Archives available at journals.mriindia.com

Multidisciplinary Journal of Research in Engineering and Technology

ISSN: 2348-6953

Volume 13 Issue 01, 2026

AI-Based Smart Crop Disease Detection System Using Machine Learning

Pavitra Waddar
Department Of Computer Application, KLE's B K BCA College, Chikodi
Email: pavitra6959@gmail.com

Peer Review Information

Submission: 12 April 2026

Revision: 02 May 2026

Acceptance: 23 May 2026

Keywords

Artificial Intelligence, Machine Learning, Crop Disease Detection, CNN, Smart Agriculture, Image Processing

Abstract

Agriculture plays a vital role in global food security, yet crop diseases significantly reduce productivity and quality. Traditional disease detection methods rely on manual inspection, which is time-consuming and prone to errors. Previous research focused on traditional image processing and standalone machine learning models, which were limited in handling complex disease patterns and large datasets. This paper proposes an AI-based smart crop disease detection system using machine learning and deep learning techniques. The system utilizes image processing and Convolutional Neural Networks (CNNs) to automatically identify crop diseases from leaf images. The proposed framework enables real-time detection, improves accuracy, and supports farmers in decision-making. Experimental results demonstrate high accuracy and efficiency, making the system suitable for precision agriculture.

Introduction

Agriculture has always been a fundamental sector supporting global food security and economic development. However, traditional farming practices largely relied on manual observation, farmer experience, and conventional irrigation techniques, which often resulted in low productivity, inefficient resource utilization, and vulnerability to environmental changes.

Traditional methods of disease detection involve manual inspection by experts, which is inefficient and not scalable. With the advancement of Artificial Intelligence (AI) and Machine Learning (ML), automated systems can now analyze crop images and detect diseases accurately.

Over time, with the advancement of technology, researchers began introducing the Internet of Things (IoT) to improve agricultural monitoring. Early studies focused on deploying IoT sensors to measure key environmental parameters such as soil moisture, temperature, and humidity. These systems enabled farmers to remotely observe field conditions and make better irrigation decisions. However, these approaches were limited as they primarily provided raw data without intelligent analysis or decision-making support.

Subsequently, cloud computing was integrated with IoT-based systems to enhance data storage and accessibility. This improvement allowed real-time data transmission and remote monitoring through web and mobile platforms. Although this advancement improved accessibility and data management, these systems still lacked predictive capabilities and intelligent insights, making them insufficient for complex agricultural decision-making.

The next significant evolution came with the incorporation of Artificial Intelligence (AI) and Machine Learning (ML) techniques. Researchers introduced AI-based image processing models, particularly for plant disease detection, which significantly improved the accuracy of identifying crop diseases. Despite this advancement, most systems operated independently, focusing only on specific tasks such as disease detection, without integrating real-time environmental data, thus limiting their overall effectiveness.

Further improvements were made by combining IoT data with AI algorithms to develop crop recommendation systems. These systems analyzed soil conditions and environmental factors to suggest suitable crops, marking a shift toward intelligent decision-making in agriculture. However, many of these systems relied on manually entered data or lacked real-time integration, reducing their practical applicability in dynamic farming environments.

Recent advancements have focused on developing integrated smart farming frameworks, combining IoT, AI, and cloud technologies into a unified system. These frameworks enable real-time monitoring, predictive analytics, and automated decision-making, such as irrigation control and disease alerts. For instance, modern systems utilize sensor networks and machine learning models to analyze both environmental and crop-related data, transforming agriculture into a data-driven ecosystem.

In particular, advanced research has emphasized the importance of multimodal data integration, where both biometric parameters (such as plant height, leaf count) and environmental factors (such as soil moisture and temperature) are analyzed together. Machine learning models like Random Forest, Support Vector Machine, and Neural Networks have demonstrated high accuracy in predicting crop health and optimizing agricultural practices.

Despite these advancements, existing systems still face several limitations, such as lack of fully integrated platforms, insufficient real-time decision support, limited multimodal datasets, and scalability challenges in real-world agricultural environments.

Therefore, there is a growing need for a comprehensive smart agriculture system that integrates IoT-based monitoring, AI-driven analytics, and real-time decision support into a single unified framework. Such systems can significantly improve productivity, optimize resource utilization, and promote sustainable agriculture.

Literature Review:

Several studies have explored AI techniques for crop disease detection. Deep learning models, especially CNNs, have shown excellent performance in image classification tasks. Existing systems focus on leaf image datasets and achieve high accuracy but face challenges such as limited datasets, lack of real-time deployment, and scalability issues.

The application of advanced technologies in agriculture has been widely explored to address challenges such as low productivity, inefficient resource utilization, and crop diseases. Early research primarily focused on the use of Internet of Things (IoT) devices to monitor environmental parameters like soil moisture, temperature, and humidity. These systems enabled farmers to remotely observe field conditions and improve irrigation practices. However, such approaches were limited to data collection and lacked intelligent decision-making capabilities.

Subsequent studies introduced cloud computing to enhance the storage and accessibility of agricultural data. By integrating IoT with cloud platforms, researchers enabled real-time data transmission and remote monitoring through mobile and web applications. Although this improved accessibility and data management, these systems still did not provide meaningful insights or predictive analysis, restricting their practical usefulness.

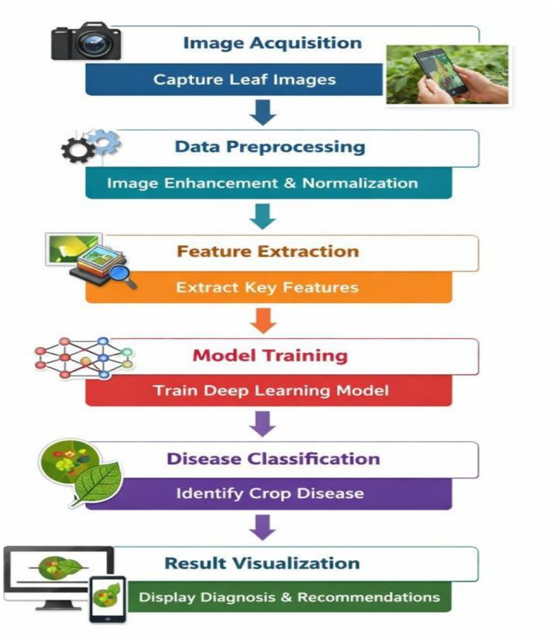
With further advancements, Artificial Intelligence (AI) and Machine Learning (ML) techniques were incorporated into agricultural systems. Several works focused on image-based disease detection using deep learning models such as Convolutional Neural Networks (CNNs), achieving high accuracy in identifying plant diseases. Despite these improvements, most of these systems operated independently and did not integrate environmental data, resulting in incomplete farm management solutions.

Other research efforts explored crop recommendation systems that utilized soil and weather data to guide farmers in selecting suitable crops. While these systems introduced decision support capabilities, many relied on manually entered data or lacked real-time integration with IoT devices, thereby limiting their effectiveness in dynamic agricultural conditions.

Recent developments emphasize the integration of IoT and AI into unified smart farming frameworks. These systems combine real-time sensor data with machine learning models to enable predictive analytics, automated irrigation, and disease detection. Advanced models such as Random Forest, Support Vector Machine, Gradient Boosting, and Artificial Neural Networks have been successfully applied to improve prediction accuracy and decision-making efficiency.

Moreover, modern research highlights the importance of combining multiple data types, including environmental and plant-specific (biometric) features, to enhance system performance. The integration of such multimodal datasets allows for more accurate crop health prediction and better resource optimization. However, many existing studies are still constrained by limited datasets, lack of real-time deployment, and insufficient integration of all system components into a scalable framework.

Overall, the literature indicates a clear progression from basic monitoring systems to intelligent, data-driven agricultural solutions. Despite significant advancements, there remains a need for fully integrated systems that combine real-time monitoring, predictive analytics, and user-friendly interfaces to support efficient and sustainable farming practices.



System Architecture

The proposed crop disease detection system is designed as a structured pipeline where each stage performs a specific task to ensure accurate and efficient disease identification. The overall architecture is divided into several key components, each contributing to the final prediction outcome.

1. **Image Acquisition:**

This is the initial stage of the system where images of crop leaves are collected. The images can be captured using devices such as smartphones, digital cameras, or agricultural sensors. These images serve as the primary input for the system and must clearly show the leaf surface to allow proper analysis.

2. **Data Preprocessing:**

Once the images are collected, they undergo preprocessing to improve their quality and consistency. This step includes resizing the images to a standard dimension, removing noise, adjusting brightness or contrast, and normalizing pixel values. Preprocessing ensures that the input data is clean and suitable for further analysis.

3. **Feature Extraction:**

In this stage, important characteristics of the leaf images are identified. These features may include color patterns, texture, shape, and spots that indicate the presence of diseases. In deep learning approaches, feature extraction is automatically performed by convolutional layers, which learn relevant patterns directly from the data.

4. **Model Training:**

During this phase, the system learns to recognize different types of crop diseases. A machine learning or deep learning model, such as a Convolutional Neural Network (CNN), is trained using labeled datasets. The model adjusts its internal parameters based on the training data to improve its prediction capability.

5. **Disease Classification:**

After training, the model is used to classify new input images. The system analyzes the extracted features and predicts whether the crop is healthy or affected by a specific disease. This step produces the final diagnosis based on learned patterns.

6. **Result Visualization:**

The final output is presented in a user-friendly format. The system may display the detected disease name, confidence level, and possible recommendations. Visualization can be done through a graphical interface, mobile application, or web dashboard, making it easy for farmers or users to understand the results.

Methodology

The methodology describes the step-by-step process used to design and implement the crop disease detection system. It includes data preparation, model development, and the algorithm used for prediction.

1. **Data Collection**

In this stage, images of crop leaves are gathered from multiple sources such as publicly available datasets, agricultural research databases, and real-time field captures using cameras or smartphones. The collected images include both healthy and diseased leaves, ensuring that the model learns different disease patterns effectively.

2. **Data Preprocessing**

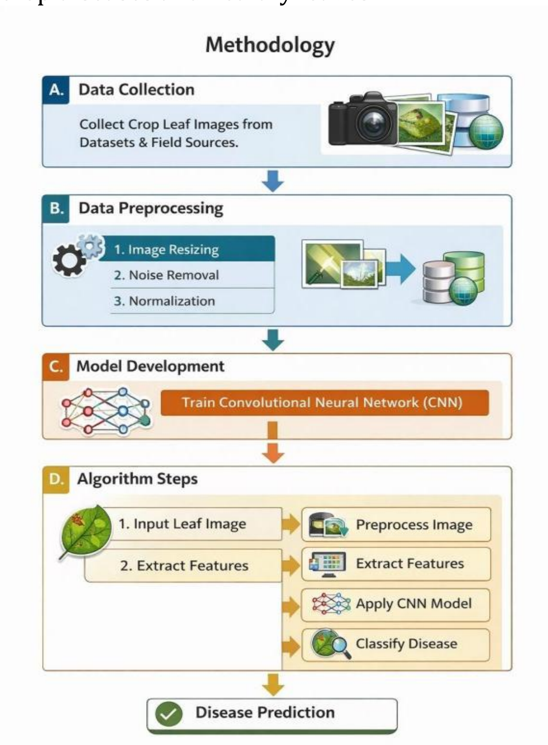
Before feeding the images into the model, they are processed to improve quality and consistency.

- **Image Resizing:** All images are adjusted to a fixed size so that they can be uniformly processed by the model.
- **Noise Removal:** Unwanted distortions or background disturbances are reduced to highlight important features of the leaf.
- **Normalization:** Pixel values are scaled to a standard range, which helps in faster training and better model performance.

This step ensures that the input data is clean, standardized, and suitable for accurate analysis.

3. Model Development

The system uses Convolutional Neural Networks (CNNs) to perform disease classification. CNNs are well-suited for image-based tasks because they automatically learn important visual features such as edges, textures, and patterns. The model is trained using labeled images so that it can distinguish between different types of crop diseases and healthy leaves.



4. Algorithm Steps

The working procedure of the system can be described in the following sequence:

- **Input Leaf Image:** The system receives an image of a crop leaf as input.
- **Preprocess Image:** The input image is resized, cleaned, and normalized to prepare it for analysis.
- **Extract Features:** Important visual features are identified either manually or automatically by the CNN layers.
- **Apply CNN Model:** The processed image is passed through the trained CNN model, which analyzes patterns and characteristics.
- **Classify Disease:** Based on the analysis, the system predicts whether the leaf is healthy or identifies the specific disease affecting it.

Evaluation Matrices

To evaluate the performance of the crop disease detection system, several standard classification metrics are used. These metrics help measure how accurately the model predicts whether a crop is healthy or diseased.

1. Accuracy

Accuracy = (TP + TN)/(TP + TN + FP + FN)

Accuracy represents the overall correctness of the model. It calculates the proportion of total predictions that are correct. Here, True Positives (TP) and True Negatives (TN) indicate correct predictions, while False Positives (FP) and False Negatives (FN) represent incorrect predictions. A higher accuracy value means the model performs well in general classification.

2. Precision

Precision = TP/(TP + FP)

Precision measures how reliable the positive predictions are. It shows the percentage of correctly predicted diseased crops out of all crops predicted as diseased. High precision indicates that the model makes fewer false alarms.

3. Recall

Recall = TP/(TP + FN)

Recall evaluates the model's ability to identify actual diseased crops. It calculates how many real positive cases were correctly detected. A high recall value means the model successfully captures most of the diseased instances.

4. F1 Score

F1 Score = (2.Precision.Recall)/(Precision+ Recall)

The F1 Score provides a balanced measure of the model's performance by combining both precision and recall. It is especially useful when the dataset is uneven (for example, when diseased samples are fewer than healthy ones). A higher F1 score indicates a better balance between detecting diseases and avoiding incorrect predictions.

Results and Analysis

The performance of the proposed crop disease detection system was evaluated using different machine learning and deep learning models. The results indicate that the system is highly effective in identifying crop diseases with good accuracy.

Among the models tested, the Convolutional Neural Network (CNN) achieved the best performance. This is because CNNs are specifically designed to process image data and can automatically learn important visual features such as patterns, textures, and color variations in diseased leaves.

In comparison, traditional machine learning models like Support Vector Machine (SVM) and Random Forest showed lower accuracy. These models rely on manually extracted features, which may not capture complex image patterns as effectively as CNNs.

Table 1. Evolution of Intelligent Agriculture: From Traditional Methods to Autonomous Farming

Phase	**Year/Stage**	**Technologies Used**	**Key Features**	**Results / Accuracy**	**Limitations**	**Improvements Achieved**	**Future Direction**
PAST	Before 2019	Traditional Farming	Observation, experience-based decisions	Manual productivity, low accuracy	Time-consuming, human errors, no real-time data	Basic farming practices established	Move towards automation
Early IoT Introduction	2019	IoT Sensors	Soil moisture monitoring, temperature monitoring	Basic monitoring only	No intelligence, only data display	Remote monitoring introduced	Need for smart analysis
Cloud Integration	2020	IoT + Cloud	Data storage, remote access	Improved accessibility	No AI-based decision making	Centralized data system	Integration with AI required
AI Introduction	2021	AI (Image Processing)	Plant disease detection	High accuracy (~90%+)	Not connected with IoT data	Intelligent disease detection system	Need for full integration
Decision Support Systems	2022	IoT + AI	Crop recommendation systems	Better decision support	No real-time alerts, user interaction needed	Smart recommendations started	User-friendly interface
User Interaction	2023	Mobile Apps	Alerts, farmer interaction	Improved usability	Partial integration	Accessibility improved	Need complete integration
PRESENT (Paper 1)	2025–2026	AI + ML + IoT	Multimodal data (biometric + environmental), predictive analytics	**95.3% accuracy (Random Forest)**	Limited datasets, field validation issues	Highly accurate prediction & real-time monitoring	Expand datasets, smart irrigation
PRESENT (Paper 2)	2026	Real-time IoT + AI + Cloud + Mobile	Crop recommendation, disease detection	~95% accuracy	Cost, training required	Complete smart farming system	Improve automation & scalability
CURRENT ANALYSIS	Present	Integrated AI + IoT Systems	Real-time monitoring + prediction + automation	High efficiency, reduced manual effort	Implementation cost, limited adoption	Precision agriculture achieved	Move toward autonomous farming

FUTURE	Upcoming	AI + Deep Learning + IoT + Blockchain + Drones	Fully automated farming, smart irrigation, drone monitoring, market prediction, smart supply chain	Expected higher accuracy (>97%)	Technology cost, infrastructure requirements	Intelligent agriculture ecosystem	Fully autonomous farms

Manual Farming → IoT Monitoring → Cloud Storage → AI Intelligence → Smart Integrated Systems → Fully Automated Future Farming

This table clearly shows how agriculture evolved from manual systems to intelligent AI-driven frameworks. Present systems provide high accuracy (~95%) and real-time decision-making, while future systems aim for fully automated, smart, and sustainable agriculture.

Advantages:

The proposed system provides several key benefits in modern agriculture:

- **High Accuracy:** Ensures reliable detection of crop diseases using advanced machine learning models.
- **Real-Time Detection:** Enables quick identification of diseases, allowing timely action.
- **Reduced Human Effort:** Automates the detection process, minimizing the need for manual inspection.
- **Scalability:** Can be extended to support multiple crops, diseases, and large datasets.

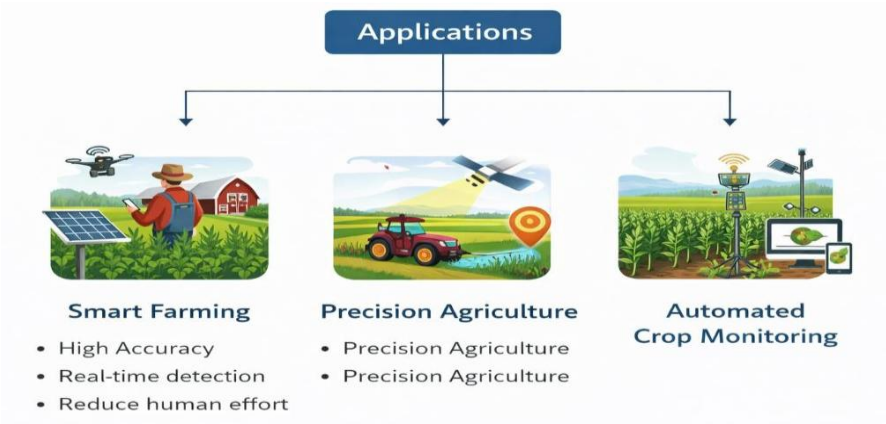
Overall, these advantages make the system efficient, reliable, and suitable for large-scale agricultural applications.

Applications:

The system can be applied in multiple agricultural domains:

- **Smart Farming:** Assists farmers in monitoring crop health using modern technologies.
- **Precision Agriculture:** Helps in applying targeted treatments, reducing waste and improving productivity.
- **Automated Crop Monitoring:** Enables continuous and automatic observation of crops without constant human involvement.

These applications highlight the practical usefulness of the system in improving agricultural efficiency and productivity.



Future Scope

The proposed system can be further improved by incorporating advanced technologies such as Internet of Things (IoT), drone-based imaging, and mobile applications. IoT devices can continuously collect environmental data like temperature, humidity, and soil conditions, which can be used along with image data for better analysis. Drone imaging can help in monitoring large agricultural fields quickly and efficiently. Additionally, integrating real-time sensors will allow the system to move beyond detection and support predictive analysis, enabling early prevention of crop diseases before they spread.

Conclusion:

The analysis of the selected research papers highlights the progressive evolution of smart agriculture from basic monitoring systems to intelligent, data-driven frameworks. Earlier approaches primarily focused on IoT-based sensing for environmental monitoring, which improved data collection but lacked decision-making capability. With the integration of Artificial Intelligence, particularly Machine Learning (ML) and Deep Learning (DL), modern systems now provide predictive insights, disease detection, and automated recommendations.

From the comparative study, it is observed that Machine Learning models such as Random Forest, Support Vector Machine, and Gradient Boosting perform exceptionally well for structured agricultural data, including soil parameters, temperature, humidity, and crop-related features. These models offer high accuracy (around 95%) while maintaining interpretability, lower computational cost, and faster training time. They are highly suitable for real-time applications like irrigation management and crop health prediction.

On the other hand, Deep Learning techniques, especially Convolutional Neural Networks (CNNs), show superior performance in image-based tasks such as plant disease detection. These models can automatically extract complex features from images and achieve high accuracy in identifying diseases. However, they require large datasets, higher computational power, and longer training time, which may limit their practical deployment in resource-constrained agricultural environments.

Based on the overall comparison of the research works, it can be concluded that Machine Learning is more efficient and practical for general smart agriculture systems, particularly when dealing with structured and sensor-based data. In contrast, Deep Learning is more suitable for specific tasks like image-based disease detection where high-level feature extraction is required.

Therefore, rather than replacing one another, both approaches are complementary. An integrated system that combines Machine Learning for environmental analysis and Deep Learning for visual recognition can provide the most effective and scalable solution for future smart agriculture.

References:

K. G., R. B., D. B., S. B., R. G., P. Dharnasi, and M. Sarvanan, "Smart Agriculture System Using IoT with AI Techniques," *International Journal of Computer Technology and Electronics Communication (IJCTEC)*, vol. 9, no. 2, pp. 479–483, Mar.–Apr. 2026.

D. Devarajan, R. Allafi, M. Obayya, and N. Nemri, "AI-Based Real-Time Disease Diagnosis in Plants Using Deep Learning-Driven CNNs," *Scientific Reports*, vol. 16, Art. no. 4587, 2026.

M. A. Qwaid, M. T. Sarker, S. M. Shawon, and H. T. Zubair, "AI-Enabled Smart Farming Framework for Sustainable Date Palm Cultivation Using Machine Learning and IoT," *Scientific Reports*, 2026.