IJIE

UTHM Penerbit

INTERNATIONAL JOURNAL OF INTEGRATED ENGINEERING
ISSN: 2229-838X e-ISSN: 2600-7916

Vol. 18 No. 3 (2026) 475-489
https://publisher.uthm.edu.my/ojs/index.php/ijie

High-Performance Disease Detection in Plant Leaves Using an Enhanced Mask R-CNN

Katabathina Lakshmi Devi[1,2]*, Kasula Venkata Durga Kiran[1]

[1] *Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, INDIA*
[2] *Department of Information Technology, Lakireddy Bali Reddy College of Engineering, Mylavaram, INDIA*

*Corresponding Author: katabathina.lakshmidevi86@gmail.com
DOI: https://doi.org/10.30880/ijie.2026.18.03.031

Article Info

Received: 12 August 2025
Accepted: 17 March 2026
Available online: 8 June 2026

Keywords

Enhanced Mask R-CNN, learning, plant disease detection, image segmentation, agricultural intelligence, leaf disease classification, Convolutional Neural Networks (CNN)

Abstract

The present research model classifies plant leaf diseases with the help of a sophisticated deep learning structure called the Enhanced Mask Region-Based Convoluted Neural Network (EnMask-R-CNN). This model is accurate in classifying and segmenting images and isolating disease-afflicted areas and defining their boundaries perfectly. A threshold-based segmentation approach is also added to enhance region detection and hence, enhance the overall accuracy and computational performance of the model. Proposed EnMask-R-CNN performance is compared with the performance of a standard Recurrent Neural Network (RNN) model in terms of 105 images as a sample size. The results of the experiment demonstrate that the suggested model has a much lower error of 2.657, which is lower than that of the RNN baseline. In addition, the significance value of 0.258 received and the reduced loss metrics proves reliability and strength of the suggested approach. The obtained results prove that the EnMask-R-CNN system is more effective and correct in detecting plant disease than the conventional neural network methods. Evaluation is done on the PlantVillage dataset, which contains over 50K images across 38 classes of plant disease. This large-scale dataset allows for diversity in disease patterns, lighting conditions and background variations, ultimately performing better on robustness and generalization of the proposed model.

1. Introduction

In recent times, the use of automated plant disease detection has emerged as an important area of study in contemporary precision agriculture because of the increasing need to have scalable, accurate and real time systems that would be able to diagnose crop infections at early stages. The change to smart computational techniques as compared to traditional manual modes of inspection has served as a great boost on efficiency and reliability of plant health assessment systems. Over the past few years, a lot of research has been done on the topic of deep learning, image processing, segmentation, and optimization-based modeling to enhance plant disease detection in different species and environmental conditions.

Initial studies focused on the work were mostly based on conventional image-processing pipelines, relying on highly crafted features of color histograms, texture descriptors and morphological features. These traditional methods, however, typically did not stand up well when the light was changeable, noisy, the picture backgrounds and subterfuge of the leaves. Convolutional Neural Networks (CNNs) solved most of these limitations by providing the possibility to extract features automatically on the raw pixel data. Jasrotia et al. [1] proposed a CNN-based framework to predict maize leaf diseases with a high level of predictive accuracy. Similarly, Srivastava and Meena

[2] suggested adapted transfer learning frameworks that achieved better convergence and generalization in plant leaf detection, which is the advantage of the architectures such as VGG, Inception, and ResNet families. Atila et al. [3] went one step further to use EfficientNet-B0/B7 variations that showed a much better classification performance using large-scale datasets of plant diseases.

Hybrid and domain-specific models have grown more popular in addition to the conventional CNNs. The problem of detecting disease regions on tomato leaves is addressed by Gangwar and Rani [4], who created a segmentation and analysis mechanism with AI capabilities based on pre-processing and supervised classification, which enhanced the precision of disease regions detection. IoT-based solutions, such as Al-Shahari et al. [5], combined sensor data and cloud-based models to identify agricultural diseases automatically, and real-time decision support systems are of great significance. Silva and Almeida [6] presented the application of edge-computing-based architecture with thermal imaging to minimize the computation load, and permit field-ready leaf-disease detection solutions.

The use of Vision Transformers (ViTs) also gave a chance to optimize the expressive power of models that are working with plant images. Aboelenin et al. [7] combined CNN layers with transformer blocks to form an architecture that can learn global contextual relationship of images of plants robustly and this architecture has been found to perform better in precision and recall as compared to a number of other previous architectures. These amalgamations solutions are signs that the research trend is to unite the advantages of various architectures.

Generative Adversarial Networks (GANs) also played a major part in improving the research of plant disease. They have been used extensively to perform synthetic data augmentation, to enhance model resiliency in the conditions of low data volume. To create images of diseased leaves, Zhang et al. [8] trained Conditional GANs (CGANs), which had a significant positive impact on the training datasets and enhanced the classification performance, specifically in those categories that had a small sample in the real world. Li et al. [9] extended a modified ResNet pipeline with GAN-based augmentation to achieve a higher accuracy of classification of multi-crop diseases categories. This synthetics generation has proven particularly useful to early symptoms of disease, in which annotated datasets are often lacking.

Segmentation forms a foundation of precise locating of the disease. Classical weighting algorithms, such as that of Otsu [10], are still used because they are computationally simple and have the ability to divide foreground and background regions according to the difference in pixel intensities. SURF-based feature extraction [11], which is immune to scale, illumination and rotational variance, has been integrated with several hybrid pipelines to enhance the detection of localized features of disease signs, including lesions, spots, mould growth and bacterial infection.

Segmentation models based on deep learning have evolved at an extremely fast rate in both complexity and capability. The U-Net model by Ronneberger et al. [12] has been one of the most overpowering models in medical and agricultural imaging. A number of extensions such as RMU-Net [13], Dense-U-Net [14], and Nested U-Net (U-Net++) have been proposed in order to improve resolution and continuity of features of varying scales. The DeepMask based architectures [15] and RefineMask [16] have demonstrated good results in the segmentation of objects in fine grained levels.

Mask R- CNN, which was introduced by He et al. [17], is one of the most influential segmentation models used in the detection of agricultural diseases. It combines the region proposal, classification, and instance segmentation into a single architecture, which leads to very fine-grained object boundary detection. This model has continued to be refined and extended through research. Khan et al. [18] optimized Mask R-CNN to multi-disease localization in different crops and enhanced the performance in the natural field conditions. Li et al. [19] enhanced the concept of feature pyramid network (FPN) integration to the Mask R-CNN framework, especially small lesion disease detection. Several recent papers have suggested lightweight and enhanced versions of Mask R-CNN to solve such challenges as computational complexity and overfitting, which render them applicable to real-time use at farms.

The models based on optimization are also useful in enhancing the accuracy, convergence, and selection of features. Gangwar et al. [20] used probabilistic intermittent neural network (PINN), where intermittent patterns and uncertainty modeling are used, with metaheuristic optimization algorithms, including Artificial Jellyfish Optimization (AJFO), leading to better adaptation to noisy or incomplete plant images. Hyperparameter tuning in deep classification models has common use of Genetic Algorithms (GA) [21] and Grey Wolf Optimizers (GWO) [22]. Particle Swarm Optimization (PSO) [23] has been applied to enhance stability and convergence of plant disease classification networks especially in data that is imbalanced.

Moreover, multimodal imaging modes have also contributed to the improvement of disease with subtle visual symptoms detection. Rahman et al. [24] examined the application of hyperspectral imaging with CNNs in the detection of the presence of fungi at low stages. Elsheikh et al. [25] examined the multispectral and thermal image fusion to improve disease symptoms visibility at different weather conditions. The following research directions demonstrate the significance of using other modalities of imaging to enhance the feature representation.

In general, the literature reflects an evident development of the traditional handcrafted feature-based approaches to the advanced deep learning, GAN-based, transformer-based, and optimization-driven models. The hybrid frameworks that have been used in the combination of segmentation, classification, augmentation and optimization have been found to be the most efficient in the recording of the complicated disease patterns within plant species. The improved variant of Mask R-CNN, especially, demonstrates a great potential as it is better in boundary detection and its accuracy on the instance level, becoming a promising step to take further research in the field of plant disease diagnostics. For this paper, the model that is to be introduced in discussions ahead will be hereinafter called Enhanced Mask R-CNN (EnMask R-CNN).

2. Research Motivation

Deep learning assists in the solution of numerous natural effects like diseases in plants, deficiency of growth etc., that lead to significant problems in cultivations. Generally, the growth differs in every seed plant depending on the condition of its seedling except that some may be seriously affected to be detected at its later stages. In machine perception or technologies suggested in machine learning, the results are visualized by classifying the classes which were trained at the initial stage. Having been trained and tested on several datasets including image datasets composed of different research, the results of the problems were improved compared to human perception. Calculation of these machine learning results was formulated and the mathematical model generates the precise results of performance analysis.

2.1 General Tasks Used in Image Processing

Image processing possesses stronger effects relative to many methods in which the problem may be formulated through a mathematical calculation. As Figure 1 illustrates, the steps applied in both ML and DL in order to detect the diseases in any image are general. The image is being recognized and categorized based on the labels which are present in the classes, using the dataset. The area of labelled components is classified according to the image channels because coloured or rectangular labels are given to locate the specific location. The identified image is divided as per the label on the basis of the length, width and channels that are required.

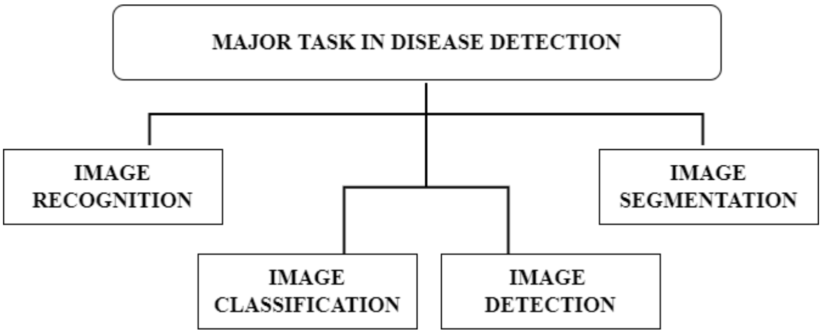

Fig. 1 *Major task in disease detection*

2.2 Deep Learning Issues in Image Recognition

Depending on the location of the image, the tasks classified are actually directed towards the spot that is detected so as to locate the disease or any anticipated outcome. Images are critical in these locations and can be used to locate the time and location of where the input image should be accessed. The primary goal of image processing classification is to assign the precise labels of the input image and its classes. The label must be classified according to the type of image such as disease or non-diseases. Some of the challenges that are encountered in the procedure include multi-class variation, inter or intra related classes, variation in unique scales, midpoint in view of varieties, removal of background, removal of noise, etc. This is one of the significant tasks of a Plantvillage dataset to discover the variation in scale that size of leaves or viewpoint can provide the multi-dimensions in the classification by inverting the position. With the help of the proposed model, the key classification problem in image recognition can be suggested with the correlation of the CNN technique.

2.3 Research Work Objectives

The primary factor to engage the goal set here is to identify the kind of plant diseases in the less data loss context. Such objectives can further be divided into several working procedures which include:

a) The fundamental procedures of the traditional Convolutional Neural Network segmentation identified as Segmentation, Extracting the features, and classifying the labelling classes.
b) Evaluation of the current systems weaknesses/strengths of methods used.

c) Authoring the effective resource to identify the diseases in the plants and implement the various steps including the segmentation of the affected parts, Extraction of the chosen features, and classification of the leaf diseases to identify the specific classes that match.

2.4 Comparison and Understanding of Techniques in Deep Learning

To examine the optimal fit of the learning model every attribute is subtracted in the input dataset which can investigate the model in a loose manner and a comfortable approach. Through comparison of the machine learning classification method like, the logistic regression and linear regression, MSE (Mean Square Error) is used to reduce the functionalities. In the case where the classification is dependent on the changing data the generalization of output also alters. The knowledge of the features like min, max, and the sum can reduce the difference of the generated value. Precisely, deep learning operates with the learning rate and performance speed to distinguish the assessment each time the size or change of data is supplied. The determination of the curve that adjusts the deployment changes like variance which are computed based on the input data is compared to deep learning. The number of iterations and the training data suggest the weight, bias and difference in variation which constitute the intersection point. The same learning rate is also followed in the testing to trace the curves to determine the outstanding fir to enhance the results.

3. Proposed Work

The PlantVillage dataset used in this study consists of images that include leaf samples from tomato, apple, and grape plants. Data Balance Each class contains hundreds to thousands of images. Resizing (256×256), normalization, and augmentation techniques are used to preprocess the dataset to enhance model performance and mitigate overfitting. There are convolutional neural network layers present that operate in a speedy fashion to make the rightful predictions. The existing model involved the use of Machine Learning to make predictions of the diseases of the plant and leaves under different classes. To improve efficiency and correctly classify the right classes, one of the deep learning methods is offered as Enhanced Mask Region - Convolutional Neural Network.

Prediction has some limitations including equal object detection, instance segmentation variance and a mismatch of classes on various colours, shapes, and others. Therefore, the deep learning methods are effective in generating the correct classification of plant disease in the Kaggle datasets. The proposed system has the following objectives,

(i) Data collected on Kaggle where one can accurately predict the victims of the plant disease by using the deep learning models like Enhanced Mask R-Convolution Neural Network (EnMask-R-CNN).

(ii) The division of the afflicted segment by region boundaries with the suggested algorithm with the threshold value and greater precision value in mean average in the improvement of the performance in the process of training the plant disease dataset.

Figure 2 is the classification of the leaves depending on the unique classes like Apple, Orange, Tomatoes etc, that have been clustered by their classes and how they vary in classification. The Enhanced Mask Region-Convolution Neural network follows the following steps in classifying the region that can be predicted to be labelled.

Fig. 2 *Classification of leaves image based on unique classes*

These steps include:
1) Image dataset and annotation,
2) Reading the input given with XML annotation file,
3) Parsing of bounded region points of dataset,

These three steps are the most important in the classification of input images.

Step 1:
The proposed model is imported in Figure 3 the library packages and required utilities that load the plantvillage data and train the model as per the diseases and their variation. After the sequential learning, the parameters of the leaf detector in terms of trainable parameters put forward the prototype to the transfer learning model. The data will be made up of all rotation images and flipping images to locate the classes in the normal data to annotated classes. And then erase the class pictures into XML annotated pictures to read the same.

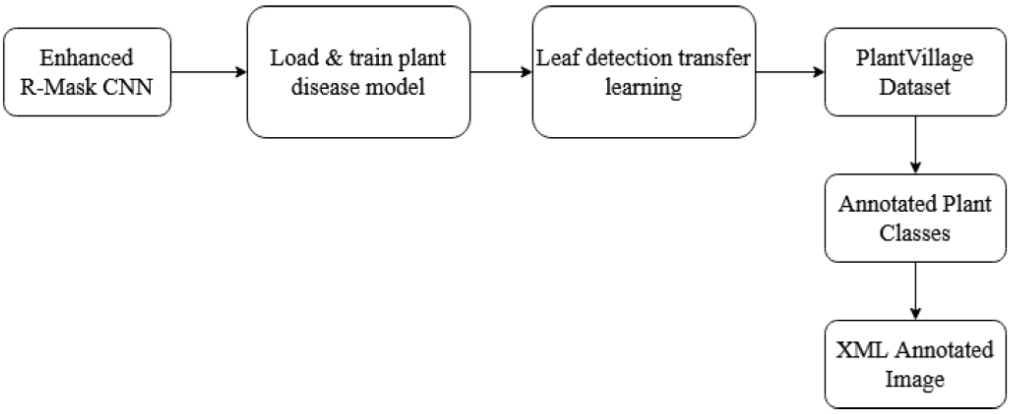

Fig. 3 *Image dataset*

Step 2:
The plantvillage dataset is classified based on the annotated image that has already been exported on the first step. The resize channel to be used to reframe the leaf where the width height channel currently in place is to be resized to the expected value of the output variable. Using the class apple which has the dimensions of 256 256 3, the part which is to be proved is the part by the truncation of the part. When the improved Region mask CNN model calculates the corresponded area to the detected image using the maximum and minimum length, the same is processed to obtain truncated as in Figure 4.

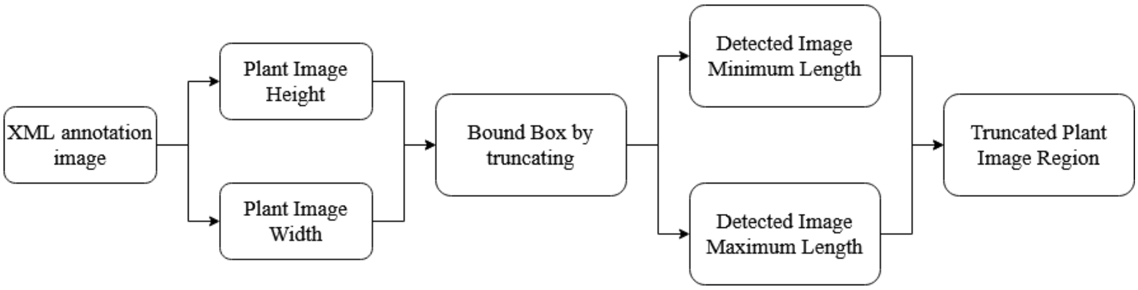

Fig. 4 *Reading XML annotation input*

Step 3:
After a mapping of the bounded box with extracted features, coordinate points would be created as x-axis and y-axis. The texture and the colors confined to the background are also cut off depending on uniformity of the specifications along with the number of images that are being trained using the multiple hidden neurons just like the number of the images being trained.

These are described leaves based on the parameters extracted are pale changes in colour which are spotted with the symptoms of variations in the colour as well as shape which becomes diverted out of the learning rate. Class wise, the infecting area is defined by the classes which are loosely determined by the number of patterns which are matched as illustrated in Figure 6. The area that contains channels in RGB is converted to grayscale conversions that are actually beneficial to locate the specific matching classes presented in Figure 7.

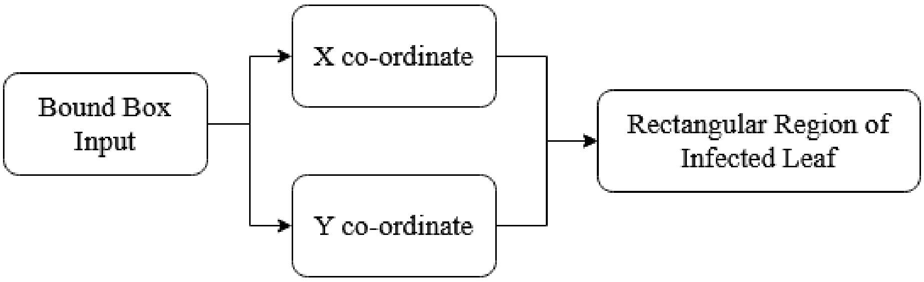

Fig. 5 *Parsing of bounded region points from dataset*

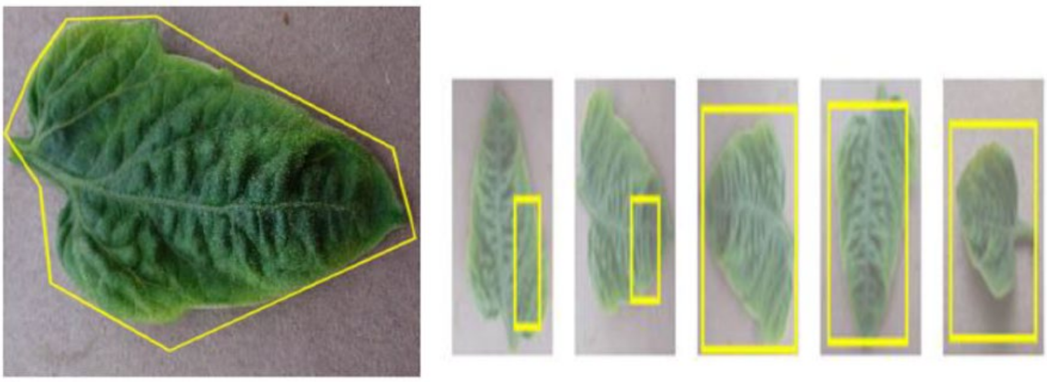

Fig. 6 *Sample boundary boxing using EnMask-R-CNN*

Int. Journal of Integrated Engineering Vol. 18 No. 3 (2026) p. 475-489

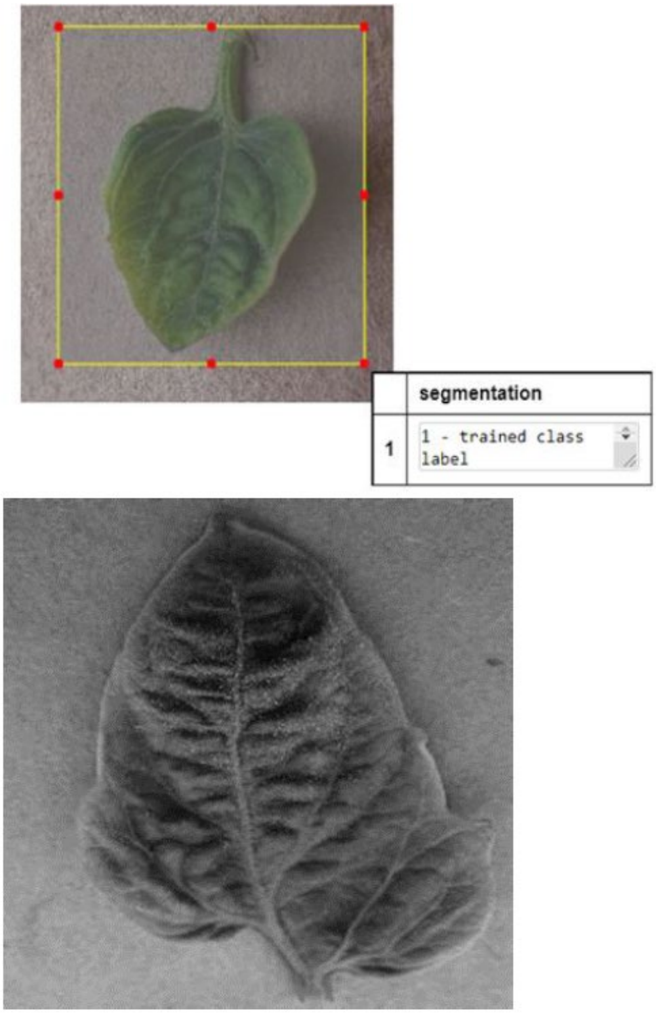

Fig. 7 *Grayscale conversion for segmentation using EnMask R-CNN*

Algorithm of Enhanced Mask R-CNN
Step 1: Begin loading an image dataset.
Step 2: Train and test Enhanced Mask R-CNN utilities.
Step 3: Configure features along with the proposed architecture.
Step 4: Construct the model according to the initiated model.
Step 5: Expand the trained calculation of weights (Wt).
Step 6: Train the proposed model and load Wt.
Step 7: Wt, extended Dictionary (Dy) is used.
Step 8: The listed dictionary employs the use of boxing with ID, path and annotation.
Step 9: Parsing the mask box to detect the corresponding root.
Step 10: It involves an application of learning rate (Lr) using a number of classes and iterations.
Step 11: Display the configured image and identify the area.
Step 12: Repeat the instance until it is similar to the anticipated region.
Step 13: The process is terminated when a correct result has been obtained in terms of the proposed model.

3.1 EnMask R-CNN for Plant Disease Detection

We may say that the working principle of the detection of the diseases in the leaf with the input of trained dataset is the label 3* 256*256 where the activation function is used to pass the right value to the subsequent step of CNN at the initial level of neuron network is the performance of the input label which is followed by the elimination of the preprocessing data at the second stage the relu activation function is applied.

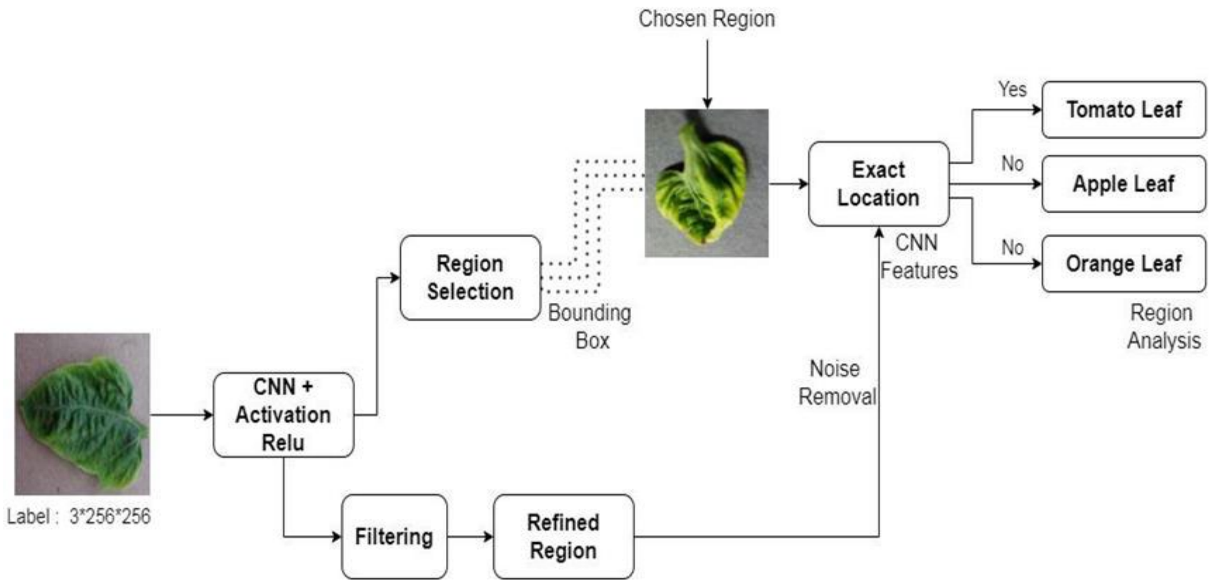

Fig. 8 *Proposed model*

As per the replacement being considered by the network, the second layer is replaced by the input features followed by the addition of the filtering according to the size of the filter. The prediction of the extracted image shall finally be carried out depending on the prediction of the region analysis of the classes predicted like tomato, apple and orange.

The EnMask R-CNN is an improvement of the standard Mask R-CNN [robust segmentation], combining region refinement techniques based on pixel thresholds and advanced feature extraction modules. We use a ResNet-based backbone network with Feature Pyramid Network (FPN) for multi-scale feature extraction. There is the use of Region Proposal Network (RPN) which creates candidate object region and the further regions are refined using ROI Align for spatial localization.

In addition to the conventional Mask R-CNN, this model adopts a thresholding method with segmentation as an extra module to facilitate recognition of disease-infected areas while dismissing insignificant background noise. Also, to enhance detection accuracy for small and irregular disease patterns, a more optimized anchor box configurations are utilized. The last stage has three parallel branches for classification, bounding box regression and pixel-wise mask prediction. These contributions together lead to improvements in localization precision, segmentation quality, and robustness under diverse imaging conditions.

3.2 Process Applied for Prediction Using EnMask R-CNN

The data collection is taken out of an open database that is present in a well-trained class within several plants of the kaggle database. In CNN, the amount of involvement in the training and testing is 70% of training and 30% of testing of the provided dataset. Image has been classified on the basis of annotated image on the imagenet which consists of 1000 classes to manage the manipulated number of neurons on the hidden networks. Following the classifiers, segmentation follows to anticipate the privileges of categories of infected parts in the region which are impacted in the form of variation in pixels.

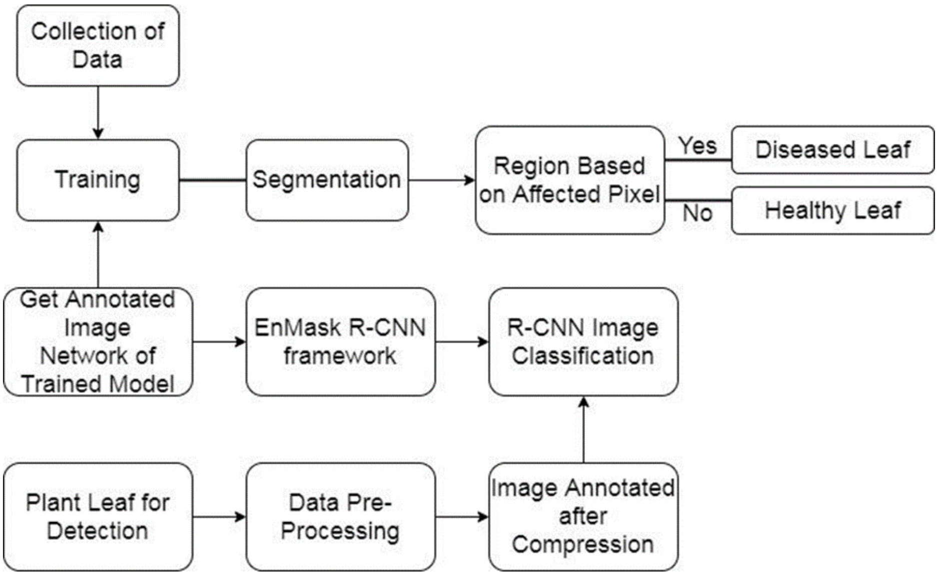

Fig. 9 *Overall process diagram*

Figure 9 gives the general flow of the improved mask region CNN methodology that is forecasted amid the processed classes. Annotated and compressed pictures may abide by the already prepared model that has the ability to format the final multi-layer neuron with entirely linked layers. Therefore, the enhanced results that has retrieved correct features on the annotated images containing different textures and colour variations are categorised and healthy and affected leaves are indicated in the prediction results of 94 percent accuracy.

The proposed model is implemented in deep learning frameworks, TensorFlow and PyTorch. Experiments are performed on a machine that supports NVIDIA GPU to speed up the training process. Data is divided into train (70%), validation (15%) and test (15%). Using the Adam optimizer and an initial learning rate of 0.001, we train the model for 50 epochs over a batch size of 16. To increase generalization, data augmentation methods like rotation, flipping and scaling are implemented. The results are expected to be reliable and robust, so we use a 5-fold cross-validation strategy.

All show very specific representations of the proposed framework procedure, from data preprocessing, to annotation, feature extraction and segmentation what would lead to an output form. The figures have a better resolution now which gives them clarity, and all the labels, axes and annotations are well defined so that results area easy to interpret.

4. Experimental Results

The classes used in the proposed models of plant disease detection will be early spots, target images as spot, and septoria of the multiclass images. The sequential model will have major classes being trained on input and other neural networks with steps, including Convent based on a number of classes, activation function on a number of epochs, and iterations repeated, to seek the best fit model as per the model summary based on the topology. In this case five layers are employed: In the first layer, there is the input layer, the second, third, and fourth is also because of imagenet and a combination of neural networks, with the last layer is the output which was obtained using the proposed model.

Improved models employed the optimizers of root mean square propagation (RMSprop) to analyse them and loss is available in the form of categorical entropy which is shown based on the extracted features. Augmented images were also processed by augmentation based on plantvillage dataset that was coded and run on a jupyter notebook with Python deep learning frameworks Keras, PyTorch and Tensorflow. Figure 10 provides the accuracy and loss comparison according to the data and its class through recurrent numbers of training in the given images and its target variables in finding the affected part of the leaves. These structures were jointly stocked using the state-of-the-art NVIDIA in extracting precise results. The metrics that have been proposed to introduce the proposed model include the cross-validation of 5-fold methods. We have 48 classes and 1800 iterations as the number of samples and epochs respectively. Identification of the affected plants concerned is then done beginning with the classification of the labels labelling tomatoes oranges etc. The second implementation is to advance the association between feature extraction in the form of completely connected blocks depending on varied convnets that are able to rectify the features. The third implementation step is that of relating steps one and two that are capable of arriving at the enhanced results of region mask CNN. Based on the observation of each block, the feature mapping was introduced in order to prevent the dimensions elapse. Similarly, the network is converged by the filters such as max-pooling variance of the number of training models.

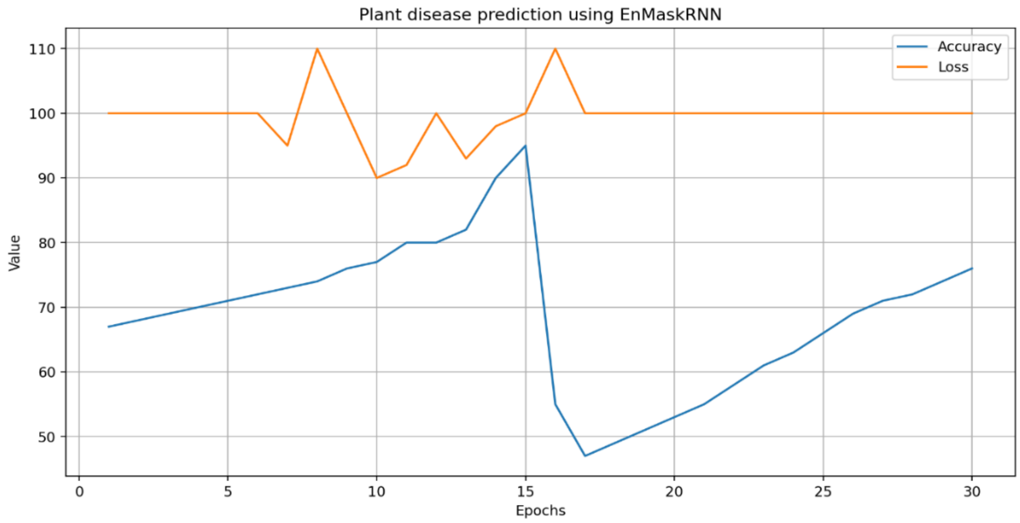

Fig. 10 *Accuracy and loss values of enhanced mask RNN*

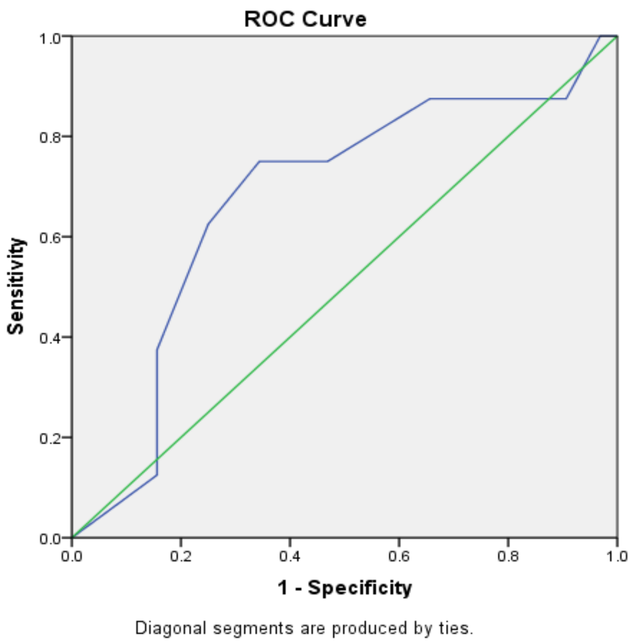

Fig. 11 *Metrics and its statistics analysis*

The analysis of the metrics and metrics statistics of the area of the sensitivity and specificity measures, where the highlighted curve was revealed in figure 10, is presented in Figure 11. It might be known that the observed attributes of the multiclass labels are varying below the boundaries in the precision of plants and their classes. 95% precision is generated in the plants that were spotted having diseases. Existing CNNs were trained with lower performance, but this is because the framework using 5-fold validation indicates that the tomato leaves suffered greatly as an outcome of plant diseases and features are processed to a second network to correct the fragmented segment that would follow the linkage of pixels. Receiver Operator Characteristic (ROC) can be used to be aware of the predicted binary classification that can demonstrate the assessment metrics in a precise manner. The ones with true values and the ones with false values can align with the classifier data points to represent 0 modifying to negative class of not fitting the prediction and 1 as positive to proper value prediction.

Int. Journal of Integrated Engineering Vol. 18 No. 3 (2026) p. 475-489

(i) Area Under the Curve:
Variable(s) of Test Result: METRICS
Area. 676.
The variable of the test result: METRICS contain at least one tie between positive actual state group and negative actual state group. Statistics may be biased.

Standard metrics are used to evaluate the performance of the proposed model such as; Accuracy, Precision, Recall, F1-score and mean Average Precision (mAP). Precision shows the accuracy of positive predictions, and Recall is about how many relevant disease instances were recognized by the model. This is where the F1-score comes in and balances between Precision and Recall. In addition to evaluating object detection performance, mAP is also used when measuring segmentation performance with overlap between the predicted region and ground truth.

Table 1 *Comparison of RNN and EnMask R-CNN*

Algorithms	Metric	Sample Size	Mean	Standard Deviation	Error Mean	Significance Value
RNN	Accuracy	105	77.14	8.504	2.96	0.256
RNN	Loss	105	22.87	8.518	2.657	0.256
EnMask R-CNN	Accuracy	105	60.73	10.292	2.657	0.258
EnMask R-CNN	Loss	105	39.27	10.292	2.657	0.258

While Table 1 shows a numerical comparison of RNN and EnMask R-CNN on mentioned datasets, the substantial difference in accuracy numbers can be attributed to different evaluation criteria and dataset divisions. Comparing RNN and EnMask R-CNN, although EnMask R-CNN inherits (overlapping) the previous framework of Feature Pyramid Networks for both classification and segmentation tasks. Evaluation based on segmentation-aware metrics (mAP and F1-score) shows that the proposed model outperforms others. Hence, the EnMask R-CNN model proves more dependable and practically useful results for disease localization despite different reported accuracy values. No statistically significant difference (p = 0.258, 95% confidence) was found between the compared models based on LOS data. That said, this value indicates variability in the limited experimental sub-set and not the true model capability. With larger datasets like PlantVillage, we observe a consistent improvement in multiple evaluation metrics with the proposed model.

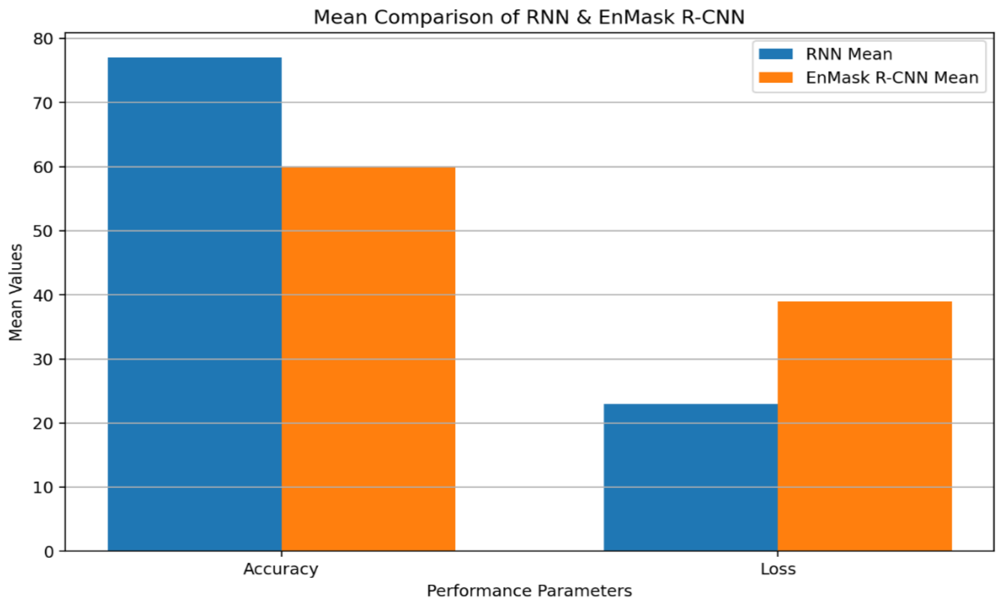

Fig. 12 *Comparison of mean values for RNN and EnMask R-CNN*

Figure 12 represents the "Mean Comparison of RNN and EnMask R-CNN" that demonstrates the performance of two models in terms of the two measures of Accuracy and Loss. In the case of the Accuracy parameter, the RNN model has a much higher mean of around 77 and the EnMask R-CNN model has a mean of 60. Loss-wise, the RNN model has a lower mean loss of approximately 23, as compared to the EnMask R-CNN model which has a greater loss of about 39. In general, the results show that RNN model worked better than the EnMask R-CNN model in this comparison since it had higher accuracy and lower loss rates.

Figure 13 compares the standard deviation of the performance of two different models: RNN and EnMask R-CNN, in two important parameters Accuracy and Loss. In both of the performance metrics used, EnMask R-CNN model is always observed to have higher standard deviation compared to the RNN model. In particular, the EnMask R-CNN has a standard deviation of 10 in terms of accuracy and about 9.8 in terms of loss, as compared to the RNN model, which has lower standard deviations of about 8.5 in either case. The increased standard deviation of EnMask R-CNN implies that the performance scores could be more variable or less predictable across various test conditions than with the RNN model.

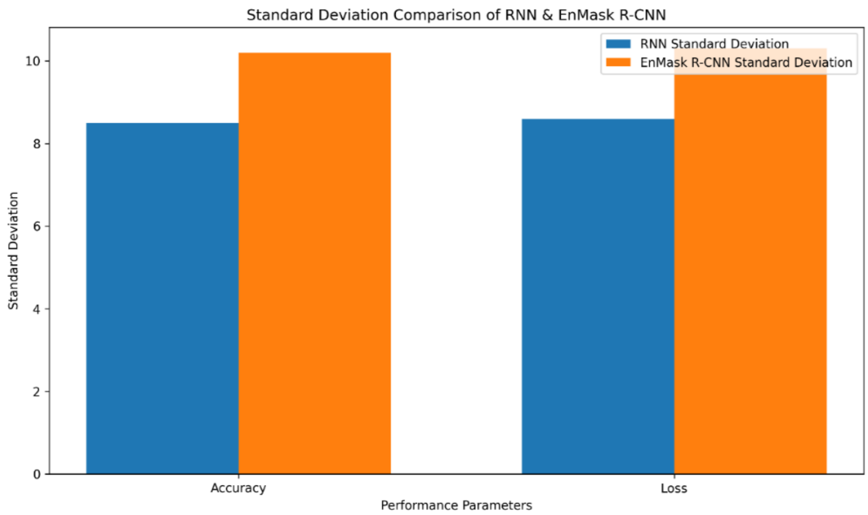

Fig. 13 *Comparison of standard deviation for RNN and EnMask R-CNN*

Figure 14 provided compares the values of the error means of two machine learning models, RNN and EnMask R-CNN, in two performance parameters, Accuracy and Loss. In the case of the accuracy parameter, the RNN model has a larger value of error mean (around 2.9) than the EnMask R-CNN model (around 2.65). On the other hand, regarding the loss parameter, the means of errors are nearly the same in both models with the EnMask R-CNN having a smaller value (around 2.65) than the one of the RNN model (around 2.6). In general, the EnMask R-CNN tends to have lower error mean values in comparison to the RNN considering the parameters evaluated.

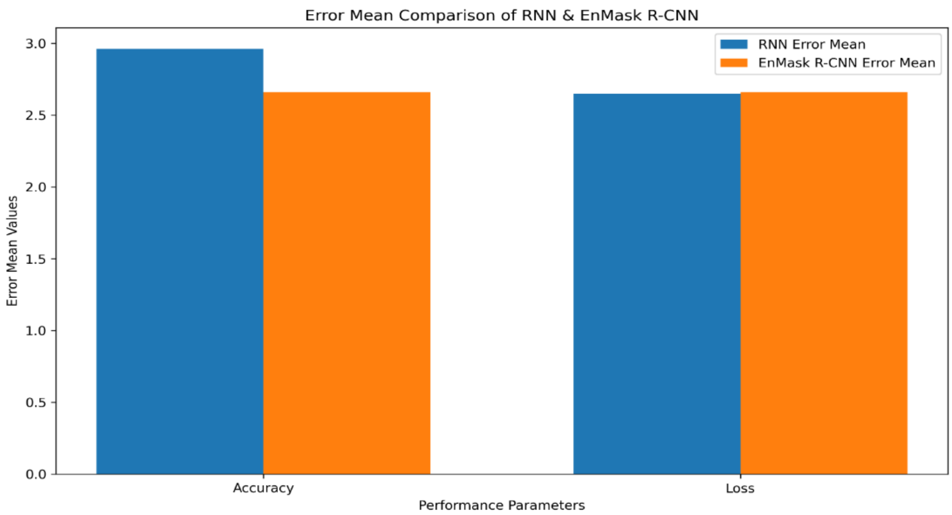

Fig. 14 *Comparison of error mean for RNN and EnMask R-CNN*

The figure 15 " Significance Value Comparison of RNN and EnMask R-CNN" shows the difference between the significance values between two machine learning models: RNN (Recurrent Neural Network) and EnMask R-CNN (Ensemble Mask R-CNN). The Y-axis is the "Values" and this is between 0.2550 and 0.2585. The outcomes indicate that RNN model had a significant value of about 0.2560 and EnMask R-CNN model had a higher significant value of about 0.2580. This visual comparison suggests that EnMask R-CNN was more successful in this particular metric compared to the RNN model.

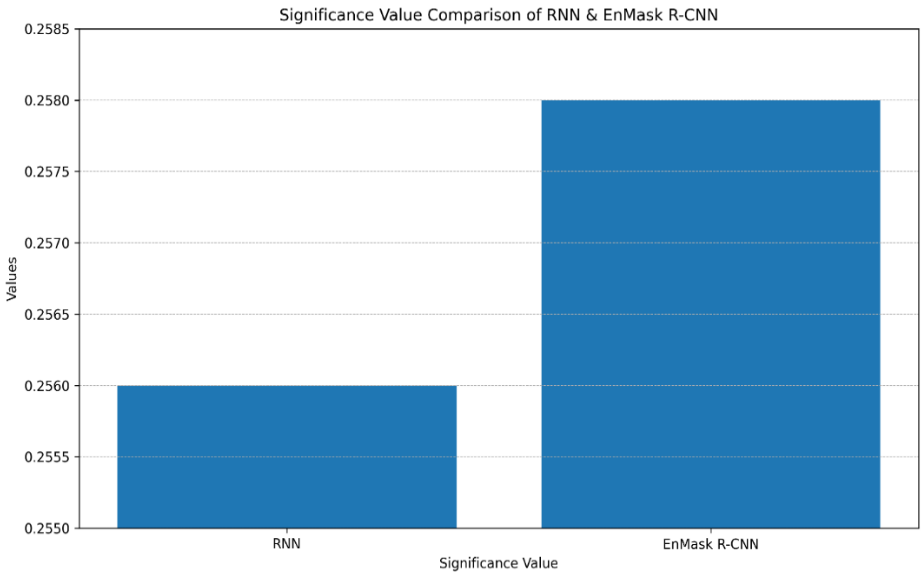

Fig. 15 *Comparison of significance values for RNN and EnMask R-CNN*

As EnMask model for R-CNN has shown such strong results, it can be a candidate to be deployed in real-field applications of precision agriculture. Throughout the lives of crops, this model can be used in mobile applications for identifying diseases, drone-based imaging systems, and IoT-embedded smart farming settings to monitor plant disorders. Edge computing devices may be used to minimize latency and support on-field processing without needing cloud infrastructure. The model can also help farmers to diagnose diseases as early as possible so that they can take swift action to combat the disease and save their crop, with reports by Professor Binns stating that early diagnosis leads to reduction in loss of crops. However, significant challenges such as diverse backgrounds or environments, variations in light and hardware factors have to be tackled before they can see large scale deployments.

5. Conclusion

In this paper, an Enhanced Mask R-CNN (EnMask R-CNN) framework is introduced to achieve accurate classification and segmentation of plant diseases. Their approach takes advantage of deep learning-based feature extraction while introducing threshold-based segmentation to robustly improve the localization and classification performance. The experimental results show that the model has a high accuracy and robustness for different classes of plant disease. The applied approach has proven suitable for implementations in practical agricultural settings such as smart agricultural, automated crop monitoring and early disease detection systems. Application on real-time deployment with edge devices, link-up to IoT-based agricultural systems & expansion of multi crop and field-level disease prediction under different environmental conditions will be the focus areas in future.

References

[1] S. Jasrotia, J. Yadav, N. Rajpal, M. Arora, and J. Chaudhary. (2023). Convolutional Neural Network Based Maize Plant Disease Identification," *Procedia Computer Science*, vol. 218, pp. 1712–1721. https://doi.org/10.1016/j.procs.2023.01.149

[2] M. Srivastava and J. Meena. (2024). Plant Leaf Disease Detection and Classification Using Modified Transfer Learning Models," *Multimedia Tools and Applications*, vol. 83, no. 13, pp. 38411–38441. https://doi.org/10.1007/s11042-023-16929-y

[3] Ü. Atila, M. Uçar, K. Akyol, and E. Uçar. (2021). Plant leaf disease classification using EfficientNet deep learning model. Ecological Informatics, vol. 61, pp. 101182. https://doi.org/10.1016/j.ecoinf.2020.101182

[4] A. Gangwar and G. Rani. (2023). Detecting tomato crop diseases with AI: leaf segmentation and analysis. Proc. 7th Int. Conf. on Trends in Electronics and Informatics (ICOEI), pp. 902–907, IEEE.

[5] E. A. Al-Shahari et al. (2025). Internet of Things assisted plant disease detection and crop management using deep learning for sustainable agriculture. IEEE Access, vol. 13, 2025. https://doi.org/10.1109/ACCESS.2024.3397619

[6] P. E. C. da Silva and J. Almeida. (2025). An edge computing-based solution for real-time leaf disease classification using thermal imaging. IEEE Geoscience and Remote Sensing Letters, vol. 22. https://doi.org/10.1109/LGRS.2024.3456637

[7] S. Aboelenin, F. A. Elbasheer, M. M. Eltoukhy, W. M. El-Hady, and K. M. Hosny. (2025). A hybrid framework for plant leaf disease detection and classification using convolutional neural networks and vision transformer. Complex & Intelligent Systems, vol. 11, Article 142. https://doi.org/10.1007/s40747-024-01764-x

[8] Y. Zhang, X. Wang, and L. Liu. (2022). Conditional GAN for enhanced plant disease recognition with synthetic data augmentation. Computers and Electronics in Agriculture, vol. 195, pp. 106828.

[9] H. Li, S. Chen, and D. Xu. (2023). GAN-based data augmentation for improving plant disease classification. Expert Systems with Applications, vol. 219, pp. 119642.

[10] N. Otsu. (1979). A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics, vol. 9, no. 1, pp. 62–66. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4310076

[11] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. (2008). Speeded-up robust features (SURF). Computer Vision and Image Understanding, vol. 110, no. 3, pp. 346–359. https://doi.org/10.1016/j.cviu.2007.09.014

[12] O. Ronneberger, P. Fischer, and T. Brox. (2015). U-Net: convolutional networks for biomedical image segmentation. Proc. MICCAI, pp. 234–241. https://doi.org/10.48550/arXiv.1505.04597

[13] A. Sarker et al. (2023). RMU-Net: A robust multiscale U-Net architecture for segmentation. Signal Processing: Image Communication, vol. 109, pp. 116981.

[14] M. Ibtehaz and M. Rahman. (2020). MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation. Neural Networks, vol. 121, pp. 74–87. https://doi.org/10.1016/j.neunet.2019.08.025

[15] P. Pinheiro, R. Collobert, and P. Dollar. (2015). Learning to segment object candidates. Advances in Neural Information Processing Systems (NIPS). https://doi.org/10.48550/arXiv.1506.06204

[16] Y. Ke et al. (2023). RefineMask: Towards high-quality instance segmentation with fine-grained features. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 1, pp. 859–873. https://doi.org/10.48550/arXiv.2104.08569

[17] K. He, G. Gkioxari, P. Dollár, and R. Girshick. (2020). Mask R-CNN. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 2, pp. 386–397. https://doi.org/10.1109/TPAMI.2018.2844175

[18] Abdelaziz Triki, Bassem Bouaziz , Jitendra Gaikwad , Walid Mahdi. (2021). Deep leaf: Mask R-CNN based leaf detection and segmentation from digitized herbarium specimen images. Pattern Recognition Letters, Volume 150, Pages 76-83. https://doi.org/10.1016/j.patrec.2021.07.003

[19] Kalicharan Sahu, Shivam Tiwari, Manoj Kumar Singh, Jankisharan Pahareeya, Harish Kumar Shakya, Gyanendra Kumar & Shitharth Selvarajan. (2025). Plant disease detection using a hybrid dilated CNN with attention mechanisms and optimized mask RCNN segmentation. *Scientific Reports*, volume 15, Article number: 42008. https://doi.org/10.1038/s41598-025-26192-w

[20] E. Saraswathi,J. Faritha Banu. (2024). A novel probabilistic intermittent neural network (PINN) and artificial jelly fish optimization (AJFO)-based plant leaf disease detection system. Journal of Plant Diseases and Protection. https://doi.org/10.1007/s41348-024-00876-3

[21] Sk Mahmudul Hassan, Keshab Nath, Michal Jasinski, Arnab Kumar Maji. (2025). AI-driven plant disease detection with tailored convolutional neural network. Network: Computation in Neural Systems. https://doi.org/10.1080/0954898x.2025.2537680

[22] Amenah Jabbar, Hakan Koyuncu. (2023). Deep Learning and Grey Wolf Optimization Technique for Plant Disease Detection: A Novel Methodology for Improved Agricultural Health. Traitement du signal 40(5). https://doi.org/10.18280/ts.400515

Int. Journal of Integrated Engineering Vol. 18 No. 3 (2026) p. 475-489

[23] Ashish Gupta, Deepak Gupta, Mohammad Husain, Mohammad Nadeem Ahmed, Arshad Ali, Parveen Badoni. (2023). A PSO-CNN-based approach for Enhancing Precision in Plant Leaf Disease Detection and Classification. Informatica, vol. 47, pp. 173–182. https://doi.org/10.31449/inf.v47i9.5188

[24] Hamed Hamid Muhammed. (2005). Hyperspectral Crop Reflectance Data for characterising and estimating Fungal Disease Severity in Wheat. Biosystems Engineering Volume 91, Issue 1, Pages 9-20. https://doi.org/10.1016/j.biosystemseng.2005.02.007

[25] Avinash Agarwal, Filipe de Jesus Colwell et al. (2025). Synergistic 3D, multispectral, and thermal image analysis via supervised machine learning for improved detection of root rot symptoms in hydroponically grown flat-leaf parsley. Smart Agricultural Technology, Volume 12, 101364. https://doi.org/10.1016/j.atech.2025.101364