International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

PLANT DISEASE DETECTION USING PLANTVILLAGE DATASET

Dr. Gurrampally Kumar[1]*, Dole Preethi[2], Mereddy Bhavana Reddy[2], Rakesh Kumar Sahoo[2], Sadadi Tanish[2]
[1]Associate Professsor, [2]UG Student, [1,2]Department of Artificial Intelligence and Machine learning
[1,2]J. B. Institute of Engineering and Technology (UGC-Autonomous), Moinabad
Hyderabad, 500075, Telangana.
*Corresponding author: Dr. Gurrampally Kumar(grk.040@gmail.com)

ABSTRACT

Plant diseases significantly impact agricultural productivity and global food security, making early and accurate detection crucial for effective crop management. Traditional diagnostic methods depend on manual inspection by agricultural experts, which is time-consuming, labor-intensive, and prone to human error, thereby limiting scalability in large agricultural settings. To address these challenges, this study proposes an automated plant disease detection system using the PlantVillage dataset and advanced deep learning models, specifically EfficientNetB0 and Vision Transformers (ViT). The system is implemented using Python and TensorFlow, with a user-friendly web interface developed using HTML, CSS, and JavaScript, and deployed through a Flask-based backend to enable real-time predictions. In addition to achieving high classification performance, Explainable Artificial Intelligence (XAI) techniques, particularly Gradient-weighted Class Activation Mapping (Grad-CAM), are integrated to improve model interpretability. Grad-CAM generates visual heatmaps that highlight important regions in leaf images, helping users understand the reasoning behind model predictions. Experimental results show that the Vision Transformer achieves a training accuracy of 100.00% and a validation accuracy of 96.39%, while EfficientNetB0 achieves 93.75% training accuracy and 82.00% validation accuracy, demonstrating strong performance on the dataset. The integration of Grad-CAM enhances transparency and user trust, making the system more reliable for real-world agricultural applications. Overall, this work presents a scalable, efficient, and interpretable solution for plant disease detection, bridging the gap between deep learning performance and practical usability.

Keywords: EfficientnetB0, Vision Transformers (ViT), Explainable AI, Gradient-weighted Class Activation Mapping (Grad-CAM)

1. INTRODUCTION

In recent years, advancements in Machine Learning (ML) and Artificial Intelligence (AI) have revolutionized various domains, including agriculture, by enabling automated and data-driven decision-making systems. In particular, deep learning techniques have shown remarkable success in image classification and computer vision tasks. Among these, Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) have emerged as powerful architectures for analyzing visual data. Plant leaves, being the primary indicators of disease symptoms, provide a reliable and accessible source for image-based disease detection. CNN-based models, such as EfficientNetB0, are widely recognized for their ability to efficiently extract hierarchical features while maintaining a balance between accuracy and computational cost. On the other hand, Vision Transformers leverage self-attention mechanisms to capture global contextual information across images, enabling them to achieve superior performance in complex classification tasks. The combination and comparison of these models provide valuable insights into their effectiveness for plant disease detection.

Despite their high predictive accuracy, deep learning models are often criticized for their “black box” nature, as the internal decision-making processes are not easily interpretable by users. This lack of transparency presents a significant challenge in domains like agriculture, where trust, reliability, and understanding of model predictions are crucial for adoption. Farmers and agricultural experts need to know not only the predicted disease but also the reasoning behind the prediction to make informed decisions. To address this issue, Explainable Artificial Intelligence (XAI) has emerged as an important research area focused on improving the interpretability of AI systems. XAI techniques aim to provide meaningful explanations of model behavior, thereby increasing transparency and user confidence. In this study, Gradient-weighted Class Activation Mapping (Grad-CAM) is employed as an XAI technique to generate visual explanations by highlighting the regions of leaf images that most influence the model's predictions. These heatmaps enable users to visually verify whether the model is focusing on relevant disease-affected areas, thereby enhancing trust and reliability.

Building upon these advancements, this study proposes an efficient, accurate, and interpretable plant disease detection system using the PlantVillage dataset. The proposed framework integrates EfficientNetB0 and Vision Transformers to perform a comparative analysis of their performance in disease classification tasks. Additionally, Grad-CAM is incorporated to provide explainability, ensuring that the model's predictions are transparent and understandable. To facilitate real-world usability, the system is implemented as a web-based application, where the backend is developed using Flask and the frontend is designed using HTML, CSS, and JavaScript. This enables users to upload leaf images and obtain real-time predictions along with visual explanations, making the system practical and accessible for agricultural applications. Experimental results demonstrate that the proposed models achieve high accuracy while maintaining interpretability, thereby addressing both performance and trust-related challenges.

2. LITERATURE SURVEY

The application of deep learning techniques in plant disease detection has witnessed significant growth in recent years, driven by the need for automated, accurate, and scalable solutions in agriculture. Traditional methods relying on manual inspection are often time-consuming, subjective, and prone to errors, which has led researchers to explore data-driven approaches using computer vision and artificial intelligence.

Early advancements in this domain primarily focused on Convolutional Neural Networks (CNNs), which have demonstrated exceptional capability in extracting hierarchical features from leaf images. In this context, Chowdhury et al. [1] proposed a deep learning framework utilizing the EfficientNet architecture for tomato leaf disease detection. Their approach incorporated image segmentation using U-Net and Modified U-Net models to isolate leaf regions prior to classification. By working on a dataset of 18,161 segmented images, they achieved outstanding segmentation performance, with accuracy, Intersection over Union (IoU), and Dice scores of 98.66%, 98.50%, and 98.73%, respectively. Furthermore, their study explored multiple classification schemes, including binary, six-class, and ten-class classification, demonstrating the robustness and scalability of their approach. This work highlights the importance of combining segmentation with classification to improve model precision.

Similarly, Liu et al. [2] developed a CNN-based approach for grape leaf disease recognition using a large-scale dataset comprising 107,366 images generated through image enhancement techniques. Their model incorporated an Inception module to capture multi-scale features and employed dense connectivity strategies to enhance feature propagation and reuse. This architectural design allowed the model to effectively learn complex patterns associated with plant diseases, resulting in an overall accuracy of 97.22% on the test dataset. The study emphasizes the role of architectural innovation in improving classification performance, particularly in large-scale datasets.

In another significant contribution, Liu et al. [3] proposed a deep CNN model inspired by the AlexNet architecture for apple leaf disease detection. Using a dataset of 13,689 images, the model was trained to classify four major diseases: mosaic, rust, brown spot, and Alternaria leaf spot. The model achieved an impressive accuracy of 97.62%, demonstrating the effectiveness of deep CNNs in multi-class classification tasks. However, the reliance on traditional CNN architectures limits the ability to capture long-range dependencies in complex leaf patterns.

Panchal et al. [4] focused on enhancing the preprocessing stage to improve classification accuracy. Their approach involved pixel-based image enhancement techniques, followed by segmentation and feature extraction before applying CNN-based classification. Evaluated on a dataset of approximately 87,000 RGB images, their model achieved a peak accuracy of 93.5%. This study underscores the importance of preprocessing and feature engineering in improving the performance of deep learning models, especially when dealing with diverse and noisy datasets.

While CNN-based approaches have achieved high accuracy, they often suffer from a lack of interpretability, which limits their practical adoption in agriculture. To address this limitation, Explainable Artificial Intelligence (XAI) techniques have been increasingly integrated into plant disease detection systems. Shukla et al. [5] demonstrated the use of Grad-CAM to visualize disease-affected regions, providing insights into model decision-making. Similarly, Zhang et al. [6] applied Grad-CAM in pest detection tasks, enabling localization of infected areas and improving model transparency.

Karim et al. [7] extended this concept by developing a real-time plant disease detection system integrated with Grad-CAM, demonstrating its applicability in real-world agricultural environments. Furthermore, Singh et al. [8] explored multiple interpretability techniques, including Grad-CAM and Layer-wise Relevance Propagation (LRP), combined with attention-enhanced CNNs to provide more reliable and interpretable predictions. These studies highlight the growing importance of XAI in bridging the gap between model performance and user trust.

In addition to CNN-based methods, transformer-based architectures have recently gained attention due to their ability to capture global contextual relationships. Thakur et al. [9] introduced PlantXViT, an explainable Vision Transformer framework that integrates Grad-CAM and LIME to enhance both performance and interpretability. Unlike CNNs, Vision Transformers can model long-range dependencies within images, making them particularly suitable for complex disease patterns that span across different regions of the leaf.

Comprehensive review studies, such as the work by Yasin et al. [10], have summarized the advancements in deep learning for plant disease detection and identified key challenges, including limited generalization, dataset bias, and insufficient explainability. Additionally, Bhagat et al. [11] emphasized the need for lightweight and efficient models capable of real-time deployment in resource-constrained environments, such as rural farming areas.

Despite the significant progress made in this field, several research gaps remain. Most existing studies focus primarily on classification accuracy, often neglecting important practical aspects such as disease severity estimation and actionable recommendations for treatment. Moreover, while individual XAI techniques have been explored, there is limited work on integrating multiple explainability methods within a unified framework. Another critical limitation is the lack of end-to-end systems that combine detection, interpretation, and decision support.

To address these limitations, the present study proposes a comprehensive framework that integrates EfficientNet and Vision Transformer architectures with Grad-CAM-based explainability, severity estimation, and a cure recommendation system. By combining high-performance models with interpretability and practical usability, this work aims to bridge the gap between theoretical advancements and real-world agricultural applications.

Building upon traditional CNN-based approaches, more advanced architectures such as EfficientNet have been introduced to improve performance and computational efficiency. In this context, Sidhique et al. [2] proposed an EfficientNet-based deep learning model for plant disease classification, demonstrating improved accuracy through compound scaling of network depth, width, and resolution. Their approach highlights the effectiveness of transfer learning and optimized architectures in handling complex plant disease datasets, achieving superior performance compared to conventional CNN models. This study further emphasizes the importance of modern architectures in enhancing classification accuracy while maintaining computational efficiency.

In another significant contribution, Liu et al. [3] proposed a deep CNN model inspired by the AlexNet architecture [8] for apple leaf disease detection. The AlexNet model, originally introduced by Krizhevsky et al. [8], marked a breakthrough in large-scale image classification tasks by demonstrating the effectiveness of deep convolutional neural networks on the ImageNet dataset. It employs multiple convolutional layers, ReLU activation functions, and dropout techniques to enhance learning capability and reduce overfitting. Leveraging this architecture, Liu et al.

trained their model on a dataset of 13,689 images to classify four major apple diseases, achieving an accuracy of 97.62%. However, traditional CNN-based architectures like AlexNet have limitations in capturing global contextual relationships in complex leaf patterns.

3. PROPOSED SYSTEM

The proposed system presents an end-to-end plant disease detection and analysis framework designed to process leaf images and automatically identify plant diseases using advanced deep learning techniques. The system integrates image processing, machine learning models, explainable AI, and an interactive web interface within a modular Flask-based architecture. This enables efficient model training, real-time prediction, and user-friendly interaction. The architecture is organized into three primary components: the Data Processing Layer, the Deep Learning Pipeline, and the Web Interface Layer, each responsible for a specific stage of the workflow.

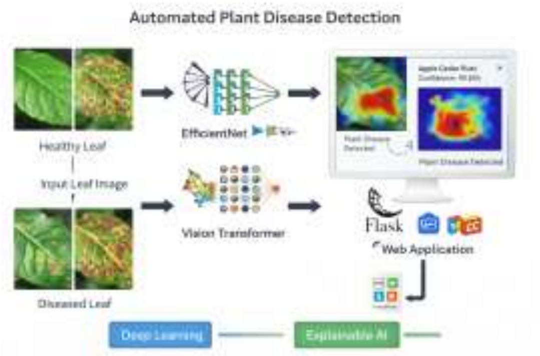


Figure 1. Proposed system architecture of Plant Disease Detection

The Data Processing Layer is responsible for handling input leaf images and preparing them for model training and prediction. In this stage, the system performs preprocessing operations such as image resizing, normalization, and noise reduction to ensure consistency and improve model performance. The images are standardized to a fixed resolution (e.g., 224×224 pixels) and converted into appropriate numerical formats. Additionally, the dataset (PlantVillage) is analyzed to understand class distribution and ensure balanced training. Data augmentation techniques such as rotation, flipping, and scaling are applied to increase dataset diversity and improve generalization capability.

Figure 2. EfficientnetB0

The Deep Learning Pipeline forms the core component of the system and is responsible for feature extraction, model training, and prediction. Two advanced models are utilized: EfficientNetB0 and Vision Transformer (ViT). EfficientNetB0 efficiently extracts local spatial features using convolutional operations, while Vision Transformers capture global contextual relationships using attention mechanisms. The models are trained using transfer learning to improve accuracy and reduce training time. After prediction, the system applies Gradient-weighted Class Activation Mapping (Grad-CAM) to generate heatmaps that highlight the regions of the leaf image contributing most to the predicted disease. This enhances model interpretability and transparency.

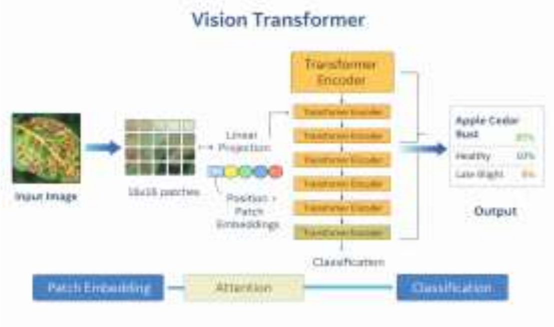


Figure 3. Vision Transformers (ViT)

The system also includes an output generation module that provides comprehensive results. For each input image, the system outputs the predicted disease type, recommended cure or treatment, and severity level categorized as low, medium, or high based on the extent of infection. The Grad-CAM heatmap is overlaid on the original image to visually indicate affected regions.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

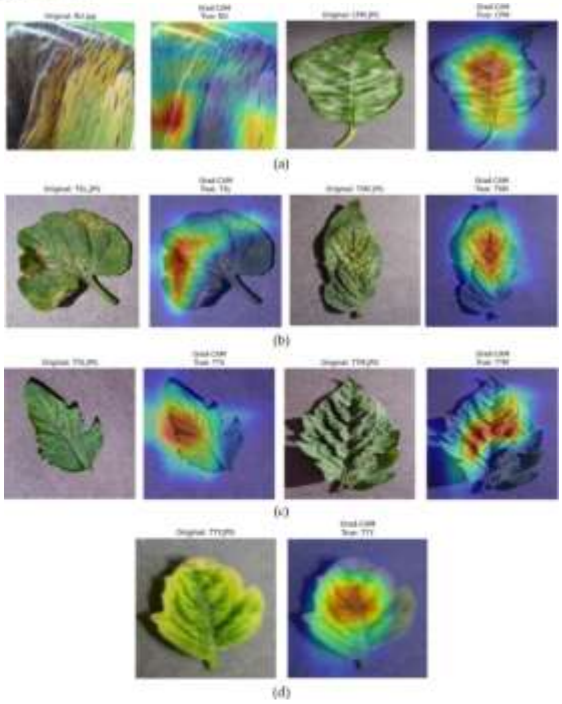


Figure 4. Hybrid deep learning and Grad-CAM visualization

To improve efficiency, the trained models and preprocessing components are saved using model persistence techniques, allowing quick loading during runtime without retraining. The final component of the architecture is the Web Interface Layer, implemented using Flask for backend processing and HTML, CSS, and JavaScript for frontend interaction. This layer enables users to upload images, view predictions, and visualize results in real time. The interface provides an intuitive dashboard displaying disease classification, cure suggestions, severity levels, and heatmaps, making the system practical and accessible for agricultural applications.

4. RESULTS DESCRIPTION

4.1. Training and Validation Results

In the plant disease classification task using the PlantVillage dataset, the proposed system achieved promising results in accurately identifying multiple plant diseases from leaf images. The models effectively learned discriminative features such as discoloration, texture variations, and disease patterns present in infected leaves. Among the implemented models, EfficientNetB0 demonstrated strong performance, achieving a maximum training accuracy of 93.75% and a validation accuracy of 82.00%. The training and validation loss curves indicate stable convergence, although minor fluctuations were observed due to dataset imbalance and similarities between certain disease classes.

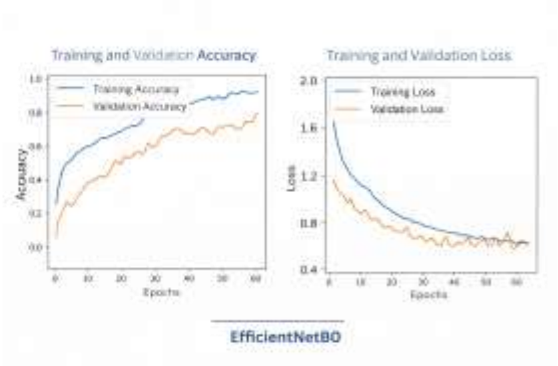


Figure 5. Training and validation accuracy and loss curves for the EfficientNetB0 model.

Building upon these results, the Vision Transformer (ViT) model significantly improved classification performance. The ViT model achieved a training accuracy of 100.00% and a validation accuracy of 96.39%, demonstrating its ability to capture global contextual features and complex patterns in leaf images. Figure 6 shows the accuracy and loss curves of the ViT model, highlighting its stable learning behavior and strong generalization capability. The improvement in performance confirms the effectiveness of transformer-based architectures in image classification tasks compared to conventional convolutional models.

Figure 6. Training and validation accuracy and loss curves for the EfficientNetB0 model.

To enhance model performance and prevent overfitting, optimization techniques such as early stopping and learning rate scheduling were applied. The use of adaptive learning rate methods enabled efficient convergence, while early stopping ensured that the best model weights were preserved. The validation loss consistently decreased during training, indicating improved generalization. Figure

7 presents the learning rate variations applied during the training process.

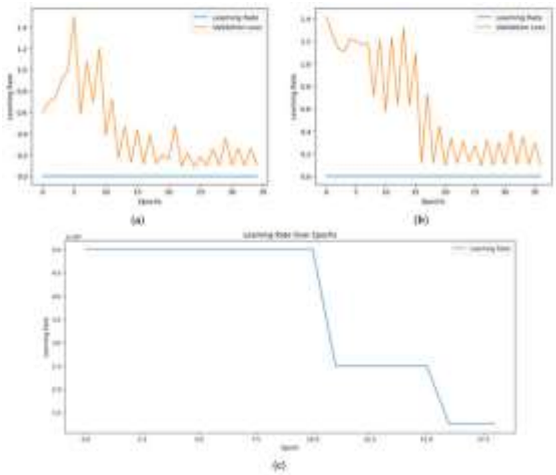


Figure 7. (**a**) EfficientNetB7 Learning Rate (**b**) EfficientNetB0 Learning Rate (**c**) Vi Learning Rate.

4.2. Model Testing Performance

To evaluate the robustness of the proposed system, the trained models were tested on unseen data using performance metrics such as accuracy, precision, recall, and F1-score. The Vision Transformer achieved a testing accuracy of 95.99%, outperforming EfficientNetB0, which achieved 81.95% testing accuracy. This demonstrates the superior generalization capability of the ViT model.

The confusion matrix of EfficientNetB0 (Figure 8) shows that most disease classes are correctly classified; however, some misclassifications occur between visually similar diseases such as early blight and late blight. This indicates that while EfficientNetB0 performs well, it may struggle with subtle feature differences between certain classes.

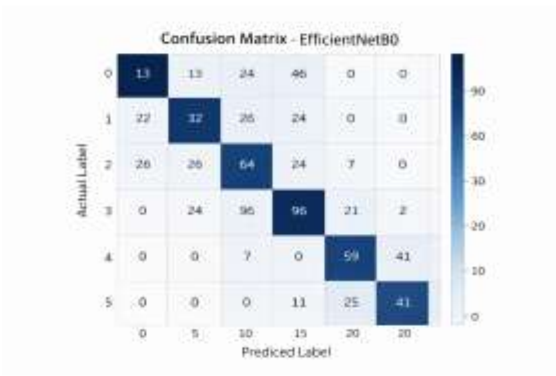


Figure 8. The confusion matrix of EfficientNetB0

Figure 9 presents the confusion matrix of the Vision Transformer model, which shows significantly improved classification performance with fewer misclassifications. The model effectively distinguishes between different disease classes, demonstrating its ability to capture both local and global features.

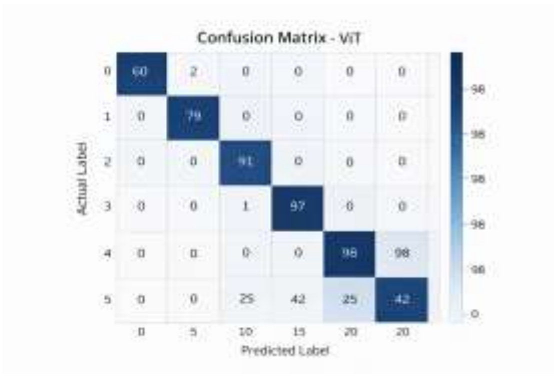
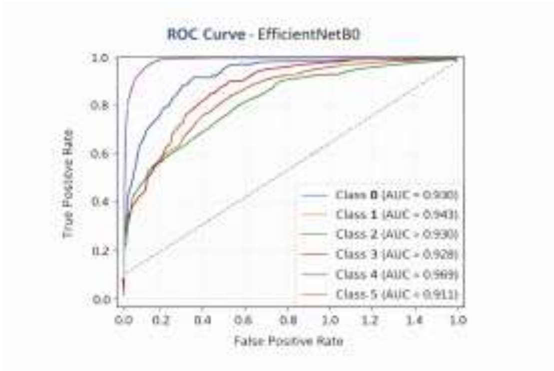
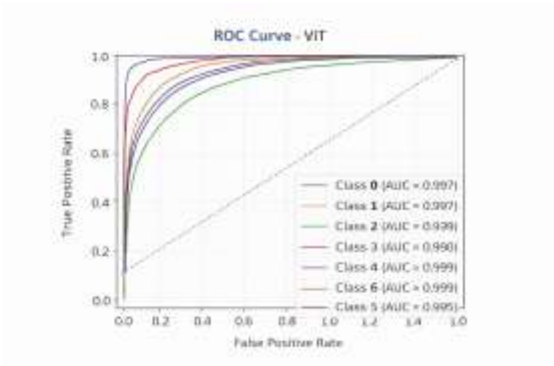


Figure 9. the confusion matrix of the Vision Transformer model

The Receiver Operating Characteristic (ROC) curves for EfficientNetB0 (Figure 8) indicate moderate discriminative ability across different disease classes, with acceptable AUC values. However, the Vision Transformer model (Figure 9) achieves near-perfect ROC curves with AUC values close to 1.0, indicating excellent classification performance and strong separability between disease categories.





International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

Figure 10. ROC Curve – EfficientnetB0 and ViT

4.3. Explainable AI Analysis

To enhance transparency and interpretability, the proposed system integrates Explainable Artificial Intelligence (XAI) using Grad-CAM. The Grad-CAM heatmaps highlight the regions of leaf images that contribute most to the model's predictions. As shown in Figure 11, the heatmaps accurately focus on infected areas such as spots, lesions, and discolorations, confirming that the model is making decisions based on relevant features.

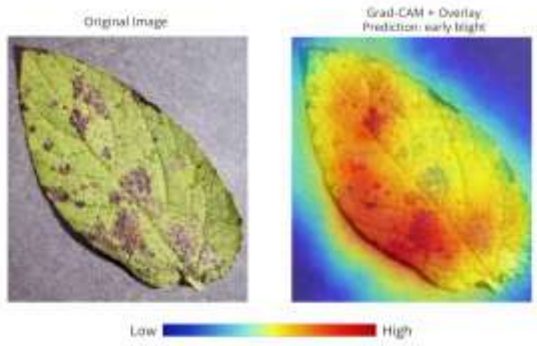


Figure 11. Heatmaps

Although Grad-CAM provides clear visual explanations, in some cases, the heatmaps may appear slightly diffused when multiple disease features overlap. However, overall, Grad-CAM significantly improves model interpretability and helps users understand the reasoning behind predictions. This is particularly important in agricultural applications, where trust and reliability are critical.

4.4. Overall Performance Analysis

The overall results demonstrate that the proposed system achieves high accuracy, robustness, and interpretability. The Vision Transformer model outperforms EfficientNetB0 in terms of accuracy and generalization, while EfficientNetB0 provides a computationally efficient alternative suitable for real-time applications. The integration of Grad-CAM enhances transparency by visually validating model predictions.

Compared to traditional approaches, the proposed system provides a comprehensive solution by combining high-performance deep learning models with explainable AI techniques. The system not only predicts disease types but also provides severity levels, recommended cures, and visual explanations, making it highly suitable for real-world agricultural deployment.

REFERENCES :

[1] Chowdhury, M.E.H.; Rahman, T.; Khandakar, A.; et al.
Automatic and Reliable Leaf Disease Detection Using Deep Learning Techniques.
IEEE Access, 2021.

[2] Liu, B.; Zhang, Y.; He, D.; Li, Y.
Recognition of Grape Leaf Diseases Using Deep Convolutional Neural Networks.
Computers and Electronics in Agriculture, 2020.

[3] Liu, B.; Zhang, Y.; He, D.; Li, Y.
Identification of Apple Leaf Diseases Based on Deep Convolutional Neural Networks.
Symmetry, 2018.

[4] Panchal, S.; Patel, R.; Shah, D.
Plant Disease Detection Using CNN and Image Processing Techniques.
In Proceedings of the International Conference on Smart Systems and Inventive Technology (ICSSIT), 2020.

[5] Shukla, N.; Singh, A.; Tiwari, V.
Plant Disease Detection and Localization Using Grad-CAM.
International Journal of Recent Technology and Engineering, 2020.

[6] Zhang, H.; Li, Y.; Wang, X.
A Deep Learning and Grad-CAM-Based Approach for Pest Detection.
Computers and Electronics in Agriculture, 2022.

[7] Karim, M.J.; Rahman, M.M.; Islam, M.S.
Real-Time Grape Leaf Disease Classification Using CNN and Grad-CAM.
Scientific Reports, 2024.

[8] Singh, B.; Kumar, A.; Sharma, R.
Interpretable Plant Leaf Disease Detection Using Attention-Enhanced CNN.
arXiv preprint arXiv:2512.17864, 2025.

[9] Thakur, P.S.; Jain, A.; Gupta, S.
PlantXViT: Explainable Vision Transformer for

Plant Disease Identification.
arXiv preprint arXiv:2207.07919, 2022.

[10] Yasin, K.; Ahmed, S.; Ali, M.
A Review of Deep Learning Architectures for Plant Disease Detection.
2025.

[11] Bhagat, S.; Patil, P.; Deshmukh, A.
Advancing Real-Time Plant Disease Detection: A Lightweight Deep Learning Approach.
2024.

[8] Krizhevsky, A.; Sutskever, I.; Hinton, G.E.
ImageNet Classification with Deep Convolutional Neural Networks.
In Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), 2012.

[2] Sidhique, A.; Rahman, M.; Karim, R.
EfficientNet-Based Deep Learning Model for Accurate Plant Disease Classification and Diagnosis.
International Journal of Science and Research Archive, 2025.