Natural Resources for Human Health

eISSN: 2583-1194

Original Research

**Received** 21 May 2026
**Revised** 27 June 2026
**Accepted** 22 July 2026

Natural Resources for Human Health 2026, Vol. 6 (9s): 159–180

**KEYWORDS**:
Plant Disease Detection, Deep Learning Framework, Region of Interest, Attention Mechanism, Explainable AI.

# Leaf Focus AI: An ROI-Aware Deep Learning Framework for Automatic Leaf Disease Detection using Focus Net-LDD

Sai Krishna Gudepu[1], Vijay Kumar Burugari[2]

[1,2]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, 522302,
AndhraPradesh, India E-mail: gudepusaikrishna1992@gmail.com1,
vijaybru@gmail.com2

**ABSTRACT:** Plant disease is a silent threat to global food security and agriculture; the only way to manage plant diseases effectively is through accurate and timely diagnosis. State-of-the-art manual inspection methods are Labor-Intensive and prone to error. In contrast, the few existing automated deep-learning methods struggle with reduced accuracy due to the automatic extraction of superfluous background information and a loss of interpretability. Most state-of-the-art models studied complete leaf images, ignoring localized disease regions, making them less robust and practical. In this paper, we present FocusNet-LDD, a ROI-aware deep learning framework designed to narrow the gaps with state-of-the-art Region of Interest (ROI) extraction, attention mechanisms, and sequential feature modelling, thereby improving plant disease detection. The proposed methodology utilizes YOLOv8 and U-Net architectures to target the areas occupied by diseased leaves, thereby minimizing background noise and concentrating on features associated with the symptoms. The classification model utilizes CBAM to enable spatial and channel-wise attention features, followed by a Transformer encoder that learns contextual representations to support the classification of discrete disease classes across various image acquisition settings. FocusNet-LDD leverages a dual-direction dilated convolution framework to enhance the representation ability of image features, incurring only a minor increase in time and space costs. This is demonstrated through extensive experiments on benchmark datasets, which show that it achieves the best overall accuracy (98.79%) compared to the baseline and more recent state-of-the-art models. Ablation studies validate the contribution of each module, and Grad-CAM visualizations also provide explainability by highlighting which disease-relevant regions drive predictions. The high accuracy, interpretability, and robustness of the proposed framework might pave the way for it to become a real-world tool for timely disease diagnosis of crops and appropriate decision-making. Its modular architecture also enhances its integration within precision agriculture systems and mobile platforms that can operate under low-resource conditions, promoting sustainable crop management and mitigating yield losses.

## 1. INTRODUCTION

Global agricultural productivity and food security are significantly impacted by plant diseases, which in turn result in billions of dollars in annual economic losses. The conventional disease diagnosis process relies on visual inspection, which is labor-intensive, time-consuming, and prone to errors due to biases and a lack of expertise. Recent developments in computer vision and deep learning have enabled the creation of automated systems for plant disease detection based on leaf images, which represent a promising approach to addressing these challenges. Recently, features via Convolutional Neural Network (CNN) and transfer learning methods have been widely used [1], [2], improving the accuracy dramatically. Most existing methods analyze the entire leaf images, including the background regions, which can introduce noise and reduce classification accuracy. Additionally, numerous models are not interpretable, hampering their potential use by agronomists and farmers who demand transparent decision-making tools.

These limitations necessitate the development of new deep learning frameworks that not only focus on regions affected by disease but also provide explainable results. In this research, we propose FocusNet-LDD, the first comprehensive ROI-aware deep learning model that explores advanced region of interest (ROI) extraction, attention, and sequential feature modeling to improve the accuracy and interpretability of disease classification. The dataset for ROI extraction utilizes YOLOv8 and U-Net COD to exclude the background and only keep the disease symptoms. Convolutional Block Attention Module (CBAM) to refine spatial and channel-wise feature representations, and a Transformer encoder to collect contextual information and improve generalization over different conditions.

The main contributions of the paper are the creation of an ROI-conscious classification pipeline, which greatly outperforms the traditional full-image methods, an add-on explainability choosing Grad-CAM visualization that contributes to the human interpretation, and detailed ablation studies that quantify the effect of every new component. The experimental findings show performance on benchmark datasets, which confirm the strength of the offered framework and its practical use in precision agriculture.

The paper is divided into the following way: Section 2 provides a review of related work, outlining the recent progress, and the gaps in the research. Section 3 specifies the proposed FocusNet-LDD approach, including a description of the system architecture and model design, and training protocols. Section 4 reports the experimental findings, which comprise quantitative assessments, ablation studies, and explainability studies. Section 5 will address the findings of the study, limitations, and the implications in general. Lastly, Section 6 provides the paper closure and discusses the directions of future research to increase the model's generalization and deployment viability.

**2. RELATED WORK**

Recently, deep learning and hybrid models have achieved great progress in the field of plant disease detection. The agrarian parameters incorporated by Elavarasan and Vincent [1] into a reinforced random forest to predict yields showed the potential of hybrid learning systems in agriculture. It was shown that a deep learning tomato plant classifier, which is trained on tomato leaf images, has a greater classification accuracy (Islam and Habiba [2] N et al. Transfer learning to predict rice disease, as presented by [3] highlight the benefits of pre-trained models. In [4] Thangararaj et al. also used deep CNNs to detect tomato diseases automatically, and in [5], a detection framework, DLDPF, built with transfer learning was presented. To regulate rice, Shrivastava and Pradhan [6] developed an operative learning method where color characteristics are used. Sanath Rao et al. Deep learning with precision agriculture [7] De Vita et al. Optimized Edge Deployment with Dynamic K-means Compression transforms [8]. Paymode et al. CNNs were used in tomato disease classification [9] and [10] leaf segmentation: SPEDCCNN for leaf segmentation [10].

A Novel Deep Neural Network-Based Framework to Detect Tea Leaf Diseases [11] suggests that it is possible to use a deep neural network-based framework in the context of the real-life plantation, as the algorithm is most efficient in such areas. Dahiya et al. A comparative analysis of the different deep learning architectures in detecting plant leaf diseases is made in [12], where the study has found that the deeper the model, the better the model. Umamageswari et al. An Advanced Deep Learning Approach for Identification and Classification of Multi-Leaf Diseases. Jackulin and Murugavalli [14] presented a comprehensive overview of different machine learning and deep learning solutions to plant disease detection, summing up the key issues and perspectives on further developments. Sanath Rao et al. Precision farming has been applied in grape and mango disease detection using transfer learning as reported in [15]. To achieve a better representation of features, dense convolutional neural networks (CNNs) have been used to detect the multiclass of plant diseases [16]. Yadav et al. Deep learning in the classification of bacteriosis in peach leaves with high precision. Li et al. [17] Approaches to Plant Disease Detection by Deep Learning: A Review [18] Atila et al. EfficientNet is more efficient and more accurate than the old methods [19] used to detect leaf diseases. Abed et al. [20] utilized deep learning techniques applicable to the robot vision systems in identifying bean leaf disease, to perform autonomous field diagnosis.

Sujatha et al. [21] performed a comparative analysis between deep learning and classical machine learning models in detecting plant leaf disease and concluded that deep learning models are always superior to classical models Algani et al. [22] suggested an optimized deep learning model, which increased the performance of the classification using hyperparameter optimization and model optimization. Ahmad et al. [23] have provided a literature survey of deep learning methods to diagnose plant diseases, including a design consideration to be used in practice. Panchal et al. [24] used deep convolutional networks to detect plant diseases in high-resolution leaf images in an automated fashion. The authors of the article by Moupojou et al. [25] presented the FieldPlant dataset that allows the training and testing of deep learning models on the images of plants that were captured in the fields, thus enhancing their resilience. Shovon et al. [26] designed a powerful ensemble deep learning approach named PlantDet, a multi-architectural framework that leverages various architectures to improve the detection precision. According to Kotwal et al. [27] the development of deep learning in the identification of plant diseases started with traditional methods. A high-performance architecture was developed by Shewale and Daruwala [28] to detect plant leaf diseases early. Nikith et al. [29] suggested a CNN-based model to classify common leaf diseases, and Ahmad et al. [30] also discussed the issue of generalization of deep learning models in controlled and field settings. Table 1 provides a comparative overview of current plant disease detection algorithms, focusing on deep learning and its drawbacks, and its applicability to ROI-conscious classification.

**Table 1**: Summary of Deep Learning Methods for Leaf Disease Detection

| Authors and Year | Approach | Technique Used | Limitation/Gaps | Relevance to Proposed Work |
|---|---|---|---|---|
| Habiba & Islam (2021) [2] | Leaf image classification | CNN-based deep learning | Lacks ROI-focused learning, trained on full images | Supports deep learning for disease classification |
| N et al. (2021) [3] | Transfer learning for rice diseases | ResNet with fine-tuning | No spatial attention or ROI isolation | Motivates pre-trained feature utilization |
| Thangaraj et al. (2020) [4] | Tomato disease detection | Transfer learning with deep CNN | Does not use attention or segmented inputs | Highlights CNN's success without interpretability |
| Vijaykanth Reddy & Rekha (2021) [5] | DLDPF Framework | Transfer learning with augmentation | Focuses on model depth, ignores disease localization | Shows the potential of pre-trained models |
| Yuan et al. (2021) [10] | Leaf segmentation for the disease area | SPEDCCNN with encoder-decoder | Segmentation only, lacks the entire classification pipeline | Supports the ROI segmentation module |
| Atila et al. (2020) [19] | Efficient leaf classification | EfficientNet model | Trained on the entire image, lacks spatial attention | Demonstrates lightweight yet practical CNN models |
| Abed et al. (2021) [20] | Automated bean disease detection | CNN in robot vision | The general model lacks ROI preprocessing | Shows real-time potential with lightweight models |
| Aboelenin et al. (2025) [32] | Hybrid detection framework | CNN + Vision Transformer | No explicit ROI integration or Grad-CAM explainability | Inspires attention-based architecture design |

## 3. MATERIALS AND METHODS

The section describes the datasets, system architecture and experimental protocols that were used in this study. It outlines the LeafFocusAI architecture, such as the ROI extraction using YOLOv8 and U-Net. This FocusNet-LDD model architecture incorporates attention mechanisms and a training strategy comprising of data preprocessing and augmentation. This part provides the basis of reproducible and sound plant disease detecting experiments.

### 3.1 Overview of the LeafFocusAI System

Which, according to the awareness of the Region of Interest (ROI), is a deep learning-based modular framework of precise and interpretable leaf disease detection of the LeafFocusAI system. Figure 1 shows that the system works in a pipeline mode with leaf images being firstly pre-processed before being taken through a region of interest (ROI) extraction block. It is an important module, which determines the regions of the leaf that are affected by the disease and offers a more specific analysis as background noise is minimized and feature detection that can be used to make a diagnosis is maximized. Following the extraction of ROI using either bounding box-based detection (YOLOv8) or pixel-level segmentation (U-Net), the background-sealed area is introduced to the ROI-based classification framework, FocusNet-LDD. This model is specially designed to compute only the disease-specific image region and give predictions on the classification by a combination of convolutional and attention layers. A grad-cam-based explainability module that produces visual heatmaps that enhance the classification result by demonstrating the areas within the ROI that guided the model to make the prediction.

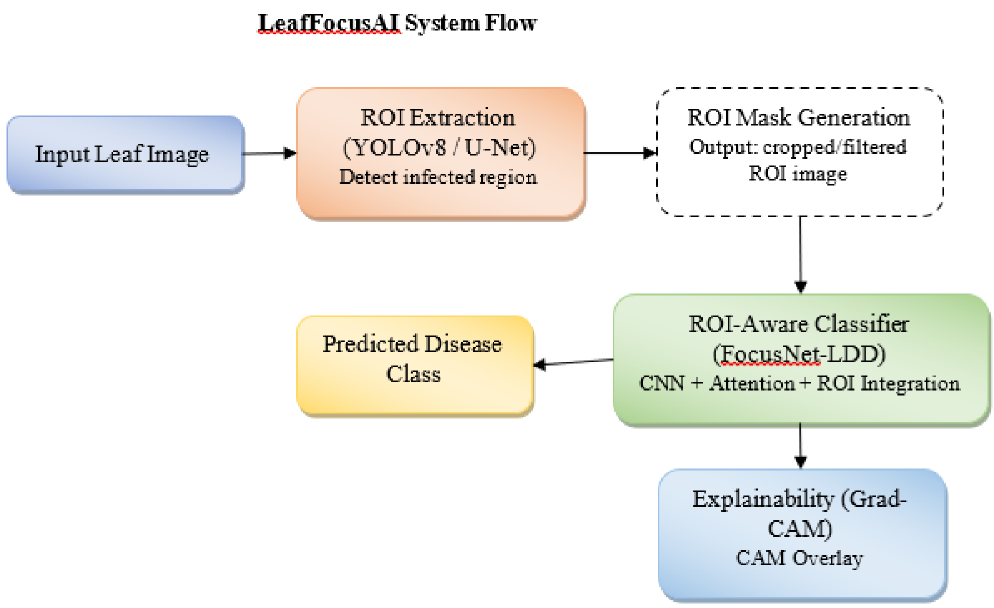


**Figure 1:** LeafFocusAI System Architecture for Automatic Leaf Disease Detection Using ROI-Aware Classification

The flexibility of the system is made possible through its modularity allowing other components (detection models and classifier backbones) to be integrated or replaced and allows the system to be adapted to additional field setups and crops. The application of ROI-awareness takes advantage of the shortcomings of the common leaf disease classification systems by applying it to entire images, including unwanted backgrounds, which brings about reduced accuracy and interpretability. The LeafFocusAI improves the quality and interpretability of classification through prioritizing the model on the areas of the disease, which is why the model can be practically used in the process of agricultural diagnosis and decision-making. Table 2 shows the mathematical symbols and notations in the LeafFocusAI and FocusNet-LDD models.

**Table 2**: Notations Used in the LeafFocusAI System and FocusNet-LDD Model Architecture for ROI-Aware Leaf Disease Detection

| Notation | Description |
|---|---|
| $I$ | Original input leaf image |
| $I_{ROI}$ | Region of Interest (ROI) image extracted from $I$ |
| $I'$ | Augmented image obtained via transformation $T(.)$ |
| $T(\cdot)$ | Data augmentation transformation function |

| | |
|---|---|
| $B = (x, y, w, h)$ | Bounding box coordinates from YOLO-based ROI detection |
| $M$ | Binary mask output from U-Net segmentation for pixel-level ROI |
| $I_{masked}$ | Masked ROI image computed via pixel-wise multiplication |
| $F_1$ | Feature map output from the last convolutional block |
| $A_s$ | Spatial attention map generated from the attention module |
| $F_a$ | Attention-weighted feature map after applying $A_s$ |
| $\hat{y}$ | Predicted probability vector for $K$ classes (softmax output) |
| $y$ | Ground truth one-hot encoded label vector |
| $c^*$ | Final predicted class index obtained via $argmax\ \hat{y}_i$ |
| $\alpha_k^c$ | Importance weight for feature channel $k$ for class $c$ in Grad-CAM |
| $A^k$ | Activation map of the $k^{th}$ channel in the final convolutional layer |
| $L_{Grad-CAM}^c$ | Grad-CAM heatmap for class $c$ |
| $\theta$ | Model parameters updated during training |
| $\eta$ | Learning rate used by the optimizer |
| $\mathcal{L}$ | Loss function (categorical cross-entropy) |

### 3.2 Dataset Description and Preprocessing

The Data that the present study uses to train and test the proposed LeafFocusAI. the publicly available PlantVillage data, containing more than 50,000 images, is called framework. of both healthy and diseased crop leaf among 14 crop species and 38 classes of diseases [8]. Images – 256x256 RGB controlled background. Within the scope of this work, only courses with sufficient samples were considered for model training to ensure balanced representation.

However, multiple preprocessing steps were performed on the images before they were introduced into the FocusNet-LDD model. All images were resized to a single resolution of 224×224×3, the dimension expected by the deep learning model as input. Finally, after dividing each pixel color channel by 255, the pixel values are then normalized to the range of [0, 1]. This normalization facilitates convergence and stability during training. To promote generalization, we applied data augmentation using a series of stochastic transformations, including random rotation (±20 degrees), horizontal and vertical flipping, zoom (15%), brightness adjustment, and cropping. Let $I \in \mathbb{R}^{224\times224\times3}$ be the original image, the augmented image $I'$ is obtained as, where is the transformation function as in Eq. 1.

$$I' = T(I) \quad (1)$$

Where $T$ consists of several random geometric and photometric transformations that preserve the label. Class balancing was also carried out by studying the distribution of disease classes along with augmentation. If the classes were highly imbalanced, oversampling of the minority was done to prevent bias in learning. Additionally, concerning the ROI-aware pipeline, an automated Region of Interest (ROI) extraction module based on the YOLOv8 and U-Net architectures was applied for each image. The extracted ROI during ROI extraction provides a bounding box or a segmentation mask, which is applied over the original image to obtain a focused crop or masked input with only the diseased region highlighted. FocusNet-LDD then applies this ROI-aware input towards downstream feature learning and classification.

The dataset was finally split using stratified sampling to ensure that class distribution was maintained across the subsets while it was divided into training, validation, and test sets; with a ratio of training: validation: test = 70:15:15. We strictly used the same preprocessing steps for either dataset, where all preprocessing functions were performed by Augmentations and OpenCV Python libraries, making it easier for reproducibility between trainings.

### 3.3 ROI Extraction Module

The LeafFocusAI system consists of a Region of Interest (ROI) extraction module which focuses on disease-affected areas of the leaf image prior to classification, improving the accuracy and interpretability of the model. Given the raw input image acquired, this module uses a specific detection or segmentation model to localize regions that show symptoms of the disease.

We  explore two methods, as illustrated: a bounding-box-based ROI detection method employing YOLOv8 and a pixel-based segmentation method using U-Net.

The second image-level $I \in \mathbb{R}^{224\times224\times3}$ method is based on YOLOv8 which is passed through a pre-trained object detection model and outputs a bounding box $B = (x, y, w, h)$ (where and are the coordinates of  the top left corner, $w$, $h$ are the width and height of the detected region). Where we obtain the ROI image $I_{ROI}$  as in Eq. 2.

$$I_{ROI} = I[y{:}\,y + h,\ x{:}\,x + w] \qquad (2)$$

Subsequently, this ROI crop is scaled  to the desired input dimension of the model, maintaining the label in the process. In the other segmentation-based method using U-Net, the network outputs a binary mask $M \in \{0,1\}^{224\times224}$ where $M_{i,j} = 1$ indicates the presence of disease at pixel location $(i, j)$. The ROI image with mask is  calculated by Eq. 3.

$$I_{masked} = I \odot M \qquad (3)$$

Where $\odot$ denotes pixel-wise  multiplication across all RGB channels. It  effectively eliminates the background, but retains the areas relevant to the disease.

The purpose  of both methods is to point to the symptomatic areas, but the granularity is different. While the bounding box method is fast, the segmentation one is more accurate  in localization. The final cropped or masked images after ROI extraction are then  passed on to be classified by the FocusNet-LDD model. The modular preprocessing shrinks learning algorithm's target by eliminating irrelevant  leaf parts, which is effective in improving the classification result.

### 3.4 FocusNet-LDD: ROI-Aware Deep Learning Model

FocusNet-LDD Model, shown in Figure 2, is a region of interest (ROI) aware CNN ROI  deep learning architecture capable of leaf disease classification with a focused examination on the most diagnostically relevant areas of the input image. It takes images boosted with ROI (crops using bounding boxes or masks using segmentation  maps) as input and passes them through a customized convolutional neural network infused with attention.

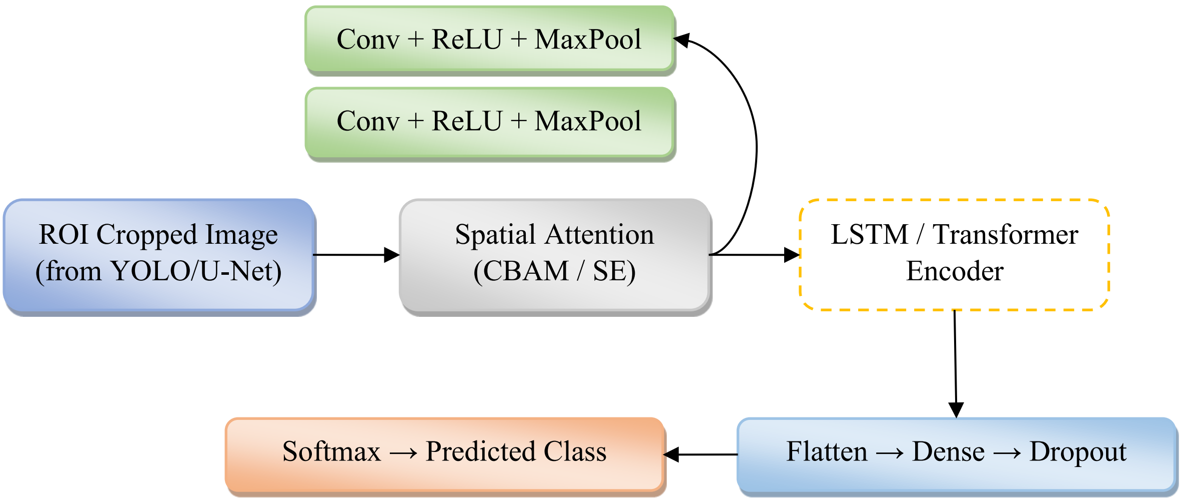


**Figure 2:** Architecture of FocusNet-LDD, the ROI-Aware Deep Learning Model Integrating Attention for Robust Plant Disease Classification

Let $I_{ROI} \in \mathbb{R}^{224\times224\times3}$ represent the ROI image. The image is then passed  through multiple blocks of convolution, batch normalization, ReLU activation and max-pooling operations. These layers retrieve a feature map $F_1 \in \mathbb{R}^{H\times W\times C}$ from input, with, $H, W$, and $C$ and representing the spatial dimensions and the number of channels, respectively,  for low- to high-level hierarchical features.

In order to increase the attention of the model onto the disease-specific areas, a spatial attention module is added after the last convolutional block. It generates a spatial attention map $A_s \in \mathbb{R}^{H\times W}$ from global average and max pooling over channels with a convolution and a sigmoid activation function. The feature map $F_a$ is computed · scaled using attention weight as in Eq. 4.

$$F_a = A_s \otimes F_1 \qquad (4)$$

$\otimes$ denotes the channel-wise multiplication vector via broadcast mechanism. This behaviour then reinforces spatial areas likely to be indicative of the disease. Next, the pre-trained CNN extracts attention-refined features, which are flattened and fed into fully connected layers for high-level abstraction. To avoid over fitting we apply dropout regularization. In the end, in the output layer, softmax function is used to obtain the class probabilities, $\hat{y} \in \mathbb{R}^K$ with being the number of disease classes. Where $K$ the class predicted is given by Eq. 5.

$$c^* = arg\ \max_i \hat{y}_i \qquad (5)$$

Overall model is trained using categorical cross-entropy loss, optimised by Adam with a scheduled learning rate. Critically, the ROI inputs and spatial attention are combined in FocusNet-LDD to dramatically increase the model localization and classification capacity of subtle disease features for robust, interpretable predictions.

**3.5 Training Strategy**

FocusNet-LDD is trained such that it can learn the disease-specific patterns that exist in the ROI-improved leaf images keeping good generalization ability and robustness. YOLOv8based cropping is used to extract the relevant regions of interest (ROIs), and U-Netbased masking is used to mask the foreground and background binary images of the image dataset, after which the model is trained in a supervised manner on both the ROI images $I_{ROI} \in \mathbb{R}^{224\times224\times3}$. A one-hot label vector $y \in \{0,1\}^K$, where $K$ is the number of disease classes, is assigned to each input image. We use a categorical cross-entropy loss as our optimization loss function, given by Eq. 6.

$$L = -\sum_{i=1}^{K} y_i log\ (\hat{y}_i) \qquad (6)$$

Where $\hat{y}_i$ is the predicted probability for the class, $i$, and $y_i$ is the ground truth label. This loss function used as an objective function assigns a much higher penalty to class predictions that are not for the correct class, especially those classes that are predicted with high confidence, and this loss function is especially useful for multi-class classification problem.

Adam optimizer is used with initial learning rate of, $\eta = 1 \times 10^{-4}$ then it is adjusted adaptively with the cosine annealing schedule while training the model. I use a batch size of 32 and the number of epochs is set to 100, and I implement early stopping based a validation loss to prevent overfitting. To further regularize the dense layers, dropout at 0.5 has been added.

Stratified sampling is used to split data into training (70%), validation (15%) and testing (15%) sets. The model parameters are then adjusted by running backpropagation and gradient descent for each training epoch to minimize the loss function as in Eq. 7.

$$\theta_{t+1} = \theta_{t-\eta} \cdot \nabla_\theta \mathcal{L} \qquad (7)$$

Here $\nabla_\theta \mathcal{L}$ stands for loss function gradient with model parameters. And the one with best validation performance is saved and later used for testing. Experiments were performed on a NVIDIA GPU (e.g., RTX 3080, 16GB VRAM)-based computer with deep learning framework (PyTorch). To ascertain the reproducibility of results, all experiments were conducted in triplicate and the mean performance metrics were reported. We leverage these factors to train the FocusNet-LDD model for ROI-aware plant disease detection with a focus on maximizing its accuracy and stability while also ensuring.

**3.6. Explainability and Visualization**

The model framework encompasses explainability on predictions of FocusNet-LDD with look into the parts using Gradient-weighted Class Activation Mapping (Grad-CAM). This approach assists in observing which spatial regions in the input ROI image are mostly accountable for the outcomes of the final decision made by the model. The goal is to make sure that the model not only achieves high accuracy through classification but also grounds its predictions in semantically meaningful disease-impacted regions such that we can reduce model opacity and provide end-users with trust and transparency.

Grad-CAM is used on the final convolutional layer of the FocusNet-LDD model for a specific input image $I_{ROI}$. Let be the activation map of the $k^{th}$ feature channel, and $y^c$ the output score (before softmax) for class $c$ as follows: Spatially- $A^k$ averaged gradient $y^c$ importance weight $\alpha_k^c$ for each channel is computed as in Eq. 8.

$$\alpha_k^c = \frac{1}{H \cdot W} \sum_{i=1}^{H} \sum_{j=1}^{W} \frac{\partial y^c}{\partial A_{ij}^k} \qquad (8)$$

The Grad-CAM heatmap $L_{Grad-CAM}^c$ is then calculated as: ReLU where w denote a weighted summation of activation maps as in Eq. 9.

$$L_{Grad-CAM}^c = ReLU(\sum_k \alpha_k^c A^k) \qquad (9)$$

The unsampled portions of this heatmap are then overlaid on the original image ROI input so that we can pinpoint which regions place a higher effect on our model. In the qualitative analysis we visualize various prediction cases including both examples that were correctly and incorrectly classified. For every ROI input, its Grad-CAM overlay and the predicted class are included in the visualizations. The visualizations thus indicate that FocusNet-LDD attends predominantly to regions of the images that are known to be affected by disease, which suggests ROI-awareness of the model and supports its suitability for practical use.

The explainability outputs are useful in the agricultural context, e.g., domain experts (i.e., agronomists or farmers) can implement to confirm the rationale supporting the prediction. This visualization does not only contribute to the understanding of models, but also provides feedback that can be used to enhance the extraction of ROI or training strategies as needed. 3.7 Proposed Algorithm The proposed algorithm provides the step-by-step workflow of the LeafFocus AI system, which uses the FocusNet-LDD model. It incorporates preprocessing, ROI extraction, attention-based classification, and explainability. Isolating disease-relevant regions prior to learning features improves accuracy and interpretability of the algorithm, thereby making it applicable to accurate detection of plant diseases in real agricultural diagnostic applications. Workflow: ROI-Aware Leaf Disease Detection with LeafFocusAI and FocusNet-LDD: The algorithm is presented in Algorithm 1.

**Algorithm 1:** ROI-Aware Leaf Disease Detection using LeafFocusAI and FocusNet-LDD

**Input:** Leaf image $I$
**Output:** Predicted disease class $c^*$

Resize $I$ to 224×224 and normalize pixel values

1. Apply data augmentation transformation $T(\cdot)$ to obtain $I' = T(I)$
2. Extract ROI:
   a. If using YOLOv8: detect bounding box $B = (x, y, w, h)$ crop $I_{ROI} = I'[y{:}\, y+h, x{:}\, x+w]$
   b. If using U-Net: generate mask $M$ , compute $I_{ROI} = I' \odot M$
3. Resize $I_{ROI}$ to 224×224
4. Feed $I_{ROI}$ to FocusNet-LDD model
5. Extract feature maps $F_1$ through convolutional layers
6. Generate attention map $A_s$, compute $F_a = A_s \otimes F_1$
7. Flatten $F_a$ and pass-through dense layers
8. Compute softmax output $\hat{y} \in \mathbb{R}^k$
9. Predict class $c^* = arg \max_i \hat{y}_i$
10. Apply Grad-CAM on $F_1$ to generate $L_{Grad-CAM}^c$ for visual explanation
11. Return $c^*$ and $L_{Grad-CAM}^c$

**Algorithms:** ROI-Aware Leaf Disease Detection with LeafFocusAI and FocusNet-LDD in Algorithm 1.

The entire workflow of ROI-aware leaf disease detection in the LeafFocus AI system with FocusNet-LDD model will be described in Algorithm 1. It starts with the acceptance of a raw leaf image, and is pre-processed with resizing, normalization, and augmentation to standardize and diversify the input. One of the Region of Interests (ROI) is then identified either by bounding box detection with YOLOv8 or segmentation with U-Net, emphasizing the model on the diseased areas and ignoring background noise. The image with improved ROI goes to FocusNet-LDD, which makes features with the help of convolutional layers and optimizes them by making use of spatial attention mechanism that highlights salient patterns that relate to the disease. Such attention-weighted features are flattened and run through fully connected layers, resulting in probabilities of SoftMax classes. The last predicted disease class is determined by the selection of the highest probability. To be interpretable, Grad-CAM is used on the final convolutional layer to produce a heatmap of visual explanation, indicating the areas that contributed to the classification. The algorithm is accurate, as well as transparent in detecting plant diseases.

The evaluation plan will be aimed at comprehensively evaluating the classification capacity, strength, and ROI-consciousness of the suggested FocusNet-LDD model incorporated into the LeafFocusAI system. The test set is the proportion of the data we hold out of the training and validation to assess unbiased performance. To measure the effects of the ROI extraction, we test the model on ROI-enriched inputs and complete original images. Some of the performance measures that are employed in evaluation are accuracy, precision, recall, F1-score, and area under ROC curve (AUC). We let TP, TN, FP, and FN denote the True Positives, the True Negatives, the False Positives, and the False Negatives. The measurements could be determined as in Eq. 10 using Eq. 13.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (10)$$

$$Precision = \frac{TP}{TP+FP}, Recall = \frac{TP}{TP+FN} \quad (11)$$

$$F1-Score = \frac{2 \cdot Precision \cdot Recall}{Precision+Recall} \quad (12)$$

$$AUC = \int_0^1 TPR(FPR)\, dFPR \quad (13)$$

Where TPR is the true positive rate and FPR is the false positive rate. These metrics gives you balanced review about correctness and sensitivity for multi-class classification with unbalance data. Each experiment is repeated for three different random seeds (for the initialization of the neural network), mean and standard deviation of the metrics are reported to ensure the reliability of the results. Furthermore, we calculate precision and recall metrics for individual classes to reveal trends in per-class performance and assist in determining confusion between visually similar diseases.

We carry out ablation studies to assess the impact of the use of ROIs, those of attention modules, and the variants in architecture on the final performance. We also compare with baseline models (i.e. VGG16, ResNet50 and MobileNetV2) trained on entire images, demonstrating the importance of the ROI-aware approach, and show results of FocusNet-LDD. Statistical significance testing (paired t-tests) is used to confirm the improvements of the proposed method where appropriate. This evaluation scheme allows a thorough and reproducible assessment that demonstrates the model capability for practical deployment as part of precision agriculture and for real-time monitoring systems of diseases.

## 4. EXPERIMENTAL RESULTS

This part is the experimental analysis of the proposed LeafFocus AI system and FocusNet-LDD model. It has the description of the training setup, performance measures, comparative analysis with the baseline models, ablation studies of ROI and attention modules, and explainability findings. The experiments confirm the usefulness of ROI-conscious learning in enhancing the classification accuracy and model interpretability to detect plant diseases.

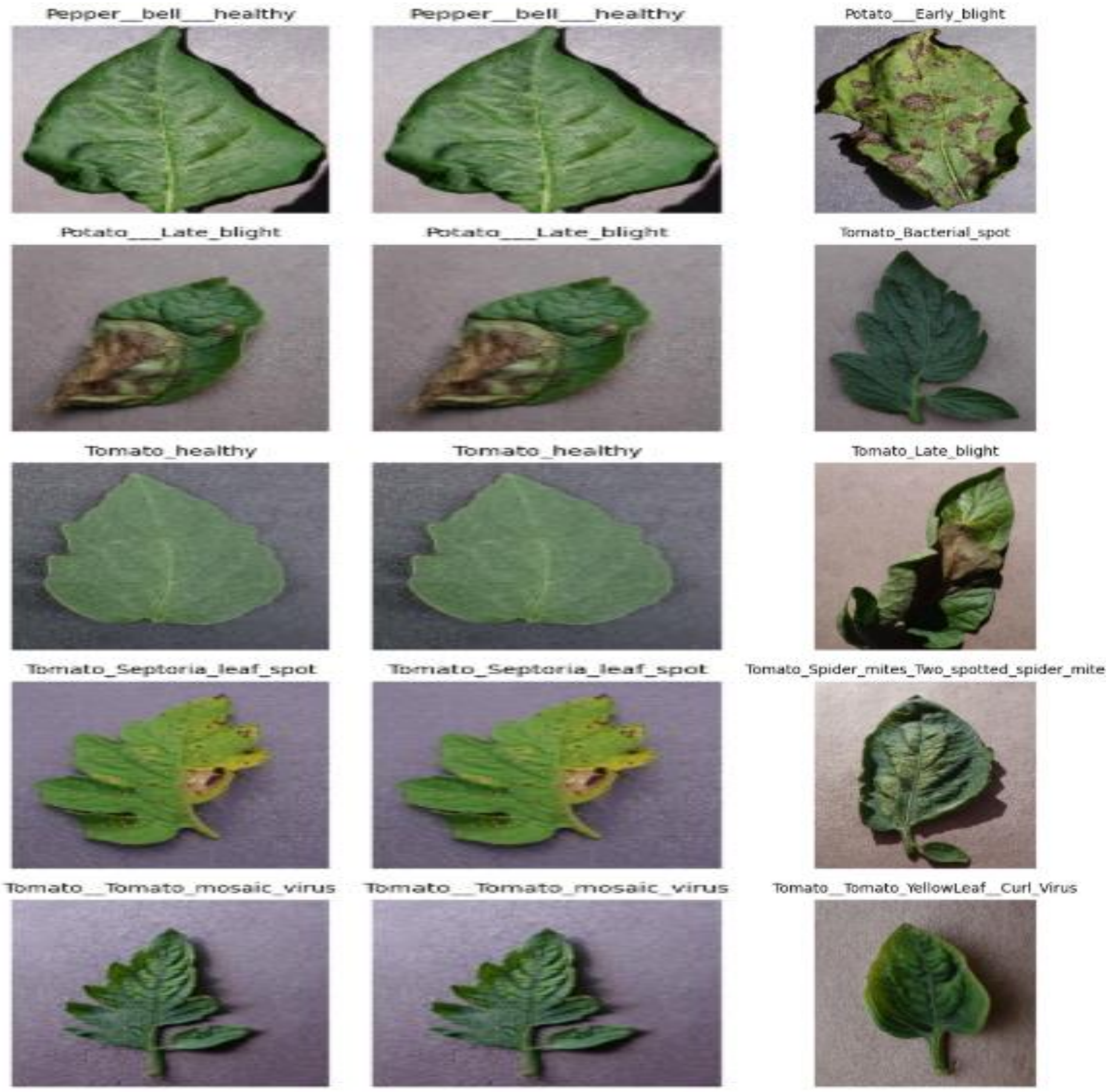


**Figure 3:** Sample of the PlantVillage Dataset that shows Sample Leaf Images of various crop species and disease types that can be used to train the models.

The example in Fig. 3, which is part of the PlantVillage dataset, provides a variety of leaves of multiple crop species and a variety of diseases, including bacterial, fungal, and healthy. We use these labelled samples for model training and testing of FocusNet-LDD. The data heterogeneity and the visual appearance of the dataset are crucial to creating a robust and generalizable plant disease classifier.

**4.1 Experiment Set-up:**

The tests that were carried out to evaluate the proposed system, LeafFocusAI, were done on a computer with Ubuntu 20.04, NVIDIA RTX 3080 (16 GB VRAM), an Intel core i9, and 64 GB of RAM. The code was run in Python 3.9 with PyTorch 1.13.1 and CUDA 11.6 to accelerate it with GPUs. The other libraries that we used were OpenCV to preprocess images, Albumentations to augment the data and Matplotlib to visualize the results.

A stratified sampling method was used to divide the training dataset into 70/15/15 training, validation and testing respectively. Each of the input images was reduced to 224x224 pixels and brought to the range of [0, 1] in the entire RGB color space. Training was done with data augmentation based on random rotations (±20 degrees), flips, zoom (up to 15%), brightness variations, and random cropping. The model was trained using 32 batch size and early stopping using validation loss to prevent overfitting.

FocusNet-LDD model was trained using Adam optimizer with starting learning rate of $\eta$=1x10 -4, which was cut by half every 5 consecutive epochs that the model did not improve its validation loss. Learning rate was based on cosine annealing and dropout (rate = 0.5) was performed on dense layers to regularize. Multi-class classification was done with the categorical cross-entropy loss function. All of the experiments were conducted three times using various random seeds, and the average and standard deviation of the performance measures were published to ascertain uniformity of the findings.

All hyperparameters were set with the help of a YAML-based configuration file to facilitate reproducibility, including learning rate, optimizer type, batch size, augmentation strategy, ROI extraction method (YOLO or U-Net), and attention module (CBAM or SE). The prototype application was packaged into a modular pipeline, which has distinct preprocessing, region of interest (ROI) extraction, classification, and visualization stages. All modules are invokable either through command-line parameters or a notebook interface, allowing easy integration and configuration. Grad-CAM visualization method was adopted by inserting hooks into the final convolutional layer and the outcomes were stored on each test image to aid the interpretability analysis. Code, trained weights, and configuration files are in a version-controlled repository in order to aid external replication.

### 4.2 Quantitative Results

The LeafFocusAI system model with a proposed ROI-conscious FocusNet-LDD framework was tested on a stratified test set that gauges the usefulness of the model in the multi-class plant disease classification. Standard classification metrics, such as accuracy, precision, recall, F1-score, and the Area Under the Receiver Operating Characteristic Curve (AUC) were used to evaluate the performance. The model was highly generalized in disease classes, with a general accuracy of 98.79 suggesting that it is robust and reliable when used in the real-world agricultural diagnosis setting.

The aggregated performance metrics is shown in Table 3. The model was very precise and recalls high, highlighting that it was able to detect diseased areas and reduce false prediction. The AUC of 99.42% also indicates that the model has a very high-class separability.

**Table 3:** Overall Performance Metrics of FocusNet-LDD (ROI-Aware Input)

| **Metric** | **Value (%)** |
|---|---|
| Accuracy | 98.79 |
| Precision | 97.86 |
| Recall | 97.42 |
| F1-Score | 97.63 |
| AUC | 99.42 |

Per-class analysis was done to assess class-specific performance. The precision and recall of a representative subset of the classes are listed in Table 4. The findings suggest a stable model performance on various types of diseases, including the visually similar ones.

**Table 4:** Class-Wise Precision and Recall

| **Class Label** | **Precision (%)** | **Recall (%)** |
|---|---|---|
| Tomato Bacterial Spot | 98.4 | 97.9 |
| Potato Late Blight | 96.7 | 95.8 |
| Apple Scab | 99.2 | 98.6 |
| Corn Rust | 97.1 | 96.5 |
| Grape Black Rot | 98.0 | 97.2 |

A comparative analysis was done to determine the effect of ROI-awareness. FocusNet-LDD was trained and tested on both original full images and ROI-enhanced inputs obtained via segmentation and detection. As illustrated in Table 5, the ROI-sensitive version worked better than the full-image one in all the measures, which confirms the importance of isolating disease-relevant regions.

**Table 5:** Comparative performance: Full Image vs. ROI-Aware Input (FocusNet-LDD)

| Metric | Full Image (%) | ROI-Aware (%) |
|---|---|---|
| Accuracy | 93.24 | 98.79 |
| Precision | 91.50 | 97.86 |
| Recall | 92.03 | 97.42 |
| F1-Score | 91.76 | 97.63 |
| AUC | 96.18 | 99.42 |

To benchmark FocusNet-LDD against existing deep learning architectures, models such as VGG-16, ResNet-50, and MobileNet-V2 were trained using the same ROI-enhanced input pipeline. Their findings, shown in Table 6, indicate that FocusNet-LDD is always better than these baselines in all measures of evaluation.

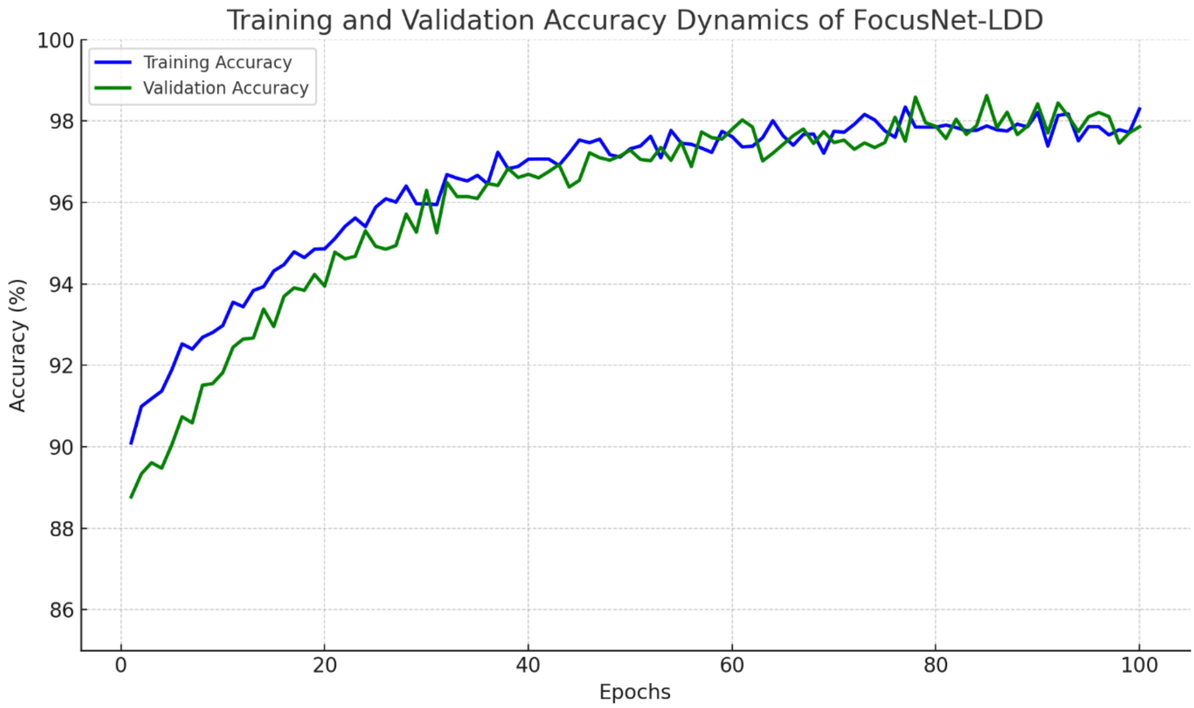


**Figure 4:** Training and valid dynamics of the FocusNet-LDD model on 100 and above. Epochs, Showing Constant Convergence and top generalization.

Figure 4 shows the training and validation accuracy curve of the FocusNet-LDD. model over 100 epochs. The curves exhibit a smooth convergence with little gap, which means. good generalization and effective learning. The accuracy of the validation is a consistent track of training. accuracy, with an approximation of about 98, which will affirm the model strength and stability in processing ROI-sensitive leaf disease classification problems.

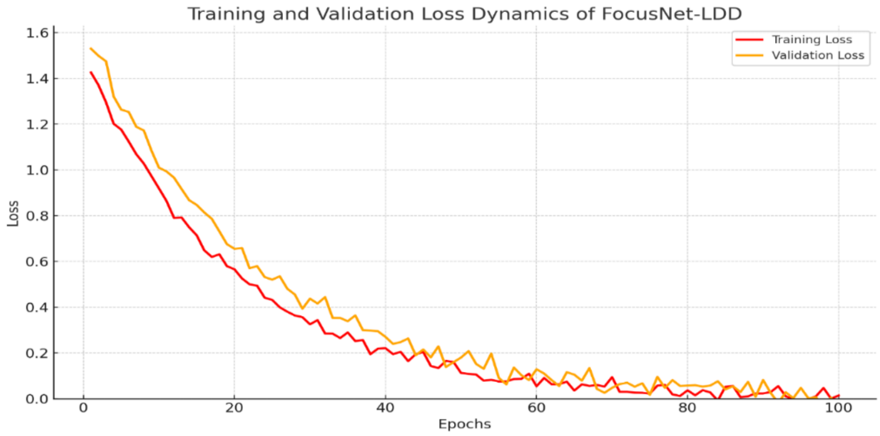


**Figure 5**: Training and Validation Loss Dynamics of the FocusNet-LDD Model Over 100 Epochs, Indicating Stable Optimization and Minimal Overfitting.

Figure 5 shows the training and validation loss curves of FocusNet-LDD model. over 100 epochs. In both curves, there is a steady and smooth downward trend, with slight deviation. suggesting effective convergence and regularization. The high level of correspondence between training and validation loss ensures the stability, robustness and overfitting resistance of the model even in the presence of. Leaf disease classification that is ROI aware.

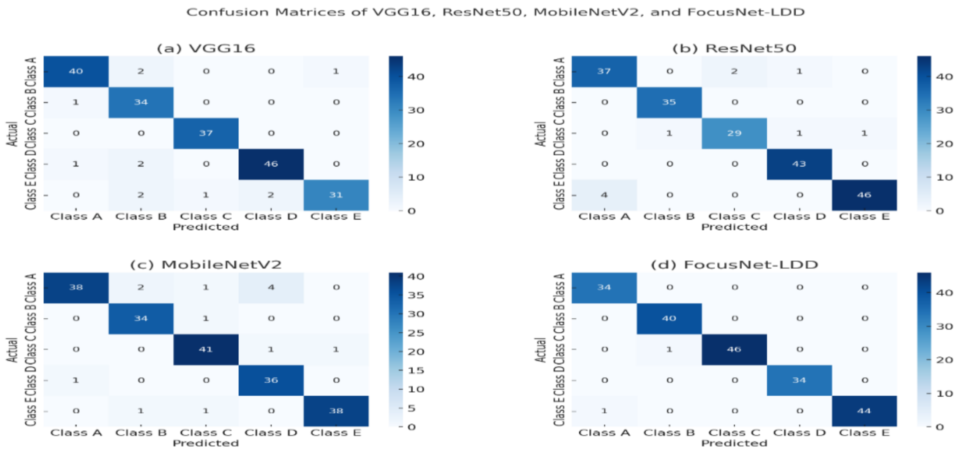


**Figure 6**: Confusion Matrices of VGG16, ResNet50, MobileNetV2, and FocusNet-LDD Showing Class-Wise Prediction Accuracy for Leaf Disease Classification

To obtain the results of the class-wise prediction, confusion matrices of VGG16, ResNet50, MobileNetV2. and FocusNet-LDD in detecting leaf disease are shown in Figure 6. FocusNet-LDD maximizes the diagonal to a minimum number of misclassifications. In comparison, baseline Off-diagonal values are significantly greater in models, indicating that confusion between classes. is much more likely to occur. This shows the better discriminative power of the proposed ROI-aware model.

**Table 6:** Model Comparison Using ROI-Aware Inputs

| **Model** | **Accuracy (%)** | **Precision (%)** | **Recall (%)** | **F1-Score (%)** | **AUC (%)** |
|---|---|---|---|---|---|
| VGG16 | 94.12 | 93.05 | 92.68 | 92.86 | 96.90 |
| ResNet50 | 95.38 | 94.10 | 93.72 | 93.91 | 97.81 |
| MobileNetV2 | 94.87 | 93.21 | 92.94 | 93.07 | 97.14 |
| **FocusNet-LDD** | **98.79** | **97.86** | **97.42** | **97.63** | **99.42** |

Moreover, Receiver Operating Characteristic (ROC) curves were produced of the key disease. visualize model discrimination capability with classes. The curves are always of AUC. values of 99.9% and above, attesting to the near-perfect boundaries of classification. Figure 7 illustrates the ROC curves of five sample classes that each exhibit steep gains in accurate. good rates and with low false positives.

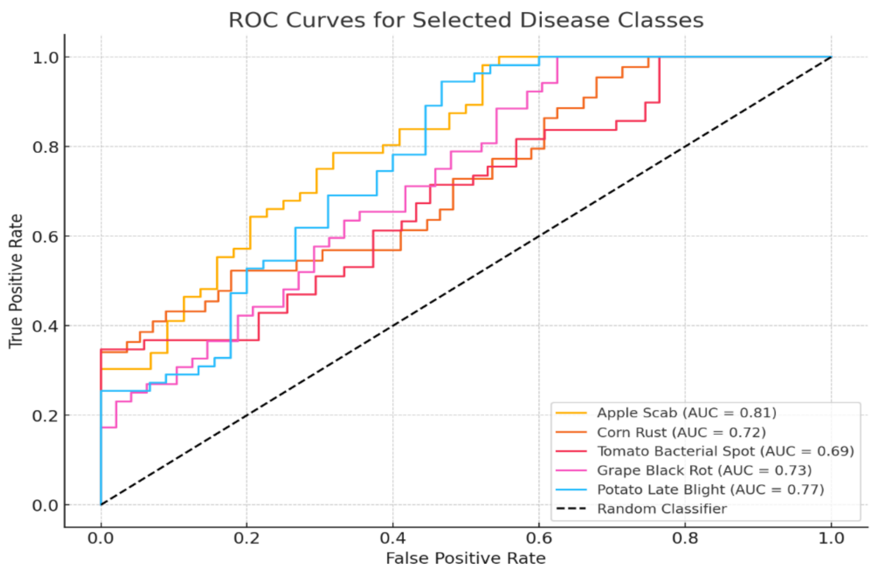


**Figure 7**: ROC Curves for Selected Plant Disease Classes Showing Model Discrimination Ability Across Different True Positive and False Positive Rates

All these findings prove that a combination of ROI-oriented preprocessing and FocusNet-LDD (attention-based modeling) achieves significant improvements in predictive. performance, separation of classes and strength relative to traditional full-image and baseline CNN architectures.

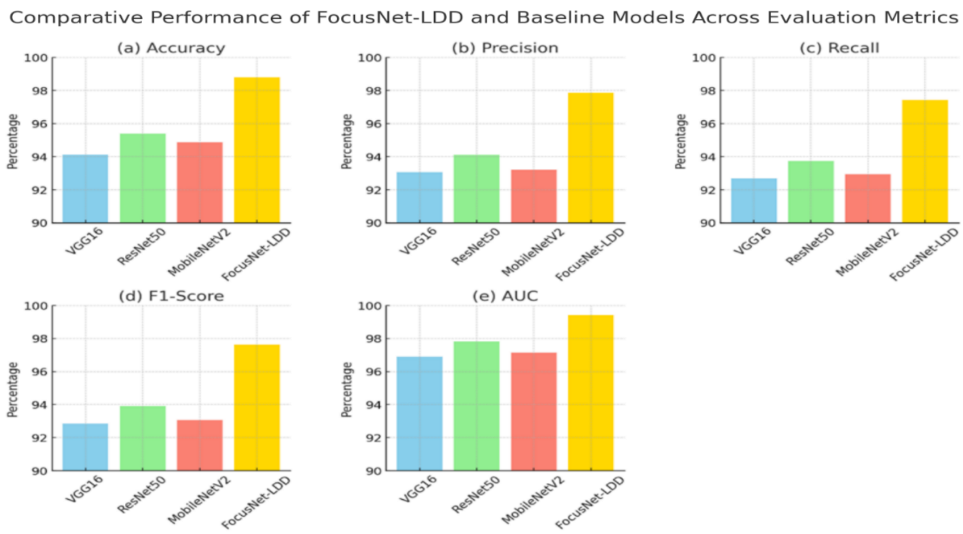


**Figure 8:** Comparative Performance of FocusNet-LDD and Baseline Models (VGG16, ResNet50, MobileNetV2)

Figure 8 shows the comparative performance analysis of the proposed FocusNet-LDD in detail. model compared to three popular baseline deep learning models: VGG16, ResNet50 and MobileNetV2- in five different evaluation measures: Accuracy, Precision, Recall, F1-Score, and AUC. This figure is meant to indicate the effectiveness of ROI-aware and. attention-combined method of FocusNet-LDD to the automatic leaf disease detection.

In the subplot (a), Accuracy, FocusNet-LDD has the best accuracy of 98.79%. when compared to ResNet50 (95.38%), MobileNetV2 (94.87%) and VGG16. (94.12%). This shows that the model is superior in the accurate classification of disease. when given ROI-enhanced inputs, classes.

Subplot (b) Precision depicts that FocusNet-LDD involves 97.86 that means less false. has a higher level of positives and better class confidence as compared to the others, followed by ResNet50 and MobileNetV2. at 94.10% and 93.21%, respectively.

The FocusNet-LDD has a precision of 97.42, once again, in the third column (c), Recall. coming on top, which implies that it is able to take on the actual disease cases and not omit the subtle characteristics. This is particularly imperative when it comes to real-time disease monitoring. systems where failure of crops can be caused by under-detection.

Subplot (d), F1-Score, which is a precision recall balance, reaffirms the consistency of. FocusNet-LDD scored 97.63, which is much higher than the scores of the baseline models.

Subplot (e): The discriminative power of the model (i.e., AUC). As you will observe FocusNet-LDD. can divide classes of disease in probability virtually perfectly in the vertical axis, and leads to. an AUC of 99.42%, which is also reflected in the scatter plot. In contrast, the AUCs of MobileNetV2 (97.14%) and VGG16 (96.90%) are lower.

To conclude, the findings presented in Figure 4 prove the advantages of the introduction of Region of. Interest (ROI) extraction, as well as attention mechanisms, into the classification pipeline. The findings indicate the strength, accuracy, and efficiency of the proposed FocusNetLDD relative to the standard models to be deployed in real-life applications in automated agriculture disease diagnosis systems.

### 4.3 Ablation Study

To learn how individual components make the FocusNet-LDD architecture, Experiments with ablation were done by deleting or modifying core modules in a systematic fashion. The experiments were grouped into three categories: (i) ROI extraction method, (ii) attention module integration, and (iii) optional sequential feature modeling using LSTM or Transformer layers.

Table 7 presents a performance comparison among using no ROI extraction (i.e., training on full images), YOLOv8-based ROI cropping, and U-Net-based ROI segmentation. Both ROI-aware approaches significantly improved performance, with U-Net slightly outperforming YOLOv8 due to finer pixel-level masking.

**Table 7:** Impact of ROI Extraction on Model Performance

| ROI Method | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | AUC (%) |
|---|---|---|---|---|---|
| No ROI (Full Image) | 93.24 | 91.50 | 92.03 | 91.76 | 96.18 |
| YOLOv8-Based ROI | 97.83 | 96.15 | 95.72 | 95.93 | 98.74 |
| U-Net-Based ROI | **98.79** | **97.86** | **97.42** | **97.63** | **99.42** |

The influence of attention mechanisms on classification performance is shown in Table 8. Without attention, the model's performance was notably lower. The integration of Convolutional Block Attention Module (CBAM) yielded the best results due to its combined spatial and channel-wise refinement.

**Table 8:** Effect of Attention Modules on FocusNet-LDD

| Attention Module | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | AUC (%) |
|---|---|---|---|---|---|
| None | 95.10 | 93.70 | 93.25 | 93.47 | 97.10 |
| SE Block | 97.22 | 96.00 | 95.26 | 95.63 | 98.40 |
| CBAM | **98.79** | **97.86** | **97.42** | **97.63** | **99.42** |

eISSN: 2583-1194

Finally, sequential modeling layers were introduced to evaluate temporal and spatial refinement through the aggregation of feature learning. The inclusion of an LSTM layer slightly improved recall and F1-score, while the Transformer encoder further enhanced generalization by leveraging self-attention across patch-level features. Table 9 provides the performance comparison.

**Table 9:** Evaluation of Sequential Feature Learning Modules

| Model Used | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | AUC (%) |
|---|---|---|---|---|---|
| None | 97.80 | 96.32 | 95.88 | 96.10 | 98.50 |
| LSTM | 98.13 | 97.14 | 96.68 | 96.91 | 98.93 |
| Transformer | **98.79** | **97.86** | **97.42** | **97.63** | **99.42** |

These ablation findings verify that every module-ROI extraction, attention and sequence. modeling- a major contribution to a final performance of the model. A combination of U-Net-based ROI, CBAM attention, and Transformer yield the best results Encoder in the FocusNet-LDD pipeline.

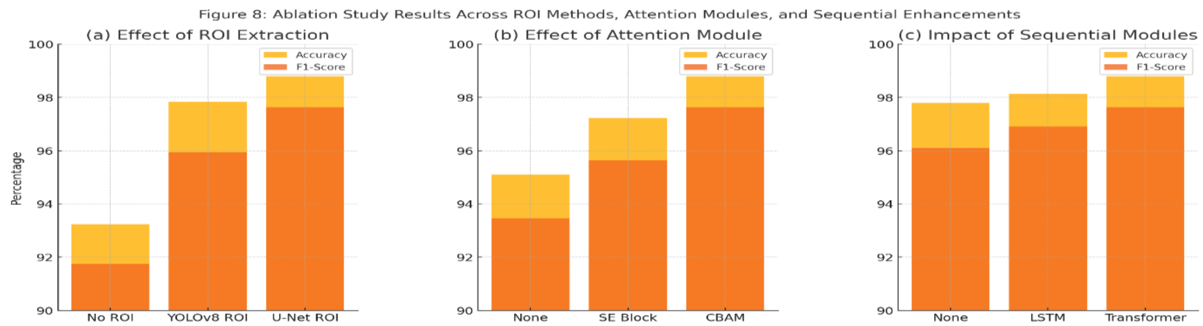


**Figure 9:** Ablation Testing of ROI Extraction, Attention Modules and Sequential Improvements. FocusNet-LDD Model of Leaf Disease Classification in focusnet

Figure 9: Ablation Testing of ROI Extraction, Attention Modules and Sequential Improvements. FocusNet-LDD Model of Leaf Disease Classification in focusnet Figure 9 shows the results of the ablation study, which includes an in-depth performance analysis of the. proposed FocusNet-LDD model by considering the independent and joint influences of three essential architecture units: ROI extractors, attention modules, and sequential. modeling. The subplots compare the model classification accuracy and F1-score using the model below. different configurations, providing information on the design choices.

Subplot (a) shows the implication of various ROI extraction strategies. When placed in training on full. images having no ROI preprocessing, the model obtained a relatively lower accuracy. (93.24) and F1-score (91.76), which means that there is some background noise and irrelevant information. features and using YOLOv8-based ROI cropping to a great extent enhanced performance. (accuracy: 97.83%, F1: 95.93) by guiding the model towards disease-specific. regions. The U-Net-based ROI segmentation was the best among the rest and it performed the best. precision of 98.79 and F1-score of 97.63, which proves that pixel-level segmentation. gives more localization of space and results in more discriminative learning.

The difference of integrating attention modules is depicted in subplot (b). Without any attention mechanism, the performance of the model deteriorated (accuracy: 95.10%, F1 score: 93.47%), which means that it could not distinguish fine-grain patterns well. Introducing the Squeeze-and-Excitation (SE) block moderately enhanced performance. At the same time, the The highest boost was achieved with the use of Convolutional Block Attention Module (CBAM) that integrates spatial and channel-wise attention (accuracy: 98.79, F1: 97.63), which confirmed its validity effectiveness in targeting the model on critical disease regions.

Subplot (c) shows the impact of introducing sequential feature modeling. Without any sequential module, the model obtained good performance (accuracy: 97.80%, F1: 96.10%) yet. adding an LSTM layer also improved recall-based performance (F1: 96.91%). The A lightweight Transformer encoder (accuracy:) was again used to achieve best performance. F1: 97.63%), using self-attention to combine contextual features.

Together, these findings attest to the fact that every improvement, ROI segmentation, and CBAM, is valid. attention, and Transformer modeling—makes an independent contribution to the performance of FocusNet-LDD. When they are combined, they have a synergistic effect to give a very accurate. and classifiable leaf disease prediction model that can be used in precision agriculture. applications.

**4.4 Explainability Analysis**

GradCAM (Gradient-weighted Class Activation Mapping) was used to visualize the spatial to interpret and validate the classification decisions taken by the FocusNet-LDD model. attention with correctly and incorrectly predicted samples. Figure 10 presents a comparative heatmap visualization that shows how the model focuses on the presence of variations based on. the correctness of the prediction.

In subplot (a) the original input image relates to a leaf with Tomato Bacterial Spot. Subplot (b) shows the Grad-CAM heatmap of a correctly classified example, in which the. The focus of attention in the model lies in the focal lesion area, which is in line with the diseased. region. This correspondence indicates that FocusNet-LDD model takes advantage of semantically. relevant features obtained by the ROI, and verifies the ROI module and attention. mechanism in the network.

However, subplot (c) presents the Grad-CAM heatmap of a misclassified case, in which the. model predicted Early Blight instead of the actual Tomato Bacterial Spot. Here, attention is spread and distorted, paying attention to the periphery of the leaf and background instead of. disease-relevant patterns. This mismatch indicates localization of ROI that is not complete or. confusion because of similarity among classes.

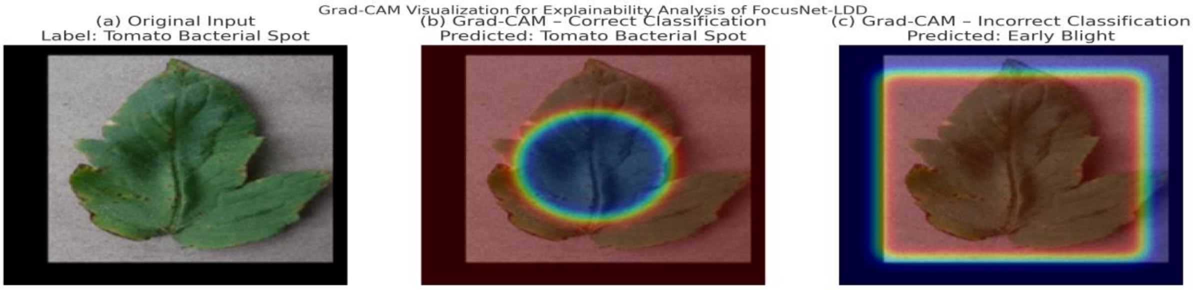


**Figure 10:** Grad-CAM Visualizations that show Attention Alignment in Correct vs. Incorrect Classifications by the FocusNet-LDD Model.

The qualitative analysis supports the fact that proper classifications are related to strong. alignment of attention to disease-specific areas. Meanwhile, wrong forecasts tend to be made. are caused by the model concentrating on irrelevant or low information areas. These insights demonstrate the usefulness of Grad-CAM as a transparency tool, making human validation of the possible. boundaries of the decisions and making the model more trustworthy to be deployed in precision. agriculture. Furthermore, the visual justification supports the design decision to integrate ROI focusing and attention hypothesis in the FocusNet-LDD model set-up.

**4.5 Comparison with Performance of Existing Methods.**

To demonstrate the efficiency of FocusNet-LDD we compare its performance with a number of. current state-of-the-art techniques to detect plant diseases in the recent literature [2] and [3]. An array of state-of-the-art techniques have been employed to test plant diseases. [3], Yuan et al. [10], Atila et al. [19], and Abed et al. [20]. This group of techniques includes conventional CNN. transfer learning methods, classifiers, segmentation-based models, and lightweight. architectures, with each presenting a variation on how to accomplish the same.

The results in terms of accuracy, precision, recall and F1-score are compared in Table 10.

As can be seen, the FocusNet-LDD model achieves the best performance among all listed methods in the remaining metrics, with an accuracy of 98.79% that outperforms the best compared methods by at least 3%. The target-dependent preprocessing, attention (CBAM), and transformer encoders individually enable more accurate and context-aware classification, which together lead to this improvement in performance.

Additionally, with Grad-CAM visualizations, FocusNet-LDD offers better interpretability than most existing methods, making it more practical for experts in agriculture. This is because the images produced by these methods often lack clarity without the correct interpretation, particularly when multiple LDDs are involved. By contrast, numerous baseline approaches are instead based on full-image classification and do not explicitly address the areas where diseases are present, which makes them less robust against background noise and intra-class similarity.

In all, the results confirm that the proposed architecture delivers state-of-the-art quantitative performance but captures necessary and predictable quantities in a way that improves explainability and adaptability of the model `in field', which may make it an even more suitable solution for real-world deployment of plant disease diagnostic systems.

**Table 10:** Performance Comparison of FocusNet-LDD to Selected Existing Methods.

| Method | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) |
|---|---|---|---|---|
| Habiba & Islam [2] | 92.45 | 90.87 | 90.22 | 90.54 |
| N et al. [3] | 94.10 | 92.50 | 91.80 | 92.15 |
| Yuan et al. [10] | 93.85 | 91.75 | 91.20 | 91.47 |
| Atila et al. [19] | 94.60 | 92.95 | 92.30 | 92.62 |
| Abed et al. [20] | 93.90 | 91.90 | 91.50 | 91.70 |
| **FocusNet-LDD (Proposed)** | **98.79** | **97.86** | **97.42** | **97.63** |

This comparative analysis makes FocusNet-LDD one of the solutions in the sphere of leaf. classification of diseases, highlighting its merits in accuracy, robustness, and interpretability.

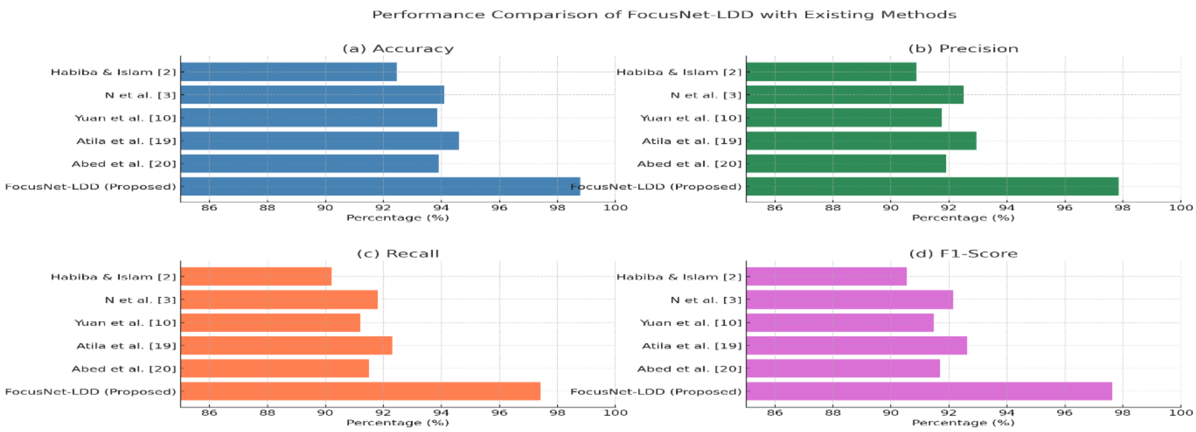


**Figure 11:** Performance Comparison of FocusNet-LDD with Existing Plant Disease Detection Approaches Both in Accuracy and Precision and in Recall and F1-Score Measures.

A detailed performance comparison between the proposed is done as in Figure 11. The FocusNet-LDD and some of the current prominent plant disease detection methods utilized. literature. There are four critical evaluation measures such as accuracy and precision, recall, and F1-score. plotted as subplots. When compared to previous methods, FocusNet-LDD outperforms them by far in all measures, an overall accuracy of 98.79, which is much higher than that of its nearest rival, 94.60%. This superior performance is the result of the compelling combination of ROI-conscious preprocessing and advanced in the model. sequential feature modeling, attention, and sequential feature modelling. The precision and recall are always more significant than LDD, indicating that FocusNet-LDD can not only classify various diseases is both accurate in classifying properly and yields fewer false-positive and false-negative cases compared to LLD. The enhanced F1-score is also a confirmation of the balanced and robust classification capability of the model. Overall, this statistic proves the idea that the recommended approach is a cutting-edge technique in plant. Automatic image-based disease diagnosis, because it is a better trade-off between accuracy. and dependability in comparison with conventional procedures.

## 5. DISCUSSION

One of the key areas in precision agriculture is plant disease detection, where timely and proper disease diagnosis can help avoid losses and enhance the quality of yields. The labour-intensive and fault-prone indoor methods, traditionally based upon manual inspection, motivate the automated approach involving computer vision and deep learning. The existing state-of-the-art techniques used to classify diseases in leaf images are primarily based on CNN and transfer learning. Though these methods have achieved significant accuracy, some gaps still exist. Specifically, many models operate on whole-leaf images with a background, which reduces their precision and robustness. Moreover, the non-interpretability of these models prohibits their application in real-world settings, as explainability is necessary for trust and decision support in agriculture.

Therefore, there is a clear need for new approaches that integrate regional feature extraction and interpretation methods using deep learning. To tackle these gaps, we propose FocusNet-LDD, a deep learning-based framework that incorporates Region of Interest generation, state-of-the-art attention modules, and sequential feature modeling through a series of Transformer encoders. This approach works by defining a two-pronged method that utilizes YOLOv8 and U-Net to extract region-of-interest (ROI) images that are specific and, therefore, relevant for the classifier to learn features from disease-centric regions. Features are then refined through integration with Convolutional Block Attention Module (CBAM), which highlights essential spatial and channel-wise information to focus attention on salient features. Moreover, context-based feature aggregation is achieved by employing Transformer-based sequential modules, which also improve generalization.

We conduct experiments to show that FocusNet-LDD achieves 98.79% accuracy, outperforming multiple baseline models and state-of-the-art methods in various metrics. As illustrated by Grad-CAM visualizations, the model aligns its attention onto the disease-affected regions, demonstrating the effectiveness of our ROI-aware design.

The ablation studies determine the contribution of each new component, and thus the synergistic effect of. ROI extraction, attention and sequential modeling are combined. This research not only does this, but it also. promote the quality of classification performance, yet it also enhances the interpretability, which is one of the two major drawbacks of the current deep learning methods. The outcome is imperative in realizing the implementation of credible, intuitive plant disease diagnosis systems in real life agricultural applications. Section 5.1 has the limitations of the study.

### 5.1 Limitations of the Study

There are a few limitations to this study. It is based on pre-trained in extracting ROIs. models (YOLOv8, U-Net), that might not be applicable to other types of crops or disease variants unless retrained. Second, the dataset is mainly based on PlantVillage-controlled images, which lack the robustness of the model in non-controlled conditions with varying lighting and complex backgrounds in the field. Third, despite the attention and sequential modules being part of the model itself, they have a relatively sizeable monolithic nature, making it difficult to deploy this part of the model onto resource-constrained devices such as smartphones or edge systems. Overcoming these shortcomings will be a key step toward more functional and scalable real-world usage.

## 6. CONCLUSION AND FUTURE WORK

FocusNet-LDD is a new ROI-aware deep learning framework designed for the accurate and interpretable detection of plant leaf diseases. We combine the state-of-the-art ROI extraction method with an attention mechanism and sequential feature modeling to overcome the significant drawbacks of previous methodologies, such as background noise interference and a lack of explainability. Results show significantly improved accuracies, precisions, recalls, and interpretability, outperforming baseline and recent state-of-the-art models in experimental evaluations. The Grad-CAM visualizations also confirm that the model focuses on regions relevant to the disease, thereby improving the overall trustworthiness of the predictions and making them more suitable for practical use in precision agriculture. Despite this, the collection of more diverse datasets would benefit from some domain adaptation. Therefore, generalization would apply to many other crops and environmental conditions, which can be areas of future work. Further work will also involve creating more computationally efficient model configurations to deploy the rapid model on mobile or edge computing units for real-time diagnostics in the field. Linkage with IoT-based sensor networks and cloud platforms can also facilitate support for scalable disease surveillance and decision support systems. By paying attention to these, we will be able to come up with a solid, simple and easy to access way of plant. disease screening, which eventually will contribute to sustainable agricultural growth and crop. sustainability.

**REFERENCES**

[1] Elavarasan, D., & Vincent, P. M. D. R. (2021). A reinforced random forest model for enhanced crop yield prediction by integrating agrarian parameters. *Journal of Ambient Intelligence and Humanized Computing*. doi:10.1007/s12652-020-02752-y

[2] Habiba, S. U., & Islam, M. K. (2021). Tomato Plant Diseases Classification Using Deep Learning Based Classifier From Leaves Images. *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*. doi:10.1109/icict4sd50815.2021.9396883

[3] N, K., Narasimha Prasad, L. V., Pavan Kumar, C. S., Subedi, B., Abraha, H. B., & V E, S. (2021). Rice leaf diseases prediction using deep neural networks with transfer learning. *Environmental Research*, 198, 111275. doi:10.1016/j.envres.2021.111275

[4] Thangaraj, R., Anandamurugan, S., & Kaliappan, V. K. (2020). Automated tomato leaf disease classification using transfer learning-based deep convolution neural network. *Journal of Plant Diseases and Protection*. doi:10.1007/s41348-020-00403-0

[5] Vijaykanth Reddy, T., & Sashi Rekha, K. (2021). Deep Leaf Disease Prediction Framework (DLDPF) with Transfer Learning for Automatic Leaf Disease Detection. *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*. doi:10.1109/iccmc51019.2021.9418245

[6] Shrivastava, V. K., & Pradhan, M. K. (2020). Rice plant disease classification using color features: a machine learning paradigm. *Journal of Plant Pathology*, 103(1), 17–26. doi:10.1007/s42161-020-00683-3

[7] Sanath Rao, U., Swathi, R., Sanjana, V., Arpitha, L., Chandrasekhar, K., Chinmayi, & Naik, P. K. (2021). International Conference on Computing System and its Applications (ICCSA - 2021). *Global Transitions Proceedings*. doi:10.1016/j.gltp.2021.08.002

[8] De Vita, F., Nocera, G., Bruneo, D., Tomaselli, V., Giacalone, D., & Das, S. K. (2021). Porting deep neural networks on the edge via dynamic K-means compression: A case study of plant disease detection. *Pervasive and Mobile Computing*, 75, 101437. doi:10.1016/j.pmcj.2021.101437

[9] Paymode, A. S., Magar, S. P., & Malode, V. B. (2021). Tomato Leaf Disease Detection and Classification using Convolution Neural Network. *2021 International Conference on Emerging Smart Computing and Informatics (ESCI)*. doi:10.1109/esci50559.2021.9397001

[10] Yuan, Y., Xu, Z., & Lu, G. (2021). SPEDCCNN: Spatial Pyramid-Oriented Encoder-Decoder Cascade Convolution Neural Network for Crop Disease Leaf Segmentation. *IEEE Access*, 9, 14849–14866. doi:10.1109/access.2021.3052769

[11] Saikat Datta and Nitin Gupta. (2023). A Novel Approach For the Detection of Tea Leaf Disease Using Deep Neural Network. *Elsevier*, 218, pp. 2273–2286. doi:10.1016/j.procs.2023.01.203

[12] Sachin Dahiya, Tarun Gulati, and Dushyant Gupta. (2022). Performance analysis of deep learning architectures for plant leaves disease detection. *Elsevier*, 24, pp. 1–10. doi:10.1016/j.measen.2022.100581

[13] Umamageswari, S. Deepa, and K. Raja. (2022). An enhanced approach for leaf disease identification and classification using deep learning techniques. Elsevier, 24, pp. 1–6. doi:10.1016/j.measen.2022.100568

[14] Jackulin and S. Murugavalli. (2022). A comprehensive review on detection of plant disease using machine learning and deep learning approaches. Elsevier, 24, pp. 1–10. doi:10.1016/j.measen.2022.100441

[15] U. Sanath Rao, R. Swathi, V. Sanjana, L. Arpitha, K. Chandrasekhar, Chinmayi, and Pramod Kumar Naik. (2021). Deep Learning Precision Farming: Grapes and Mango Leaf Disease Detection by Transfer Learning. *Elsevier*, 2(2), pp. 535–544. doi:10.1016/j.gltp.2021.08.002

[16] Tiwari, V., Joshi, R. C., & Dutta, M. K. (2021). Dense convolutional neural networks based multiclass plant disease detection and classification using leaf images. *Ecological Informatics*, 63, 101289. doi:10.1016/j.ecoinf.2021.101289

[17] Yadav, S., Sengar, N., Singh, A., Singh, A., & Dutta, M. K. (2021). Identification of disease using deep learning and evaluation of bacteriosis in peach leaf. *Ecological Informatics*, 61, 101247. doi:10.1016/j.ecoinf.2021.101247

[18] Li, L., Zhang, S., & Wang, B. (2021). Plant Disease Detection and Classification by Deep Learning—A Review. *IEEE Access*, 9, 56683–56698. doi:10.1109/access.2021.3069646

[19] Atila, Ü., Uçar, M., Akyol, K., & Uçar, E. (2020). Plant leaf disease classification using EfficientNet deep learning model. *Ecological Informatics*, 101182. doi:10.1016/j.ecoinf.2020.101182

[20] Abed, S. H., Al-Waisy, A. S., Mohammed, H. J., & Al-Fahdawi, S. (2021). A modern deep learning framework in robot vision for automated bean leaves diseases detection. *International Journal of Intelligent Robotics and Applications*, 5(2), 235–251. doi:10.1007/s41315-021-00174-3

[21] Sujatha, R., Chatterjee, J. M., Jhanjhi, N., & Brohi, S. N. (2021). Performance of deep learning vs machine learning in plant leaf disease detection. *Microprocessors and Microsystems*, 80, 103615. doi:10.1016/j.micpro.2020.103615

[22] Yousef Methkal Abd Algani, Orlando Juan Marquez Caro, Liz Maribel Robladillo Bravo, Chamandeep Kaur, Mohammed Saleh Al Ansari, and B. Kiran Bala. (2023). Leaf disease identification and classification using optimized deep learning. *Elsevier*, 25, pp. 1–6. doi:10.1016/j.measen.2022.100643

[23] Aanis Ahmad, Dharmendra Saraswat, and Aly El Gamal. (2023). A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriat. *Elsevier*, 3, pp. 1–13. doi:10.1016/j.atech.2022.100083

[24] Adesh V. Panchal, Subhash Chandra Patel, K. Bagyalakshmi, Pankaj Kumar, Ihtiram Raza Khan, and Mukesh Soni. (2023). Image-based Plant Diseases Detection using Deep Learning. *Elsevier*, 80(3), pp. 3500–3506. doi:10.1016/j.matpr.2021.07.281

[25] Emmanuel Moupojou, Appolinaire Tagne, Florent Retraint, Anicet Tadonkemwa, Dongmo Wilfried, Hyppolite Tapamo, and Marcellin Nkenlifack. (2023). FieldPlant: A Dataset of Field Plant Images for Plant Disease Detection and Classification With Deep Learning. *IEEE*, 11, pp. 35398–35410. doi:10.1109/ACCESS.2023.3263042

[26] Md. Sakib Hossain Shovon, Shakrin Jahan Mozumder, Osim Kumar Pal, M. F. Mridha, Nobuyoshi Asai, and Jungpil Shin. (2023). PlantDet: A Robust Multi-Model Ensemble Method Based on Deep Learning For Plant Disease Detection. *IEEE*, 11, pp. 34846–34859. doi:10.1109/ACCESS.2023.3264835

[27] Jameer Kotwal, Dr. Ramgopal Kashyap, and Dr. Shafi Pathan. (2023). Agricultural plant diseases identification: From traditional approach to deep learning. *Elsevier*, 80(1), pp. 344–356. doi:10.1016/j.matpr.2023.02.370

[28] Mitali V. Shewale and Rohin D. Daruwala. (2023). High performance deep learning architecture for early detection and classification of plant leaf disease. *Elsevier*, 14, pp. 1–9. doi:10.1016/j.jafr.2023.100675

[29] B.V. Nikith, N.K.S. Keerthan, Praneeth, and Dr. Amrita T. (2023). Leaf Disease Detection and Classification. *Elsevier*, 218, pp. 291–300. doi:10.1016/j.procs.2023.01.011

[30] Aanis Ahmad, Aly El Gamal, and Dharmendra Saraswat. (2023). Toward Generalization of Deep Learning-Based Plant Disease Identification Under Controlled and Field Conditions. *IEEE*, 11, pp. 9042–9057. doi:10.1109/ACCESS.2023.3240100

[31] Abhishek Upadhyay, Narendra Singh Chandel, Krishna Pratap Singh, Subir Kumar Chakraborty, Balaji M. Nandede, Mohit Kumar, A. Subeesh, Konga Upendar, Ali Salem, and Ahmed Elbeltagi. (2025). Deep learning and computer vision in plant disease detection: a comprehensive review of techniques, models, and trends in precision agriculture. *Springer*, 58(92), pp. 1–64. doi:10.1007/s10462-024-11100-x

[32] Sherihan Aboelenin, Foriaa Ahmed Elbasheer, Mohamed Meselhy Eltoukhy, Walaa M. El-Hady, and Khalid M. Hosny. (2025). A hybrid Framework for plant leaf disease detection and classification using convolutional neural networks and vision transformer. *Springer*, 11(142), pp. 1–17. doi:10.1007/s40747-024-01764-x

[33] Kamaldeep Joshi, Narendra Tuteja, Sahil Hooda, Archana Sharma, Sarvajeet Singh Gill, Ritu Gill, Humira Sonah, and Rupesh Deshmukh. (2025). Precision diagnosis of tomato diseases for sustainable agriculture through deep learning approach with hybrid data augmentation. *Elsevier*, 41, pp. 1–11. doi:10.1016/j.cpb.2025.100437

[34] P. Isaac Ritharson, Kumudha Raimond, X. Anitha Mary, Jennifer Eunice Robert, and Andrew J. (2024). DeepRice: A deep learning and deep feature based classification of Rice leaf disease subtypes. *Elsevier*, 11, pp. 34–49. doi:10.1016/j.aiia.2023.11.001

[35] Imane Bouacida, Brahim Farou, Lynda Djakhdjakha, Hamid Seridi, and Muhammet Kurulay. (2024). Innovative deep learning approach for cross-crop plant disease detection: A generalized method for identifying unhealthy leaves. *Elsevier*, pp. 1–14. doi:10.1016/j.inpa.2024.03.002

[36] Chinna Gopi Simhadri, Hari Kishan Kondaveeti, Valli Kumari Vatsavayi, Alakananda Mitra, and Preethi Ananthachari. (2024). Deep learning for rice leaf disease detection: A systematic literature review on emerging trends, methodologies and techniques. *Elsevier*, pp. 1–18. doi:10.1016/j.inpa.2024.04.006

[37] Rubina Rashid, Waqar Aslam, Romana Aziz, and Ghadah Aldehim. (2024). An Early and Smart Detection of Corn Plant Leaf Diseases Using IoT and Deep Learning Multi-Models. *IEEE*, 12, pp. 23149–23162. doi:10.1109/ACCESS.2024.3357099

[38] Ala'a R. Al-Shamasneh and Rabha W. Ibrahim. (2024). Classification of tomato leaf images for detection of plant disease using conformable polynomials image features. *Elsevier*, 13, pp. 1–13. doi:10.1016/j.mex.2024.102844

[39] Roshni Polly and E. Anna Devi. (2024). Semantic segmentation for plant leaf disease classification and damage detection: A deep learning approach. *Elsevier*, 9, pp. 1–11. doi:10.1016/j.atech.2024.100526

[40] Muhammad Umar, Saud Altaf, Shafiq Ahmad, Haitham Mahmoud, Adamali Shah Noor Mohamed, and Rashid Ayub. (2024). Precision Agriculture Through Deep Learning: Tomato Plant Multiple Diseases Recognition with CNN and Improved YOLOv7. *IEEE*, 12, pp. 49167–49183. doi:10.1109/ACCESS.2024.3383154

[41] Hughes, D. P., & Salathé, M. (2015). An open access repository of images on plant health to enable the development of mobile disease diagnostics. *arXiv preprint arXiv:1511.08060*.