IOSR Journal of Business and Management (IOSR-JBM)
e-ISSN: 2278-487X, p-ISSN: 2319-7668 PP 56-60
www.iosrjournals.org

Disease detection in Plant Leaves utilizing Machine Learning techniques: A comprehensive review

Amrita Kumari, Sakshi Devi, Aachal Patwari, and Ashish Saini
Computer Science & Engineering
Quantum University Roorkee, India

Abstract- *Agriculture is the prime source behind India's economic progress. Global food security and agricultural productivity are seriously threatened by plant diseases. These diseases affect the efforts, plans and economy of a country. This is very essential that these diseases should be identified accurately, effectively managed and controlled. A number of researchers are working in this area utilizing several Machine Learning (ML) techniques for prompt detection of leaf diseases. This paper presents the work done in this area highlighting the techniques/algorithms used for detecting diseases in plant. The paper also examines the usage of Convolution Neural Networks (CNNs), k-Nearest Neighbor (KNN), YOLOv5, YOLOv7, and other cutting-edge models in particular area. An in-depth analysis of past research is summarized.*

I. INTRODUCTION

Agriculture is said to be one of the primary area and source behind India's economic progress [1]. When selecting a crop, the farmer keeps in mind some factors like the type of soil, the climatic condition of the area, and the economic value of the product. Growing populations, everchanging weather patterns and political ambiguity forced the agricultural sectors to search for creative ways to increase food output. The primary reason affecting agricultural production is the loss due to diseases in crops, which leads to the approximate reduction in the output by 20-30 percent. Plant illnesses arise from the disruption of the plants' regular physiological processes, which results in the distinctive symptoms. The quality and productivity of the plants are significantly impacted by several diseases in plants [2], specifically on the leaves.

Plants are known to be significant because it is the most important source of energy for humans due to their nutritional, medicinal, and other qualities. Plant diseases can damage the plant at any point during crop farming, due to which farmers face a loss in yield and market value [3]. But still at many places, the primary means of diagnosing plant diseases depend on the visual observations of experts or experienced farmers. But such approach resulted in several shortcomings in large agricultural setups due to the lack of sufficient experts required for manual observation and it also consumes a lot of time [4]. This has led to the development of smart agricultural systems that results in easy and efficient identification of diseases in crops.

For plants, several AI-based technologies are currently used for disease classification and diagnosis. K-NN, logistic regression, support vector machines (SVMs), decision trees, and convolutional neural networks (CNNs) are among few of the most often utilized techniques.

Deep CNN is a kind of deep learning (DL) algorithm which is much useful to detect the nature of disease. A knowledge-sharing strategy called transfer learning cut down the amount of time, data, and computing power needed to build deep learning models. Many fields have employed transfer learning for making the prediction more accurate.

As, the physical examination do not yield trustworthy results for large plantation fields, developing an automated, accurate, and economical method for diagnosing plant diseases is crucial. The diseases in leaf being the reason for more loss, this study focuses on the techniques used for identifying leaf diseases. Such techniques include transfer learning models constructed with multiple CNN architectures, including VGG19, InceptionV3, ResNet50 and ResNet152V2 among others. This paper provides a deep analysis of several machine learning (ML) and DL models used for the detection of leaf diseases in various plants highlighting their accuracy and research gap to the best of authors' knowledge.

II. LITERATURE REVIEW

Several research works have been conducted which highlights the necessity of machine learning algorithms for identification of diseases in the plant leaves. Some of the previous literatures (mostly of 2024 and 2025) [1-21] have been studied and presented here for quick reference to the researchers working in this area.

In order to identify pumpkin leaf disease, Khandaker et al. [2] investigated the use of Explainable AI (XAI) techniques. ResNet50, ResNet101, DenseNet121, DenseNet169, DenseNet201, Xception, and InceptionResNetV2 were studied among the pre-trained CNN architectures. ResNet50 performed the best out of all of these. The model's decision-making process was visually represented and its interpretability was enhanced through the use of XAI techniques such as Grad-CAM, Grad-CAM++, Score-CAM, and Layer-CAM. ResNet50 outperformed the other models that were assessed, achieving the greatest accuracy of 90.5% along with comparable precision, recall, and F1-score metrics.

Using a modified depthwise CNN combined with squeeze-and-excitation (SE) blocks and improved residual skip connections, Ashurov et al. [4] presented a sophisticated deep learning method for plant disease identification. The model was trained on a large dataset that included several plant species and disease categories. It showed an accuracy of 98% and an F1 score of 98.2%. The findings highlight the model's effectiveness in identifying diseases in agricultural settings, facilitating prompt and well-informed crop protection decision-making.

In order to classify plant diseases, Roumeliotis et al. [5] investigated the combination of multimodal big language models, more especially, GPT-4o with CNNs. On apple leaf photos, their optimized GPT-4o model outperformed ResNet-50, which had an accuracy of 96.88%, with an accuracy of up to 98.12%. But GPT-4o's zero-shot performance was noticeably worse, underscoring the need for little instruction.

MobilePlantViT, a hybrid Vision Transformer (ViT) architecture created for resource-efficient plant disease categorization, was presented by Tonmoy et al. [6]. MobilePlantViT outperformed larger models such as MobileViTv1 and MobileViTv2, attaining test accuracies which ranged from 80% to over 99% across a variety of datasets with only 0.69 million parameters.

A customized CNN model was created by Oni and Prama [7] for the real-time detection of tomato leaf disease. Their approach outperformed models such as YOLOv5, MobileNetV2, and ResNet18, achieving an astounding accuracy of 95.2%. This demonstrated how customized CNN architectures can be used for particular crop disease detection applications.

In order to effectively capture multiscale spatial characteristics, Bhagat et al. [8] created a Lite-MDC architecture, which integrated Multi-kernel Depthwise Separable Convolutions (MDsConv). With just 2.2 million parameters, this design drastically lowers computing complexity. A 62% decrease over the conventional VGG16 architecture has been observed. The Lite-MDC model performed 94.14% accurately on the recently released pigeon pea dataset. It also performed well on additional datasets, including the Apple Leaf dataset (97.2%), Cassava dataset (86.4%), and PlantVillage dataset (99.78%). Additionally, the model demonstrated its appropriateness for real-time applications by achieving an inference speed of 34 frames per second (FPS) on an NVIDIA P100 GPU.

With an emphasis on improving the YOLOv8 architecture, the paper by Chin et al. [9] offers an improved deep learning method for detection and classification of plant disease. The researchers included a Coordinate Attention (CA) technique to enhance feature extraction and integrated the GhostNet module into the YOLOv8 backbone to lower computing complexity. Transfer learning was also used to improve model performance by utilizing prior information. When paired with the CA mechanism and transfer learning, the optimized YOLOv8 model obtained a detection accuracy of 72.2% on a small-scale plant disease dataset. With GhostNet and CA integration, the model achieved 69.3% accuracy without transfer learning.

Balaji et al. [10] used CNNs like MobileNet, InceptionNet, ResNet, and ResNeXt to provide a deep learning-based method for identification and categorization of plant disease. In order to minimize crop damage, the study highlighted the importance of early disease diagnosis in agriculture. Due to its cardinality feature, which makes it possible to learn intricate patterns, ResNeXt outperformed other models in the evaluation with an accuracy of 98.2%. The study emphasized how effective and lightweight CNN architectures may be used to diagnose plant diseases in real time, allowing for prompt adjustments to farming methods.

A deep learning-based method for managing the diseases in plants has been presented in [11] with an emphasis on the early identification and categorization of leaf diseases in potato, corn, and apple crops. By fine-tuning and adjusting hyperparameters, the researchers created an Enhanced CNN (E-CNN) architecture. To assess this model's efficacy, it was contrasted with other ML and pre-trained DL models. The refined E-CNN model detected fungal infections in all of the crops under study with a noteworthy accuracy of 98.17%.

The use of ML and DL techniques in crop disease identification has also been thoroughly examined in the study by Ngugi et al. [12]. It presented how well several ML and DL algorithms like CNNs, recurrent neural networks (RNNs), capsule networks (CapsNet) and vision transformers (ViT) perform in classifying and estimating the severity of diseases. The paper highlighted the shift from conventional techniques to sophisticated DL approaches, pointing out increases in efficiency and accuracy. It also discussed about issues including the necessity for different datasets, model interpretability, and data quality.

A thorough analysis of ML and DL methods for plant disease identification has also been provided by Jackulin and Murugavalli [13]. It highlighted how well CNNs work, frequently achieving accuracy over 95% in identifying plant illnesses from leaves' images. The function of image preprocessing and more conventional machine learning techniques like SVM and Random Forest has also been studied. It highlighted important issues including sparse datasets and different image circumstances. To improve agricultural sustainability, the authors support real-time detection systems and hybrid models.

The study by Demilie [14] assesses a number of ML and DL techniques, including CNNs, SVMs, and YOLOv7, for identifying plant diseases. The study discovered that lightweight models like MobileNetV2 and dilated CNNs were quite successful, attaining up to 99.5% accuracy using datasets like PlantVillage. For better classification, it also investigated hybrid models that combine triplet and cross-entropy loss. The study demonstrated how DL outperforms more conventional methods. They also discussed real-world applications such as smartphone integration for real-time detection.

The incorporation of IoT, ML and DL for detecting plant diseases at an early stage utilizing leaf analysis is reviewed in the study by Prasad and Thyagaraju [15]. It emphasizes how accurately CNNs and RNNs detect illnesses in their early stages so that prompt action can be taken. The application of transfer learning models boosts classification performance. The study put a strong emphasis on predicting accuracy and real-time monitoring, despite the lack of explicit accuracy measurements. The identification of soil-borne illnesses and the use of smart technologies to support sustainable agricultural methods are also covered in the article.

Table 1. Review of Previous Research Articles Showing Different ML/DL Techniques Used and Accuracy obtained.

Author's Name	Algorithm/Model Name	Accuracy Obtained	Research Gap
Khandaker, M. A. A., et al. (2025) [2]	CNN	ResNet50 obtained 90.5%	Small dataset
Ashurov, A. Y., et al. (2025) [4]	Depthwise CNN with squeeze and excitation	98%	Accuracy to be enhanced
Roumeliotis K. I., et al. (2025) [5]	CNN, LLM	98.12%	Multimodal approach integration with fine-tuning for better accuracy
Tonmoy, M. R., et al. (2025) [6]	Hybrid ViT	80% to over 99% across a variety of datasets	Generalization across different plant species and diseases
Oni, M. K., & Prama, T. T. (2025) [7]	Optimized CNN	95.2%	lack of data augmentation methods and difficulties managing intricate backdrops in practical situations.
Bhagat, S., et al. (2024) [8]	Lightweight CNN	94.14%	Requirement for additional verification for devices with limited resources.
Chin, P. W., et al. (2024) [9]	Deep learning models	72.2%	model robustness and real-time performance needs to be enhanced
Balaji, G. N., et al. (2025) [10]	CNN, VGG-16, ResNet-50	98.2% (plant disease)	model robustness across diverse plant species
Iftikhar, M., et al. (2024) [11]	Enhanced CNN	98.17%	Dataset to be expanded
Demilie, W. B. (2024) [14]	Various machine learning models	99.7% (fully connected CNN) 99.98% (ACO-CNN model) 99.66% (DenseNet169) 99.99% (Xception)	lack of diverse and representative datasets
Khalid, M. M., & Karan, O. (2024) [20]	Standard CNN and MobileNet	89% (CNN) 96% (MobileNet)	Need for real-time plant disease detection
Joshi, B., & Bhavsar, H. (2024) [21]	Enhance-Nightshade-CNN	95.23% for tomato, 97% for bell pepper and potato, and 100% for eggplant.	Larger dataset validation for Nightshade crop disease detection
Naeem, A. B., et al. (2023) [22]	VGG-16	98%	Need for better performance on cotton leaf diseases
Soeb, M. J. A., et al. (2023) [23]	YOLOv7 (YOLO-T)	96.3% (on tea leaves)	Limited cross-plant adaptability
Sladojevic S, et al. [25]	Deep Neural Networks (DNN)	An overall accuracy of 96.3% across all classes, with individual class accuracies ranging from 91% to 98%	Generalization across different plant species
Hu, Y., et al. (2023) [26]	Lightweight one-stage CNN with knowledge distillation	92.5% (on maize leaves)	Optimization for real-world, large-scale deployment
Towfek, S. K., & Khodadadi, N. (2023) [27]	Deep CNN with metaheuristic optimization	95.1% (on various crops)	Real-time deployment in diverse conditions

A comprehensive bibliometric analysis of research trends in apple leaf disease detection using ML is presented by Bonkra et al. [16]. The study offers insightful information about the development of research trends and the increasing significance of machine learning in agricultural diagnostics. Simhadri et al. [17] offers a thorough analysis which emphasizes on the application of DL models for rice disease diagnosis, such as YOLOv5x (accuracy up to 94.65%), MobileNetV2 (accuracy 84.21%), and EfficientNet-B0 (accuracy 99.8% accuracy). Among the methods covered are data augmentation, bespoke CNNs, and transfer learning. The study put a strong emphasis on mobile deployment, real-time detection, and model efficiency.

A thorough sugarcane leaf dataset with 6,748 tagged photos in 11 classes including 9 disease categories and healthy/dried leaves has been presented by Thite et al. [18]. A 50MP smartphone camera was used to take the photos in actual field settings, and the resolution was set to 768 x 1024. The dataset is meant to enable ML and DL applications for plant disease identification, however no particular deep learning model or accuracy is mentioned. Its importance for agricultural diagnostics, validation, and model training has been emphasized in the paper.

The paper by Omaye et al. [19] provided a thorough analysis with focus on ML approaches for crop disease detection. The authors emphasized how CNNs are the most effective method for recognizing plant diseases with high accuracy. The efficiency of diverse supervised and unsupervised machine learning models across a range of crops and disease types has also been noted in the review. The study mentioned a research deficit concerning prevention, control, and recovery, even though diagnostic capacities have been thoroughly investigated. It highlighted the necessity of end-to-end and more integrated machine learning frameworks. In order to improve practical deployment in agriculture, the study also recommends larger datasets and real-time applications.

The use of DL models, particularly CNNs and MobileNet, for the early and precise detection of plant illnesses is examined in the study by Khalid and Karan [20]. The CNN model obtained an F1-score of 96%, accuracy of 89%, and precision and recall of 96%. On the other hand, the MobileNet model showed a slightly lower precision (90%) and recall (89%), as well as a slightly lower F1-score (89%), although having a higher accuracy of 96%. Grad-CAM for explainable AI was also included in the study, offering visual perceptions into the decision-making process of the model. These results highlighted how deep learning, especially lightweight architectures like MobileNet, can improve plant disease detection techniques and advance sustainable farming methods.

The study by Joshi and Bhavsar [21] offered a CNN model that has been optimized for the purpose of identifying diseases in nightshade crops, including eggplants, tomatoes, bell peppers, and potatoes. In addition to introducing unique models like Nightshade-CNN and Enhanced Nightshade-CNN that provide a balance between computational efficiency and accuracy, the paper assessed popular architectures such as AlexNet, VGG, and GoogleNet. With lower layer counts and filter sizes, these customized models were able to attain disease classification accuracies of 93% to 95%, which qualified them for use in contexts with limited resources. The study highlighted that it is crucial to optimize model parameters to identify diseases effectively and accurately. This would promote better crop management and increase agricultural productivity.

A DL method for identifying cotton leaf illnesses using VGG-16 CNN architecture has been discussed by Naeem et al. [22]. Four categories, Fusarium Wilt, Cotton Leaf Curl Virus, Bacterial Blight and healthy leaves were used to train the model. The researchers used Adam and RMSProp optimizers to tune the model and used transfer learning to improve performance. The VGG-16 model classified cotton leaf diseases with an overall accuracy of 98%. However, the study emphasized the need for additional research to evaluate the model's efficacy in actual agricultural contexts and to enhance model generality across various environmental situations.

A sophisticated deep learning method for identifying tea leaf illnesses has been presented by Soeb et al. [23]. The researchers used a dataset of 4,000 manually annotated photos gathered from four well-known tea gardens in Bangladesh to train the system using the YOLOv7 model. With a detection accuracy of 97.3%, recall of 96.4%, precision of 96.7%, F1-score of 0.965 and mean average precision (mAP) of 98.2%, the model successfully recognized five different tea leaf illnesses. Comparative studies showed that YOLOv7 performed better in terms of accuracy and detection speed than earlier models like YOLOv5, CNNs, and AX-RetinaNet. The study highlighted that YOLOv7 can improve real-time disease monitoring, which can help with prompt interventions and lower financial losses in tea production.

A comprehensive analysis of the use of DL techniques to agricultural problems is presented by Kamilaris and Prenafeta-Boldú [24]. The authors mentioned that CNNs can be used for tasks including yield prediction, categorization, and crop disease detection. The study emphasized how DL models outperform conventional techniques, particularly when managing intricate, unstructured agricultural data. Table 1 presents previous research articles showing different ML/DL techniques used along with the model accuracy obtained.

III. CONCLUSION

This review article highlights the research conducted for identifying diseases in leaves of several plants highlighting mostly the research conducted in 2024 and 2025. Several machine learning and deep learning models have been mentioned which aim to improve the efficiency for detecting illness in plants. The paper also presents the accuracy obtained by several models along with the gap identified to the best of authors knowledge. Future work may include expanding the research and designing a model which aims to achieve higher accuracy addressing the gaps and taking care of the real time dataset.

REFERENCES

[1]. Kumari, A., Saini, A., & Kumar, A., (2024). A review on Agricultural Practices for Long-term Sustainability. *Engineering a Sustainable Future: Role of Science and Technology for Achieving SDGs Volume-I*, Walnut Publication. ISBN: 979-8-89171-176-1.

[2]. Khandaker, M. A. A., Raha, Z. S., Islam, S., & Muhammad, T. (2025). Explainable AI-Enhanced Deep Learning for Pumpkin Leaf Disease Detection: A Comparative Analysis of CNN Architectures. *arXiv preprint arXiv:2501.05449*.

[3]. Kumari, A., Anand, N., Saini, A., Kumar, A., & Sharma, S. (2025). Predictive Modeling of Maize Leaf Diseases Using Machine Learning Techniques. *International E-Conference on Emerging Applications of Material Science and Technology*, Feb. 14-15.

[4]. Ashurov, A. Y., Al-Gaashani, M. S., Samee, N. A., Alkanhel, R., Atteia, G., Abdallah, H. A., & Saleh Ali Muthanna, M. (2025). Enhancing plant disease detection through deep learning: a Depthwise CNN with squeeze and excitation integration and residual skip connections. *Frontiers in Plant Science*, *15*, 1505857.

[5]. Roumeliotis K. I., Sapkota, R., Karkee, M., Tselikas. N. D., Nasiopouloset, D. K. (2025). Plant Disease Detection through Multimodal Large Language Models and Convolutional Neural Networks. *arXiv preprint arXiv:2504.20419*.

[6]. Tonmoy, M. R., Hossain, M. M., Dey, N., & Mridha, M. F. (2025). MobilePlantViT: A Mobile-friendly Hybrid ViT for Generalized Plant Disease Image Classification. *arXiv preprint arXiv:2503.16628*.

[7]. Oni, M. K., & Prama, T. T. (2025). Optimized Custom CNN for Real-Time Tomato Leaf Disease Detection. *arXiv preprint arXiv:2502.18521*.

[8]. Bhagat, S., Kokare, M., Haswani, V., Hambarde, P., Taori, T., Ghante, P. H., & Patil, D. K. (2024). Advancing real-time plant disease detection: A lightweight deep learning approach and novel dataset for pigeon pea crop. *Smart Agricultural Technology*, *7*, 100408.

[9]. Chin, P. W., Ng, K. W., & Palanichamy, N. (2024). Plant disease detection and classification using deep learning methods: A comparison study. *Journal of Informatics and Web Engineering*, *3*(1), 155-168.

[10]. Balaji, G. N., Parthasarathy, G., & Kovendhan, A. K. P. (2025). Using Deep Learning for Plant Disease Detection and Classification.

[11]. Iftikhar, M., Kandhro, I. A., Kausar, N., Kehar, A., Uddin, M., & Dandoush, A. (2024). Plant disease management: A fine-tuned enhanced CNN approach with mobile app integration for early detection and classification. *Artificial Intelligence Review*, *57*(7), 167.

[12]. Ngugi, H. N., Ezugwu, A. E., Akinyelu, A. A., & Abualigah, L. (2024). Revolutionizing crop disease detection with computational deep learning: a comprehensive review. *Environmental Monitoring and Assessment*, *196*(3), 302.

[13]. Jackulin, C., & Murugavalli, S. J. M. S. (2022). A comprehensive review on detection of plant disease using machine learning and deep learning approaches. *Measurement: Sensors*, *24*, 100441.

[14]. Demilie, W. B. (2024). Plant disease detection and classification techniques: a comparative study of the performances. *Journal of Big Data*, *11*(1), 5.

[15]. Prasad, S. R., & Thyagaraju, G. S. (2024). Leaf analysis based early plant disease detection using Internet of Things, Machine Learning and Deep Learning: A comprehensive review. *Journal of Integrated Science and Technology*, *12*(2), 734-734.

[16]. Bonkra, A., Pathak, S., Kaur, A., & Shah, M. A. (2024). Exploring the trend of recognizing apple leaf disease detection through machine learning: a comprehensive analysis using bibliometric techniques. *Artificial Intelligence Review*, *57*(2), 21.

[17]. Simhadri, C. G., Kondaveeti, H. K., Vatsavayi, V. K., Mitra, A., & Ananthachari, P. (2024). Deep learning for rice leaf disease detection: A systematic literature review on emerging trends, methodologies and techniques. *Information Processing in Agriculture*.

[18]. Thite, S., Suryawanshi, Y., Patil, K., & Chumchu, P. (2024). Sugarcane leaf dataset: A dataset for disease detection and classification for machine learning applications. *Data in Brief*, *53*, 110268.

[19]. Omaye, J. D., Ogbuju, E., Ataguba, G., Jaiyeoba, O., Aneke, J., & Oladipo, F. (2024). Cross-comparative review of Machine learning for plant disease detection: Apple, cassava, cotton and potato plants. *Artificial Intelligence in Agriculture.*

[20]. Khalid, M. M., & Karan, O. (2024). Deep learning for plant disease detection. *International Journal of Mathematics, Statistics, and Computer Science*, *2*, 75-84.

[21]. Joshi, B., & Bhavsar, H. (2025) Efficient Nightshade Crop Leaf Disease Identification: An Optimized CNN Approach with Comprehensive Time and Space Complexity Analysis. *International Journal of Educational Research*, 143-152.

[22]. Naeem, A. B., Senapati, B., Chauhan, A. S., Kumar, S., Gavilan, J. O., & Abdel-Rehim, W. M. F. (2023). Deep learning models for cotton leaf disease detection with VGG-16. *International Journal of Intelligent Systems and Applications in Engineering*, *11*(2), 550-556.

[23]. Soeb, M. J. A., Jubayer, M. F., Tarin, T. A., Al Mamun, M. R., Ruhad, F. M., Parven, A., ... & Meftaul, I. M. (2023). Tea leaf disease detection and identification based on YOLOv7 (YOLO-T). *Scientific reports*, *13*(1), 6078.

[24]. Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and electronics in agriculture*, *147*, 70-90.

[25]. Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016). Deep neural networks based recognition of plant diseases by leaf image classification. *Computational intelligence and neuroscience*, *2016*(1), 3289801.

[26]. Hu, Y., Liu, G., Chen, Z., Liu, J., & Guo, J. (2023). Lightweight one-stage maize leaf disease detection model with knowledge distillation. *Agriculture*, *13*(9), 1664.

[27]. Towfek, S. K., & Khodadadi, N. (2023). Deep convolutional neural network and metaheuristic optimization for disease detection in plant leaves. *Journal of Intelligent Systems and Internet of Things*, *10*(1), 66-75.