TRANSFORM LEARNING WITH VISION TRANSFORMERS FOR ACCURATE PLANT DISEASE DIAGNOSIS IN FIELD IMAGES

S. Sankar Ganesh[1*], Bathina Nikitha[2], Thakur Yokshitha Singh[2], Barapati Sadvika[2]

[1]Associate Professor, [2]UG Student, [1,2]Department of Computer Science and Engineering (AI & ML),

[1,2]Kommuri Pratap Reddy Institute of Technology, Ghanpur, Ghatkesar, 501301, Telangana, India.

*Correspondence: S. Sankar Ganesh (sankar@kpritech.ac.in)

ABSTRACT

Plant diseases significantly affect agricultural productivity and food security by reducing both crop yield and quality. Early detection and accurate diagnosis are essential to prevent large-scale crop losses and support sustainable farming practices. With rapid advancements in artificial intelligence and computer vision, automated plant disease detection systems have become valuable tools for identifying plant diseases using leaf images. These intelligent systems help farmers and agricultural experts detect diseases quickly and make timely crop management decisions. Traditional plant disease detection methods mainly rely on manual inspection by agricultural experts or laboratory-based analysis. These methods are time-consuming, require specialized knowledge, and are not practical for large-scale agricultural monitoring. Conventional machine learning approaches have attempted to automate disease detection, but they often depend on handcrafted features and limited datasets. As a result, they struggle to capture complex visual patterns in plant leaf images and may produce unreliable results under varying environmental conditions. To overcome these limitations, this research proposes a Transformer-Driven Hybrid Pipeline for Scalable Plant Disease Diagnostics, integrating deep learning, machine learning, and explainable artificial intelligence (XAI). The system first applies an XAI module to verify whether the uploaded image contains a valid plant leaf and extracts contextual attributes such as plant type, health condition, and dominant color. After validation, deep feature extraction is performed using DenseNet121, which captures spatial features, and a Vision Transformer (ViT) that model's global contextual relationships. The extracted DenseNet121 features are processed using classifiers including Perceptron, Nearest Centroid Classifier (NCC), and an ensemble-based DenseNet121 with Ensemble Neighbor Model (ENM) that combines K-Nearest Neighbor (KNN), and Radius Neighbors Classifier (RNC). Additionally, DeepPercepNet (DPN) integrates ViT feature embeddings with a Perceptron classifier to enhance prediction performance. The system detects multiple plant diseases across crops such as apple, cherry, corn, grape, peach, pepper, and strawberry, including healthy leaf conditions. A Tkinter-based interface enables interactive image uploads and predictions, while model performance is evaluated using accuracy, precision, recall, F1-score, confusion matrix, and ROC analysis.

Keywords: Plant Disease Detection, Agricultural Productivity, Artificial Intelligence (AI), DenseNet121, Vision Transformer (ViT).

1. INTRODUCTION

Plant diseases cause considerable economic loss in the global agricultural industry [1]. A current challenge in agriculture is the development of reliable methods for detecting plant diseases and plant stress [2]. Existing disease detection methods mainly involve the manual and visual assessments of crops for visible disease indicators. These methods are time-consuming and demanding, depending on the crop field area. Manual detection depends on the apparent disease or stress symptoms, which mostly manifest in the middle to late stages of infection [3]. Visual assessments depend on the ability and reliability of the assessor, who is prone to human error and subjectivity. Improved technologies for plant disease detection and stress identification beyond visual appearance are required to reduce yield loss and improve crop protection [4]. Generally, a plant develops a disease when it is continuously disturbed by certain causal agents, which cause abnormal physiological processes that disrupt the typical structure, growth, function, or other activities of the plant. Plant diseases are caused by microorganisms, such as certain viruses, bacteria, fungi, nematodes, or protozoa. Diseases enter plants through wounds or natural openings or are carried and inserted by vectors such as insects [5].

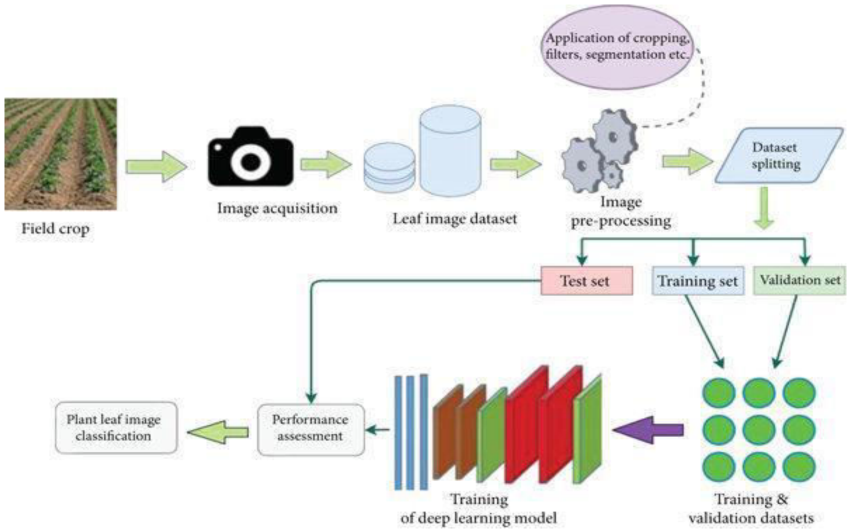


Fig. 1: Approach to plant disease detection and classification

Plants respond to stress by undergoing biophysical and biochemical changes, such as a decrease in the amount of chlorophyll in their leaves or alterations in the structure of their leaf cells as presented in fig. 1. Decreased chlorophyll pigments have been shown to drastically impair a leaf's ability to absorb light [6]. Different light absorption patterns of leaves indicate plant stress or disease infection. For example, brownish-yellow spots on the upper side of a leaf are indicators of downy mildew disease, which commonly affects several plants [7].

2. RELATED WORK

The development of automated plant disease detection systems has evolved significantly with the integration of computer vision and deep learning techniques. Earlier approaches focused on basic image processing and manual feature extraction, which lacked robustness in real-world scenarios. Recent advancements emphasize deep learning-based models that improve classification accuracy and enable real-time detection in agricultural environments.

2.1 Deep Learning-Based Detection Models

Modern object detection frameworks have played a crucial role in plant disease classification. Ghafar A et al. [8] proposed a methodology using the YOLOv8 architecture for real-time disease detection. The model was trained on a custom dataset and evaluated on both test and unseen images, demonstrating strong generalization and high classification accuracy in real-world conditions. Similarly, Abasi AK et al. [11] developed a deep learning-based model that achieved an accuracy of 91.4% with minimal overfitting. Their approach outperformed transfer learning models such as Inception-v3 and EfficientNet-B2, highlighting its effectiveness in precision agriculture.

2.2 Hyperspectral Imaging and Advanced Data Acquisition

Beyond conventional RGB images, advanced sensing techniques have been explored for improved disease detection. Kuswidiyanto LW et al. [9] provided a comprehensive overview of hyperspectral imaging combined with deep learning algorithms. Their study incorporated UAV-based data acquisition and analyzed various workflows for plant disease detection. This approach enhanced the ability to capture detailed spectral information, improving classification performance in complex agricultural environments.

2.3 Feature Selection and Ensemble Learning Approaches

Efficient feature selection and classification methods are essential for handling large datasets. Sengupta S et al. [10] utilized rough set theory for feature selection on infected rice plant images. The selected features were processed using ensemble classification techniques within a map-reduce framework, improving computational efficiency and scalability. This approach demonstrated the importance of combining feature optimization with distributed computing for disease prediction.

2.4 Data Augmentation and Generative Techniques

Data scarcity remains a major challenge in plant disease detection. Sun R et al. [12] addressed this issue by proposing a generative adversarial network-based approach to synthesize lesion images. The method incorporated edge-smoothing and image pyramid techniques to generate realistic images with varying lesion sizes. This improved dataset diversity and enhanced the overall accuracy of classification models.

2.5 Large-Scale Training and Optimization Techniques

Model performance is heavily influenced by training strategies and hyperparameter tuning. Pandian JA et al. [13] developed a deep convolutional neural network trained over multiple GPUs for 1000 epochs. A coarse-to-fine random search method was used to identify optimal hyperparameters, resulting in improved training efficiency and prediction accuracy. This study emphasized the importance of computational optimization in deep learning models.

2.6 Object Detection Frameworks and Localization

Accurate localization of plant diseases is essential for effective diagnosis. Saleem MH et al. [14] applied multiple deep learning meta-architectures, including SSD, RCNN, and RFCN, using the TensorFlow object detection framework. Their approach focused on both classification and localization tasks, achieving improved mean average precision through optimization techniques.

2.7 Hybrid Models for Disease Classification

Hybrid models combining deep learning and machine learning techniques have shown promising results. Altalak M et al. [15] proposed a system integrating CNN, CBAM, and SVM for early detection of tomato leaf diseases. The use of attention mechanisms improved feature extraction, while SVM enhanced classification performance. The model demonstrated effective results across multiple disease classes.

2.8 Research Gap

Although significant progress has been made in plant disease detection using deep learning, several limitations still exist. Many existing approaches rely on large datasets and high computational resources, limiting their practical deployment in real-time agricultural environments. Additionally, most models focus either on classification or localization, with limited integration of both functionalities in a lightweight framework. There is also a lack of efficient models tailored for specific crops such as sugarcane under varying environmental conditions. The proposed system addresses these challenges by developing an optimized, accurate, and scalable model for real-time plant disease detection with improved generalization capability.

3. PROPOSED SYSTEM

The proposed system is a Transformer-driven hybrid pipeline for plant disease diagnostics that combines deep learning feature extraction, machine learning classifiers, and explainable AI. Plant leaf images are first processed using DenseNet121 and ViT models to extract meaningful visual features. These features are classified using multiple machine learning models including Perceptron, NCC, and ENM, while the proposed DPN (ViT with Perceptron) improves classification performance. The system integrates a Tkinter-based graphical interface with secure authentication using SHA-256 hashing for user and admin access. Additionally, an Explainable AI module powered by the Explainable API analyzes plant attributes such as plant type, health status, and dominant color. The final system provides prediction visualization, evaluation metrics, and report generation for effective plant disease diagnosis is illustrated in Fig. 2.

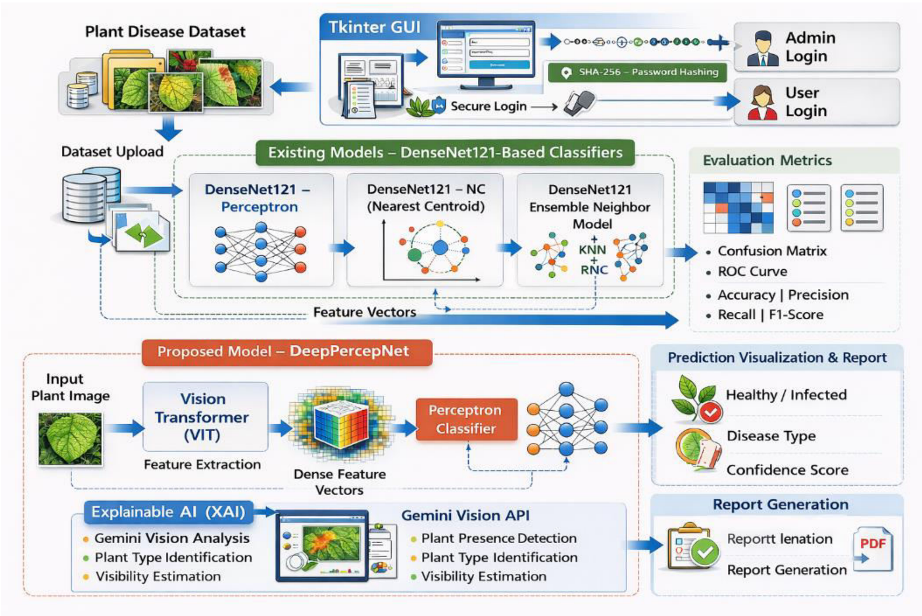


Fig. 2: Proposed workflow of system architecture for classifying plant diseases.

1. Dataset Upload: The system begins by uploading the plant disease dataset through the GUI. Images are organized by class labels representing different plant conditions or disease categories. This dataset forms the foundation for training and evaluating the classification models.

2. DenseNet121 Feature Extraction: DenseNet121, a deep convolutional neural network pretrained on ImageNet, is used to extract high-level visual features from plant leaf images. These deep feature vectors capture important patterns such as texture, color variations, and disease symptoms that are useful for classification.

3. Existing Model Classification: The extracted features are passed to multiple machine learning classifiers including Perceptron, NCC, and the DenseNet121 ENM. These models provide baseline performance comparisons and help evaluate the effectiveness of different classification strategies.

4. Proposed Model DPN: The proposed model integrates ViT feature extraction with a Perceptron classifier. ViT captures global contextual relationships in plant images through attention mechanisms, producing dense feature embeddings that improve disease detection accuracy.

5. Explainable AI: An Explainable AI module analyzes the uploaded plant image using the Gemini Vision API. This module extracts interpretable attributes such as plant presence, plant type, health status, visibility level, and dominant color to provide additional insights alongside the classification result.

6. Prediction Visualization and Report Generation: The final system displays prediction results along with evaluation metrics including accuracy, precision, recall, F1-score, ROC curve, and confusion matrix. The predicted class label and analysis results are visualized on the image, and a report is generated for user interpretation.

DPN Model

DPN is the proposed classification model designed to improve plant disease detection by combining transformer-based feature extraction with a linear classification mechanism. In this approach, the ViT extracts deep feature embeddings from plant leaf images by analysing relationships between image patches using self-attention mechanisms. These embeddings capture global contextual information and represent the visual characteristics of the input image. The extracted feature vectors are then provided to a Perceptron classifier, which learns a linear decision boundary to classify different plant disease categories. During training, the Perceptron updates its weights and bias values based on classification errors to improve prediction performance. This hybrid architecture leverages the powerful feature representation capability of transformers and the efficiency of a lightweight linear classifier. The internal workflow of the DPN model is illustrated in Fig. 3.

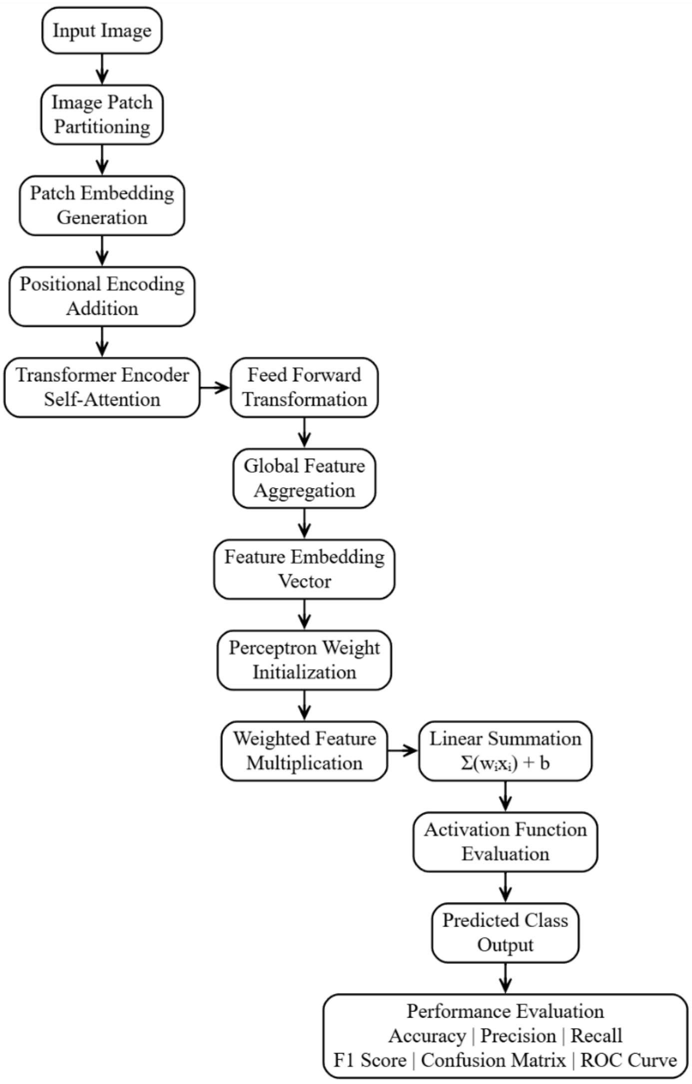


Fig. 3: Internal workflow of DPN

ViT Feature Embedding Generation: The ViT processes the input image and divides it into patches. Through transformer encoder layers and self-attention mechanisms, the model extracts deep feature embeddings representing the global visual structure of the plant leaf.

Feature Vector Preparation: The generated feature embeddings are organized into a structured feature vector. This vector serves as the input representation for the Perceptron classifier.

Linear Combination Computation: The Perceptron multiplies each feature value with its corresponding weight and adds a bias term to compute a linear activation score.

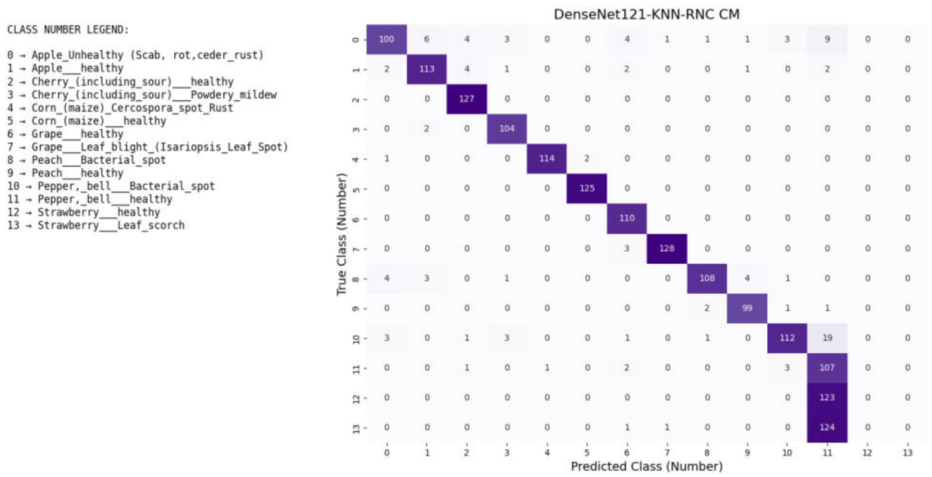
Activation Function Evaluation: The activation function evaluates the computed score and determines the predicted class label based on the decision boundary.

Error Evaluation and Parameter Update: During training, the predicted label is compared with the actual label. If a classification error occurs, the Perceptron updates its weight and bias parameters to improve prediction accuracy.

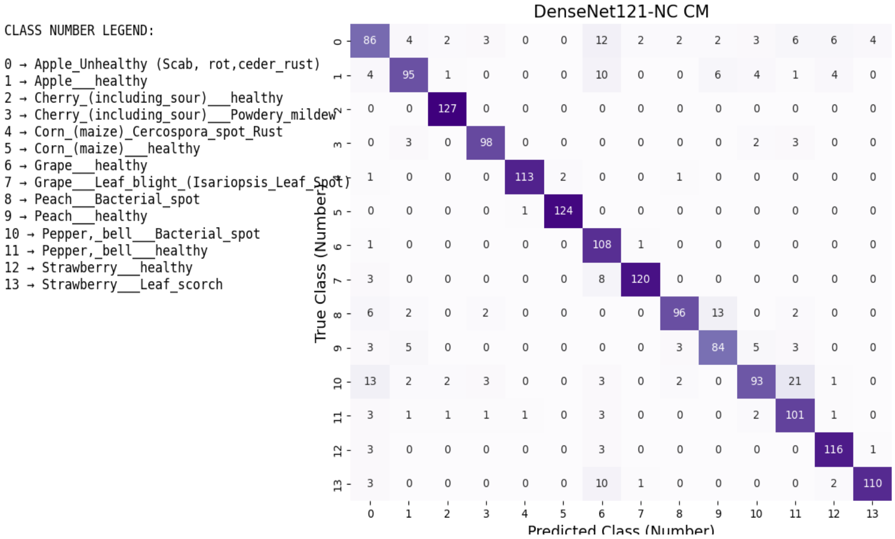
Final Classification Output: After training convergence, the optimized Perceptron classifier produces the final prediction, identifying the plant disease class associated with the input feature vector.

4. RESULTS DESCRIPTION

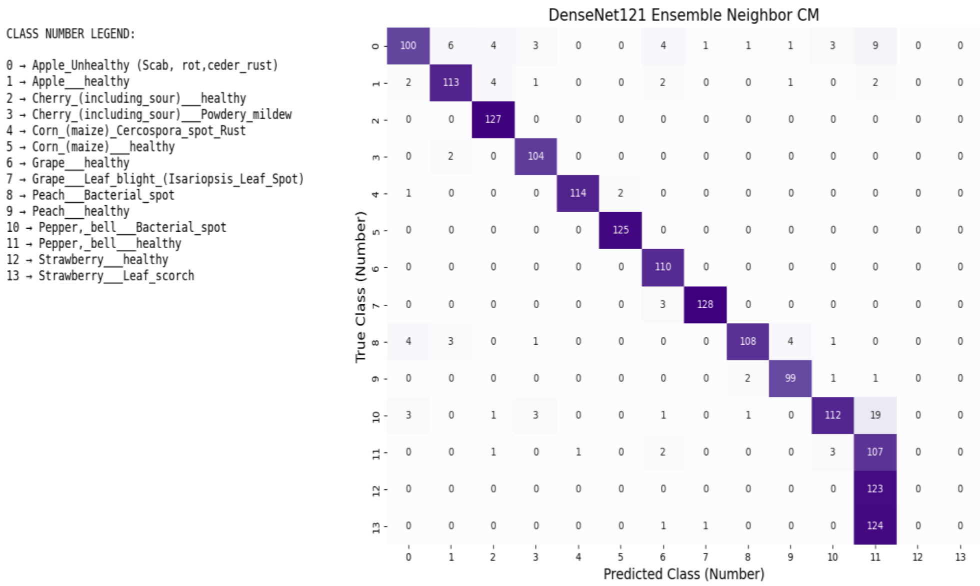
Fig. 4 depicts the confusion matrices obtained from three different classification approaches used for plant disease identification based on DenseNet121 extracted features across fourteen plant classes. Subfigure (a) presents the Perceptron classifier results where several classes achieve high correct predictions such as Cherry healthy (127), Pepper bell healthy (125), Peach healthy (128), Strawberry healthy (123), and Strawberry leaf scorch (124), while other classes such as Apple healthy (113), Grape healthy (114), and Pepper bell bacterial spot (108) also show strong classification accuracy. However, minor misclassifications are observed for a few classes, for example Apple unhealthy samples are predicted as other classes with values such as 6, 4, and 3 instances, and Pepper bell bacterial spot shows some confusion with neighboring classes. Subfigure (b) illustrates the Nearest Centroid classifier where correct predictions are slightly reduced for several categories including Apple unhealthy (86), Apple healthy (95), Corn cercospora rust (98), Peach bacterial spot (96), and Pepper bell bacterial spot (84), while some classes such as Cherry powdery mildew (127) and Peach healthy (124) still maintain high correct classification. Subfigure (c) shows the DenseNet121 ENM which improves the classification performance with higher diagonal values including Apple unhealthy (100), Apple healthy (113), Cherry healthy (127), Corn cercospora rust (104), Grape healthy (114), Pepper bell healthy (125), Peach healthy (128), Strawberry healthy (123), and Strawberry leaf scorch (124).



(a)



(b)



(c)

Fig. 4: Confusion Matrices of (a) Perceptron (b) Nearest Centroid (c) DenseNet121 Ensemble Model

Fig. 5 depicts the confusion matrix of the proposed DPN model for multi-class plant disease classification. The matrix shows highly concentrated diagonal values indicating accurate predictions across all fourteen plant categories. Several classes achieve perfect or near-perfect classification results including Apple unhealthy (121), Apple healthy (120), Corn cercospora rust (121), Grape leaf blight (121), Peach bacterial spot (121), Pepper bell bacterial spot (125), Pepper bell healthy (121), Strawberry healthy (126), and Strawberry leaf scorch (123). Other categories such as Cherry powdery mildew (115), Peach healthy (124), and Pepper bell healthy (122) also demonstrate strong prediction accuracy. The absence of off-diagonal values indicates that the model produces minimal or no misclassification between classes.

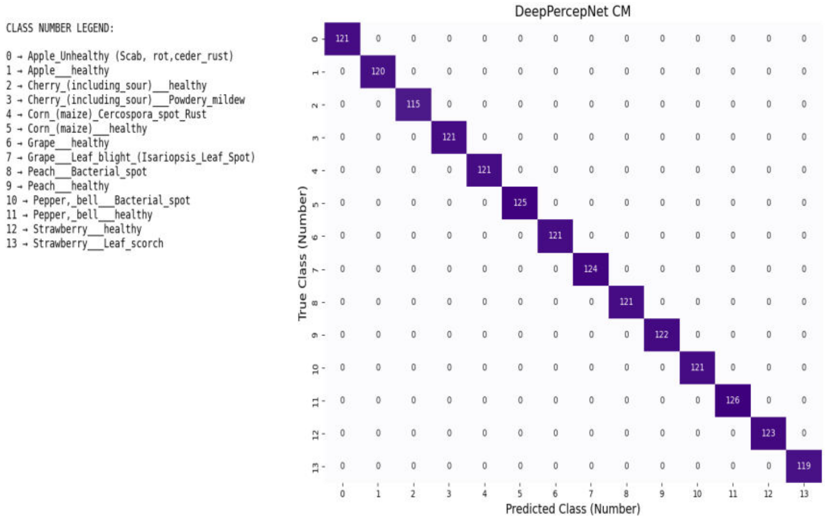


Fig. 5: Confusion Matrix of DPN

Fig. 6 depicts the test image uploading stage where a plant leaf image is selected from the dataset for prediction. In this step, the user browses the dataset directory and chooses a specific leaf image belonging to a particular plant class. Once the image is selected, the system loads the input sample and initiates the prediction process using the trained DPN model. This stage allows the framework to perform real-time disease diagnosis on unseen plant leaf images.

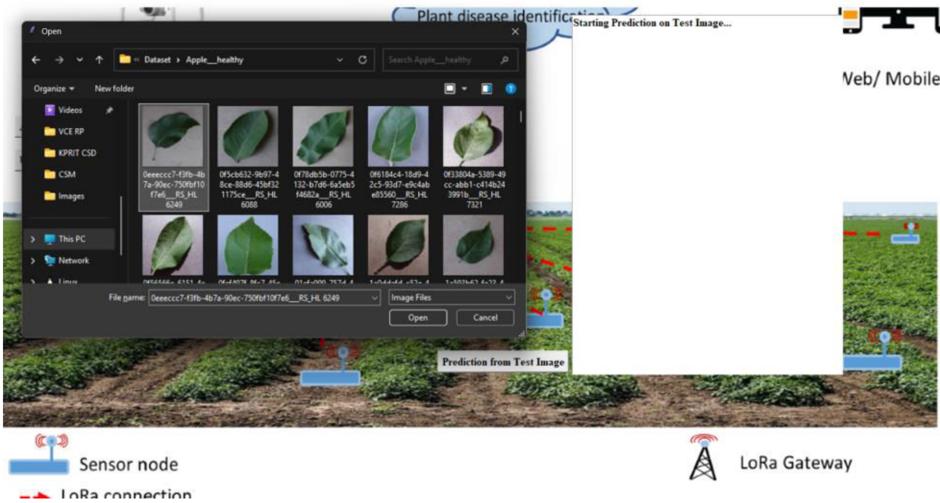
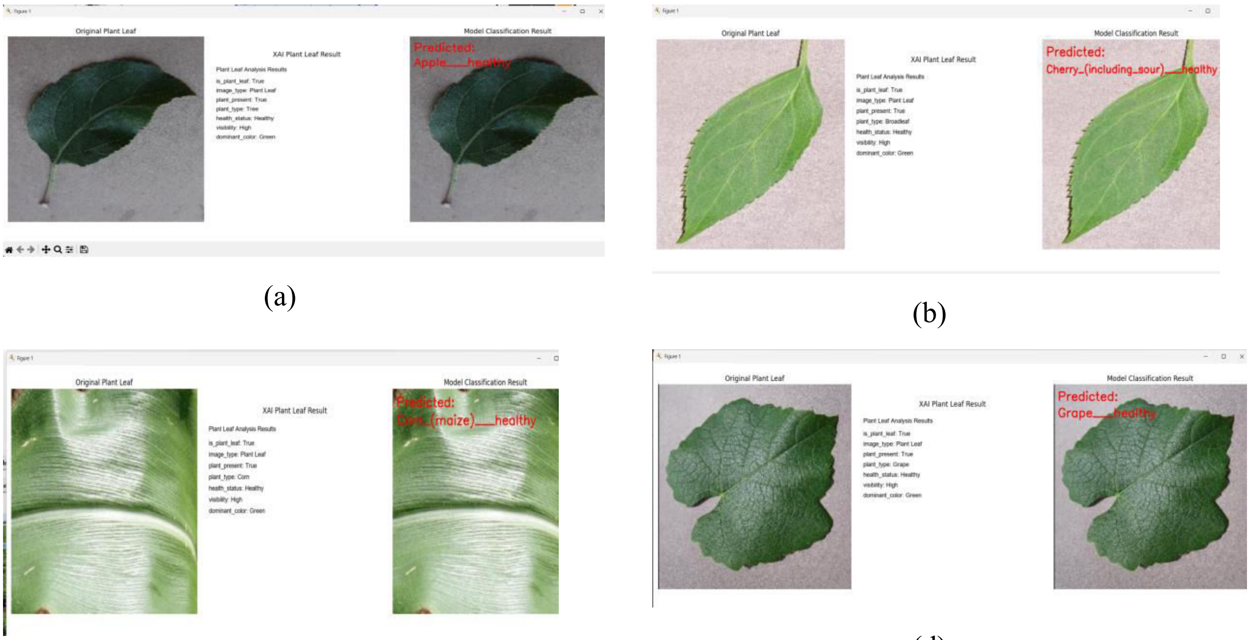


Fig. 6: Uploading the test image screen

Fig. 7 illustrates sample prediction results generated by the DPN model along with the XAI analysis for different plant leaf images. Subfigure (a) presents the prediction result for an apple leaf where the model classifies the image as Apple healthy, and the XAI analysis confirms that the input is a plant leaf with plant presence marked as true, plant type identified as tree, health status detected as healthy, visibility reported as high, and the dominant color identified as green. Subfigure (b) shows the prediction result for a cherry leaf where the system predicts the class Cherry healthy, and the XAI module reports plant presence as true, plant type as broadleaf, health status as healthy, high visibility, and green as the dominant color in the image. Subfigure (c) depicts the classification of a corn leaf image where the model predicts Corn healthy, while the XAI output indicates that the image contains a plant leaf with plant presence true, plant type corn, health status healthy, high visibility, and dominant color green. Subfigure (d) presents the prediction for a grape leaf image classified as Grape healthy, and the XAI explanation verifies plant presence as true, plant type grape, healthy leaf condition, high image visibility, and green as the dominant color.



(a) (b) (c) (d)

Fig. 7: Sample Test Prediction of (a) apple (b) Cherry (c) Corn (d) Grape Images

Comparative Analysis

Table 1 presents the overall performance comparison of all evaluated models using accuracy, precision, recall, and F1-score metrics. The DenseNet121–Perceptron model achieved an accuracy of 93.88%, with 94.65% precision, 94.21% recall, and 93.83% F1-score, indicating strong classification capability across the plant disease categories. The DenseNet121–NC model recorded slightly lower performance with 86.53% accuracy, 87.04% precision, 86.91% recall, and 86.57% F1-score, showing moderate effectiveness in disease detection. The DenseNet121 ENM achieved 79.23% accuracy, 76.06% precision, 79.90% recall, and 76.57% F1-score, which represents the lowest performance among the baseline models due to higher misclassification in certain classes. In contrast, the proposed DPN model achieved 100% accuracy, 100% precision, 100% recall, and 100% F1-score, demonstrating perfect classification performance across all plant disease categories. The results clearly indicate that the proposed DPN model significantly outperforms the existing models in terms of accuracy and reliability for plant disease detection.

Table 1: Performance comparison of Algorithms.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Perceptron	93.88	94.65	94.21	93.83
NCC	86.53	87.04	86.91	86.57
ENM	79.24	76.06	79.90	76.57
DPN	100.00	100.00	100.00	100.00

Table 3 presents the class-wise evaluation metrics of the DenseNet121-Perceptron model for plant disease classification. The model achieved an overall accuracy of 93.88%, with 94.65% precision, 94.21% recall, and 93.83% F1-score, indicating strong and balanced predictive performance. Several classes such as Cherry healthy (100% recall), Corn healthy (99%), Corn Cercospora rust (99%), Grape leaf blight (100%), and Strawberry leaf scorch (100%) were detected with very high accuracy, showing that the model effectively learned distinctive disease features. However, comparatively lower recall values were observed for Apple unhealthy (61%) and Pepper bell bacterial spot (80%), suggesting that a small portion of diseased samples were misclassified. Overall, the DenseNet121–Perceptron model demonstrates reliable classification capability with high precision and recall across most plant disease categories.

5. Conclusion

This research developed a Plant Disease Diagnostics that integrates deep learning, machine learning, and explainable AI techniques for automated plant disease detection. The system uses DenseNet121 and ViT models for feature extraction, combined with classifiers such as Perceptron, NCC, and the DenseNet121 ENM. Experimental results show that the Perceptron model achieved 93.88% accuracy, while the NCC model obtained 86.53% accuracy and the DenseNet121 ENM achieved 79.24% accuracy. The proposed DPN model achieved the highest performance with 100% accuracy, precision, recall, and F1-score, demonstrating the effectiveness of transformer-based feature representations. The integration of an XAI module ensures that only valid plant leaf images are analysed before classification. The developed GUI-based interface enables easy interaction for image upload and prediction visualization.

REFERENCES

[1] Moghadam, P.; Ward, D.; Goan, E.; Jayawardena, S.; Sikka, P.; Hernandez, E. Plant Disease Detection Using Hyperspectral Imaging. In Proceedings of the 2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA), Sydney, Australia, 29 November–1 December 2017; IEEE: Sydney, Australia, 2017; pp. 1–8.

[2] Lowe, A.; Harrison, N.; French, A.P. Hyperspectral Image Analysis Techniques for the Detection and Classification of the Early Onset of Plant Disease and Stress. Plant Methods 2017, 13, 80.

[3] Khan, M.J.; Khan, H.S.; Yousaf, A.; Khurshid, K.; Abbas, A. Modern Trends in Hyperspectral Image Analysis: A Review. IEEE Access 2018, 6, 14118–14129.

[4] Nagasubramanian, K.; Jones, S.; Singh, A.K.; Sarkar, S.; Singh, A.; Ganapathysubramanian, B. Plant Disease Identification Using Explainable 3D Deep Learning on Hyperspectral Images. Plant Methods 2019, 15, 98

[5] Agrios, G.N. Plant Pathogens and Disease: General Introduction. Encyclopedia of Microbiology, 3rd ed.; Academic Press: Cambridge, MA, USA.

[6] Gates, D.M.; Keegan, H.J.; Scshleter, J.C.; Weidner, V.R. Spectral Properties of Plants. Appl. Opt.2020, 4, 11.

[7] Salcedo, A.F.; Purayannur, S.; Standish, J.R.; Miles, T.; Thiessen, L.; Quesada-Ocampo, L.M. Fantastic Downy Mildew Pathogens and How to Find Them: Advances in Detection and Diagnostics. Plants 2021, 10, 435.

[8] Ghafar A, Chen C, Atif Ali Shah S, Ur Rehman Z, Rahman G. Visualizing Plant Disease Distribution and Evaluating Model Performance for Deep Learning Classification with YOLOv8. Pathogens. 2024; 13(12):1032. https://doi.org/10.3390/pathogens13121032

[9] Kuswidiyanto LW, Noh H-H, Han X. Plant Disease Diagnosis Using Deep Learning Based on Aerial Hyperspectral Images: A Review. Remote Sensing. 2022; 14(23):6031. https://doi.org/10.3390/rs14236031

[10] Sengupta S, Dutta A, Abdelmohsen SAM, Alyousef HA, Rahimi-Gorji M. Development of a Rice Plant Disease Classification Model in Big Data Environment. Bioengineering. 2022; 9(12):758. https://doi.org/10.3390/bioengineering9120758

[11] Abasi AK, Makhadmeh SN, Alomari OA, Tubishat M, Mohammed HJ. Enhancing Rice Leaf Disease Classification: A Customized Convolutional Neural Network Approach. Sustainability. 2023; 15(20):15039. https://doi.org/10.3390/su152015039

[12] Sun R, Zhang M, Yang K, Liu J. Data Enhancement for Plant Disease Classification Using Generated Lesions. Applied Sciences. 2020; 10(2):466. https://doi.org/10.3390/app10020466

[13] Pandian JA, Kumar VD, Geman O, Hnatiuc M, Arif M, Kanchanadevi K. Plant Disease Detection Using Deep Convolutional Neural Network. Applied Sciences. 2022; 12(14):6982. https://doi.org/10.3390/app12146982

[14] Saleem MH, Khanchi S, Potgieter J, Arif KM. Image-Based Plant Disease Identification by Deep Learning Meta-Architectures. Plants. 2020; 9(11):1451. https://doi.org/10.3390/plants9111451

[15] Altalak M, Uddin MA, Alajmi A, Rizg A. A Hybrid Approach for the Detection and Classification of Tomato Leaf Diseases. Applied Sciences. 2022; 12(16):8182. https://doi.org/10.3390/app12168182