Advances in Leaf Disease Detection: A Comprehensive Review of Image-Based Techniques and Future Opportunities

Pawan Kumar Verma[1], Divya Midhun[2], and Hemalatha[3]

[1]Sharda University, Greate Noida, Uttar Pradesh, India, abes.pawan@gmail.com
[2]Lincoln University College, Selangor Darul Ehsan, Malaysia, divya@lincoln.edu.my
[3]Panimalar Engineering College , Chennai, Tamil Nadu, India, Pithemalatha@gmail.com

Abstract

Detection of leaf disease plays an important role in sustainable agricultural productivity by providing early diagnosis and intervention. Traditional methods involve manual inspection, which are time consuming, labor intensive, and prone to errors. Therefore, there is a need for automated image-based techniques based on computer vision and deep learning for the identification of disease. This paper presents a comprehensive analysis of recent advancements in leaf disease detection techniques focusing on classical machine learning approaches, deep learning architectures, and hybrid models. The challenges of real-world deployment, such as environmental changes, dataset imbalance, and the lack of Artificial Intelligence (AI) model explanability, are covered in this paper. In the end of this paper, future research opportunities are discussed to build a robust AI model for leaf disease detection with improved results. This paper aims to establish the foundation for more effective, scalable, and interpretable leaf disease detection systems by synthesizing state-of-the-art techniques and outlining potential research directions for researchers to fill the gap between research advances and practical applications in agriculture.

Introduction

Agriculture serves as a foundation for global food production. It supports a large number of people and constitutes a solid basis for economic stability. However, the threat of leaf diseases and pest infestations is compromising both the quality and quantity of crops. Diseases such as fungal, bacterial, and viral pathogens are increased by environmental pressures, climate change, and accidental infections [1]. Rapid and accurate identification and classification of these infections are the keys to reducing large losses in agriculture and saving the global food budget. Pest infestations alone reduce productivity by 30 - 33% annually [2].

Despite various advancements in agriculture, most farmers across the globe are using traditional methods to detect diseases, which are slow, prone to errors, and inefficient [3]. In the traditional method, various undetected infected leaves are still present that cause extreme economic losses and food shortages [4]. This issue is mainly observed in developing countries where advanced agricultural services and tools are limited.

The Sustainable Development Goals (SDGs) of the United Nations (UN) state that the "Zero Hunger" initiative requires technological input in agriculture to ensure food security for the world by 2030 [5]. The World Health Organization (WHO) states that the growing population has required the development of high-yielding, disease-resistant crops. The absence of strong, scalable, and efficient disease detection systems can increase hunger and malnutrition worldwide.

Researchers are developing automated models for the disease detection process with the help of AI, Machine Learning (ML), and Deep Learning (DL) for the initiative "Zero Hunger". These automated models allow early detection of the disease and less dependence on chemical pesticides while increasing yield production [6, 7, 8, 9].

Recent advances in the identification and classification of plant disease have increasingly relied on AI to improve accuracy. Specifically, Convolutional Neural Networks (CNNs) demonstrate high precision in the detection or classification of various diseases by analyzing complex leaf patterns. The use of Vision Transformers (ViTs) and hybrid AI models has improved the classification task in challenging environmental conditions, such as lighting variations. Furthermore, the integration of Internet of Things (IoT) devices has transformed large-scale agricultural monitoring that enables real-time disease detection across large farmlands. Edge-AI and mobile applications are also used for disease detection that empower farmers to use mobile phone smartphone cameras for immediate disease identification and motivate immediate action. These technological advances are intended to achieve sustainable agriculture and minimize yield losses.

With the increasing demand for leaf disease detection, this article provides a comprehensive review of recent research, methodologies, and existing challenges. The purpose of this article is to fill the gap between technological advancements in leaf disease detection methodologies and their practical implementation. This will be valuable for researchers, AI developers, and farmers in reducing crop losses and improving global food security.

This article is organized into five sections. Section 2 explains the literature survey on the detection of plant leaf disease. This section is divided into three further sub-sections named Traditional approaches, Machine Learning Based Approaches, Deep Learning based approach, and hybrid approaches. Section 3 highlights the challenges that occur in the detection process. Section 4 enlists the future possibilities in the detection process. At the end conclusion, section 5 concludes the complete article.

Literature Survey

Plant leaf diseases significantly affect agricultural productivity and lead to economic losses and food security problems throughout the world. Early leaf detection of these diseases is the most vital stage in reducing these impacts. Researchers have proposed different methodologies using traditional image processing techniques, ML, and DL approaches to the identification and classification of leaf diseased. This section explains the existing studies on leaf disease detection techniques.

Traditional Approaches

Early researchers conducted research using traditional methods of image processing for the classification and detection of plant leaf diseases. For feature extraction in leaf images, the researchers applied Gray Level Co-occurrence Matrix (GLCM), Scale-Invariant Feature Transform (SIFT), and Histogram of Oriented Gradients (HOG) techniques. These methods relied on hand-crafted feature engineering that requires domain knowledge for optimal performance. Kumar and Singh [10] reviewed some of the existing techniques related to leaf disease detection that emphasize the problem of classical techniques due to sensitivity to the intensity of light illumination and background clutter. These approaches provided an excellent understanding of the plant disease detection framework; however, since they depended on manually chosen characteristics, the accuracy could not match those of more modern approaches.

Machine Learning-Based Approaches

With the advancement in AI, it has gained more prominence in leaf disease detection. Support Vector Machine (SVM), Decision Tree (DT), K-Nearest Neighbors (KNN), and Random Forest (RF) are some of the ML algorithms that have been used extensively for classification tasks. These models are based on features extracted from the leaf images to identify diseased regions and classify them accordingly.

Sethy et al. [11] proposed a deep feature-based classification model based on SVM for the detection of rice leaf disease. The results show that Convolutional Neural Network (CNN) combined with SVM produced better accuracy compared to traditional ML classifiers. Similarly, Sujatha et al. [12] have compared the performance of the DL and ML models. The results have been found that DL outperforms the traditional ML algorithms with high performance in the extraction of features and classification.

Despite the excellent results of ML models, their reliance on hand-selected features makes it difficult to transfer them to new plant species or environmental conditions. In addition, ML models are heavily based on hand-crafted domain-specific features and less amenable to learning new data with minimal retraining. Such restrictions have led to the pursuit of DL based methods, where hierarchical representations of raw image data are learned automatically.

Deep Learning-Based Approaches

The solution of the problem has evolved due to DL, as it has removed the barriers to manual extraction of features and encouraged end-to-end learning from the raw input image data. CNNs have been the most predominant DL architecture used for leaf disease classification because of their spatial hierarchies of features that capture input images. Many research studies have developed CNN-based models for accurate and efficient identification of diseases in plants.

Tiwari et al. [13] proposed the DL based approach for the diagnosis of multiple plant diseases. They observed that increasing the depth of the network and the number of connections reduces the appearance of the vanishing gradient, leading to high-performance classification for deep networks. Like Karthik et al. [14] focused on attention based deep residual network for disease in tomato leaves. This proposed model achieved the accuracy of 98% in the validation set by 5-fold cross-validation process.

Another advancement in deep learning-based plant disease detection is the use of lightweight architectures for deployment on edge devices such as smartphones and embedded systems. Mazumder et al. [15] developed a robust transfer learning-based model capable of detecting diseases in multiple plant species using a smaller data set. Their model leveraged pre-trained networks to enhance feature extraction while reducing computational overhead.

Although CNN-based models have significantly improved the accuracy of disease classification, the challenges remain in handling complex backgrounds and variations in leaf orientation. Recent research has explored hybrid approaches and advanced architectures to address these issues.

Hybrid and Advanced Architectures

Researchers have addressed hybrid approaches to overcome the limitations of standalone deep learning models. Hybrid models combine several models for better performance. Integrating deep learning with advanced architectures of neural networks improves accuracy and robustness and enhances interpretability.

Aldakheel et al. [16] proposed a hybrid real-time plant disease detection model that combines CNNs along with YOLOv4. The presented YOLOv4 was shown to obtain a better object detection capability while accurately localizing the diseased regions in an image of leaves. Similarly, Wang et al. [17] presented MGA-YOLO, a lightweight one-stage apple leaf disease detection network. By reducing the computations, their solution was suitable for real-time agricultural applications while maintaining excellent accuracy.

The integration of deep learning with optimization techniques is another technique. Mukhopadhyay et al. [18] used multi-objective image segmentation to improve the detection of tea leaf diseases. This improves the segmentation accuracy by optimizing several criteria at the same time and could better differentiate between diseased and healthy regions.

Some recent attention has been given to transformer-based architectures in

Tab. 1: Literature Survey

S. No.	Ref.	Year	Result	Methodology	Dataset Used	Limitation
1	[14]	2020	Improved precision with attention mechanisms	Attention-based CNN	Tomato leaf disease dataset	Overfitting on small datasets
2	[25]	2020	Enhanced disease severity classification	Spectral disease indices with CNN	Corn dataset	Requires specialized equipment
3	[11]	2020	High performance with deep features	SVM with deep features	Rice leaf disease dataset	Dependency on feature extraction
4	[13]	2021	High accuracy with DenseNet	Dense Convolutional Neural Networks	PlantVillage dataset	Computationally expensive
5	[18]	2021	Enhanced segmentation accuracy	Multi-objective image segmentation	Tea leaf disease dataset	Complex model implementation
6	[12]	2021	Better performance with deep learning	Comparison of ML vs DL	General plant datasets	ML underperformed against DL
7	[10]	2022	Varied performance depending on technique	Various traditional ML methods	Diverse leaf disease datasets	Manual feature selection required
8	[21]	2022	High accuracy for apple leaf diseases	CNN-based classification	Apple leaf disease dataset	Limited generalization across species
9	[17]	2022	Lightweight model with high precision	MGA-YOLO for apple leaves	Apple dataset	Limited scalability
10	[23]	2022	UAV imagery improved detection	Deep learning with UAV-based RGB images	Sugarcane dataset	Affected by environmental conditions
11	[26]	2022	Improved accuracy with multispectral imaging	Machine learning with UAV multispectral images	Sugarcane dataset	Expensive data collection
12	[27]	2022	Precise segmentation with smart spraying	Leaf segmentation with deep learning	Apple orchard dataset	Limited real-world validation
13	[19]	2022	Improved lesion segmentation	Deep learning with lesion detection	Plant lesion dataset	Limited dataset diversity
14	[22]	2023	Accurate detection of tiny targets	HSSNet deep learning model	Apple dataset	High computational requirement
15	[16]	2024	High accuracy with YOLOv4	YOLOv4-based object detection	Custom dataset	Limited to specific leaf diseases
16	[15]	2024	Robust performance with limited images	Transfer learning with MobileNetV2	Multiple plant disease datasets	Dependency on pre-trained networks
17	[24]	2024	Mobile-based real-time detection	CNN-based Android application	Mulberry leaf disease dataset	Limited to smartphone processing power
18	[20]	2024	Comprehensive AI-based disease detection	AI-driven plant disease classification	Multiple plant datasets	Lack of standardization

the field of plant disease detection. Shoaib et al. [19] proposed a deep learning-based model that added an additional mechanism of attention to improve the identification of diseases. Their work showed that transformer-based architectures will pay much attention to capture long-range dependencies in images and significantly improve the classification accuracy for complex datasets in plant diseases.

These hybrid and advanced architectures show that the combination of more than one methodology increases the overall robustness of leaf disease detection systems. However, their optimization for actual use in real-world scenarios, specifically concerning scalability, generality, and deployment on resource constrained devices remains challenging.

The summary of various approaches in terms of methodology, dataset used, results, and limitations is summarized in Table 1.

Challenges

Despite outstanding progress in machine learning and deep learning-based leaf disease detection, there are still a couple of other problems. With regard to the diversity of the dataset, the datasets available so far have been based on controlled environmental conditions. Those models are usually poor at generalization when moving from controlled environments. Research like that of Jafar et al. [20] highlighted that data sets provided in recent years often lack variations in lighting, background complexity, and disease severity, which makes them prone to overfitting.

Another serious challenge faced in the task is class imbalance. Some plant diseases are rare, so the dataset contains much fewer images of infrequent diseases than frequent ones. This imbalance affects the model performance because deep learning algorithms are biased toward the majority class. Vishnoi et al. [21] have attempted to overcome this bias by augmenting images of apple leaf disease with synthetic data generation techniques. Their results indicate that data augmentation improves classification accuracy, but continues to pose challenges in biological realism in synthetic images.

In addition to data issues, the main issue for real-time deployment of deep learning models in agricultural scenarios is computational complexity. Complex architectures like CNNs, ViTs, and even hybrid models have very high computation demands for both training and inference. Gao et al. [22] proposed an end-to-end network called HSSNet, used for tiny-target detection of apple leaf diseases against complex backgrounds. This method might increase detection accuracy but required significant computational power, thus hampering its use in mobile and edge computing applications.

Environmental factors also impact the robustness of the model. Changes in illumination, occlusions, and variation in leaf orientations can degrade the accuracy of the model. Amarasingam et al. [23] worked on UAV-based RGB imagery for the detection of white leaf disease in sugarcane crops. They pointed out the changing environmental conditions' impact on the model's performance. Their research suggested that the use of multispectral imaging and domain adaptation techniques could enhance the generalizability of the model.

Future directions

Recent advances in AI, edge computing, and multimodal data fusion have opened new avenues to improve leaf disease detection. Researchers are increasingly exploring federated learning and decentralized AI to enhance model generalization while preserving data privacy. Salam et al. [24] proposed a CNN-based smart Android application for mulberry leaf disease detection, demonstrating the potential of mobile-based AI solutions in precision agriculture. Their study suggested that integrating lightweight deep learning models with mobile computing could bridge the gap between research and practical deployment.

Another promising direction is the use of hyperspectral and multispectral imaging combined with deep learning. Meng et al. [25] introduced spectral disease indices to detect southern corn rust, utilizing spectral bands beyond the visible spectrum to improve the accuracy of the classification. Their findings highlighted that integrating spectral data with CNNs improves the detection of disease severity levels, making it a valuable approach for large-scale agricultural monitoring.

Moreover, the application of generative adversarial networks (GANs) for data augmentation has gained traction in addressing dataset limitations. Narmilan et al. [26] explored the use of GAN-generated synthetic multispectral images for the detection of white leaf disease in sugarcane. Their study demonstrated that GAN-based augmentation improved model robustness, particularly for rare disease classes.

Incorporating explainable AI (XAI) techniques is also a growing research focus. As deep learning models become more complex, understanding their decision-making processes is crucial for adoption in agriculture. Storey et al. [27] introduced a leaf disease segmentation model that enhances interpretability, making AI-based decisions more transparent for farmers and agronomists.

The following research opportunities are formulated from a literature survey in leaf disease detection:

1. *High-Accuracy Deep Learning Models for Leaf Disease Detection*: After the success of CNNs and attention-based networks [13, 14], researchers can further refine deep learning architectures to improve classification accuracy.

2. *Real-Time Disease Detection with Lightweight Architectures*: Real-time applications require computational efficiency [17, 15] therefore researchers are required to develop lightweight models for classification.

3. *Improving Generalization Across Plant Species and Environmental Variability*: Most existing models struggle with data set diversity and variations in lighting and background [20, 21]. Researchers can explore domain adaptation, data enhancement, transfer learning, and self-supervised learning techniques to improve model adaptability.

4. *XAI for Model Interpretability*: Since DL models are often considered black-box systems, integrating XAI methods [27, 19] will help to understand the importance of characteristics and make model decisions transparent to farmers and agricultural experts. Future studies can focus on integrating saliency maps, SHAP values, or attention-based visualizations to enhance model transparency.

5. *Addressing Class Imbalance in Leaf Disease Classification*: Many datasets contain an uneven distribution of disease classes, leading to biased model predictions [21, 26]. Researchers can investigate synthetic data generation, cost-sensitive learning, and novel loss functions to improve model fairness and accuracy.

6. *Optimizing Leaf Disease Detection for Edge Computing and IoT Devices*: Deploying deep learning models on resource-constrained devices is a growing area of interest [24, 15]. Researchers can explore model quantization, pruning, and federated learning to make disease detection more efficient for real-world agricultural applications.

Future research should focus on improving the efficiency of the model for real-time applications, integrating IoT-based disease monitoring systems, and improving the generalization between species. Balafas et al. [28] emphasized the need for unified frameworks that combine various AI techniques, ensuring scalability and adaptability in diverse agricultural environments.

Conclusion

The field of leaf disease detection has undergone significant advancements, evolving from traditional image processing techniques to sophisticated deep learning models capable of identifying and classifying plant diseases with high accuracy. Traditional approaches are based on handcrafted characteristics and provide initial insight into disease classification. However, it was limited by their inability to generalize across diverse datasets. Machine learning methods improved classification performance but required extensive feature engineering, while deep learning models revolutionized the field by automating feature extraction and enhancing classification accuracy.

Hybrid approaches and advanced architectures, such as YOLO-based models, attention mechanisms, and transformer-based networks, have further refined disease detection capabilities. These models have demonstrated improved precision, real-time processing, and robustness in complex agricultural environments. However, challenges remain in terms of dataset diversity, computational complexity, class imbalance, and environmental variability, which impact the real-world deployment of these models.

Recent innovations in federated learning, multispectral imaging, GAN-based augmentation, and XAI have opened new research avenues, offering potential solutions to existing challenges. Future research should focus on optimizing deep learning models for edge computing, improving cross-species generalization, and developing AI-driven decision support systems for farmers. The integration of IoT and cloud-based platforms could further enhance the scalability and accessibility of leaf disease detection solutions in precision agriculture.

As advancements continue, interdisciplinary collaboration between computer scientists, agronomists, and domain experts will be crucial in developing practical, efficient, and scalable plant disease detection systems. The ongoing research efforts highlight the potential of AI-driven methodologies in transforming modern agriculture, ensuring early disease detection, reducing crop losses, and promoting sustainable farming practices.

References

[1] T. Kodama and Y. Hata, "Development of classification system of rice disease using artificial intelligence", in 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Miyazaki, Japan, 2018, pp. 3699–3702, doi: 10.1109/SMC.2018.00626.

[2] A. Kumar, S. Sarkar, and C. Pradhan, "Recommendation system for crop identification and pest control technique in agriculture", in 2019 IEEE International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2019, pp. 0185–0189, doi: 10.1109/ICCSP.2019.8698099.

[3] T. A. Shaikh, T. Rasool, and F. R. Lone, "Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming", Computers and Electronics in Agriculture, vol. 198, p. 107119, 2022, doi: 10.1016/j.compag.2022.107119.

[4] D. Munjal, L. Singh, M. Pandey, and S. Lakra, "A systematic review on the detection and classification of plant diseases using machine learning", International Journal of Software Innovation (IJSI), vol. 11, pp. 1–255, 2023, doi: 10.4018/IJSI.315657.

[5] L. Keeling, H. Tunón, G. O. Antillón, C. Berg, M. Jones, L. Stuardo, J. Swanson, A. Wallenbeck, C. Winckler, and H. Blokhuis, "Animal welfare and the United Nations sustainable development goals", Frontiers Veterinary Sci., vol. 6, p. 336, Jan. 2019.

[6] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection", Frontiers in Plant Science, vol. 7, p. 1419, Sep. 2016.

[7] D. Shah, V. Trivedi, V. Sheth, A. Shah, and U. Chauhan, "ResTS: Residual deep interpretable architecture for plant disease detection", Information Processing in Agriculture, vol. 9, no. 2, pp. 212–223, Jun. 2022.

[8] A. Picon, M. Seitz, A. Alvarez-Gila, P. Mohnke, A. Ortiz-Barredo, and J. Echazarra, "Crop conditional convolutional neural networks for massive multi-crop plant disease classification over cell phone acquired images taken on real field conditions", Computers and Electronics in Agriculture, vol. 167, Dec. 2019, Art. no. 105093.

[9] S. Huang, G. Zhou, M. He, A. Chen, W. Zhang, and Y. Hu, "Detection of peach disease image based on asymptotic non-local means and PCN-NIPELM", IEEE Access, vol. 8, pp. 136421–136433, 2020.

[10] A. Kumar and R. Singh, "The Different Techniques for Detection of Plant Leaves Diseases", International Journal of Artificial Intelligence, vol. 9, no. 1, pp. 1–7, 2022.

[11] P. Sethy, N. Barpanda, A. Rath, and S. Behera, "Deep feature based rice leaf disease identification using support vector machine", Comput. Electron. Agric., vol. 175, p. 105527, 2020.

[12] R. Sujatha, J. Chatterjee, N. Jhanjhi, and S. Brohi, "Performance of deep learning vs machine learning in plant leaf disease detection", Microprocess. Microsystems, vol. 80, p. 103615, 2021.

[13] V. Tiwari, R. C. Joshi, and M. K. Dutta, "Dense convolutional neural networks based multiclass plant disease detection and classification using leaf images", Ecological Informatics, vol. 63, p. 101289, 2021.

[14] R. Karthik, M. Hariharan, S. Anand, P. Mathikshara, A. Johnson, and R. Menaka, "Attention embedded residual CNN for disease detection in tomato leaves", Applied Soft Computing, vol. 86, p. 105933, 2020.

[15] M. K. A. Mazumder et al., "A robust and light-weight transfer learning-based architecture for accurate detection of leaf diseases across multiple plants using less amount of images", Frontiers in Plant Science, vol. 14, p. 1321877, 2024.

[16] E. A. Aldakheel, M. Zakariah, and A. H. Alabdalall, "Detection and identification of plant leaf diseases using YOLOv4", Frontiers in Plant Science, vol. 15, p. 1355941, 2024.

[17] Y. Wang, Y. Wang, and J. Zhao, "MGA-YOLO: A lightweight one-stage network for apple leaf disease detection", Frontiers in Plant Science, vol. 13, p. 927424, 2022.

[18] S. Mukhopadhyay, M. Paul, R. Pal, and D. De, "Tea leaf disease detection using multi-objective image segmentation", Multimedia Tools and Applications, vol. 80, pp. 753–771, 2021.

[19] M. Shoaib et al., "A deep learning-based model for plant lesion segmentation, subtype identification, and survival probability estimation", Frontiers in Plant Science, vol. 13, p. 1095547, 2022.

[20] A. Jafar, N. Bibi, R. A. Naqvi, A. Sadeghi-Niaraki, and D. Jeong, "Revolutionizing agriculture with artificial intelligence: plant disease detection methods, applications, and their limitations", Frontiers in Plant Science, vol. 15, p. 1356260, 2024.

[21] V. K. Vishnoi, K. Kumar, B. Kumar, S. Mohan, and A. A. Khan, "Detection of apple plant diseases using leaf images through convolutional neural network", IEEE Access, vol. 11, pp. 6594–6609, 2022.

[22] X. Gao et al., "HSSNet: An end-to-end network for detecting tiny targets of apple leaf diseases in complex backgrounds", Plants, vol. 12, no. 15, p. 2806, 2023.

[23] N. Amarasingam, F. Gonzalez, A. S. A. Salgadoe, J. Sandino, and K. Powell, "Detection of White Leaf Disease in Sugarcane Crops Using UAV-Derived RGB Imagery with Existing Deep Learning Models", Remote Sensing, vol. 14, no. 23, p. 6137, 2022.

[24] A. Salam et al., "Mulberry Leaf Disease Detection Using CNN-based Smart Android Application", IEEE Access, 2024.

[25] R. Meng et al., "Development of spectral disease indices for southern corn rust detection and severity classification", Remote Sensing, vol. 12, no. 19, p. 3233, 2020.

[26] A. Narmilan, F. Gonzalez, A. S. A. Salgadoe, and K. Powell, "Detection of white leaf disease in sugarcane using machine learning techniques over UAV multispectral images", Drones, vol. 6, no. 9, p. 230, 2022.

[27] G. Storey, Q. Meng, and B. Li, "Leaf disease segmentation and detection in apple orchards for precise smart spraying in sustainable agriculture", Sustainability, vol. 14, no. 3, p. 1458, 2022.

[28] V. Balafas, E. Karantoumanis, M. Louta, and N. Ploskas, "Machine Learning and Deep Learning for Plant Disease Classification and Detection", IEEE Access, vol. 11, pp. 114352–114377, 2023.