International Journal of Research in Engineering, Science and Management
Volume 7, Issue 12, December 2024
https://www.ijresm.com | ISSN (Online): 2581-5792

Hybrid Deep Learning Approach for Automated Plant Disease Detection in Precision Agriculture

Rutika Pandurang Shinde[1*], Sagar Bharat Shah[2], Rishit Belani[3], Naveen Chowdary Nune[4], Arihant Jha[5], Harshal Pawar[6], Mohammed Mehataab Naazneen[7], Jaskeerat Singh[8]

[1]*Student, Department of Computer Science, University of Florida, Gainesville, USA*
[2]*Student, Department of Business Analytics, University of Cincinnati, Cincinnati, USA*
[3]*Student, Department of Computer Engineering, Bharati Vidyapeeth College of Engineering, Pune, India*
[4]*Student, Department of Computer Science and Engineering, SRM University, Amaravati, India*
[5]*Student, Department of Electronics and Communication Engineering, Manipal Institute of Technology, Manipal, India*
[6]*Student, Department of Information Technology, Rajiv Gandhi Institute of Technology, Mumbai, India*
[7]*Student, Department of Computer Science and Engineering, RVR & JC College of Engineering, Guntur, India*
[8]*Student, Department of Computer Science and Engineering, Ujjain Engineering College, Ujjain, India*

***Abstract*: Plant pests and diseases pose a significant challenge to agricultural productivity and economic stability. Traditional detection methods, which rely on manual inspection by farmers or experts, are often time-consuming, expensive, and susceptible to human error. This paper introduces a novel automated plant disease detection system utilizing a hybrid machine learning (ML) approach based on deep convolutional neural networks (CNNs). Leveraging advancements in computer vision, our model demonstrates enhanced precision in plant protection and extends the application of computer vision to precision agriculture. The methodology encompasses image collection and database creation, with validation by agricultural experts, followed by the development and training of a deep CNN framework. The proposed system effectively distinguishes healthy leaves from diseased ones and separates leaves from their environmental background, achieving an overall accuracy of 96.77%. This solution offers a reliable, efficient, and scalable tool for plant disease recognition, catering to the needs of both amateur gardeners and professional agriculturists.**

***Keywords*: Plants, Machine Learning, Automated, Disease.**

1. Introduction

Efficient plant disease management is crucial for sustainable agriculture, as unchecked pest infestations can significantly reduce crop yields and cause substantial economic losses. Traditional detection methods often rely on manual inspection by farmers or plant pathologists, which are not only time-intensive and costly but also prone to human error. Additionally, improper pesticide application due to misdiagnosis can lead to pathogen resistance, complicating disease management further.

Timely and accurate diagnosis of plant diseases is essential for precision agriculture, enabling targeted interventions that conserve resources while promoting crop health. While some diseases present visible symptoms that allow for manual detection, others exhibit signs that are too subtle or delayed for effective intervention. Automated detection systems incorporating image processing and machine learning techniques offer a promising solution to these limitations.

Recent advances in computer vision and deep learning, particularly convolutional neural networks (CNNs), have revolutionized image classification and recognition tasks. These advancements provide a robust foundation for developing automated, accurate, and scalable plant disease detection models. Leveraging these tools, it is possible to create systems that assist both amateur gardeners and professional agriculturists in identifying and managing plant diseases more effectively.

This paper introduces a hybrid machine learning approach that utilizes deep CNNs for automated plant disease recognition. We detail the methodology, encompassing image collection, preprocessing, database creation, and the training and fine-tuning of the CNN model. The proposed approach emphasizes simplicity and precision, enabling the system to distinguish between healthy and diseased leaves, as well as effectively separating leaves from their environmental background. This work aims to contribute to the field of precision agriculture by providing a scalable and efficient solution for plant disease management.

2. Related Work

The detection and management of plant diseases have been long-standing challenges in agriculture. Researchers and practitioners have explored various methods, from traditional manual approaches to sophisticated machine learning and deep learning systems. This section reviews these methodologies, emphasizing their strengths, limitations, and potential.

A. Traditional Detection Approaches

Conventional methods of plant disease detection primarily involve visual inspection by farmers or plant pathologists. These methods rely on observable symptoms such as

*Corresponding author: mr.ms2862@gmail.com

discoloration, spots, or fungal growth on leaves, stems, and fruits. However, these approaches have notable challenges:

- *Dependency on Expertise*: Accurate detection requires trained personnel, limiting the scalability of this approach in regions with a shortage of agricultural experts.
- *Time Consumption*: Manual inspection of large farms is labor-intensive and time-consuming, delaying interventions.
- *Error Susceptibility*: Diseases with subtle or overlapping symptoms are often misdiagnosed, leading to incorrect treatments or pathogen resistance.

Despite these drawbacks, traditional methods remain widely used due to their accessibility in resource-constrained settings. However, their limitations underscore the need for automated and scalable solutions.

B. Machine Learning-Based Approaches

The advent of machine learning (ML) has enabled data-driven plant disease detection. Early ML models utilized handcrafted features extracted from leaf images, such as texture, color, and shape, combined with classifiers like Support Vector Machines (SVM), k-Nearest Neighbors (KNN), and decision trees.

- *Support Vector Machines (SVM):* SVMs have been employed to classify healthy and diseased leaves by learning decision boundaries in high-dimensional feature spaces. Studies have reported accuracies above 80% for specific crops but often struggled with multi-class classification or complex backgrounds.
- *k-Nearest Neighbors (KNN):* KNN models are simple and effective for small datasets, but their performance degrades with large datasets due to high computational requirements.
- *Decision Trees:* These models offer interpretability but are prone to overfitting and may require ensemble techniques like Random Forests or Gradient Boosting to achieve competitive performance.

While these methods improved upon traditional detection, their reliance on feature engineering and limited generalization across diverse datasets highlighted the need for more advanced techniques.

C. Deep Learning Advancements

Deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized image-based plant disease detection. CNNs automatically extract hierarchical features from raw image data, eliminating the need for manual feature engineering.

- *Architectures:* Popular CNN architectures like AlexNet, VGG, ResNet, and Inception have been adapted for plant disease detection. These models have achieved accuracies exceeding 90% on various benchmark datasets, demonstrating their ability to generalize across different crops and disease types.
- *Scalability*: Deep learning models have shown remarkable scalability, handling large, diverse datasets effectively. They also outperform traditional ML models in scenarios with complex backgrounds or overlapping symptoms.
- *Challenges:* Despite their success, deep learning models require substantial computational resources and large annotated datasets, limiting their deployment in resource-constrained settings.

Several studies have explored lightweight CNN architectures to address these challenges, enabling deployment on mobile or edge devices for real-time detection.

D. Hybrid Models

Hybrid models combine the strengths of traditional ML, deep learning, and other technologies to address specific challenges in plant disease detection.

- *IoT Integration:* IoT-enabled systems use sensors and cameras to capture real-time data, which is then processed by machine learning models for disease detection. These systems facilitate continuous monitoring and timely interventions.
- *Edge Computing:* By deploying lightweight models on edge devices, such as drones or mobile phones, hybrid approaches enable real-time detection without relying on cloud infrastructure.
- *Fusion Techniques:* Combining CNNs with traditional ML classifiers has been explored to enhance interpretability and reduce computational complexity. For example, CNNs can extract features, which are then classified using SVM or Random Forest.

Hybrid models represent a promising direction, offering a balance between accuracy, resource efficiency, and scalability. They are particularly relevant for real-world applications in precision agriculture.

E. Traditional Detection Methods

Manual inspection has been the primary method for identifying plant diseases for decades. Zhang et al. (2015) explored the role of visual inspection by farmers and agricultural pathologists to detect visible symptoms such as leaf spots, discoloration, and fungal growth. While this approach is accessible and requires no advanced technology, it is inherently subjective and prone to errors. The study emphasized that accurate diagnosis depends on the expertise of the inspector, which is not always available in rural or resource-constrained areas. Furthermore, the method is time-intensive, especially for large-scale agricultural operations, and often fails to identify diseases with subtle symptoms or those in their early stages. Zhang et al. (2015) concluded that traditional methods, though still widely used, are inadequate for modern precision agriculture, emphasizing the urgent need for automated and scalable detection systems.

F. Early Machine Learning Models

The introduction of machine learning (ML) techniques in agriculture marked a significant advancement in disease detection by automating the classification process. Patil and Kumar (2016) employed Support Vector Machines (SVMs) to classify tomato leaf diseases based on manually extracted features such as texture, shape, and color. The SVM model achieved an accuracy of approximately 85%, demonstrating the potential of ML in plant disease detection. However, their approach faced challenges in handling large datasets due to the high computational cost of SVM and its sensitivity to hyperparameter tuning.

Similarly, Ali et al. (2018) utilized k-Nearest Neighbors (KNN) for rice disease classification. The simplicity of KNN made it suitable for small-scale datasets, but its performance degraded with high-dimensional feature spaces and larger datasets, highlighting scalability issues. Jadhav et al. (2017) explored decision trees and ensemble methods like Random Forests for apple leaf disease classification, reporting an accuracy of 87%. While decision trees are interpretable, they are prone to overfitting, requiring ensemble techniques to improve generalization. These studies demonstrated the effectiveness of ML but also revealed limitations, such as dependency on manual feature engineering and difficulties in distinguishing subtle disease patterns, which necessitated the transition to deep learning.

G. Deep Learning Advancements

Deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized image-based plant disease detection. Unlike traditional ML, CNNs automatically extract hierarchical features from raw image data, eliminating the need for manual feature engineering. Mohanty et al. (2016) were among the first to apply deep learning to plant disease detection, using standard CNN architectures like AlexNet and GoogLeNet. Their model achieved an impressive accuracy of 99.35% on the PlantVillage dataset, showcasing the robustness of CNNs in handling diverse crops and diseases.

Building on this, Ferentinos (2018) utilized transfer learning with pretrained models like InceptionV3 to reduce training time and improve accuracy. Transfer learning enabled the model to generalize across datasets with limited labeled data, making it a practical solution for real-world applications. Karthik et al. (2020) proposed a lightweight CNN architecture for mobile-based plant disease detection, achieving an accuracy of 92% with reduced computational overhead. These studies highlight the scalability, accuracy, and efficiency of CNNs, making them the gold standard for plant disease detection. However, challenges such as high computational demands and the need for large annotated datasets remain.

H. Hybrid and IoT-Based Models

Recent advancements have focused on integrating machine learning with Internet of Things (IoT) devices and edge computing for real-time monitoring and detection. IoT-enabled systems use sensors and cameras to capture real-time data, which is then processed by ML models. For instance, Pires et al. (2021) developed an IoT-based system that combined drone imagery with CNNs to detect plant diseases in vineyards. The system facilitated large-scale monitoring and timely intervention, addressing the scalability issues of traditional methods.

Another approach explored by Li et al. (2020) involved deploying lightweight CNN models on edge devices such as drones or mobile phones. This hybrid approach allowed real-time disease detection without relying on cloud infrastructure, reducing latency and ensuring data privacy. Such innovations bridge the gap between high-performance ML models and practical, field-deployable systems, making them highly relevant for precision agriculture.

3. Methodology

This section provides a comprehensive overview of the hybrid approach and workflow of the proposed system for plant disease detection. It covers the processes of data collection, preprocessing, model architecture design, training strategies, and fine-tuning techniques.

A. Data Collection

Data collection forms the foundation of the proposed system, emphasizing diversity and quality to ensure robust and reliable model performance. The dataset comprised a mix of publicly available sources and self-collected images. The PlantVillage dataset, a widely recognized benchmark for agricultural research, served as the primary source, offering a vast collection of high-resolution, annotated images of plant leaves categorized by health and disease types. To complement this, additional images were captured using high-resolution cameras in real-world farm settings and agricultural research labs. These images represented diverse environmental conditions, including variations in lighting, background complexity, and angles. Collecting such diverse data ensures the model's applicability to real-world scenarios where factors like uneven lighting and cluttered backgrounds can challenge detection accuracy.

Moreover, agricultural experts meticulously reviewed and validated the dataset to ensure precise labeling and minimize errors. This process involved excluding low-quality images and ambiguous cases where diseases overlapped or were indistinct. The resulting dataset included multiple crop types, disease stages, and environmental variations, making it representative of practical agricultural conditions. This well-curated dataset laid the groundwork for building a model that could generalize effectively across different crops and environments.

B. Preprocessing

The preprocessing stage was critical in optimizing the dataset for training deep learning models. This step addressed inconsistencies in image quality, reduced noise, and enhanced the dataset's overall reliability. First, all images were resized to a uniform dimension of 224x224 pixels to ensure compatibility with the chosen convolutional neural network (CNN) architecture. Uniform image dimensions are essential for maintaining a consistent input structure, simplifying the computational requirements during model training.

To improve model performance, image enhancement techniques were employed, including adjustments to brightness, contrast, and sharpness. These adjustments highlighted key disease features such as lesions, discoloration, and fungal growth. Additionally, histogram equalization was applied to underexposed images to improve feature visibility. Noise reduction techniques, including Gaussian filtering, were used to remove shadows, background artifacts, and other irrelevant elements, further isolating the plant leaves for effective feature extraction.

Data augmentation played a vital role in expanding the dataset and reducing overfitting. Augmentation techniques included random rotations, flipping, cropping, zooming, and Gaussian noise addition. These transformations increased dataset diversity, enabling the model to become invariant to real-world variations such as different orientations and environmental lighting. Normalization of pixel values to a range of [0,1] was applied to standardize the input data and accelerate convergence during training. Together, these preprocessing steps ensured that the dataset was clean, diverse, and well-prepared for the next stage of model development.

C. *Model Architecture*

The proposed system utilized a hybrid approach based on a deep convolutional neural network (CNN), designed to leverage the strengths of advanced feature extraction capabilities. A pretrained CNN model, such as ResNet-50, was selected due to its proven performance in image classification tasks. Transfer learning was employed to adapt this model to the specific task of plant disease detection. The use of transfer learning allowed the system to leverage the knowledge embedded in the pretrained network, significantly reducing training time and computational resources while maintaining high accuracy.

The architecture consisted of multiple convolutional layers for hierarchical feature extraction, pooling layers for dimensionality reduction, and fully connected layers for classification. The final layer used a softmax activation function to output probabilities for each disease class, enabling multi-class classification. To further enhance the system's performance, a traditional machine learning classifier, such as Random Forest or Support Vector Machine (SVM), was integrated into the architecture. This hybrid approach combined the robust feature extraction capabilities of CNNs with the interpretability and flexibility of traditional classifiers, improving accuracy in edge cases and challenging scenarios. Regularization techniques, such as dropout, were incorporated to prevent overfitting, ensuring the model's reliability and robustness.

D. *Training and Fine-Tuning*

The training process involved optimizing the model's performance through systematic parameter tuning and evaluation. The dataset was split into training, validation, and test sets in a 70:20:10 ratio to ensure sufficient data for model development and evaluation. Training was conducted using a supervised learning approach, with labeled data guiding the model to learn patterns and features associated with plant health and diseases.

The Adam optimizer was employed for its adaptive learning rate, ensuring efficient convergence during training. Cross-entropy loss was used as the objective function, suitable for multi-class classification tasks. Regularization techniques, including dropout and L2 regularization, were applied to minimize overfitting and improve the model's generalization to unseen data. Early stopping was implemented to halt training when validation performance plateaued, preventing unnecessary overtraining.

Hyperparameter optimization was conducted to fine-tune key parameters such as learning rate, batch size, and the number of layers in the fully connected network. Grid search and random search techniques were employed to identify the optimal configuration for the model. Training was performed on a GPU-enabled environment, significantly accelerating the computational process and enabling the handling of large datasets efficiently.

By the end of this stage, the model achieved a high level of accuracy and robustness, demonstrating its ability to generalize across various crop types and environmental conditions. The training process, combined with the hybrid approach and carefully curated dataset, resulted in a scalable and efficient plant disease detection system suitable for real-world applications.

4. Experimental Results

This section provides an in-depth analysis of the results obtained during the testing phase of the proposed plant disease detection system. The evaluation was carried out using well-established metrics to ensure a comprehensive assessment of the model’s performance and reliability.

A. *Evaluation Metrics*

The evaluation of the proposed system was performed using a set of carefully selected metrics that collectively reflect the model's performance across various dimensions. These metrics were chosen based on their relevance to plant disease detection tasks, ensuring an accurate representation of the system's strengths and weaknesses.

1) *Accuracy*

Accuracy measures the proportion of correctly classified samples out of the total number of samples. It is a widely used metric for providing an overall performance overview. The model achieved an accuracy of 96.77%, reflecting its high effectiveness in identifying healthy and diseased leaves.

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Samples}}$$

2) *Precision*

Precision calculates the proportion of correctly identified diseased samples (true positives) out of all samples predicted as diseased. This metric is crucial for applications where false positives could lead to unnecessary interventions, such as pesticide misuse. The model’s precision was calculated for each class, with an average precision of 95.6% across all disease types.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

3) Recall (Sensitivity)

Recall measures the model's ability to correctly identify all diseased samples. It focuses on minimizing false negatives, ensuring that no diseased plants are overlooked, which is critical for disease management. The proposed system achieved an average recall of 94.2%, indicating its effectiveness in detecting subtle or early-stage symptoms.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

4) F1 Score

The F1 Score is the harmonic mean of precision and recall, providing a single metric that balances the trade-offs between these two. It is especially useful in scenarios with imbalanced datasets, where accuracy alone may not be sufficient to judge performance. The system achieved an F1 score of 94.9%, demonstrating its balanced ability to identify diseased and healthy samples.

$$\text{F1 Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

5) ROC-AUC (Receiver Operating Characteristic - Area Under Curve)

The ROC-AUC evaluates the model's performance across all classification thresholds by plotting the true positive rate (recall) against the false positive rate. The proposed system achieved an AUC score of 0.98, signifying excellent capability to distinguish between healthy and diseased samples.

$$\text{AUC} = \int_0^1 \text{TPR}(\text{FPR})\, d\text{FPR}$$

The AUC value indicates that the model performs exceptionally well in distinguishing among multiple disease classes, even under varying environmental conditions and challenging scenarios.

B. Justification for Metric Selection

The chosen metrics address the multifaceted requirements of plant disease detection:

- Accuracy provides a holistic view of the system's performance but does not account for class imbalance.
- Precision ensures that false positives are minimized, a critical factor when dealing with sensitive agricultural scenarios.
- Recall ensures that all diseased samples are identified, which is crucial for preventing disease spread.
- F1 Score balances precision and recall, providing a reliable measure of performance when these metrics are in conflict.
- ROC-AUC offers insights into the model’s discriminative power across different thresholds, ensuring robustness and generalizability.

C. Performance Comparison

The proposed plant disease detection system was evaluated against traditional machine learning models and standard convolutional neural network (CNN) architectures to demonstrate its superiority. Traditional approaches such as Support Vector Machines (SVM), k-Nearest Neighbors (KNN), and Random Forest (RF) were used as baselines. SVM achieved an accuracy of 84.2% but struggled with multi-class classification and high-dimensional feature spaces, making it computationally expensive for large datasets. KNN, with an accuracy of 79.5%, showed significant performance degradation as the dataset size increased and exhibited poor scalability in handling complex backgrounds. Random Forest performed slightly better, achieving 87.6% accuracy; however, it required extensive feature engineering and displayed limitations in distinguishing diseases with overlapping symptoms or subtle differences. These results highlight the shortcomings of traditional machine learning models in addressing the complexities of plant disease detection.

In addition to traditional approaches, the system’s performance was compared with well-established CNN architectures such as AlexNet, VGG-16, and ResNet-50. AlexNet achieved an accuracy of 90.1%, benefiting from its simplicity and faster training but struggled to handle images with complex environmental backgrounds effectively. VGG-16 performed better, achieving 92.3% accuracy due to its hierarchical feature extraction capabilities; however, its computational cost was high due to the large model size. ResNet-50, the base model for the proposed hybrid system, achieved 94.8% accuracy by overcoming vanishing gradient issues and effectively learning deep features. However, its performance slightly dropped in scenarios with noisy datasets or highly diverse environmental conditions.

The proposed hybrid system outperformed all these models, achieving an accuracy of 96.77%. By integrating the robust feature extraction capabilities of CNNs with traditional classifiers like Random Forest for edge cases, the hybrid approach improved precision and robustness. This combination allowed the system to effectively handle complex scenarios, such as overlapping symptoms, challenging lighting conditions, and noisy backgrounds. The results validate the hybrid system's ability to address limitations faced by both traditional and deep learning-based models, making it a reliable tool for real-world applications in precision agriculture.

D. Error Analysis

While the proposed system achieved high accuracy and robustness, some errors were observed during testing. Misclassifications were primarily concentrated in images with overlapping disease symptoms or poor quality due to environmental factors. For instance, diseases with similar visual characteristics, such as leaf blight and bacterial spot, were occasionally confused by the model. These errors can be attributed to the subtle differences in symptoms that even human experts may find challenging to distinguish. Additionally, images with excessive noise, such as shadows, poor lighting, or occlusions caused by overlapping leaves, contributed to false positives and false negatives.

Another significant challenge was posed by diseases in their

Table 1
Performance comparison across models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Support Vector Machine (SVM)	84.2	82.5	80.3	81.4
k-Nearest Neighbors (KNN)	79.5	78.9	77.2	78.0
Random Forest (RF)	87.6	86.3	85.1	85.7
AlexNet	90.1	89.4	88.7	89.0
VGG-16	92.3	91.8	91.1	91.4
ResNet-50 (Base Model)	94.8	94.3	93.7	94.0
Proposed Hybrid System	96.77	95.6	94.2	94.9

early stages, where symptoms were either too subtle or not fully developed. In these cases, the model struggled to distinguish between healthy leaves and those showing early signs of infection. Furthermore, the presence of unrelated artifacts in the background, such as soil, debris, or other plants, occasionally led to incorrect classifications.

To address these issues, several steps were taken to improve the model's performance. First, data augmentation techniques were enhanced to simulate real-world variations, making the model more robust to environmental factors. Second, additional preprocessing steps, such as advanced noise reduction and shadow removal, were implemented to minimize the impact of poor image quality. Third, the dataset was expanded with more samples of similar diseases and early-stage symptoms to improve the model's ability to distinguish subtle differences. Finally, the hybrid approach was refined by incorporating a secondary traditional classifier to handle edge cases where CNN predictions were less confident, thereby improving overall reliability.

The analysis of errors and the subsequent improvements highlight the importance of addressing practical challenges in plant disease detection, ensuring that the system is robust, accurate, and ready for deployment in diverse agricultural environments.

5. Discussion

This section provides an interpretation of the experimental results and contextualizes their significance in relation to the objectives of the study. The discussion highlights the strengths and limitations of the proposed system, its practical applications, and the ethical considerations associated with its deployment.

A. Strengths

The proposed hybrid plant disease detection system demonstrates several strengths, making it a viable solution for real-world applications. The high accuracy of 96.77% achieved by the model underscores its effectiveness in distinguishing between healthy and diseased plants. By leveraging a hybrid approach that combines the feature extraction capabilities of convolutional neural networks (CNNs) with traditional classifiers, the system offers enhanced robustness, particularly in handling edge cases where symptoms are subtle or overlap.

Scalability is another significant advantage of the system. The use of a pretrained CNN architecture such as ResNet-50 ensures that the model can be adapted to various crop types and diseases with minimal retraining. This flexibility is crucial for applications in diverse agricultural settings. Additionally, the simplicity of deployment, enabled by converting the model to lightweight formats and integrating it with user-friendly interfaces, makes it accessible to a wide range of users, including small-scale farmers and agricultural researchers.

The system also demonstrates robustness in diverse environmental conditions. The preprocessing steps, including advanced noise reduction and data augmentation, have equipped the model to handle variations in lighting, shadows, and background artifacts effectively. This ensures consistent performance in real-world scenarios where environmental factors can introduce noise and complexity.

B. Limitations

Despite its strengths, the system has certain limitations that warrant further investigation and improvement. One significant challenge is the issue of unbalanced datasets. While efforts were made to ensure class balance during data collection, some rare diseases had limited representation, potentially impacting the model's ability to generalize across these classes.

Real-time implementation poses another hurdle. While the model performs well in controlled settings, deploying it for real-time monitoring in large-scale farms requires integration with hardware systems such as drones or IoT-enabled devices. The computational demands of CNNs, even when optimized, may limit their applicability on resource-constrained devices.

Generalization to unseen data remains an area of concern. Although the system achieves high accuracy on the test dataset, its performance on entirely new datasets, particularly those from different geographical regions or with novel disease symptoms, needs further evaluation. Incorporating more diverse datasets and employing techniques such as domain adaptation could address this limitation.

C. Practical Applications

The proposed system has significant potential for practical applications in precision agriculture. Its high accuracy and scalability make it an ideal tool for real-time monitoring of crops, enabling early detection of diseases and targeted interventions. By integrating the system with IoT devices such as drones and sensors, farmers can automate disease surveillance over large areas, reducing manual labor and improving efficiency.

Additionally, the system can be deployed in agricultural research settings to analyze the effectiveness of pest control measures and develop disease-resistant crop varieties. The intuitive interface allows amateur gardeners and small-scale farmers to leverage the technology without requiring technical expertise. Furthermore, the system can assist agricultural extension services in disseminating timely disease alerts and recommendations to farmers.

D. Ethical Considerations

As with any AI-based system, ethical considerations play a vital role in ensuring the equitable and responsible deployment of the proposed solution. Accessibility to small-scale farmers is a critical factor. Efforts should be made to provide affordable versions of the system, possibly through government or NGO support, to ensure that technological benefits are not limited to large-scale agricultural enterprises.

Bias in the training dataset is another ethical concern. Since the model's performance depends on the data it was trained on, any bias in the dataset, such as underrepresentation of specific diseases or crops, could result in unfair outcomes. Addressing this requires continuous updates to the dataset, incorporating diverse samples from different regions and agricultural contexts.

Lastly, considerations around data privacy and transparency are essential. Farmers and users must be informed about how their data is collected, stored, and used. Providing explainable AI (XAI) features to help users understand the model's predictions could enhance trust and adoption.

6. Conclusion

This research presents a hybrid approach to automated plant disease detection that combines the powerful feature extraction capabilities of convolutional neural networks (CNNs) with the interpretability and precision of traditional classifiers. The proposed system achieves a high accuracy of **96.77%**, demonstrating its effectiveness in distinguishing healthy plants from those affected by various diseases, even under challenging environmental conditions.

The significance of this work lies in its contribution to precision agriculture, where timely and accurate disease detection is critical for optimizing crop health and minimizing economic losses. By leveraging advancements in deep learning and data preprocessing techniques, the system not only enhances the scalability of plant disease detection but also simplifies deployment for a wide range of users, from small-scale farmers to large agricultural enterprises.

The hybrid approach addresses limitations faced by traditional and standalone CNN models, particularly in handling edge cases and subtle symptoms. This research provides a robust foundation for integrating machine learning into agricultural practices, improving efficiency, and reducing dependency on manual inspection.

Future work could explore expanding the system's capabilities to support multi-crop disease detection, further increasing its versatility. Additionally, integrating the model with edge computing platforms or IoT devices could enable real-time applications, enhancing the system's practical utility in large-scale farming operations.

7. Future Work

Building on the findings of this research, several directions for future improvement and extension are proposed to enhance the system's functionality, scalability, and accessibility.

A. Expanding the Dataset

One critical area for improvement involves expanding the dataset to include more crops and a broader range of diseases. Collecting data from diverse geographical regions and varying environmental conditions will ensure the system's robustness and generalizability. Rare or underrepresented diseases could be prioritized to address existing class imbalances, making the model more inclusive and applicable globally.

B. Exploring Advanced Architectures

Future work could investigate the use of other deep learning architectures, such as transformers, which have shown remarkable performance in image processing tasks. Lightweight models, such as MobileNet or efficient transformers, could also be explored to create mobile-friendly applications. These architectures could reduce computational overhead, enabling the deployment of the system on resource-constrained devices.

C. Real-Time Disease Detection

Integrating the model with edge computing platforms or drones presents an exciting opportunity for real-time disease detection and monitoring. By deploying lightweight versions of the model on IoT-enabled devices, the system could provide real-time feedback to farmers, automating disease surveillance and reducing response times. This could be particularly beneficial for large-scale farms and remote areas where manual inspection is impractical.

D. Explainable AI (XAI)

Incorporating explainable AI techniques into the system could provide users with insights into the model's decision-making process. This transparency would not only increase user trust but also help identify and rectify potential biases in the model. Visualizing the features or regions of the image that influenced the model's predictions could make the system more accessible to non-experts, such as small-scale farmers.

E. Multi-Crop and Multi-Disease Detection

Expanding the system to support simultaneous detection of multiple diseases in a single crop or across different crop types could significantly enhance its utility. This would involve building a multi-task learning framework capable of handling diverse agricultural challenges with a single, unified model.

References

[1] Al-Hiary, H., Bani-Ahmad, S., Reyalat, M., Braik, M., & ALRahamneh, Z. (2011). Fast and accurate detection and classification of plant diseases. *International Journal of Computer Applications*, 17(1), 31-38.

[2] Amara, J., Bouaziz, B., & Algergawy, A. (2017). A deep learning-based approach for banana leaf diseases classification. *Proceedings of the Datenbanksysteme für Business, Technologie und Web (BTW) Workshop*, 79-88.

[3] Barbedo, J. G. A. (2016). A review on the main challenges in automatic plant disease identification based on visible range images. *Biosystems Engineering*, 144, 52-60.

[4] Brahimi, M., Arsenovic, M., Laraba, S., Sladojevic, S., Boukhalfa, K., & Moussaoui, A. (2018). Deep learning for plant diseases: Detection and saliency map visualisation. *Human and Machine Learning*, 93-117.

[5] Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, 145, 311-318.

[6] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, 7, 1419.
[7] Patil, S. A., & Kumar, M. A. (2016). Tomato leaf disease detection using SVM classifier. *Proceedings of the IEEE International Conference on Emerging Trends in Engineering, Technology and Science (ICETETS)*, 1-5.
[8] Pires, R. C., Silva, F. M., Pereira, R. D., & Oliveira, R. A. (2021). IoT-based systems for precision agriculture: Integrating drones and machine learning for plant disease monitoring. *Sensors*, 21(4), 1458.
[9] Arsenovic, M., Sladojevic, S., Andelic, N., & Stefanovic, D. (2019). Deep learning-based automatic detection of plant diseases. *Remote Sensing Applications: Society and Environment*, 19, 100318.
[10] Karthik, R., Krishnaveni, V., & Menaka, R. (2020). Mobile-based plant leaf disease detection using deep learning. *International Journal of Computer Vision and Machine Intelligence*, 10(2), 13-21.
[11] Lu, J., Hu, J., Zhao, G., Mei, F., & Zhang, C. (2017). An in-field automatic wheat disease diagnosis system. *Computers and Electronics in Agriculture*, 142, 369-379.
[12] Zhang, S., Huang, W., Zhang, C., Dong, W., & Shi, Y. (2020). Plant disease detection using deep learning and feature fusion. *Multimedia Tools and Applications*, 79, 26371-26383.
[13] Brahimi, M., Boukhalfa, K., & Moussaoui, A. (2017). Deep learning for tomato diseases: Classification and symptoms visualization. *Applied Artificial Intelligence*, 31(4), 299-315.
[14] Singh, V., & Misra, A. K. (2017). Detection of plant leaf diseases using image segmentation and soft computing techniques. *Information Processing in Agriculture*, 4(1), 41-49.
[15] Hughes, D. P., & Salathé, M. (2015). An open access repository of images on plant health to enable the development of mobile disease diagnostics. *arXiv preprint arXiv:1511.08060*.
[16] 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, 1-11.
[17] Liu, B., Zhang, Y., He, D., & Li, Y. (2017). Identification of apple leaf diseases based on deep convolutional neural networks. *Symmetry*, 10(1), 11.
[18] Li, Y., Nie, J., & Chao, X. (2020). Multi-scale convolutional neural networks for plant leaf disease identification. *Agricultural and Forest Meteorology*, 287, 107930.
[19] Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019). A comparative study of fine-tuning deep learning models for plant disease identification. *Computers and Electronics in Agriculture*, 161, 272-279.
[20] Abbas, N., Naz, S., & Razzak, I. (2021). Plant disease classification using convolutional neural networks. *Computers and Electronics in Agriculture*, 184, 106067.
[21] Zhang, X., Qiao, Y., Meng, F., Fan, C., & Zhang, M. (2018). Identification of maize leaf diseases using improved deep convolutional neural networks. *IEEE Access*, 6, 30370-30377.
[22] Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147, 70-90.
[23] Yalcin, H., & Razavi, S. (2016). Plant disease detection using deep learning algorithms. *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, 1483-1487.
[24] Fuentes, A., Yoon, S., Kim, S. C., & Park, D. S. (2017). A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors*, 17(9), 2022.
[25] Mohapatra, S. D., Rautaray, S. S., & Nath, R. K. (2021). Machine learning and deep learning techniques for plant disease detection: A review. *Information Processing in Agriculture*, 8(4), 518-532.