International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 15 Issue 1, January-2026

Smart Agriculture Through Machine Learning: A Systematic Review Of Plant Disease Detection Models

Karthik Saiju[1], Yedukrishnan. S[2], Amal AP[3], Rahul T[4], Indu V Nair[5]

[1,2,3,4] B-Tech Students, Department of Electronics and Communication Engineering, UKF College of Engineering and Technology, Kollam, India
[5]Assistant Professor, Department of Electronics and Communication Engineering, UKF College of Engineering and Technology, Kollam, India

ABSTRACT

The application of machine learning in smart agriculture has gained considerable attention as a promising solution for automated plant disease detection and crop health monitoring. Several review studies across diverse research platforms have explored data-driven approaches that utilize image analysis, environmental sensing, and predictive modelling to support precision farming. This systematic review consolidates findings from these studies to examine the evolution of plant disease detection models, ranging from traditional machine learning classifiers to advanced deep learning architectures. Emphasis is placed on the role of convolutional neural networks, transfer learning strategies, and ensemble models, which consistently demonstrate improved accuracy and adaptability across multiple crop types and disease conditions. Furthermore, recent reviews highlight a growing trend toward integrating machine learning models with IoT-enabled sensing systems and edge computing platforms to enable real-time, field-level decision making Recent review studies also emphasize the increasing importance of lightweight and resource-efficient machine learning models for deployment in real-world agricultural environments. As many farming applications rely on mobile devices, unmanned aerial vehicles, and edge-based systems, researchers have focused on optimizing model complexity while maintaining acceptable detection accuracy.

Index Terms: Intelligent Farming Systems, Data-Driven Agriculture, Learning-Based Crop Analysis, Automated Plant Health Assessment, Disease Identification in Crops, Vision-Based Agricultural Monitoring, Neural Network Models for Agriculture, Image-Guided Disease Recognition, Computational Plant Pathology, Smart Crop Management, AI-Enabled Agricultural Practices, Sustainable Crop Protection.

INTRODUCTION

Agriculture remains a cornerstone of global economic development and food security, yet crop productivity is continually threatened by plant diseases that cause significant yield losses each year. Early and accurate identification of plant diseases is critical to minimizing damage, reducing unnecessary pesticide usage, and improving overall crop health. Traditionally, disease detection has relied on visual inspection by farmers or agricultural experts, a process that is time-consuming, subjective, and often infeasible in large-scale farming environments. These limitations have driven researches are encouraged to investigate automated, technology-enabled approaches that can effectively address the needs of modern systems agricultural practice. With the advancement of smart agriculture, machine learning (ML) and deep learning (DL) techniques have emerged as powerful tools for automating plant disease detection. Existing studies demonstrate that ML-based systems can analyze visual symptoms from leaf images and classify diseases with promising accuracy, offering scalable alternatives to manual diagnosis. Early approaches employed handcrafted feature extraction combined with classical classifiers such as support vector machines and decision trees, while more recent research has shifted toward convolutional neural networks that automatically learn discriminative features from large image datasets. By critically examining datasets, algorithms, evaluation metrics, and deployment considerations reported in prior studies, this review aims to provide a structured understanding of the current research landscape.

Problem Statement

Plant diseases are a persistent and severe challenge to global agricultural productivity, directly threatening food security, farmer livelihoods, and economic stability. Traditional disease identification methods—such as manual field inspection and laboratory diagnosis—are laborious, slow, and vulnerable to human error, particularly in resource-limited environments where expert agronomists may be unavailable. These conventional techniques often fail to provide the timely, scalable, and accurate detection needed to mitigate crop loss effectively, making automated approaches increasingly vital in modern agriculture. In response, research has increasingly focused on the application of machine learning (ML) and deep learning (DL) for automated plant disease detection, aiming to provide real-time, accurate, and cost-effective solutions. While these models show promise— especially convolutional neural networks (CNNs) and other DL architectures—their performance often remains constrained to controlled laboratory datasets and fails to extend effectively to complex, real-world agricultural environments. Most existing systems demonstrate high accuracy only under ideal conditions with uniform lighting, simple backgrounds, and well-curated image sets, leading to a significant gap between research outcomes and practical field deployment. A core challenge lies in the quality and diversity of training data. Many models rely on publicly available datasets that are limited in size, imbalance, and representation of real environmental variations such as lighting fluctuations, occlusions, seasonal changes, and multiple disease symptoms occurring simultaneously on a single plant. Additionally, the increasing complexity and computational demands of advanced ML/DL architectures pose further limitations. High-performing models often require substantial computational resources for training and inference, hindering their deployment on mobile devices or low-cost edge hardware commonly used in rural agricultural settings.

RELATED WORKS

[1] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. Frontiers in Plant Science, 7, 1419.

It has conducted a pioneering study on using deep learning for plant disease detection through image analysis. They utilized convolutional neural networks (CNNs), specifically AlexNet and GoogLeNet, to automatically classify images of plant leaves into healthy and diseased categories. The study employed the Plant Village dataset, which consisted of over 54,000 high-quality images covering 14 crop species and 26 different diseases. Their approach eliminated the need for manual feature extraction, as the CNNs could automatically learn and extract relevant features from leaf images. The models achieved high classification accuracy, demonstrating the potential of deep learning to provide fast, scalable, and reliable disease detection solutions. This work not only established a benchmark for subsequent research but also highlighted the feasibility of deploying AI-based systems for early disease diagnosis in smart agriculture.

[2] Ferentinos (2018). Examined the potential of convolutional neural networks in classifying plant diseases from leaf images, highlighting the advantages of deep learning over conventional approaches.

It has conducted a detailed study on the application of deep learning models for plant disease detection and diagnosis. The research focused on evaluating several state-of-the-art convolutional neural network (CNN) architectures, including AlexNet, VGG, and GoogLeNet, to classify images of plant leaves into healthy and diseased categories. The dataset used comprised thousands of high-resolution images covering multiple crops and a variety of common plant diseases. The study demonstrated that deep learning models could automatically extract complex and discriminative features from leaf images without the need for manual feature engineering, achieving significantly higher accuracy compared to traditional machine learning methods. Additionally, the research highlighted the influence of network depth, architecture choice, and data augmentation techniques on model performance. Ferentinos emphasized the practical implications of deploying CNN-based models in smart agriculture, noting their potential for rapid, scalable, and automated disease diagnosis in real- world farming scenarios. This work provided important insights into model selection, dataset preparation, and performance optimization, laying a foundation for future research in AI-driven plant disease detection.

[3] Akhter and Sofi (2024). Investigated decision-support mechanisms in precision agriculture by integrating Internet of Things (IoT) technologies with machine learning techniques, with particular emphasis on disease management in rice crops.

This research analyzed how IoT-enabled devices can be effectively coupled with machine learning algorithms techniques for decision-making in precision agriculture, specifically targeting rice disease management. The study highlighted how IoT-enabled sensors can continuously monitor key environmental parameters such as temperature, humidity, soil moisture, and leaf wetness, providing real-time data streams from the crop field. These data were then analyzed using machine learning algorithms to detect early signs of diseases, predict potential outbreaks, and recommend timely interventions such as irrigation adjustments or pesticide application. The proposed system aims to reduce manual monitoring efforts, minimize crop losses, and optimize the use of agricultural inputs. By integrating IoT with AI-driven analytics, the research demonstrates a practical pathway for enhancing the efficiency, accuracy, and sustainability of precision agriculture, especially for high- value crops like rice.

[4] S. Condran, M. Bewong, M. Z. Islam, L. Maphosa, and L. Zheng, "Machine learning in precision agriculture: A survey on trends, applications and evaluations over two decades," IEEE Access, vol. 10, pp. 73786–73803, 2022, doi: 10.1109/ACCESS.2022.3188649.

It has conducted an extensive survey examining the evolution and impact of machine learning in precision agriculture over the past two decades. The study systematically reviewed various applications, including crop yield prediction, plant disease and pest detection, soil and water quality monitoring, and resource optimization. It analyzed both traditional machine learning techniques, such ranging from classical algorithms such as support vector machines and decision trees to advanced deep learning models, including convolutional and recurrent neural networks. The survey also assessed evaluation methodologies, performance metrics, and dataset characteristics used in prior studies. By highlighting trends in algorithm adoption, data collection strategies, and practical deployment challenges, the authors identified key research gaps such as limited real-world testing, lack of standardized datasets, and underexplored integration of multimodal data. The study emphasizes the importance of combining robust ML techniques with IoT and sensor technologies to enhance the efficiency, sustainability, and scalability of precision agriculture systems

[5] A. Devikarani, G. V. D. Jyothi, J. Divya Lalitha Sri, and B. S. Kiruthika Devi, "Towards smart agriculture using machine learning algorithms," in Proc. 2019 Int. Conf. on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2019, pp. 1–6, doi: 10.1109/ICACCS.2019.8728416.
It investigated the use of machine learning (ML) algorithms to facilitate smart agriculture and enhance decision-making processes in farming. The study reviewed and implemented various ML techniques for multiple agricultural tasks, including crop yield prediction, disease detection, soil and moisture monitoring, and irrigation management. The authors emphasized how ML models can process large volumes of sensor and image data collected from farms to provide actionable insights, reduce manual monitoring, and enable timely interventions. The paper also discussed practical challenges, such as the dependence on high-quality and diverse datasets, the need for computational resources for training and inference, and the difficulty of generalizing models across different crops and environmental conditions. Overall, the work highlighted the potential of integrating machine learning with IoT and data-driven techniques to create intelligent, automated, and sustainable agricultural systems, laying a foundation for future research in smart farming.

Research Gaps Identified In Existing Works

One of the most consistently noted gaps in plant disease detection research is the lack of large, diverse, and representative datasets. Despite the significant progress demonstrated by Mohanty et al. (2016) in applying deep learning models like AlexNet and GoogLeNet to the PlantVillage dataset, a key limitation lies in the lack of real-world variability in the training data. The dataset consists of high- quality leaf images captured under controlled conditions, which means that models trained on it may not perform well in field environments where lighting, background clutter, overlapping leaves, and other environmental factors introduce noise. This indicates a clear gap in developing robust models capable of generalizing to diverse and complex real-world agricultural scenarios. Ferentinos (2018) extended the evaluation of deep learning models by assessing multiple CNN architectures on larger datasets and employing transfer learning to improve accuracy. However, the study also highlighted that despite high accuracy on benchmark datasets, most models are limited in handling multiple simultaneous diseases or rare disease instances, which are common in practical farming. Additionally, the computational demands of deeper networks can restrict deployment on edge devices or mobile platforms, pointing to a gap in resource-efficient and field-deployable solutions.

The work by Akhter and Sofi (2024) integrated IoT sensors with machine learning for rice disease detection, addressing real-time monitoring and decision-making. Yet, this approach largely focuses on a single crop and specific environmental parameters, leaving limited generalization to other crops or broader disease types. There is also a need for comprehensive frameworks that combine visual image-based analysis with sensor data for multi-crop disease detection, indicating a gap in multimodal and scalable solutions for precision agriculture. Condran et al. (2022) surveyed machine learning applications in agriculture over two decades, identifying trends, evaluation methods, and challenges. Their survey reveals that although AI techniques have evolved substantially, there is a lack of standardized, large- scale, and publicly available datasets that can support benchmarking and comparison across studies. Furthermore, limited attention has been given to explainability and interpretability of models, which are crucial for farmer adoption and trust in AI-driven recommendations. Finally, Devikarani et al. (2022) discussed practical applications of machine learning for crop management, disease detection, and resource optimization.Collectively, these studies indicate that while deep learning and IoT-based approaches have shown strong potential, there remains a research gap in developing robust, scalable, interpretable, and multimodal plant disease detection systems that are applicable across diverse crops, diseases, and environmental conditions in real-world agriculture.

COMPARATIVE ANALYSIS

The application of machine learning (ML) and deep learning (DL) in smart agriculture, particularly for plant disease detection, has evolved significantly over the past decade. Early studies, such as Mohanty et al. (2016), primarily focused on applying convolutional neural networks (CNNs) to controlled datasets like PlantVillage, demonstrating that deep learning could achieve high accuracy in classifying plant diseases without manual feature extraction. However, these models were largely limited to ideal laboratory conditions and did not account for real-world field variability, which restricts their practical deployment. Building on this foundation, Ferentinos (2018) compared multiple deep learning architectures, including AlexNet, VGG, and GoogLeNet, highlighting the importance of network depth and architecture selection in improving disease classification accuracy. This study further confirmed that deep learning models outperform traditional machine learning algorithms, yet it also emphasized that model performance depends heavily on dataset quality and preprocessing techniques. Subsequent research, such as Akhter and Sofi (2024), introduced the integration of Internet of Things (IoT) with machine learning to provide real-time monitoring and decision-making in precision agriculture, particularly for rice disease management. Combining IoT-enabled data collection with ML/DL methods shows promise for overcoming these limitations by enabling real- time monitoring and predictive analytics, thus providing actionable insights for farmers. Overall, the comparative analysis indicates a clear trend: research is moving from controlled image-based classification toward integrated, data-driven, real-time smart agriculture systems. While progress has been significant, key gaps remain in model generalization to diverse environments, deployment on resource-limited devices, and interpretability for practical adoption. Addressing these challenges is critical for developing robust, scalable, and field-ready plant disease detection solutions.

OBJECTIVES

The main objective of this study is to systematically review the application of machine learning and deep learning techniques for plant disease detection in smart agriculture. It aims to analyze state-of- the-art algorithms and architectures, examining their effectiveness in identifying various crop diseases and understanding the factors that influence model performance. A critical aspect of the review is to evaluate the datasets, preprocessing methods, and feature extraction techniques used in prior studies, highlighting their role in model accuracy and generalization. The study also seeks to identify challenges and limitations in current approaches, including issues related to real-world data variability, computational requirements, multi-disease detection, and interpretability of models. Furthermore, it aims to compare traditional machine learning methods with advanced deep learning models to assess their suitability for practical agricultural applications. Another important goal is to explore the integration of emerging technologies, such as IoT and edge computing, with machine learning models to enable real-time disease monitoring, early warning systems, and automated decision-making. Finally, the review intends to provide insights and recommendations for future research, focusing on developing robust, scalable, and field-ready plant disease detection systems that can be effectively deployed across diverse agricultural environments.

FUTURE SCOPE

The future scope of machine learning-based plant disease detection in smart agriculture is vast and promising. One key area is the development of robust and generalized models capable of performing accurately under real-world field conditions, including variations in lighting, background, and multiple simultaneous disease symptoms. Expanding datasets to cover a wider range of crops, diseases, and geographic regions will enhance model adaptability and applicability. Integration of multimodal data from IoT sensors, drones, and satellite imagery offers another avenue for advancement, enabling predictive analytics and early warning systems that combine visual, environmental, and soil-related information for more precise disease management.

The deployment of lightweight and efficient models on mobile and edge devices can make these technologies accessible to smallholder farmers, providing real-time disease detection and decision support without the need for high computational resources. There is also significant potential in incorporating explainable AI (XAI) to make models more transparent and interpretable, thereby increasing trust among farmers and agronomists. Additionally, combining machine learning with automation technologies, such as robotic sprayers and precision irrigation systems, can create fully integrated smart farming solutions that optimize resource use, reduce pesticide dependency, and increase crop yield. Finally, future research can focus on sustainable and scalable deployment strategies, including cloud-based platforms and community-driven datasets, which will facilitate widespread adoption of AI-powered plant disease detection systems. Overall, these advancements will contribute to the evolution of precision agriculture, enhancing productivity, sustainability, and food security globally.

CONCLUSION

Machine learning and deep learning techniques have shown remarkable potential in transforming plant disease detection within the framework of smart agriculture. Through automated analysis of plant images and integration with sensor data, these technologies offer faster, more accurate, and scalable solutions compared to traditional manual methods. Studies reviewed in this systematic analysis demonstrate that convolutional neural networks, advanced deep learning architectures, and IoT-enabled decision-making systems can significantly enhance disease diagnosis, crop monitoring, and resource management. Despite notable improvements, several unresolved issues remain, including the necessity for diverse and representative datasets, model generalization under real-world field conditions, computational efficiency, and interpretability of predictions. Addressing these gaps is essential to enable practical, real-time deployment of AI-powered disease detection systems in agriculture. The future of smart agriculture lies in developing integrated, data-driven solutions that combine machine learning, IoT, edge computing, and automation technologies to create robust, accessible, and sustainable farming systems. By overcoming current limitations and leveraging emerging technologies, AI-based plant disease detection has the potential to improve crop productivity, reduce limit agricultural losses, optimize the use of available resources, and help ensure food security worldwide.

REFERENCES

[1] Mohanty et al., "Using Deep Learning for Image-Based Plant Disease Detection" (2016) — Introduced CNNs (AlexNet, GoogLeNet) on the PlantVillage dataset (54k images) to classify 14 crops / 26 diseases, reporting very high accuracy and demonstrating feasibility of image-based diagnosis.

[2] Ferentinos, "Deep learning models for plant disease detection and diagnosis" (Computers & Electronics in Agriculture, 2018) — Systematic evaluation of many CNN architectures on large datasets (≈87k images); shows transfer learning and deeper nets give strong accuracy

[3] R. Akhter and S. A. Sofi, "IoT- and machine learning-based decision support for precision agriculture with emphasis on rice disease management," in Proceedings of the 2024 International Conference on Decision Aid Sciences and Applications (DASA), IEEE, 2024.

[4] S. Condran, M. Bewong, M. Z. Islam, L. Maphosa, and L. Zheng, "A two-decade review of trends, applications, and evaluation methods of machine learning in precision agriculture," IEEE Access, vol. 10, pp. 73786–73797, 2022.

[5] A.Devikarani, G. V. D. Jyothi, J. Divya Lalitha Sri, and B. S. Kiruthika Devi, "Towards smart agriculture using machine learning algorithms," in Proc. 2022 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), IEEE, 2022.

[6] A. Kumar, A. Raj, K. Singh, R. Sarkar, and B. T. V. Murthy, "A solar-powered multipurpose agricultural robot incorporating leaf disease identification," in Proceedings of the 2023 International Conference on Smart Systems for Applications in Electrical Sciences (ICSSES), IEEE, 2023.

[7] F. Weber, G. Rosa, F. Terra, A. Oldoni, and P. Drews-Jr., "A cost-efficient mobile and computer vision–based system for optimized pesticide application," in Proceedings of the 2018 Latin American Robotic Symposium (LARS), the 2018 Brazilian Symposium on Robotics (SBR), and the 2018 Workshop on Robotics in Education (WRE), IEEE, 2018.

[8] V. Riyanto et al.(2025) offer a thorough review of deep learning approaches employed for plant disease diagnosis, outlining model types, datasets, evaluation strategies, and future research opportunities.

[9] E. Hirani, V. Magotra, J. Jain, and P. Bide, "Plant Disease Detection Using Deep Learning," in 2021 6th Int. Conf. Converg. Technol. (I2CT), pp. 1–4, 2021, doi: 10.1109/I2CT51068.2021.9417910.

[10] J. G. A. Barbedo, "A review of key challenges in automated plant disease detection based on Visible Range Images," Biosystems Eng., vol. 144, pp. 52-60, 2016.

[11] M. B. Ahmad Supian, H. Madzin, and E. Albahari, "Plant Disease Detection and Classification Using Image Processing Techniques a review," in Proc. 2nd Int. Conf. Appl. Eng. (ICAE), Batam, Indonesia, 2019.

[12] Arsenovic et al., "Studies on dataset expansion and augmentation (2019–2021) for plant disease detection, involving expansions of the PlantVillage dataset with more images and in-field photographs, and employing GANs and data augmentation techniques to enhance model generalization to real-world conditions.

[13] ETASR / Madhu et al., "Comprehensive analysis of YOLO-based deep learning for cotton leaf disease" (2025) — Focused YOLO study on cotton, combining preprocessing and postprocessing to raise detection accuracy in dense foliage.

[14] GitHub repositories and applied implementations, such as mzakariah's plant-disease-detection- using-yolov4, provide reproducible codebases, training pipelines, dataset annotation formats for YOLO, and inference scripts for YOLOv4, v5, and v8 on PlantVillage datasets.

[15] N. Sulaiya and S. Banerjee, "Review of convolutional neural network approaches for plant leaf disease detection," Int. J. Adv. Res. Multidiscip. Trends, vol. 2, no. 1, pp. 642–653, 2025.