

# Mask R-CNN Powered Deep Learning Model for Accurate Diagnosis and Classification of Plant Leaf Diseases

**Dr. S. Madhanmohan**

<sup>1</sup>Assistant Professor, Department of Computer Applications,  
Arignar Anna Government Arts College, Villupuram.  
(Deputed from Annamalai University)  
Email: nsmadhanmohan@gmail.com

## Abstract

Timely and accurate identification of plant leaf diseases is crucial in the agricultural sector, where plant health directly impacts crop yield and food security. Traditional methods for diagnosing plant diseases predominantly rely on manual inspection and expert judgment, which are often time-consuming, labor-intensive, and not scalable for large-scale monitoring. As agricultural demands continue to rise, there is a pressing need for automated, intelligent systems that can assist in the early detection and classification of plant diseases with high precision and reliability. In this paper, we propose a novel Mask R-CNN Powered Deep Learning Model for Accurate Diagnosis and Classification of Plant Leaf Diseases (MRPDL-PLD). The proposed model integrates advanced image processing techniques with deep learning architectures to create an end-to-end pipeline capable of identifying infected regions and classifying disease types from raw leaf images. To enhance the quality of the input data, a preprocessing stage using Adaptive Filtering (AF) is employed, effectively reducing image noise and preserving critical features necessary for accurate segmentation and classification.

The core of the proposed system is a Mask Region-Based Convolutional Neural Network (Mask R-CNN), which excels in both instance segmentation and object detection. This architecture allows for precise localization of diseased regions on the leaf surface, enabling a more detailed understanding of the severity and spread of infections. Feature extraction is performed on these segmented regions to generate high-quality feature vectors, which are then used to classify the specific type of disease affecting the plant. To further optimize model performance, we incorporate the ADAM optimizer for adaptive learning rate adjustment, which significantly enhances convergence speed and model accuracy. The MRPDL-PLD model is trained and evaluated on publicly available benchmark datasets consisting of diverse plant species and disease categories. Extensive

simulation experiments demonstrate that our model not only achieves superior classification accuracy and segmentation performance compared to existing contemporary approaches but also generalizes well across different types of plant diseases.

**Keywords:** Plant Leaf Disease, Leaf Region, Disease Diagnosis, Mask RCNN, ADAM, Adaptive Filtering, Segmentation, Deep Learning

## 1. Introduction

The global population is projected to reach approximately 9.1 billion by the year 2050, as reported by the Food and Agriculture Organization (FAO) of the United Nations. To meet the nutritional demands of this rapidly growing population, global food production must increase by nearly 70% by mid-century [1]. However, achieving this goal presents several critical challenges. The scarcity of arable land, limited access to clean and sufficient water resources, and the prevalence of plant diseases collectively hinder the ability to scale agricultural productivity. Among these challenges, plant diseases play a particularly detrimental role, significantly reducing both the quality and quantity of crop yields. These agricultural losses not only have a direct negative impact on farmers' income but also contribute to higher food prices for consumers and pose a threat to food security, especially in developing regions. In such regions, limited access to timely disease management and preventative interventions exacerbates the problem, often resulting in prolonged food shortages, hunger, and malnutrition [2]. Traditionally, the identification of plant diseases has relied heavily on manual inspection, wherein agricultural experts physically examine crops in the field. However, this method is labor-intensive, time-consuming, and increasingly impractical for large-scale farming operations. Moreover, due to limited availability of skilled labor and the vastness of cultivated lands, disease detection is often delayed, leading to the rapid spread of infections and increased crop loss [3].

In response to these limitations, the agricultural research community has increasingly focused on developing computer-aided technologies to support automated crop monitoring and disease diagnosis. Recent advances in deep learning and computer vision have paved the way for high-performing diagnostic models that provide fast, accurate, and scalable solutions for plant health assessment [4][5]. Several studies have explored

the use of convolutional neural networks (CNNs) and transfer learning models such as AlexNet and ResNet-50 for disease classification from leaf images [6][7]. While these methods have demonstrated promising results, many are limited by their inability to localize disease-affected areas with precision. This limitation has directed attention toward instance segmentation techniques such as Mask R-CNN, which excels in simultaneously detecting and segmenting specific regions of interest within images [8]. In this context, the article proposes a novel Mask R-CNN Powered Deep Learning Model for Accurate Diagnosis and Classification of Plant Leaf Diseases (MRPDL-PLD). The proposed model introduces an enhanced detection pipeline that incorporates Adaptive Filtering (AF) during the preprocessing stage to denoise and enhance the quality of leaf images. This step is crucial for preserving disease-relevant features and improving downstream analysis [9].

Following preprocessing, a Mask Region-Based Convolutional Neural Network (Mask R-CNN) is employed to segment the infected areas of the leaf with high precision. From these segmented regions, discriminative feature vectors are extracted to facilitate robust classification of various plant disease types. Furthermore, the performance of the model is fine-tuned using the ADAM optimization algorithm, which ensures efficient parameter updates and accelerates convergence. Comprehensive simulations and evaluations are conducted on standard plant leaf disease datasets to demonstrate the effectiveness of the proposed approach. The experimental results confirm that our method outperforms several existing models in terms of accuracy, segmentation quality, and computational efficiency [10].

## **2. The Related Work**

Pantazi XE et al. [11] presents a robust approach to automated plant disease detection using Local Binary Pattern (LBP) texture features and One-Class Classifiers (OCCs). The study focuses on developing classifiers trained solely on one disease class at a time, enabling effective disease identification without requiring exhaustive multiclass training data. A novel conflict resolution mechanism is also introduced to handle cases where multiple OCCs detect the same image. Tested on 46 different plant-condition combinations, the system achieved an impressive overall success rate of 95%, demonstrating high generalization across different crop species. The method's reliance on texture features makes it computationally efficient, while its design supports

scalability in real-world agricultural monitoring systems. This work significantly contributes to precision agriculture by offering a low-cost, reliable, and adaptable solution for early plant disease detection. Liu et al. [12] proposes a novel deep learning approach for enhancing the accuracy of prostate cancer diagnosis using ultrasound imaging. The proposed framework integrates an improved S-Mask R-CNN for precise segmentation of prostate regions in ultrasound images, effectively addressing challenges such as speckle noise and low signal-to-noise ratios. Subsequently, the segmented images are processed using an Inception-v3 network, which has been modified to replace its original classification module with a new network model composed of forward and back propagation mechanisms. This design leverages transfer learning strategies to enhance feature extraction and classification accuracy. The experimental results demonstrate that the integrated model achieves superior performance in both segmentation and classification tasks, outperforming traditional diagnostic methods and offering a promising tool for computer-aided diagnosis of prostate cancer. Lichun Zhang et al. [13] introduces MedMusNet, a novel deep learning model designed to improve the detection and segmentation of clinically significant prostate cancer using B-mode micro-ultrasound images. The model incorporates mask-enhanced deep supervision to effectively reduce noise and improve feature consistency across image frames, addressing challenges such as subtle tumor boundaries common in ultrasound imaging. Tested on a large dataset of over 10,000 images, MedMusNet achieved notable improvements in detection accuracy and segmentation quality compared to baseline models, successfully identifying 76% of significant cancers. The results also showed promising performance surpassing human experts in lesion and patient-level analysis, highlighting its potential as a valuable tool to assist clinicians in prostate cancer diagnosis and biopsy guidance. This work represents an important advancement in leveraging deep learning to enhance non-invasive prostate cancer diagnostics.

Sulaiman Vesal et al. [14] presents a novel deep learning model, ProsDectNet, designed to improve prostate cancer detection using B-mode transrectal ultrasound images. The model tackles the common challenges of low sensitivity and image noise in traditional TRUS-guided biopsies by combining lesion detection and patch classification within a multi-task architecture. Trained on radiologist-labeled and biopsy-confirmed datasets, and enhanced with uncertainty minimization techniques, ProsDectNet demonstrated superior performance compared to expert clinicians, achieving an ROC-AUC of 82%,

sensitivity of 74%, and specificity of 67% on patient data. This study highlights ProsDectNet's potential as a reliable computer-aided diagnostic tool to improve biopsy accuracy and prostate cancer diagnosis, especially in environments where MRI is less accessible. Yi Wang et al. [15] introduces a novel deep learning model aimed at enhancing prostate segmentation in 3D transrectal ultrasound (TRUS) images. The challenge of accurately delineating the prostate boundary in TRUS images is addressed through the development of a 3D deep neural network equipped with attention modules. These modules leverage multi-level features from different layers of the convolutional neural network to refine the segmentation process. By selectively focusing on relevant features and suppressing non-prostate noise, the model improves the accuracy of prostate delineation. Experimental results demonstrate the effectiveness of this approach, achieving satisfactory segmentation performance on challenging 3D TRUS volumes. The proposed attention mechanism offers a general strategy to aggregate multi-level deep features, which can be applied to other medical image segmentation tasks. The code for this method is publicly available, providing a valuable resource for further research and application in the field.

### **3. The Proposed Model**

This work introduced a novel MRPDL-PLD technique for classifying plant diseases. The MRPDL-PLD method uses a variety of processes at different stages, including AF-based pre-processing, Mask RCNN-based segmentation and classification. The complete workflow of the MRPDL-PLD model is illustrated in Figure 1, which outlines each step of the processing pipeline, from raw image input to final disease diagnosis.

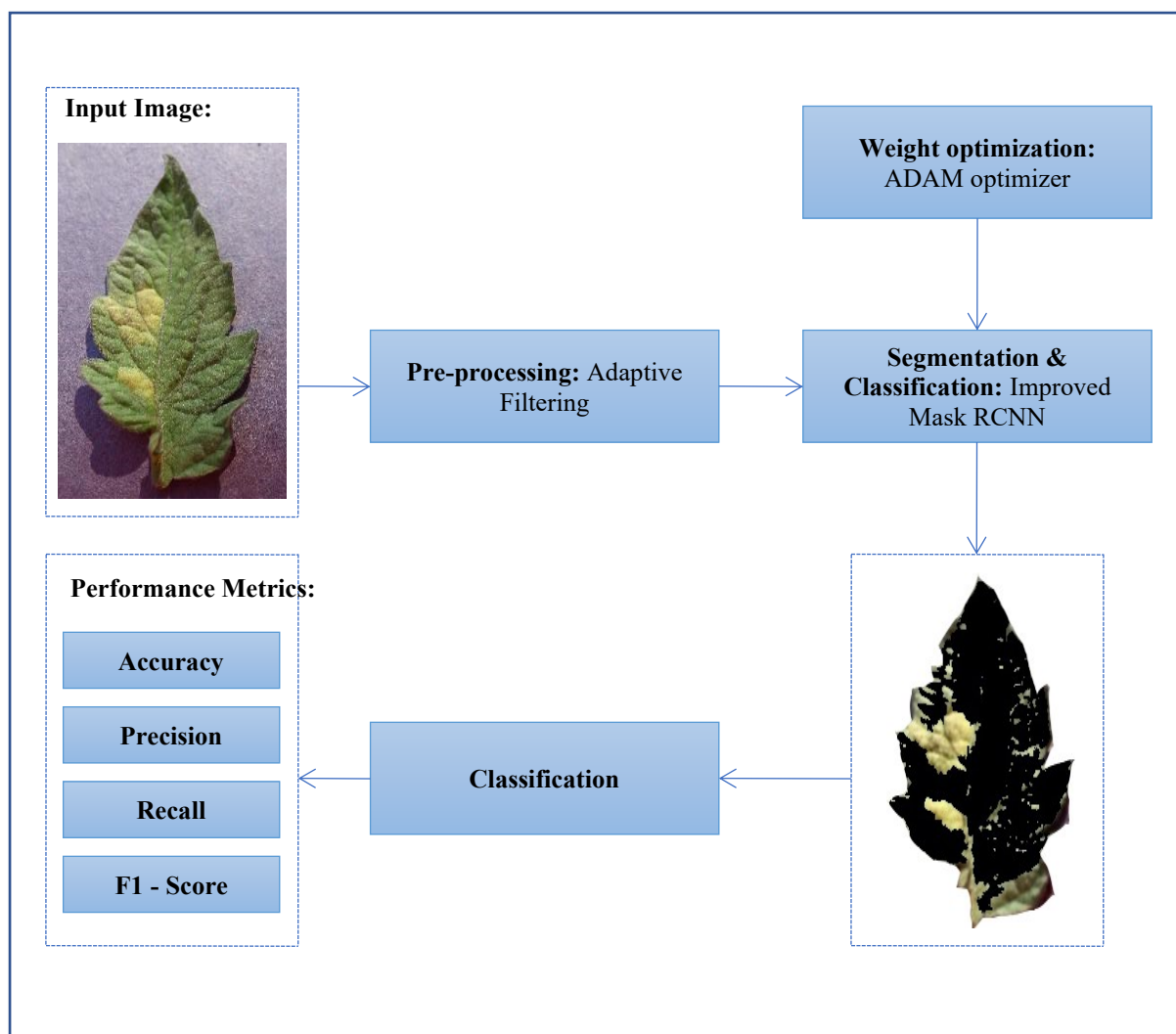


Fig. 1. Architecture of the proposed MRPDL-PLD technique

### 3.1 Image Pre-processing

The initial stage of the proposed MRPDL-PLD framework involves noise removal and image enhancement, which is accomplished using the AF technique. This method is specifically designed to enhance or restore plant leaf images while preserving their essential visual quality and structural details. Let  $f(n_1, n_2)$  denote the original input image. The local luminance mean, represented by  $f_L(n_1, n_2)$ , is computed by applying a low-pass filter to  $f(n_1, n_2)$ , effectively smoothing the image and capturing the overall brightness trends.

To extract local contrast, the low-frequency component  $f_L(n_1, n_2)$  is subtracted from the original image, yielding the high-frequency series  $f_H(n_1, n_2) = f(n_1, n_2) - f_L(n_1, n_2)$ . This series highlights subtle variations and textures, which are crucial for accurate

disease detection. To enhance or suppress contrast as needed,  $f_H(n_1, n_2)$  is multiplied by a scaling factor  $k(f_L)$ , a scalar function dependent on the local luminance  $f_L(n_1, n_2)$ . The adjusted local contrast is denoted by,

$$f'_H(n_1, n_2) = k(f_L) \cdot f_H(n_1, n_2) \quad (1)$$

The value of  $k(f_L)$  determines the contrast transformation behaviour, when  $k(f_L) > 1$ , the local contrast is amplified, when  $k(f_L) < 1$ , the contrast is suppressed. In parallel, the local luminance mean is modified through a nonlinear point-wise transformation, resulting in  $f'_L(n_1, n_2)$ . This nonlinear function is typically chosen based on application requirements and is designed to ensure that the dynamic range of the processed image remains close to the limits of the recording or display medium. Finally, the enhanced image  $g(n_1, n_2)$  is reconstructed by combining the modified local luminance and contrast components:

$$g(n_1, n_2) = f'_L(n_1, n_2) + f'_H(n_1, n_2) \quad (2)$$

This method allows for the adaptive enhancement of brightness and contrast, tailored to local image characteristics. As a result, the output image is not only visually enhanced but also more suitable for subsequent segmentation and classification tasks, contributing to the overall performance of the MRPD-PLD system.

### 3.2 Mask R-CNN based Segmentation and Classification

To accurately identify and classify plant leaf diseases, this study employs a novel deep learning approach termed MRPD-PLD. This section provides a comprehensive overview of the Mask R-CNN architecture and its role in the disease diagnosis pipeline. The Mask R-CNN framework consists of three primary stages: feature extraction, object detection, and segmentation, each playing a critical role in the accurate identification of diseased leaf regions. The process begins with input images of plant leaves, which are first passed through a Convolutional Neural Network backbone for feature extraction. These features serve as a rich representation of spatial and textural patterns that are crucial for distinguishing between healthy and diseased regions. Next, the Region Proposal Network (RPN) is used to generate a set of candidate regions also known as region proposals that potentially contain objects of interest, in this case, infected areas on the leaf. The RPN employs a sliding window approach across the feature map, applying anchor boxes of

multiple scales and aspect ratios to capture objects of varying sizes. For each region, object scores and bounding box coordinates are predicted.

Following region proposal, the Region of Interest Alignment (ROIAlign) operation is applied to extract fixed-size feature maps from these proposals. Unlike traditional ROI pooling, ROIAlign avoids spatial quantization, thereby preserving precise spatial locations crucial for accurate pixel-level segmentation. In the segmentation stage, a fully convolutional network (FCN) branch is employed to generate a binary mask for each detected object, effectively distinguishing foreground from background. This mask is aligned with the bounding box and predicts per-pixel labels within that region. The result is a fine-grained segmentation of the diseased portion of the leaf. The final output includes the predicted class label, bounding box, and segmentation mask for each identified region of interest. These outputs collectively enable both classification and localized detection of plant leaf diseases, offering a robust and interpretable solution. During training, the learning rate is initialized randomly at 0.01, and the network weights are updated iteratively using backpropagation with ADAM optimization algorithm. The weight update rule used in this context is defined as:

$$\omega_{t+1} = \omega_t - \eta \cdot \nabla L(\omega_t) \quad (3)$$

where,  $\omega_t$  denotes the weights at iteration  $t$ ,  $\eta$  is the learning rate, and  $\nabla L(\omega_t)$  represents the gradient of the loss function with respect to the weights.

### 3.3 ADAM Optimization

In the proposed methodology, the ADAM (Adaptive Moment Estimation) optimizer is employed for adaptive learning rate selection, ensuring efficient and stable convergence during model training. The learning rate is dynamically adjusted based on the loss function of the convolutional neural network, allowing the model to fine-tune its parameters effectively over successive iterations. ADAM is particularly well-suited for deep learning applications involving large datasets, high-dimensional parameter spaces, and complex non-convex optimization landscapes, which are typical in plant disease detection tasks. It combines the advantages of two widely used optimization techniques: AdaGrad, which works well with sparse gradients, and RMSProp, which handles non-stationary objectives. ADAM achieves this by utilizing first-order gradients while incorporating estimates of both the first and second moments of the gradients.

One of ADAM's key strengths is its low memory requirement and its ability to perform well even with noisy or sparse gradients, making it a practical choice for training deep convolutional networks. The first moment estimate (mean of the gradients) is computed as follows:

$$m_t = \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t \quad (4)$$

Where,  $m_t$  is the first moment (mean) at time step  $t$ ,  $\beta_1$  is the exponential decay rate for the first moment,  $g_t$  is the gradient of the loss function with respect to model parameters at time step  $t$ ,  $m_{t-1}$  is the first moment from the previous time step.

The 2nd momentum is expressed by,

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) \cdot g_t^2 \quad (5)$$

Where,  $v_t$  is the second moment (variance) estimate at time step  $t$ ,  $\beta_2$  is the exponential decay rate for the second moment,  $g_t$  is the gradient of the loss function at time step  $t$ ,  $v_{t-1}$  is the second moment from the previous time step,  $g_t^2$  denotes the element-wise square of the gradient.

This approach ensures that recent gradients have a higher influence on the learning process while still preserving the historical trend of updates. The ADAM optimizer thus plays a vital role in accelerating convergence and improving the accuracy and generalization ability of the proposed MRPD-PLD model.

#### 4. Results and Discussion

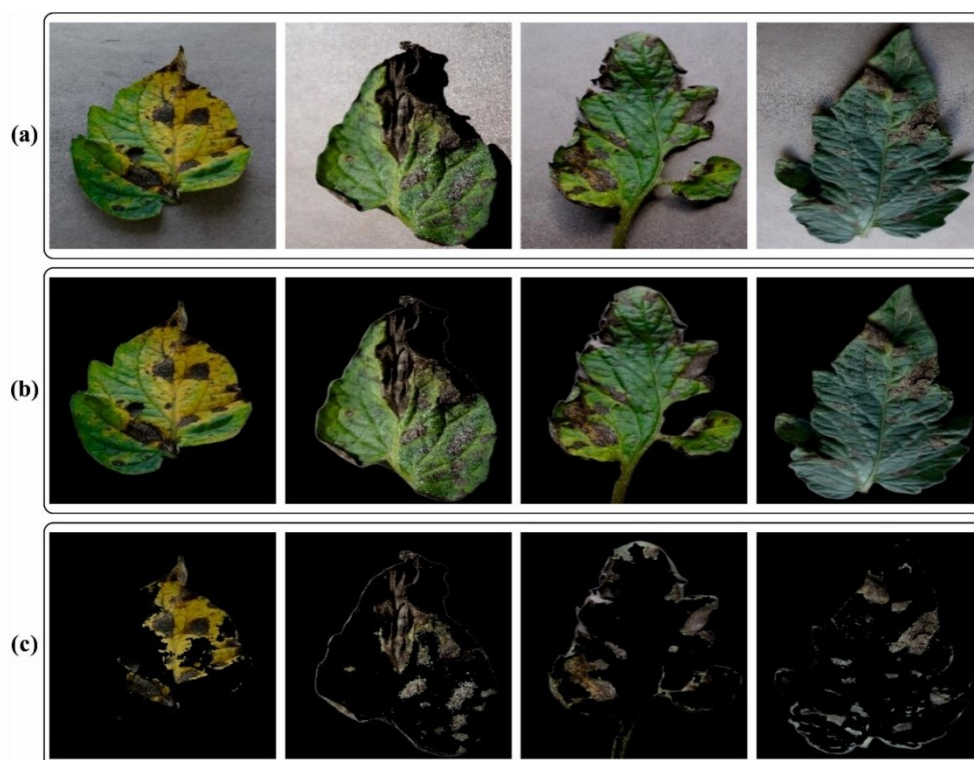
The proposed model is implemented and simulated using Python-based deep learning tools and libraries. To evaluate its performance, the model is tested on the widely recognized PlantDoc benchmark dataset [16], which comprises 5,452 annotated images categorized into four distinct disease classes. This dataset provides a diverse and realistic set of plant leaf images captured under natural conditions, making it highly suitable for validating disease detection algorithms. A detailed breakdown of the dataset, including the number of images per class and their respective labels, is presented in Table 1.

**Table 1:** Detail on Dataset

Classes	No. of Images
Early_Blight	1000

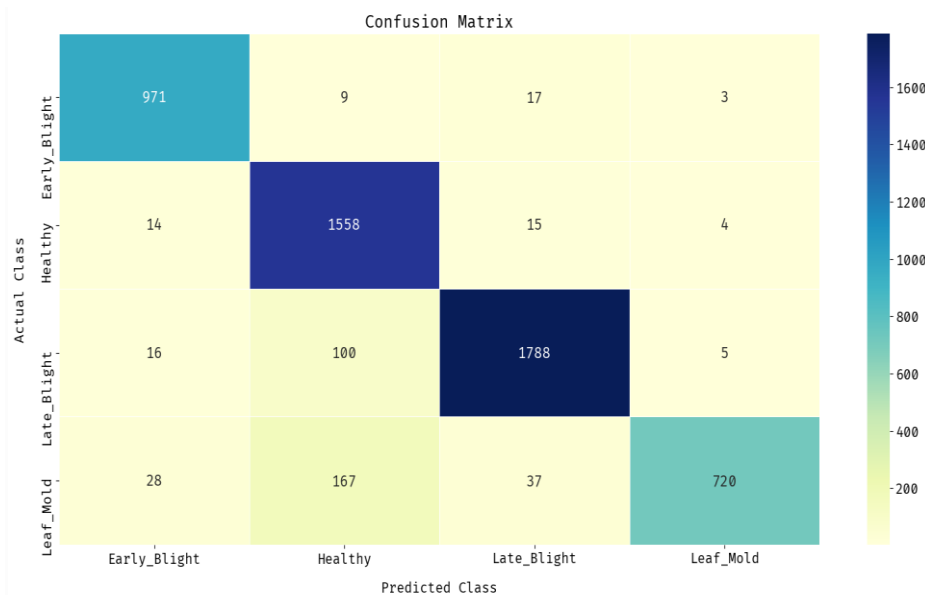
Late_Blight	1909
Leaf_Mold	952
Healthy	1591
<b>Total Images</b>	<b>5452</b>

Fig. 2a represents the sample original images and its pre-processed versions are provided in Fig. 2b. Then, the segmented images are illustrated in Fig. 2c.



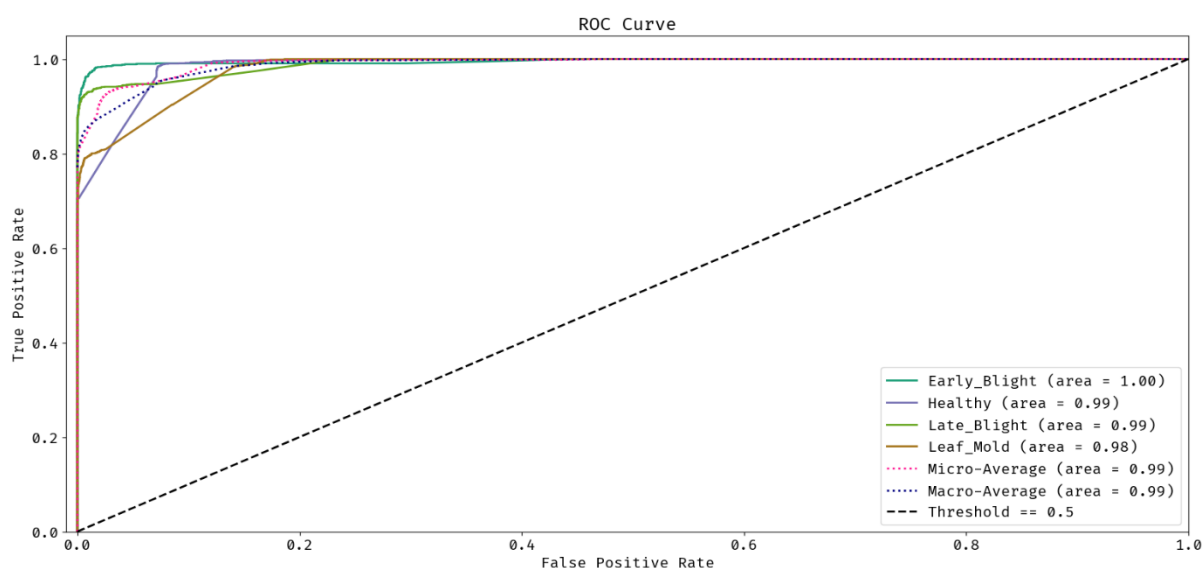
**Fig. 2.** a) Original Images b) Pre-processed Images c) Segmented Images

Figure 3 presents the confusion matrix generated by the proposed MRPD-PLD model for the classification of plant leaf diseases. As depicted, the model demonstrates strong performance across all four categories. Specifically, it correctly identifies 971 images as Early Blight, 1,558 images as Healthy, 1,788 images as Late Blight, and 720 images as Leaf Mold. These results indicate the model's capability to accurately distinguish between healthy and diseased leaf conditions, reflecting its robustness and reliability in real-world plant disease diagnosis scenarios.



**Fig. 3.** Confusion Matrix for Testing Set

Figure 4 illustrates the Receiver Operating Characteristic (ROC) analysis of the proposed MRPD-PLD model for the accurate diagnosis of plant leaf diseases. The ROC curves provide a comprehensive view of the model's discriminative performance across the four target classes. As observed, the MRPD-PLD approach exhibits exceptionally high classification accuracy, achieving an Area Under the Curve (AUC) of 1.00 for Early Blight, 0.99 for Healthy, 0.99 for Late Blight, and 0.98 for Leaf Mold. These near-perfect AUC values underscore the model's high sensitivity and specificity, confirming its robustness and effectiveness in distinguishing between healthy and diseased leaf samples with minimal false positives and negatives.

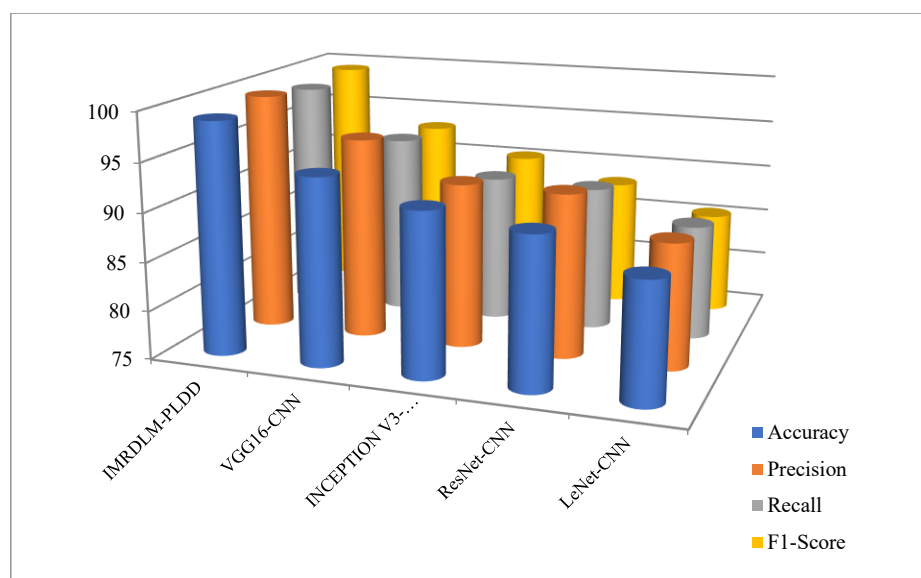


**Fig. 4.** ROC analysis of MRPDL-PLD model

A comprehensive comparative performance analysis is presented in Table 2 and Figure 5, highlighting the superior accuracy of the proposed MRPDL-PLD model in plant leaf disease classification. The experimental evaluation involved benchmarking against several well-established deep learning architectures, including VGG16-CNN, InceptionV3-CNN, ResNet-CNN, and LeNet-CNN. The respective classification accuracies achieved by these models were 94.16%, 91.82%, 90.54%, and 87.37%. In contrast, the proposed MRPDL-PLD model significantly outperformed these existing approaches by achieving an impressive accuracy of 98.86%. This notable performance improvement demonstrates the effectiveness of the integrated Mask R-CNN framework, combined with adaptive preprocessing and optimized training strategies, in capturing complex disease patterns with higher precision. The consistent and superior results clearly establish the MRPDL-PLD model as a more accurate and reliable solution for automated plant disease diagnosis compared to traditional CNN-based models.

**Table 2:** Accuracy analysis of MRPDL-PLD approach with recent algorithms

<b>Methods</b>	<b>Accuracy (%)</b>
MRPDL-PLD	98.86
VGG16-CNN	94.16
INCEPTION V3-CNN	91.82
ResNet-CNN	90.54
LeNet-CNN	87.37



**Fig. 5.** Comparative analysis of MRPD-PLD approach with recent algorithms

## 5. Conclusion

In this study, a novel MRPD-PLD system was developed for the automated classification and diagnosis of plant leaf diseases. The proposed approach begins with an AF-based pre-processing stage, which enhances the visual quality of leaf images by effectively reducing noise and improving contrast. These pre-processed images are then fed into the MRPD-PLD model, where a Mask R-CNN-based architecture performs precise segmentation and detection of diseased regions. Subsequently, the model extracts discriminative feature vectors from the segmented regions, which are used to accurately classify the type of disease present. To further optimize classification performance, the parameters of the Mask R-CNN are fine-tuned using the ADAM optimizer, ensuring adaptive learning and faster convergence during training.

To validate the effectiveness of the proposed system, extensive simulation experiments were carried out on benchmark datasets. The experimental results clearly demonstrate that the MRPD-PLD model significantly outperforms existing deep learning architectures in terms of classification accuracy and robustness. Looking ahead, the performance of the MRPD-PLD framework can be further enhanced by integrating feature reduction and dimensionality reduction techniques, which may improve computational efficiency and generalization across diverse plant species and environmental conditions. This advancement has the potential to contribute

substantially to precision agriculture by enabling faster, cost-effective, and scalable plant disease diagnostics.

## References

1. J. Schmidhuber and F. N. Tubiello, "Global food security under climate change", *Proceedings of the National Academy of Sciences*, vol. 104, no. 50, pp. 19703–19708, 2007.
2. S. P. Mohanty, D. P. Hughes, M. Salathé, "Using Deep Learning for Image-Based Plant Disease Detection", *Frontiers in Plant Science*, vol. 7, no. 1419, pp. 1–10, 2016.
3. S. Mustofa, M. M. H. Munna, Y. R. Emon, G. Rabbany, M. T. Ahad, "A comprehensive review on Plant Leaf Disease detection using Deep learning", *arXiv preprint*, arXiv:2308.14087, 2023.
4. D. Ward, P. Moghadam, N. Hudson, "Deep Leaf Segmentation Using Synthetic Data", *arXiv preprint*, arXiv:1807.10931, 2018.
5. M. Kumar, N. S. Chandel, D. Singh, L. S. Rajput, "Soybean Disease Detection and Segmentation Based on Mask-RCNN Algorithm", *Journal of Experimental Agriculture International*, vol. 45, no. 5, pp. 63–72, 2023.
6. N. Selvam, J. K. Joy, "Plant Leaf Disease Detection with Multivariable Feature Selection Using Deep Learning AEN and Mask R-CNN in PLANT-DOC Data", *Biosciences Biotechnology Research Asia*, vol. 21, no. 4, pp. 1–10, 2024.
7. Y. Pan, "Leaf disease detection and classification with a CNN-LSTM hybrid model", *Applied and Computational Engineering*, vol. 104, pp. 7–20, 2024.
8. K. He, G. Gkioxari, P. Dollár, R. Girshick, "Mask R-CNN", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 386–397, 2020.
9. H. Ajra, M. A. Majid, M. S. Islam, D. Abdullah, "Leaf Disease Detection in Plant Care using CNN Architecture: AlexNet and ResNet-50 Models", *International Journal on Advanced Science, Engineering and Information Technology*, vol. 15, no. 1, pp. 283–292, 2025.
10. A. Rahman, M. H. A. Foisal, M. H. Rahman, M. R. Miah, M. F. Mridha, "Deep-CNN for Plant Disease Diagnosis Using Low Resolution Leaf Images", *Smart Innovation, Systems and Technologies*, vol. 269, pp. 359–368, 2022.

11. Pantazi XE, Moshou D, Tamouridou, AA "Automated leaf disease detection in different crop species through image features analysis and One Class Classifiers", *Comput Electron Agric*, vol. 156(96), pp. 104, 2019.
12. Liu, Zhiyong, Chuan Yang, Jun Huang, Shaopeng Liu, Yumin Zhuo, and Xu Lu., "Deep learning framework based on integration of S-Mask R-CNN and Inception-v3 for ultrasound image-aided diagnosis of prostate cancer", *Future Generation Computer Systems*, vol. 114, pp. 358-367, 2021.
13. Lichun Zhang, Steve Ran Zhou, Moon Hyung Choi, et al., "Mask Enhanced Deeply Supervised Prostate Cancer Detection on B-mode Micro-Ultrasound," *arXiv preprint*, vol. 2412.10997, 2024.
14. Sulaiman Vesal, Indrani Bhattacharya, Hassan Jahanandish, et al., "ProsDectNet: Bridging the Gap in Prostate Cancer Detection via Transrectal B-mode Ultrasound Imaging," *arXiv preprint*, vol. 2312.05334, 2023.
15. Yi Wang, Haoran Dou, Xiaowei Hu, et al., "Deep Attentive Features for Prostate Segmentation in 3D Transrectal Ultrasound," *Medical Image Analysis*, vol. 57, pp. 186–197, 2019.
16. Singh, D., Jain, N., Jain, P., Kayal, P., Kumawat, S. and Batra, N., 2020. PlantDoc: a dataset for visual plant disease detection. In *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD* (pp. 249-253).