectral information across multiple wavelength bands, enabling the detection of plant stress and diseases at pre-symptomatic stages [11]. This capability is particularly important for preventing the spread of diseases before visible symptoms appear.

Unmanned Aerial Vehicles (UAVs) equipped with advanced sensors provide large-scale monitoring of agricultural fields, offering real-time insights into crop health and enabling precision agriculture practices [13]. Additionally, Internet of Things (IoT)-based systems integrate environmental sensors to collect data on temperature, humidity, soil moisture, and other factors that influence disease development. These systems, when combined with AI models, enable predictive analytics for disease outbreak prevention [19].

Smartphone-based applications powered by lightweight DL models allow farmers to diagnose plant diseases instantly by capturing leaf images, making advanced technology accessible even in developing regions [3], [10].

D. Challenges and Research Directions

Despite significant progress, several challenges hinder the widespread adoption of AI-based plant disease detection systems. One major limitation is the lack of generalization of models trained on controlled datasets such as PlantVillage when applied to real-world field conditions [4]. Variations in lighting, background noise, occlusion, and environmental conditions significantly impact model performance.

To overcome these challenges, researchers are exploring advanced techniques such as federated learning, few-shot learning, and domain adaptation. Federated learning enables decentralized training while preserving data privacy, making it suitable for collaborative agricultural environments [17], [18]. Few-shot learning approaches aim to improve model performance with limited labeled data, addressing the issue of dataset scarcity [15]. Additionally, Generative Adversarial Networks (GANs) are used to generate synthetic datasets, enhancing model robustness and reducing overfitting [25].

Future research should focus on developing lightweight, energy-efficient, and scalable models that can operate in real-time under diverse environmental conditions. The integration of AI with edge computing, IoT, and remote sensing technologies will be instrumental in building intelligent agricultural systems capable of early disease detection, decision support, and automated crop management.

E. Motivation and Objectives of this Review

The rapid growth of Artificial Intelligence (AI) in agriculture has created a need for a comprehensive understanding of existing techniques and their effectiveness in early plant disease detection. This review is motivated by the increasing demand for accurate, real-time, and scalable solutions to address challenges in modern agriculture. The key objectives of this study are as follows:

- To analyze existing AI-based techniques for plant disease detection:
  This review examines various Machine Learning (ML) and Deep Learning (DL) approaches, including CNNs, Vision Transformers, and hybrid models, to understand their performance and limitations in disease classification tasks [1], [2], [12].
- To evaluate advancements in early disease detection technologies: The study explores emerging technologies such as hyperspectral imaging, UAV-based monitoring, and IoT systems that enable early and pre-symptomatic disease detection in crops [11], [13], [19].
- To identify challenges in real-world deployment:
  Despite high accuracy in controlled

environments, many models fail to generalize in field conditions. This review highlights issues related to dataset bias, environmental variability, and computational constraints [4], [15].

BACKGROUND

A. Traditional and Machine Learning-Based Approaches:

Early plant disease detection methods primarily relied on manual inspection and traditional image processing techniques. Farmers and agricultural experts visually examined plant leaves to identify symptoms, which often led to delayed and inaccurate diagnoses. With the introduction of Machine Learning (ML), automated systems began to utilize handcrafted features such as color, texture, and shape for classification tasks [7]. Techniques such as Support Vector Machines (SVM), Random Forest, and k-Nearest Neighbors (k-NN) were widely used for disease detection.

These approaches required extensive feature engineering and were limited in handling complex and large-scale datasets. Studies have shown that ML models are less effective compared to deep learning models in capturing intricate patterns in plant leaf images [1]. Despite these limitations, ML- based systems laid the foundation for the development of more advanced AI-driven disease detection techniques.

B. Deep Learning-Based Plant Disease Detection:

Deep Learning (DL) has significantly improved the accuracy and efficiency of plant disease detection systems. Convolutional Neural Networks (CNNs) such as VGG16, ResNet, DenseNet, and EfficientNet automatically learn hierarchical features from images, eliminating the need for manual feature extraction [1], [12], [21]. These models have demonstrated high accuracy in classifying plant diseases across various datasets, particularly under controlled conditions.

Recent advancements include the integration of attention mechanisms and ensemble learning techniques, which further enhance feature representation and classification performance [5], [9]. Vision Transformer (ViT)-based models have also emerged as powerful alternatives to CNNs by capturing global contextual information using self-attention mechanisms [2], [20]. Additionally, object detection models such as YOLOv4 enable real-time disease localization, making them suitable for practical agricultural applications [8]. These developments highlight the growing importance of deep learning in modern plant disease detection systems.

C. Emerging Technologies for Early Detection:

To image-based approaches, several emerging technologies have been introduced to enable early and accurate detection of plant diseases. Hyperspectral imaging (HSI) allows the detection of plant stress at pre-symptomatic stages by analyzing spectral information across different wavelength bands [11]. This technique is highly effective in identifying diseases before visible symptoms appear, providing a significant advantage in early intervention.

Unmanned Aerial Vehicles (UAVs) equipped with multispectral and thermal sensors offer large-scale crop monitoring capabilities, enabling real-time disease detection and precision agriculture practices [13]. Furthermore, Internet of Things (IoT)-based systems collect environmental data such as temperature, humidity, and soil conditions, which can be integrated with AI models for predictive disease analysis [19]. Federated learning and edge computing technologies are also being explored to ensure data privacy and enable decentralized model training across multiple agricultural sites [17], [18]. These advancements demonstrate the shift toward intelligent, scalable, and real-time plant disease detection systems.

LITRATURE REVIEW

A. Machine Learning and Deep Learning-Based Approaches:

Early research in plant disease detection focused on traditional machine learning techniques that relied on handcrafted features such as color, texture, and shape. Methods like Support Vector

Machines (SVM), k-Nearest Neighbors (k-NN), and Random Forest classifiers were widely used for classification tasks [7]. However, these approaches required manual feature extraction and lacked robustness when applied to complex datasets.

With the evolution of deep learning, Convolutional Neural Networks (CNNs) became the dominant approach due to their ability to automatically extract hierarchical features from images. Sujatha et al. [1] demonstrated that deep learning models such as VGG-16 outperform traditional ML methods, achieving an accuracy of 89.5% for citrus plant disease detection. Furthermore, advanced models such as Efficient DenseNet achieved up to 97.2% accuracy for potato leaf disease classification by addressing data imbalance using reweighted loss functions [1].

Additional studies have shown that transfer learning significantly improves model performance. Johri et al. [12] evaluated multiple pre-trained models, including EfficientNet, ResNet, and DenseNet, with EfficientNetB3 achieving 99.96% accuracy. Similarly, Pandian et al. [21] proposed a deep residual network (ResNet197), achieving 99.58% accuracy across multiple plant species. These findings highlight the superiority of deep learning models in plant disease classification tasks.

B. Advanced Deep Learning Architectures and Hybrid Models:

Recent research has focused on improving model performance through advanced architectures and hybrid techniques. Vision Transformers (ViTs) have emerged as a powerful alternative to CNNs by capturing global dependencies using self-attention mechanisms. Murugavalli et al. [2] proposed the PLA-ViT framework, which outperforms traditional CNN-based models in disease classification, localization, and computational efficiency. Similarly, Borhani et al. [20] demonstrated that lightweight ViT models can achieve competitive accuracy while maintaining lower computational costs.

Attention-based CNN models have also shown significant improvements in feature representation. Karthikeyan et al. [5] introduced an attention-based Squeeze- and-Excitation CNN model (CNN-SEEIB), achieving 99.79% accuracy on the PlantVillage dataset. Ensemble learning approaches further enhance model robustness by combining multiple classifiers. Shafik et al. [9] proposed PDDNet models using ensemble techniques, achieving accuracies of 96.74% and 97.79%. Additionally, He et al. [14] developed an ensemble learning method (ELCDR) that improved classification accuracy across multiple crop datasets.

Object detection models such as YOLOv4 and Faster R-CNN have also been widely used for real-time disease detection and localization. Aldakheel et al. [8] reported an accuracy of 99.99% using YOLOv4 on the PlantVillage dataset, demonstrating its effectiveness in detecting plant diseases in large-scale datasets. Gong et al. [16] further highlighted the capability of YOLOv3 and Faster R-CNN in detecting multiple diseases in complex field environments.

C. Emerging Intelligent Systems and Real-World Applications:

Recent studies emphasize the integration of AI with advanced technologies to enable early and real-time plant disease detection. Hyperspectral imaging (HSI) has been widely used for detecting plant stress at pre- symptomatic stages. Zhang et al. [11] demonstrated that HSI can detect bacterial leaf spot in tomatoes with high accuracy, even before visible symptoms appear. Similarly, Seralathan et al. [6] introduced hyperspectral vegetation indices (MLVI and H-VSI), enabling early stress detection 10–15 days before traditional methods.

Unmanned Aerial Vehicles (UAVs) and remote sensing technologies have also gained attention for large-scale agricultural monitoring. Zhu et al. [13] highlighted the effectiveness of UAV-based systems equipped with multispectral and thermal sensors for real-time disease detection. Additionally, IoT-based systems enable continuous monitoring of environmental conditions, improving disease prediction and prevention [19].

To address challenges related to data privacy and limited datasets, federated learning and data augmentation techniques have been introduced. Kabala et al. [17], [18] proposed federated learning frameworks that allow decentralized model training while preserving data privacy. Bi et al. [25] utilized Generative Adversarial Networks (GANs) to generate synthetic plant leaf images, improving model performance by increasing dataset diversity. Furthermore, Yang et al. [15] introduced cross-domain few-shot learning techniques to improve generalization across different datasets and environmental conditions.

The literature indicates that while significant progress has been made in plant disease detection using AI, challenges such as dataset bias, real-world variability, and computational constraints remain. Future research should focus on developing robust, scalable, and real-time systems that can operate efficiently in diverse agricultural environments.

PROPOSED METHODOLOGY

A. Data Acquisition and Dataset Preparation:

The first step involves collecting plant leaf images from publicly available datasets such as PlantVillage, as well as real-world field images to improve model generalization. The dataset includes both healthy and diseased plant samples across multiple crop categories. Since controlled datasets often lack environmental variability, additional field images are incorporated to address real-world challenges [4].

To enhance dataset diversity, data augmentation techniques such as rotation, flipping, scaling, and brightness adjustment are applied. Additionally, synthetic image generation using Generative Adversarial Networks (GANs) is employed to overcome data scarcity and improve model robustness [25].

B. Image Preprocessing and Segmentation:

In this stage, raw images are preprocessed to improve quality and remove noise. Techniques such as resizing, normalization, contrast enhancement, and background removal are applied. Image segmentation methods are used to isolate the leaf region from the background, ensuring that only relevant features are analyzed.

Advanced segmentation models such as UNet and DeepLabV3+ are utilized to identify diseased regions at the pixel level, enabling precise localization of infection areas [23]. This step enhances the accuracy of subsequent classification by focusing on critical features.

C. Feature Extraction and Model Selection:

The proposed system utilizes deep learning models for automatic feature extraction. Pre-trained Convolutional Neural Networks (CNNs) such as VGG16, ResNet50, and EfficientNet are employed using transfer learning techniques to extract high-level features from leaf images [1], [12]. These models are fine-tuned on the dataset to adapt to specific plant disease patterns.

Additionally, Vision Transformer (ViT)- based models are incorporated to capture global contextual information, improving detection accuracy in complex scenarios [2]. Attention mechanisms are integrated into CNN architectures to enhance feature representation by focusing on important regions of the image [5].

D. Disease Classification and Detection:

The extracted features are passed to classification layers for disease prediction. The system employs a hybrid approach combining CNN-based classifiers and Support Vector Machines (SVM) to improve classification performance [1]. For real-time detection and localization, object detection models such as YOLOv4 are used to identify multiple diseases within a single image [8].

Ensemble learning techniques are also applied by combining predictions from multiple models to improve overall accuracy and robustness [9], [14]. This approach ensures reliable performance across diverse datasets and environmental conditions.

E. Model Training and Optimization:

The model is trained using labeled datasets with appropriate loss functions such as cross-entropy loss. Optimization techniques such as Adam optimizer and learning rate scheduling are used to improve convergence. To handle class imbalance, weighted loss functions are applied during training [1].

Performance evaluation is conducted using metrics such as accuracy, precision, recall, and F1-score. Cross-validation techniques are used to ensure model reliability and prevent overfitting. Hyperparameter tuning is performed to achieve optimal model performance.

F. Deployment and Real-Time Application:

The final model is deployed as a user- friendly application that can run on smartphones or edge devices. Lightweight models such as MobileNetV2 are used for efficient on-device inference, enabling real- time disease detection for farmers [3].

The system can also be integrated with IoT-based sensors and UAV platforms for large-scale monitoring and predictive analysis [13], [19]. Additionally, cloud-based deployment allows centralized data processing and continuous model updates, ensuring scalability and adaptability in real- world agricultural environments.

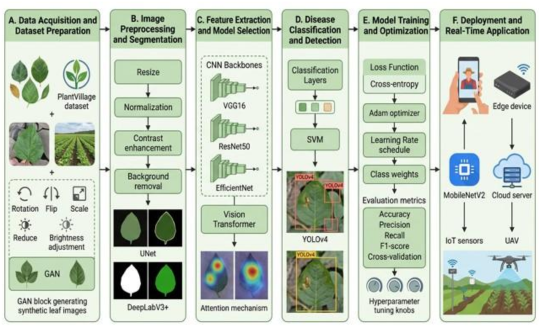

FIGURE 1: Complete workflow chart.

COMPARATIVE ANALYSIS OF EXISTING APPROACHES

The evolution of plant disease detection methods shows a clear transition from traditional machine learning (ML) approaches to advanced deep learning (DL) techniques. Early ML-based models relied heavily on handcrafted features such as color, texture, and shape, using classifiers like Support Vector Machines (SVM) and Random Forest [7]. While these methods provided moderate accuracy, they were limited by their dependency on manual feature extraction and inability to handle complex image variations. In contrast, DL models such as Convolutional Neural Networks (CNNs) have demonstrated superior performance by automatically learning hierarchical features from raw images. Studies such as Sujatha et al. [1] highlight that models like VGG-16 significantly outperform traditional ML approaches, achieving higher classification accuracy and robustness across different plant diseases.

Further advancements in deep learning architectures have introduced models with improved accuracy and efficiency. Pre- trained models such as EfficientNet, ResNet, and DenseNet have achieved remarkable results through transfer learning, with some studies reporting accuracies exceeding 99% [12], [21]. Additionally, attention-based models and hybrid approaches combining CNNs with ML classifiers have further enhanced performance by focusing on important regions of leaf images and improving feature representation [5], [1]. Vision Transformer (ViT)-based models have also emerged as strong alternatives, capturing global contextual relationships using self-attention mechanisms. Murugavalli et al. [2] demonstrated that ViT-based frameworks outperform conventional CNN models in disease localization and classification tasks, particularly in complex datasets.

Despite high accuracy in controlled environments, many models face challenges in real-world applications. Datasets such as PlantVillage often contain images captured under ideal laboratory conditions, which limits model generalization when applied to field data [4]. To address this issue, object detection models like YOLOv4 and Faster R-CNN have been introduced for real-time disease detection in complex environments, achieving high accuracy and enabling localization of multiple diseases [8],

[16]. Moreover, segmentation-based approaches such as UNet and DeepLabV3+ provide pixel-level classification, allowing precise identification of infected regions and improving decision-making for disease treatment [23]. These methods enhance practical applicability but often require higher computational resources.

Recent research has focused on integrating advanced technologies and improving model robustness. Techniques such as ensemble learning, federated learning, and Generative Adversarial Networks (GANs) have been proposed to overcome challenges related to data scarcity, privacy, and variability. Ensemble models combine multiple classifiers to achieve better accuracy and generalization [9], [14], while federated learning enables decentralized model training without sharing sensitive data [17], [18]. Additionally, GAN-based data augmentation improves dataset diversity and reduces overfitting [25]. Emerging technologies such as hyperspectral imaging and UAV-based monitoring further enhance early disease detection capabilities by enabling pre-symptomatic analysis and large-scale crop monitoring [11], [13]. Overall, while significant progress has been made, the need for scalable, efficient, and real-time systems remains a key focus for future research in AI-based plant disease detection.

CHALLENGES AND RESEARCH GAPS

A. Lack of Generalization in Real-World Conditions:

One of the major challenges is the poor generalization of models trained on controlled datasets such as PlantVillage. These datasets contain images captured under ideal laboratory conditions with uniform backgrounds, which do not represent real-world agricultural environments. As a result, models often fail to perform accurately when exposed to variations in lighting, background noise, and environmental conditions [4], [16]. This creates a significant gap between laboratory performance and field applicability.

Studies have shown that models trained on single-domain datasets struggle when applied to cross-domain scenarios, leading to performance degradation in practical use cases [15]. Addressing this issue requires the development of robust domain adaptation and generalization techniques capable of handling diverse agricultural conditions.

B. Limited Availability and Diversity of Datasets:

Another critical issue is the lack of large, diverse, and annotated datasets for training robust models. Many studies rely on limited datasets that do not cover variations in crop types, disease stages, and environmental conditions. Although data augmentation and synthetic data generation using GANs have been proposed to address this issue, the generated data may not fully capture real-world complexity [25], [1].

This limitation affects model accuracy and scalability across different agricultural scenarios.

The imbalance in dataset classes further impacts model performance, causing biased predictions toward dominant classes [9]. Therefore, there is a need for large-scale, well-annotated, and balanced datasets collected from real agricultural fields.

C. High Computational Complexity and Resource Constraints:

Advanced deep learning models such as Vision Transformers and large CNN architectures require significant computational resources for training and inference. This makes them unsuitable for deployment on low-power devices commonly used in rural and resource-constrained areas. While lightweight models like MobileNetV2 have been introduced for mobile applications, there is still a need for efficient models that balance accuracy and computational cost [3], [2], [20].

Real-time disease detection systems demand fast processing speeds, which are difficult to achieve with complex architectures. This creates a gap in developing optimized models that can operate efficiently on edge devices without compromising performance.

D. Challenges in Early and Pre-Symptomatic Detection:

Most existing models focus on detecting diseases after visible symptoms appear on plant leaves. However, early detection at pre-symptomatic stages remains a major challenge. Techniques such as hyperspectral imaging have shown potential in detecting stress before visible symptoms, but their high cost and complexity limit widespread adoption [11], [6]. There is a need for cost-effective and scalable solutions for early disease detection.

Integrating early detection techniques with AI models requires high-quality spectral data and advanced processing methods, which are not easily accessible to farmers [13]. Future research should focus on simplifying these technologies for practical agricultural use.

E. Data Privacy and Security Concerns:

With the increasing use of cloud-based and IoT-integrated systems, data privacy and security have become important concerns. Agricultural data collected from farms may include sensitive information, making centralized data storage risky. Federated learning has been proposed as a solution to enable decentralized training without sharing raw data, but challenges related to communication overhead and model synchronization still exist [17], [18].

Ensuring secure data transmission and preventing unauthorized access remain critical issues in large-scale deployments [19]. Developing secure and efficient frameworks is essential for building trust in AI-based agricultural systems.

F. Lack of Robustness to Environmental Variability:

Environmental factors such as varying light conditions, occlusion, weather changes, and complex backgrounds significantly affect model performance. Many models struggle to maintain accuracy under such conditions, leading to unreliable predictions in real- world scenarios [4], [8]. This limitation reduces the reliability of AI systems in practical farming environments.

Techniques such as cross-domain learning and ensemble methods have been proposed to improve robustness, further research is required to develop adaptive models that can handle dynamic environmental changes effectively [14], [15]. Enhancing model resilience to environmental variability remains a key research gap.

VII. FUTURE DIRECTIONS

The rapid evolution of Artificial Intelligence (AI) and related technologies has opened new opportunities for enhancing plant disease detection systems. Despite significant progress, several areas require further research to improve accuracy, scalability, and real-world applicability. This section outlines key future directions for advancing AI-based early plant disease detection.

A. Development of Robust and Generalized Models:

Future research should focus on developing models that can generalize effectively across diverse agricultural environments. Current models often suffer from performance degradation when applied to real-world conditions due to variations in lighting, background, and environmental factors [4], [15]. To address this, advanced techniques such as domain adaptation, cross-domain learning, and self-supervised learning can be explored to improve model robustness.

Incorporating large-scale and diverse datasets collected from real agricultural fields will enhance model reliability. Hybrid architectures combining CNNs and Vision Transformers can also be further optimized to balance local feature extraction and global context understanding [2], [20].

B. Integration of Edge Computing, IoT, and Smart Agriculture Systems:

The integration of AI with edge computing and IoT technologies is expected to play a crucial role in the future of precision agriculture. Lightweight models such as MobileNetV2 can be deployed on edge devices for real-time disease detection, reducing dependency on cloud-based systems [3]. IoT-enabled sensor networks can provide continuous monitoring of environmental parameters such as temperature, humidity, and soil conditions, enabling predictive disease analysis [19].

UAV-based remote sensing systems equipped with multispectral and hyperspectral sensors can enable large- scale monitoring of crops, allowing early detection and timely intervention [13]. The combination of these technologies will lead to the development of intelligent, automated agricultural systems.

C. Advancements in Early Detection and Data-Efficient Learning:

Early detection of plant diseases before visible symptoms appear remains a critical research area. Future work should focus on integrating hyperspectral imaging and advanced AI techniques to detect subtle changes in plant physiology [11], [6]. Developing cost-effective solutions for early detection will make these technologies more accessible to farmers.

Data-efficient learning techniques such as few-shot learning, federated learning, and GAN-based data augmentation can address challenges related to limited datasets and data privacy [15], [17], [25]. These approaches will enable the development of scalable and privacy-preserving models capable of performing effectively in real- world scenarios with minimal labeled data.

VIII. CONCLUSION

The advancement of Artificial Intelligence (AI) has significantly transformed the field of plant disease detection, enabling accurate, efficient, and automated identification of diseases from plant leaf images. This study reviewed various Machine Learning (ML) and Deep Learning (DL) approaches, highlighting the superior performance of deep learning models such as CNNs, EfficientNet, and Vision Transformers in disease classification tasks [1], [12], [2]. These models have demonstrated high accuracy and robustness, especially when combined with techniques such as transfer learning, attention mechanisms, and ensemble learning. The integration of object detection and segmentation methods has further enhanced the ability to localize and quantify disease-affected regions, making these systems more practical for real-world applications [8], [23].

Despite these advancements, several challenges remain in deploying AI-based plant disease detection systems in real- world agricultural environments. Issues such as poor generalization of models trained on controlled datasets, limited availability of diverse datasets, and high computational requirements continue to hinder practical implementation [4], [25]. Additionally, early detection of diseases at pre-symptomatic stages remains a critical challenge, although emerging technologies such as hyperspectral imaging and UAV- based monitoring show promising results [11], [13]. Addressing these challenges requires the development of robust, lightweight, and scalable models that can operate efficiently under varying environmental conditions.

The future of early plant disease detection lies in the integration of AI with emerging technologies such as IoT, edge computing, and federated learning, which can enable real-time, secure, and decentralized disease monitoring systems [17], [19]. Furthermore, advancements in data-efficient learning techniques and the availability of large-scale real-world datasets will play a crucial role in improving model performance and generalization [15]. By addressing current limitations and leveraging technological innovations, AI-based plant disease detection systems have the potential to revolutionize precision agriculture, enhance crop productivity, and contribute to global food security.

ACKNOWLEDGMENT

The author expresses sincere gratitude to **Ms. Saroj Singh**, Assistant Professor, Department of Computer Science and Engineering, Babu Banarasi Das Institute of Technology and Management, Lucknow, India, for her invaluable guidance, constructive feedback, and continuous encouragement throughout the preparation of this review paper. Author also extends thanks to the Department of Computer Science and Engineering for providing a supportive academic environment and institutional resources that significantly contributed to the successful completion of this work.

REFERENCES

[1] Sujatha et al., "Advancing plant leaf disease detection integrating machine learning and deep learning," Scientific Reports, vol. 15, 2025.

[2] Murugavalli et al., "Plant leaf disease detection using vision transformers for precision agriculture," Scientific Reports, vol. 15, 2025.

[3] Mhembere et al., "Low-cost smartphone-based plant disease diagnosis for Zimbabwean farmers," International Journal of Computer Applications, vol. 187, no. 32, 2025.

[4] Richter et al., "Assessing the performance of domain-specific models for plant disease detection: a comprehensive benchmark of transfer learning on open datasets," Scientific Reports, vol. 15, 2025.

[5] Karthikeyan et al., "Enhanced plant disease classification with attention- based squeeze-and-excitation CNN (CNN-SEEIB)," Frontiers in Artificial Intelligence, vol. 8, 2025.

[6] Seralathan et al., "MLVI-CNN: a hyperspectral stress detection framework with machine-learning-based vegetation indices," Frontiers in Plant Science, vol. 16, 2025.

[7] Jafar et al., "Revolutionizing agriculture with artificial intelligence: plant disease detection and management," Frontiers in Plant Science, vol. 15, 2024.

[8] Aldakheel et al., "Detection and identification of plant leaf diseases using YOLOv4," 2024.

[9] Shafik et al., "Using transfer-learning-based plant disease classification with ensemble early fusion and lead voting methods," BMC Plant Biology, vol. 24, 2024.

[10] Babatunde et al., "A novel smartphone application for early detection of habanero plant diseases using a modified VGG16 deep transfer learning model," Scientific Reports, vol. 14, p. 1423, 2024.

[11] Zhang et al., "Hyperspectral imaging analysis for early detection of tomato bacterial leaf spot disease," Scientific Reports, vol. 14, p. 27666, 2024.

[12] Johri et al., "Advanced deep transfer learning techniques for efficient cotton plant disease detection," 2024.

[13] Zhu et al., "Intelligent agriculture: deep learning in UAV-based remote sensing for crop disease monitoring," 2024.

[14] He et al., "A novel ensemble learning method for crop leaf disease recognition (ELCDR)," 2024.

[15] Yang et al., "From laboratory to field: cross-domain few-shot learning for plant disease detection," 2024.

[16] Gong et al., "An analysis of plant disease identification based on deep learning: comparative evaluation of YOLOv3 and Faster R-CNN in field environments," 2023.

[17] Kabala et al., "Image-based crop disease detection with federated learning," 2023.

[18] Kabala et al., "Secure and decentralized plant disease detection via federated learning," 2023.

[19] Hnatiuc et al., "Intelligent grapevine disease detection using IoT sensor networks," 2023.

[20] Borhani et al., "A deep-learning- based approach for automated plant disease classification using Vision Transformer," 2022.

[21] Pandian et al., "An improved deep residual convolutional neural network (ResNet197) for plant leaf disease detection," 2022.

[22] Khan et al., "End-to-end semantic leaf segmentation framework for plant disease identification," 2022.

[23] Li et al., "Detection of potato foliage diseases using instance, classification, and semantic segmentation models," 2022.

[24] Atila et al., "Plant leaf disease classification using EfficientNet deep learning model,"

[25] Bi et al., "Improving image-based plant disease classification with generative adversarial network and label smoothing under limited training data," Frontiers in Plant Science, vol. 11, 2020