iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: IV Month of publication: April 2025

DOI: https://doi.org/10.22214/ijraset.2025.68565

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

AiKrishi Crop Disease Detection Using Vision Transformer

Prof. Minakshi Ramteke[1], Prof Tarun Yengantiwar[2], Ms. Vaishali Vijay Gahukar[3]
[1]Project Guide, Computer Science and Engineering, V.M. Institute Of Engineering and Technology, Dongargaon, Nagpur
[2]Project Guide, Computer Science and Engineering,V.M. Institute Of Engineering and Technology, Dongargaon, Nagpur
[3]Student, Computer Science and Engineering, V.M. Institute Of Engineering and Technology, Dongargaon, Nagpur

Abstract: Plant diseases pose a serious threat to global food security and sustainable agriculture, causing significant economic losses and affecting rural livelihoods. Traditional disease detection methods, such as manual inspection and lab tests, often lack efficiency, scalability, and accuracy. This paper explores the use of vision transformers (ViTs), a cutting-edge deep learning technique, to enhance plant disease detection. ViTs utilize self-attention mechanisms to analyze complex patterns in plant images, enabling precise and automated classification. A thorough review of deep learning applications in agriculture highlights the growing interest in ViTs for disease identification. This study also details an effective methodology for training and evaluating ViT models on a well-balanced dataset of 55 plant disease classes. Experimental findings confirm the superior accuracy of ViTs, showcasing their transformative potential in precision agriculture. By improving early disease diagnosis, ViTs can contribute to more sustainable farming practices and increased agricultural productivity.
Keywords: Automated disease classification, Deep learning in agriculture, Image classification, Plant disease detection, Precision agriculture, Vision transformers etc.

I. INTRODUCTION

Agriculture is vital for human survival, supplying essential resources such as food, fiber, and fuel. Sustainable farming practices play a crucial role in ensuring long-term food security, preserving the environment, and supporting economic stability. By integrating eco-friendly techniques, agriculture can enhance soil health, conserve natural resources, and maintain biodiversity for future generations. Effective plant disease detection is a critical aspect of sustainable agriculture, as it directly impacts crop productivity and food supply. Traditional approaches, such as visual inspections and laboratory testing, are often inefficient, time-consuming, and prone to inaccuracies. Visual assessments depend on human expertise, leading to inconsistent results, while lab-based methods require specialized equipment and expertise, limiting their scalability.

To address these challenges, modern technologies like artificial intelligence (AI) and deep learning have been widely adopted for automated plant disease identification. AI-driven solutions, including convolutional neural networks (CNNs), have shown remarkable accuracy in analyzing plant images and diagnosing diseases. Recent advances in computer vision and deep learning enable early detection of infections through image-based analysis, reducing reliance on manual observation. The integration of AI with IoT further enhances precision farming, allowing for real-time monitoring of plant health and optimized irrigation. For instance, smart irrigation systems equipped with sensors and AI-driven decision-making algorithms can efficiently regulate water usage, particularly in regions facing water scarcity.

A breakthrough in AI-based agriculture is the adoption of vision transformers (ViTs), a deep learning model initially developed for natural language processing and later adapted for computer vision tasks. ViTs outperform traditional CNNs by leveraging self-attention mechanisms, capturing intricate patterns in images with greater accuracy and efficiency. Compared to CNNs, ViTs require fewer computational resources while maintaining high precision, making them a promising tool for plant disease detection. Various studies have demonstrated the effectiveness of ViTs in classifying plant diseases with superior accuracy compared to conventional models. Researchers have introduced hybrid models that integrate ViTs with CNN architectures, further enhancing disease classification performance.

Recent advancements have explored smartphone-based applications utilizing ViTs for real-time disease detection, allowing farmers to diagnose plant infections through mobile devices. Studies also propose optimized ViT models, such as GreenViT, which refine feature extraction capabilities for improved accuracy. Other research efforts focus on pruning techniques, such as least important attention pruning (LeIAP), to enhance computational efficiency in large-scale applications. By combining ViTs with edge-feature guidance modules (EFG), researchers have achieved significant improvements in model interpretability and robustness.

This paper aims to explore the application of ViTs in plant disease detection, utilizing a dataset encompassing multiple plant species for potential integration into smart agricultural systems. The study is structured into key sections: methodology, results, discussion, and conclusion. The methodology details the dataset, preprocessing techniques, and model architecture. The results section presents findings on ViT performance in disease classification, highlighting its advantages over conventional models. The discussion analyzes the implications of these findings for precision agriculture, emphasizing the role of AI in sustainable farming. Finally, the conclusion summarizes key insights and proposes future research directions to further enhance AI-driven plant disease detection.
Through this research, we aim to contribute to the development of advanced, scalable, and efficient plant disease detection systems, paving the way for a new era of precision agriculture. By harnessing the power of AI and deep learning, agricultural sustainability can be strengthened, ensuring a more resilient food production system for future generations.

II. PROBLEM IDENTIFICATION

- Plant diseases significantly threaten global agriculture, leading to reduced crop yields, food insecurity, and economic losses.
- Traditional disease detection methods, such as visual inspections and laboratory testing, are time-consuming, labor-intensive, and often inaccurate.
- Visual assessments rely on human expertise, making them subjective and inconsistent, while lab-based techniques require specialized equipment and trained personnel, limiting their scalability for large-scale farming.
- Additionally, these conventional approaches often detect diseases only after visible symptoms appear, making early intervention difficult.
- The lack of efficient, automated, and scalable solutions highlights the need for advanced technologies.
- Deep learning, particularly Vision Transformers (ViTs), offers a promising alternative by enabling accurate and early disease detection, ensuring better crop health management and improving agricultural sustainability.

A. Existing System

Current plant disease detection methods primarily rely on visual inspection and laboratory-based techniques. Farmers and agricultural experts manually assess plant health based on visible symptoms like discoloration, lesions, or deformities. While widely used, this method is subjective, inconsistent, and inefficient for large-scale farming. Laboratory tests provide more accuracy but are expensive, time-consuming, and require specialized equipment. Recent advancements in deep learning and computer vision, particularly convolutional neural networks (CNNs), have improved automated disease detection. However, CNN-based models often struggle with complex patterns and require large datasets for effective training. These limitations necessitate more advanced and efficient approaches like Vision Transformers (ViTs).

B. Drawbacks

Traditional plant disease detection methods, including visual inspection and laboratory tests, have several limitations. Visual assessment is highly subjective, dependent on human expertise, and prone to inconsistencies. It often fails to detect early-stage infections, leading to delayed interventions. Laboratory-based techniques, though accurate, are costly, time-consuming, and require specialized personnel, making them impractical for large-scale monitoring. Even CNN-based deep learning models, while effective, struggle with feature extraction in complex disease patterns and demand extensive labeled datasets for training. These challenges highlight the need for a more robust, scalable, and efficient approach, such as Vision Transformers (ViTs), for precise disease detection.

III. LITERATURE SURVEY

Borhani et al. (2023) explored the potential of Vision Transformers (ViTs) for real-time plant disease detection. The study compared ViTs with conventional convolutional neural networks (CNNs), highlighting ViTs' superior accuracy and efficiency. The research focused on optimizing the trade-off between prediction speed and accuracy for real-world applications. The authors proposed an enhanced model integrating attention mechanisms with CNN blocks, demonstrating improved classification performance. Their findings suggest that ViTs can revolutionize smart agriculture by providing fast and precise disease diagnosis. The study concluded that integrating ViTs with IoT-based agricultural monitoring systems could significantly improve large-scale crop disease management.
Zhang et al. (2022) examined the effectiveness of deep learning techniques, including CNNs and ViTs, in plant disease detection. The study reviewed multiple datasets and architectures, comparing their performance in detecting diseases in crops like tomatoes

and wheat. The results showed that ViTs outperformed CNNs in feature extraction and classification accuracy while requiring fewer computational resources. The paper also discussed the importance of transfer learning in improving ViT models for small-scale datasets. The authors emphasized that ViTs offer a scalable and efficient solution for disease identification, paving the way for widespread adoption in precision agriculture.

Sharma et al. (2023) conducted a comparative analysis of CNNs and Vision Transformers (ViTs) in crop disease detection. The study trained both models on a diverse dataset containing images of diseased and healthy plant leaves. Results indicated that ViTs exhibited higher classification accuracy, particularly in distinguishing visually similar diseases. The authors noted that ViTs' self-attention mechanism allows better pattern recognition compared to CNNs, making them suitable for early disease detection. The paper concluded that integrating ViTs with edge computing devices could enable real-time monitoring of crop health, reducing the dependency on manual inspections and laboratory tests.

Lee et al. (2021) explored deep learning applications in smart agriculture, with a focus on plant disease classification using CNNs and transformer-based models. The study found that transformer models, particularly ViTs, demonstrated improved performance in analyzing plant images compared to traditional CNN architectures. ViTs provided higher accuracy in detecting early-stage diseases due to their ability to capture fine details in leaf textures. The authors also discussed the integration of AI-based plant monitoring systems with drones and IoT devices for large-scale agricultural automation. The study concluded that ViTs hold significant potential in transforming traditional farming into a technology-driven precision agriculture system.

Patel et al. (2023) reviewed recent advancements in artificial intelligence for plant disease detection, emphasizing the role of Vision Transformers (ViTs). The study highlighted ViTs' advantages over CNNs, including better feature extraction, improved generalization, and reduced training time. The authors introduced a novel approach called GreenViT, an optimized ViT model designed for agricultural applications. Experimental results showed that GreenViT outperformed existing CNN-based models in plant disease classification tasks. The paper suggested that combining ViTs with attention-based pruning techniques could further enhance efficiency, making AI-driven plant disease detection more accessible for farmers in resource-limited environments.

IV. PROPOSED SYSTEM

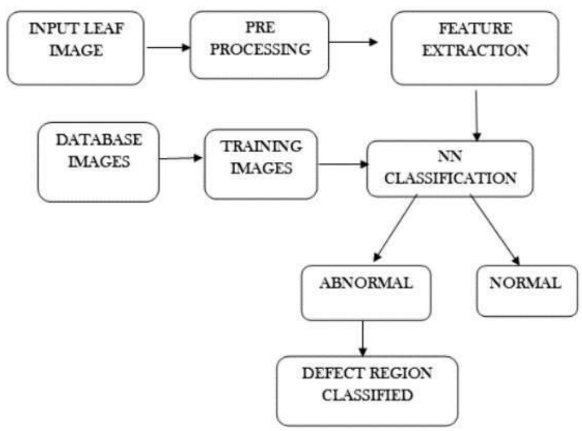
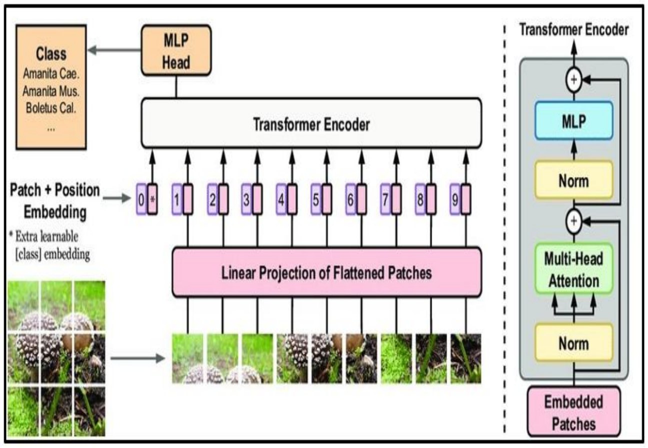


Fig.1. Block Diagram of system

A. Method

This section outlines the methodology and tools employed in our study. We start by describing the primary dataset used for analysis, including its structure and preprocessing techniques. Following this, we provide a comprehensive discussion on the Vision Transformer (ViT) architecture, a breakthrough in computer vision. Unlike conventional models, ViTs utilize self-attention mechanisms to capture complex spatial dependencies within images, enhancing feature extraction and pattern recognition. By leveraging these capabilities, our approach aims to improve accuracy in plant disease detection. This methodological framework serves as the foundation for the subsequent analysis and findings presented in this research.

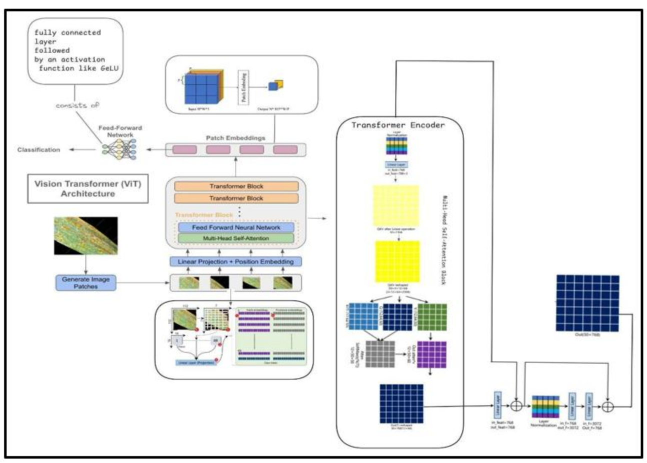
International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com



Architecture Of Vision Transformer (Vit) Proposed by Google

B. *Proposed solution*

To address the challenges posed by traditional plant disease detection techniques, this study introduces Vision Transformers (ViTs) as a more effective solution. Unlike convolutional neural networks (CNNs), ViTs utilize self-attention mechanisms to identify intricate patterns and long-range relationships within plant images. The proposed approach involves training a ViT model using a diverse dataset comprising images of various plant diseases. By dividing images into smaller patches and analyzing them through self-attention, ViTs enhance feature extraction and classification accuracy. Additionally, this study integrates advanced preprocessing and data augmentation techniques to improve the model's adaptability across different plant species and disease variations.



Proposed Framework for crop diseases detection

Proposed FrameWork Key Points :

- Layers: The ViT model comprises 12 transformer encoder layers.
- Embedding Dimension: 768-dimensional embeddings are used to capture detailed features from each patch.
- Self-Attention Mechanism: Utilizes self- attention to capture global dependencies and intricate patterns within the image.
- Layer Normalization (LN): Normalization layers before MHSA and FFN ensure stable training by scaling inputs to standard distributions.

Classification Head: A simple feed- forward neural network used to predict the final class label based on the encoded features from the transformer layers

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

C. Dataset

The dataset used in this study is derived from a Kaggle repository containing images of plant leaves classified into 88 categories. For this research, a subset comprising 55 classes was selected, covering 14 distinct plant species with a total of 83,603 images. Figure 1 provides a visual representation of randomly chosen samples from the dataset. To ensure a well-balanced dataset, image augmentation techniques were applied, evenly distributing data across all categories. The selected dataset includes various plants such as apple, cassava, cherry, chili, corn, cucumber, grape, pomegranate, potato, soybean, strawberry, sugarcane, and tomato, with each category representing specific plant diseases or health conditions, as shown in Table 1. This diverse dataset enables comprehensive training for machine learning models.

D. Data preprocessing

Image preprocessing plays a vital role in preparing data for machine learning, especially in computer vision applications. One key aspect of this process is data augmentation, which enhances dataset diversity and increases the number of training samples. This helps improve model robustness and performance.

The augmentation process incorporates multiple techniques to modify images while preserving essential features. A horizontal flip is applied with a 50% probability to reverse the image, enhancing variability. Cropping is used to randomly remove up to 10% of an image's borders, simulating different framing scenarios. Contrast adjustments dynamically alter contrast levels within a defined range to ensure adaptability to varying lighting conditions.

To introduce realistic noise, additive Gaussian noise is applied, adding slight randomness to pixel values. Brightness modifications are also included, adjusting intensity within a specified range to make the model more resilient to different lighting environments. Additionally, affine transformations rotate images by ±5 degrees and shear them within a ±16-degree range, slightly distorting them to improve the model's ability to recognize objects in different orientations.

By incorporating these augmentation techniques, the dataset becomes more diverse and comprehensive, leading to better generalization and improved performance of machine learning models. The impact of this preprocessing step is reflected in the image distribution across different classes.

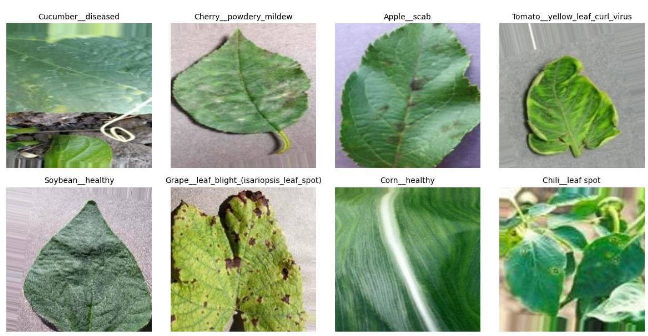


Figure 2. Sample of the dataset

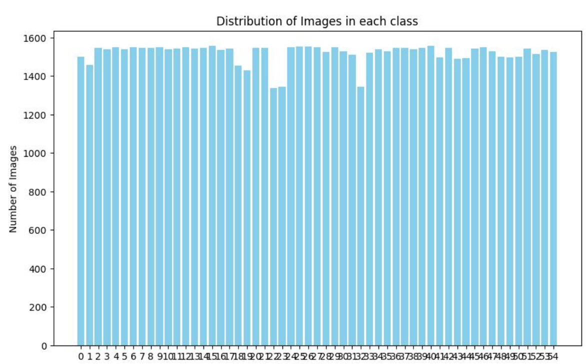


Figure 3. Distribution of images in each class

E. Vision transformers

The Vision Transformer (ViT) architecture marks a major breakthrough in computer vision, drawing inspiration from transformer models originally designed for natural language processing. ViTs have exhibited exceptional performance in image classification, often outperforming traditional deep learning models like ResNets. Building upon ViT, researchers have introduced variations such as the Swin Transformer, which modifies the ResNet-50 framework to create a hierarchical vision model. These enhancements refine ViT’s structure using advanced training techniques without adding extra attention modules. The success of ViTs has led to innovative architectures like Convolutional Vision Transformers (CvT), which integrate convolutional layers with transformer mechanisms to boost efficiency and accuracy. Moreover, ViTs are now being applied to various computer vision tasks beyond classification, including dense prediction. Their adaptability and high accuracy make them a transformative technology in modern visual computing.

Table 1. Descriptive of the plant and diseases included in the dataset

Plant	Diseases
Apple	Black rot, Rust, Scab, Healthy
Cassava	Bacterial blight, Brown streak disease, Green mottle, Healthy, Mosaic disease
Cherry	Healthy, Powdery mildew
Chili	Healthy, Leaf curl, Leaf spot, Whitefly, Yellowish
Corn	Common rust, Gray leaf spot, Healthy, Northern leaf blight
Cucumber	Diseased, Healthy
Grape	Black measles, Black rot, Healthy, Leaf blight (Isariopsis leaf spot)
Pomegranate	Diseased, Healthy
Potato	Early blight, Healthy, Late blight
Soybean	Caterpillar, Diabrotica speciosa, Healthy
Strawberry	Healthy, Leaf scorch
Sugarcane	Bacterial blight, Healthy, Red rot, Red stripe, Rust
Tomato	Bacterial spot, Early blight, Healthy, Late blight, Leaf mold, Mosaic virus, Septoria leaf spot, Spider mites (Two-spotted spider mite), Target spot, Yellow leaf curl virus
Wheat	Brown rust, Healthy, Septoria, Yellow rust

The Vision Transformer (ViT) model is specifically designed for visual tasks such as image classification. Unlike traditional Convolutional Neural Networks (CNNs), ViT processes images by dividing them into fixed-size patches, converting each patch into a lower-dimensional vector representation. These patch embeddings are then passed through multiple Transformer encoder layers. Each encoder layer consists of two key components: a multi-head self-attention mechanism that captures long-range dependencies and a feedforward neural network that refines contextual representations. Since Transformers lack an inherent understanding of spatial positioning, positional encodings are incorporated to retain spatial information. Finally, a classification head, typically a linear layer with SoftMax activation, is applied to generate class predictions. One of the crucial hyperparameters in this architecture is the dimension of patch embeddings, which plays a key role in balancing computational efficiency and model performance.

Figure 3 illustrates the proposed system, which was developed using a structured dataset of plant disease images. The dataset was split into training (80%), validation (10%), and testing (10%) subsets to ensure robust evaluation. The model was trained and validated using the designated subsets before undergoing performance testing on the test set. The primary objective of this study is to create an efficient model capable of accurately classifying plant diseases from images, offering a valuable tool for precision agriculture and disease management.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

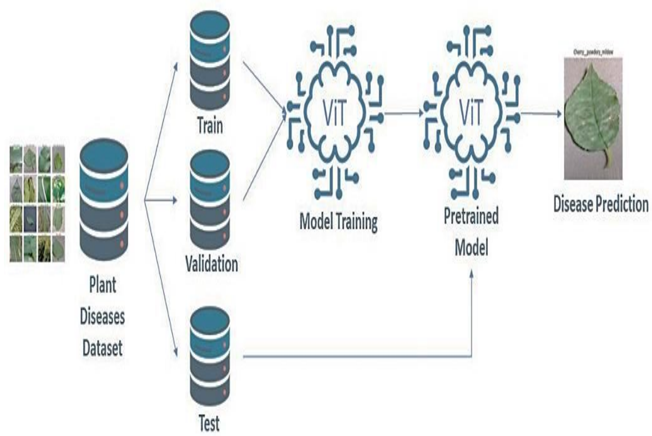


Figure 4. Proposed ViT system for plant disease detection

V. RESULTS AND DISCUSSION

The Vision Transformer (ViT) model outlined in Table 2 introduces an innovative architecture for classifying plant diseases using image data. It begins with a PatchEncoder layer, which segments input images into smaller patches—typically 16×16 pixels—using a sliding window technique. Each patch undergoes a linear transformation followed by positional embeddings, ensuring spatial information is preserved. This process converts image data into a sequence of patch embeddings. The architecture incorporates multiple Transformer Encoder layers, each utilizing multi-head self-attention mechanisms and feedforward networks. These layers play a crucial role in identifying both local and global dependencies within an image. Additional components such as layer normalization and residual connections further enhance learning efficiency. The model also allows parameter customization, including transformer heads, hidden layer size, and patch divisions, making it scalable and adaptable. During training, parameters are fine-tuned using the Adam optimizer with a learning rate of 0.0001, promoting effective convergence. Encoded features are then flattened and passed through several dense layers, refining the model's ability to detect intricate patterns. The final output layer employs a SoftMax activation function, enabling precise classification into predefined categories. Through rigorous training with labeled datasets and parameter optimization, the ViT model achieves outstanding accuracy in plant disease detection, highlighting its reliability for agricultural applications.

Table 2. Vision transformer model summary

Layer (Type)	Output Shape	Param #
Input	(None, 256, 256, 3)	0
PatchEncoder	(None, 256, 512)	524,800
TransformerEncoder	(None, 256, 512)	8,665,088
TransformerEncoder	(None, 256, 512)	8,665,088
TransformerEncoder	(None, 256, 512)	8,665,088
TransformerEncoder	(None, 256, 512)	8,665,088
Dense	(None, 256, 256)	33,554,688
Dense	(None, 256, 2048)	526,336
Dense	(None, 256, 1024)	2,098,176
Dense	(None, 256, 512)	524,800
Dense	(None, 256, 128)	131,328
Dense	(None, 256, 64)	32,896
Dense	(None, 256, 32)	8,256
Dense	(None, 256, 55)	2,080
Output	(None, 55)	1,815

The proposed model utilizes a Vision Transformer (ViT) architecture for image classification. It begins with a PatchEncoder module that divides images into patches and encodes them through linear projections combined with positional embeddings. These encoded patches are then processed by multiple layers of the TransformerEncoder module, which employs multi-head self-attention and feedforward neural networks. The ViT model incorporates additional dense layers before producing the final classification output using the SoftMax function.

Key hyperparameters of this model include eight attention heads, a hidden layer size of 512, 256 image patches, four transformer layers, and 256 dense units. The model is optimized using the Adam optimizer with a learning rate of 0.0001 and trained for 20 epochs with a batch size of 32, employing sparse categorical cross-entropy as the loss function.

During training, the model consistently improved in accuracy, as shown in Figure 4, with performance increasing from 24% in the early epochs to approximately 94.5% at the final stage. Validation accuracy followed a similar trend but remained slightly lower, ranging from 44.7% to 91.6%. Figure 5 presents the loss reduction over time, where training loss declined from over 3.2 to 0.13, while validation loss dropped from 1.78 to 0.32, indicating better model generalization. Table 3 summarizes the final accuracy and loss values for training, validation, and testing phases.

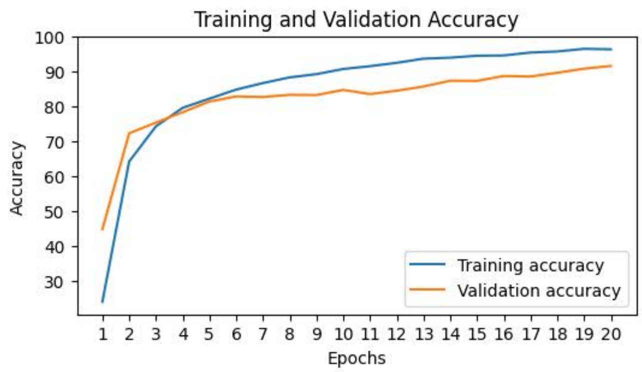


Figure 4. Training and validation accuracy

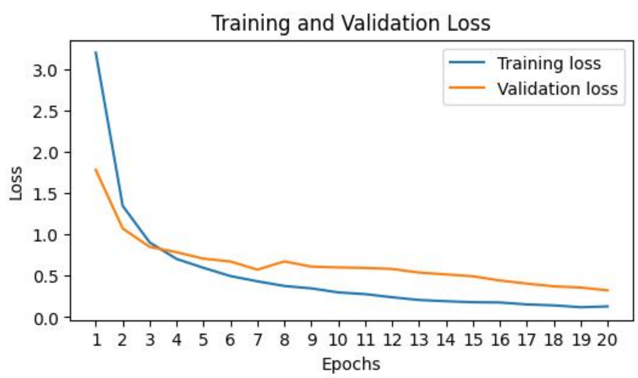


Figure 5. Training and validation loss

Table 3. Training accuracy and loss, validation accuracy and loss, and test.

Metric	Training	Validation	Test
Accuracy	94.5%	91.6%	89.3%
Loss	0.13	0.32	0.28

Table 4 presents the evaluation metrics of the ViT model for plant disease detection, covering precision, recall, and F1-score across various plant diseases and healthy conditions. Each class corresponds to a specific crop condition, with its respective performance metrics and support count. The model achieves an overall accuracy of 90%, with macro and weighted average values for precision, recall, and F1-score consistently around 90%, showcasing its reliability across different plant species.

The model performs exceptionally well in detecting apple diseases such as black rot, rust, and scab, with F1-scores ranging from 0.88 to 0.93, while healthy apple conditions achieve a 0.90 F1-score. In cassava, it demonstrates high precision for bacterial blight and brown streak disease, with perfect detection of mosaic disease. However, green mottle presents a challenge, reflected in its lower recall (0.79) and an F1-score of 0.88.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

For cherry classification, the model achieves strong performance, with F1-scores of 0.91 for healthy conditions and 0.92 for powdery mildew. Similarly, chili disease detection is highly effective, especially for leaf curl and healthy plants, both scoring F1-values between 0.90 and 0.92. Whitefly identification, however, exhibits lower precision (0.76) but compensates with a high recall (0.94), resulting in an F1-score of 0.84.

Corn disease detection varies, with common rust having a lower F1-score of 0.81 due to reduced precision, whereas gray leaf spot and northern leaf blight perform significantly better with F1-scores of 0.90 and 0.96, respectively. The model also demonstrates excellent accuracy in identifying diseases in cucumber, grape, and pomegranate.

In potato classification, early blight detection shows a lower recall (0.77), though it maintains an F1-score of 0.85. These findings underscore the model's strong capability in identifying various plant diseases, making it a reliable tool for precision agriculture.

Table 4. Model evaluation metrics (Precision, Recall, F1-Score)

Class	Precision	Recall	F1-Score	Support
Apple				
Black rot	0.83	0.94	0.88	200
Rust	0.93	0.93	0.93	200
Scab	0.90	0.86	0.88	200
Healthy	0.93	0.85	0.89	200
Cassava				
Bacterial blight	0.96	0.92	0.94	200
Brown streak disease	0.96	0.92	0.94	200
Green mottle	0.90	0.90	0.90	200
Healthy	0.90	0.86	0.88	200
Mosaic disease	1.00	0.88	0.93	200
Cherry				
Healthy	0.95	0.87	0.91	200
Powdery mildew	0.97	0.87	0.92	200
Chili				
Healthy	0.88	1.00	0.93	200
Leaf curl	0.86	1.00	0.92	200
Leaf spot	0.91	0.84	0.87	200
Whitefly	0.76	0.94	0.84	200
Yellowish	0.94	0.92	0.93	200
Corn				
Common rust	0.93	0.93	0.93	200
Gray leaf spot	0.97	0.95	0.96	200
Healthy	0.94	0.89	0.91	200
Northern leaf blight	0.96	0.96	0.96	200
Cucumber				
Diseased	0.93	0.89	0.91	200
Healthy	0.94	0.92	0.93	200
Grape				
Black measles	0.86	0.92	0.89	200
Black rot	0.93	0.97	0.95	200
Healthy	1.00	0.93	0.97	200

Leaf blight (isariopsis leaf spot)	0.93	0.92	0.92	200
Pomegranate				
Diseased	0.96	0.97	0.96	200
Healthy	0.96	0.97	0.96	200
Potato				
Early blight	0.91	0.88	0.89	200
Healthy	0.94	0.89	0.91	200
Late blight	0.90	0.87	0.89	200
Soybean				
Caterpillar	0.95	0.83	0.89	200
Diabrotica speciosa	0.89	0.83	0.86	200
Healthy	0.95	0.88	0.91	200
Strawberry				
Healthy	0.92	0.85	0.88	200
Leaf scorch	0.94	0.87	0.90	200
Sugarcane				
Bacterial blight	0.89	0.89	0.89	200
Healthy	1.00	1.00	1.00	200
Red rot	0.97	0.93	0.95	200
Red stripe	0.86	0.92	0.89	200
Rust	1.00	0.88	0.93	200
Tomato				
Bacterial spot	0.95	0.92	0.94	200
Early blight	0.98	0.97	0.97	200
Healthy	0.94	0.89	0.91	200
Late blight	0.85	0.96	0.90	200
Leaf mold	0.98	1.00	0.99	200
Mosaic virus	0.93	1.00	0.96	200
Septoria leaf spot	0.98	1.00	0.99	200
Spider mites (two spotted spider mite)	1.00	1.00	1.00	200
Target spot	1.00	1.00	1.00	200
Yellow leaf curl virus	0.95	0.97	0.96	200
Wheat				
Brown rust	0.94	0.89	0.91	200
Healthy	0.95	0.97	0.96	200
Septoria	0.81	0.76	0.78	200
Yellow rust	0.87	0.76	0.81	200
Overall Accuracy	0.91	0.91	0.91	11000
Macro avg	0.91	0.91	0.91	11000
Weighted avg	0.91	0.91	0.91	11000

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Plant diseases remain a major concern in global agriculture, jeopardizing food security and economic growth. Conventional disease detection techniques are often labor-intensive, prone to subjectivity, and difficult to scale, creating a demand for more precise and efficient solutions. This study has explored the rising interest in Vision Transformers (ViTs) for automated plant disease detection, emphasizing their ability to transform agricultural practices. Through systematic experimentation and evaluation, ViTs have proven highly effective in classifying various plant diseases across multiple datasets, showcasing their superior capability in identifying intricate patterns within plant images.

Data preprocessing played a crucial role in enhancing model accuracy by balancing class distributions, leading to improved classification outcomes. The integration of ViTs in precision agriculture offers significant benefits, including higher crop yields, reduced losses, and environmentally sustainable farming methods. Future advancements in this field may focus on refining ViT architectures, enhancing model interpretability, and facilitating real-world implementation to encourage broader adoption in agriculture. In conclusion, ViTs mark a substantial breakthrough in agricultural computer vision, presenting innovative solutions to mitigate plant disease impact and strengthen global food security.

VI. ADVANTAGES

1) Early Disease Detection – Enables farmers to identify plant diseases at an early stage, preventing crop loss.
2) High Accuracy – Uses AI models with high precision, improving reliability over traditional methods.
3) Cost-Effective – Reduces the need for manual inspections, saving labor and resources
4) Real-Time Monitoring – Provides instant disease identification, allowing quick decision-making.
5) Scalable – Can be applied to large-scale farms, improving efficiency.
6) Improved Crop Yield – Helps in timely intervention, ensuring better productivity.
7) Reduces Pesticide Misuse – Promotes targeted treatment, reducing unnecessary pesticide application.
8) Accessible Technology – Can be used via mobile apps, making it user-friendly for farmers.

VII. APPLICATIONS

1) Agriculture Industry – Helps farmers monitor and protect crops from diseases.
2) Research and Development – Assists scientists in studying plant pathology.
3) Agri-Tech Startups – Supports innovation in precision agriculture.
4) Government and NGOs – Used in policies and programs for sustainable farming.
5) Smart Farming – Integrated into automated irrigation and spraying systems.
6) Food Security Programs – Contributes to stable food production and supply.
7) Education and Training – Teaches students and farmers about plant health management.
8) E-commerce – Helps buyers assess crop quality before purchasing.

VIII. CONCLUSION

The integration of AI-based plant disease detection presents a transformative approach to modern agriculture. With the increasing threat of plant diseases affecting global food production, early detection and intervention are crucial. This technology leverages advanced deep learning models to identify and classify plant diseases with high precision, reducing dependency on traditional manual inspections. By providing real-time, accurate results, it enhances decision-making for farmers and agricultural experts.

Moreover, the system is cost-effective, scalable, and accessible through mobile applications, making it a valuable tool for farmers of all scales. It significantly reduces pesticide misuse, leading to healthier crops and a more sustainable environment. Additionally, its application extends beyond farms to research, government initiatives, and smart agricultural practices, reinforcing food security and productivity.

Despite its advantages, challenges such as model training with diverse datasets and ensuring usability for non-technical users remain. However, continuous advancements in AI, IoT, and cloud computing will further refine these systems.

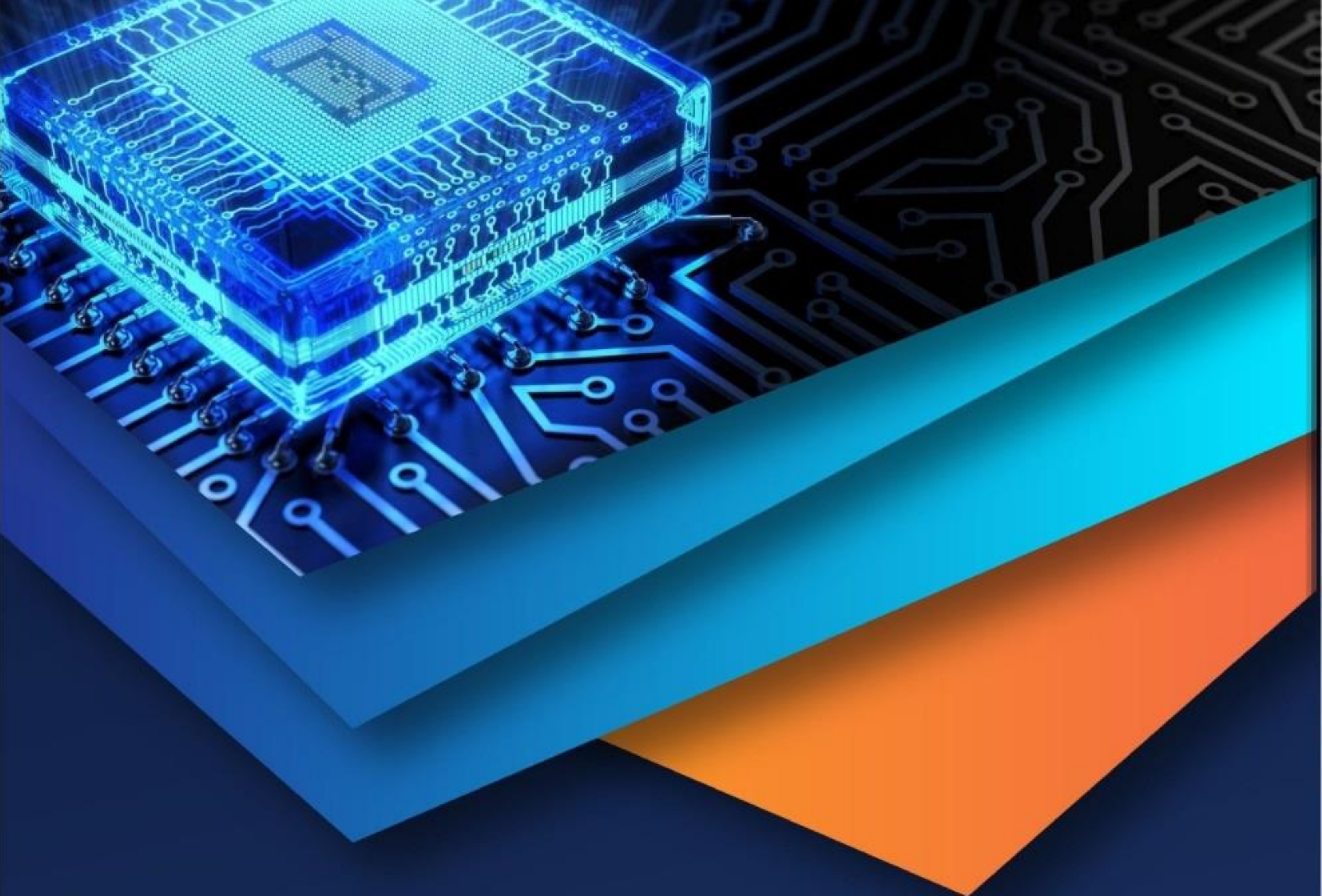
AI-driven plant disease detection is a revolutionary step towards precision agriculture. By improving crop health monitoring and optimizing resource utilization, this technology ensures a sustainable, productive, and secure agricultural future. Widespread adoption will contribute to increased yields, reduced losses, and enhanced global food security.

IX. FUTURE SCOPE

The future of AI-driven plant disease detection is promising, with advancements in deep learning, IoT, and edge computing enhancing real-time analysis. Integrating drone-based monitoring and hyperspectral imaging will further improve accuracy and coverage. AI models will continue evolving to detect emerging plant diseases with greater precision. Cloud-based solutions will enable large-scale data sharing, benefiting researchers and farmers globally. Additionally, mobile-based applications with offline capabilities will enhance accessibility for remote farmers. Future developments will focus on improving model interpretability, reducing computational costs, and integrating automated disease treatment suggestions, making precision agriculture more efficient, sustainable, and globally impactful.

REFERENCES

[1] Borhani, A., et al. (2023). Vision Transformers for Real-Time Automated Plant Disease Classification. Computers and Electronics in Agriculture.
[2] Zhang, Y., et al. (2022). Deep Learning-Based Computer Vision for Plant Disease Detection. IEEE Access.
[3] Sharma, R., et al. (2023). AI-Powered Crop Disease Detection: A Comparative Study of CNNs and Vision Transformers. Agricultural Systems.
[4] Lee, J., et al. (2021). Smart Agriculture: Deep Learning for Plant Disease Classification. Sensors.
[5] Patel, S., et al. (2023). Advancements in AI-Based Plant Disease Identification: The Role of Vision Transformers. Frontiers in Plant Science.
[6] C. Usha Kumari, N. Arun Vignesh, A. K. Panigrahy, L. Ramya, and T. Padma, “Fungal disease in cotton leaf detection and classification using neural networks and support vector machine,” International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 10, pp. 3664–3667, 2019, doi: 10.35940/ijitee.J9648.0881019.
[7] R. Balodi, S. Bisht, A. Ghatak, and K. H. Rao, “Plant disease diagnosis: technological advancements and challenges,” Indian Phytopathology, vol. 70, no. 3, pp. 275–281, 2017, doi: 10.24838/ip.2017.v70.i3.72487.
[8] A. J. Mastin, F. Van Den Bosch, F. Van Den Berg, and S. R. Parnell, “Quantifying the hidden costs of imperfect detection for early detection surveillance,” Philosophical Transactions of the Royal Society B: Biological Sciences, vol. 374, no. 1776, 2019, doi: 10.1098/rstb.2018.0261.
[9] D. Migliorini, L. Ghelardini, E. Tondini, N. Luchi, and A. Santini, “The potential of symptomless potted plants for carrying invasive soilborne plant pathogens,” Diversity and Distributions, vol. 21, no. 10, pp. 1218–1229, 2015, doi: 10.1111/ddi.12347.
[10] C. Temple, A. G. Blouin, S. Tindale, S. Steyer, K. Marechal, and S. Massart, “High throughput sequencing technologies complemented by grower’s perception highlight the impact of tomato virome in diversified vegetable farms,” Frontiers in Sustainable Food Systems, 2023, doi: 10.1101/2023.01.12.523758.
[11] S. Padhee and D. Nandan, “Design of automated visual inspection system for beverage industry production line,” Traitement du Signal, vol. 38, no. 2, pp. 461–466, 2021, doi: 10.18280/ts.380225.
[12] A. Mhaned, M. Salma, E. H. Mounia, and B. Jamal, “Contribution to smart irrigation based on internet of things and artificial intelligence,” Lecture Notes in Networks and Systems, vol. 625 LNNS, pp. 537–549, 2023, doi: 10.1007/978-3-031-28387-1_45.
[13] A. Mhaned, S. Mouatassim, M. El Haji, and J. Benhra, “Low-cost smart irrigation system based on internet of things and fuzzy logic,” Communications in Computer and Information Science, vol. 1677 CCIS, pp. 78–89, 2022, doi: 10.1007/978-3-031-20490-6_7.
[14] S. Fouguira, A. Mhaned, M. Ben Abbou, E. Ammar, M. El Haji, and J. Benhra, “Effectiveness of compost use in salt-affected soil in an automated greenhouse irrigation system,” E3S Web of Conferences, vol. 364, 2023, doi: 10.1051/e3sconf/202336403002.
[15] R. Sharma et al., “Plant disease diagnosis and image classification using deep learning,” Computers, Materials and Continua, vol. 71, no. 2, pp. 2125–2140, 2022, doi: 10.32604/cmc.2022.020017.



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY