International Journal of Innovative Research and Reviews, 9(2) 45-50

International Journal of Innovative Research and Reviews
ISSN: 2636-8919
Website: www.injirr.com

RESEARCH ARTICLE

A Comparative Analysis of Multi-Class and Binary Papaya Leaf Classification Using Optimized Darknet Architectures

Rabiye KILIÇ[1], Hilal Kübra SAĞLAM[2,3]

[1] *Department of Computer Engineering, Faculty of Engineering, Atatürk University, Erzurum, Türkiye*
[2] *Department of Electrical and Electronics Engineering, Faculty of Engineering, Atatürk University, Erzurum, Türkiye*
[3] *Department of Nano-Science and Nano-Engineering, Institute of Science and Technology, Ataturk University, Erzurum, Türkiye*

** Corresponding author E-mail: rabiyekilic@atauni.edu.tr*

ARTICLE INFO

Received : 07.19.2025
Accepted : 11.20.2025
Published : 12.15.2025

Keywords:
Papaya leaf disease
Deep learning
CNN
Darknet-19
Darknet-53
Multi-class classification
Binary classification
Plant disease detection

ABSTRACT

Accurately determining the health status of papaya leaves is of great importance for early detection of diseases and the continuity of sustainable agricultural production. In this study, the performance of two improved Darknet architectures, Darknet-19 and Darknet-53, was evaluated on both the five-class disease recognition problem and the Healthy–Unhealthy binary classification scenario. The used BDPapayaLeaf dataset contains a total of 2,159 high-resolution leaf images belonging to five categories: Anthracnose, Bacterial Spot, Curl, Ring Spot and Healthy. Various training methods such as data augmentation, label smoothing, dropout and cosine-based learning rate planning were applied to increase the generalization ability of the models. Looking at the five-class classification results, while the Darknet-19 model achieved 79.81% accuracy, the Darknet-53 model with a deeper structure reached 84.78% accuracy. This result shows that residual layers provide a significant advantage especially in capturing complex disease patterns. When binary classification is performed, the accuracy rates increased significantly; Darknet-19 achieved 97.20% accuracy, while Darknet-53 achieved 98.14% accuracy. The findings demonstrate that optimized Darknet architectures offer a reliable, effective, and applicable approach to automatically identifying papaya leaf diseases.

Contents

Cite this article Kılıç R, Sağlam HK. A Comparative Analysis of Multi-Class and Binary Papaya Leaf Classification Using Optimized Darknet Architectures. *International Journal of Innovative Research and Reviews (INJIRR)* (2025) 9(2) 45-50

Link to this article: http://www.injirr.com/article/view/248

Copyright © 2025 Authors.
This is an open access article distributed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits unrestricted use, and sharing of this material in any medium, provided the original work is not modified or used for commercial purposes.

1. Introduction

Papaya has medicinal importance due to its high nutritional value and is widely cultivated. This tropical fruit holds significant commercial value in the market. However, its economic advantage is significantly reduced by leaf diseases. Therefore, accurate disease detection methods are crucial. Traditional methods rely on visual expert judgment. However, high human error rates and slow analysis compromise the detection process. Furthermore, if a disease is not detected in its early stages, it misses the target. To overcome these challenges, researchers have focused on recent advances in deep learning (DL). From image analysis to smart farming, deep learning has demonstrated impressive results. Numerous deep learning studies have been conducted for papaya disease detection [1]. It has been demonstrated that computer-aided disease diagnosis and classification using machine learning (ML) models can improve papaya production to increase papaya quality and yield [2]. Currently, systems using deep learning methods can achieve low cost and high accuracy [3].

However, this approach requires a large number of samples to achieve successful results, making it challenging. Because there are few annotated datasets in the literature, development in this area has been low. It is believed that the existing datasets need to be expanded to accommodate computer vision techniques [4]. With the rise of big data technologies, DL technology has created a new platform for expanding the scope of agricultural activities [5]. Furthermore, to implement smart agriculture, farmers need access to decision-making tools and automation improvements to increase their productivity, efficiency, and profitability [6].

Image recognition systems use computer vision algorithms to detect papaya leaf diseases. These algorithms are trained on the database's training set. After training, the image recognition system can classify new images of papaya leaves as healthy or diseased. Therefore, image recognition systems are a valuable tool for papaya farmers who want to protect their crops from disease. Methods such as Convolutional Neural Networks (CNN) often face difficulties distinguishing diseases with similar symptoms. To address these disadvantages, the following are being developed: a smartphone-based diagnostic system, a mobile application using user-generated photos, and comprehensive functionality testing [7]. Food quality can be improved by using network models enhanced with comprehensive databases. Furthermore, disease detection can be tested in autonomous systems, resulting in cost savings. Furthermore, providing mobile detection devices to farmers who lack expert information could significantly alleviate these problems [8]. One study, using Artificial Neural Networks, focused on images of diseases affecting papaya captured using various digital cameras, such as smartphones [9]. Furthermore, machine vision-based recognition could help create an online system that recognizes disease-related defects by analyzing images taken using a mobile handheld device to assist farmers remotely [10].

Accurately capturing the complex patterns and features in papaya leaf images would facilitate disease classification [11]. Such studies are advancing the field of precision agriculture by providing a crop-specific artificial intelligence system for the sustainable monitoring of papaya diseases [12].

DarkNet-19 is a deep convolutional neural network architecture used for object detection, known for its efficiency and speed. While not designed for image classification, it can be adapted for classification tasks by removing layers specific to object detection. It is efficient in terms of memory usage and computational resources. DarkNet-53, like DarkNet-19, is a deep convolutional neural network architecture designed for image classification tasks, offering advantages in efficiency and speed. It is widely used in the computer vision community for tasks such as image recognition, object classification, and highly discriminative feature extraction [13].

In this study, a comparison was made of the disease classification success rates of papaya leaves using the DarkNet-19 and DarkNet-53 models.

2. Methods

In this study, a deep learning-based approach was developed to determine the health status of papaya leaves and examined under two different evaluation scenarios: (i) five-class disease recognition and (ii) binary classification as healthy/unhealthy.

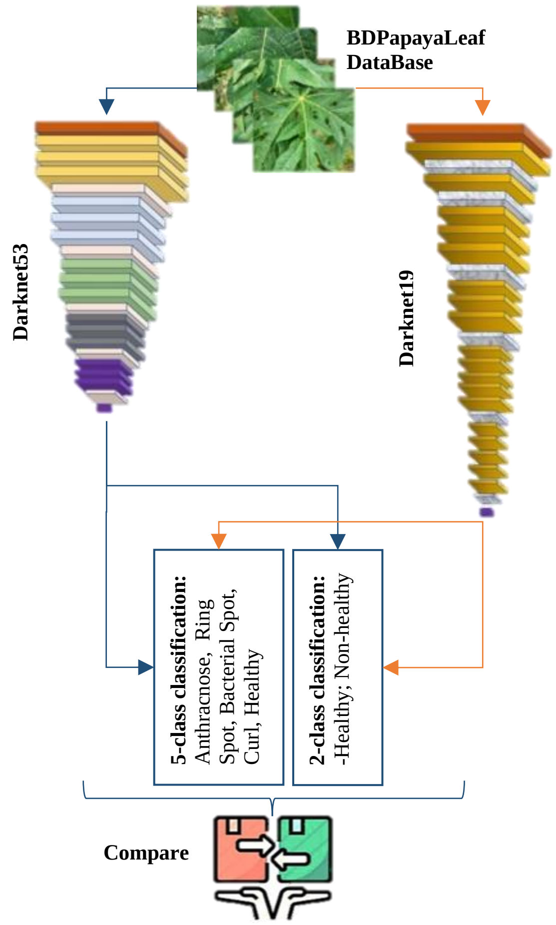

Figure 1 Architectures of proposed model

The methodological process begins with a description of the dataset's characteristics. The class distribution in the dataset, the proportions of leaf images by category, and the random partitioning strategy into training, validation, and test subsets are discussed in detail; this structure forms the basis of the experimental design. Following the dataset introduction, the two convolutional neural network architectures used in this study (Figure 1), Darknet-19 and Darknet-53, are explained in detail, and how these models were adapted and developed for papaya leaf classification is explained. The training process is comprehensively described, emphasizing the use of advanced optimization techniques such as data augmentation, label smoothing, dropout regularization, and cosine-based learning rate planner to improve the model's generalization to limited data. These strategies were integrated into the training process to improve both the model's stability and performance.

2.1. Dataset

The experiments in this study were conducted using the BDPapayaLeaf dataset [2], available in the literature. The dataset consists of a total of 2,159 papaya leaf images and is divided into five main classes: Anthracnose, Bacterial Spot, Curl, Ring Spot, and Healthy. Figure 2 provides examples of the dataset.

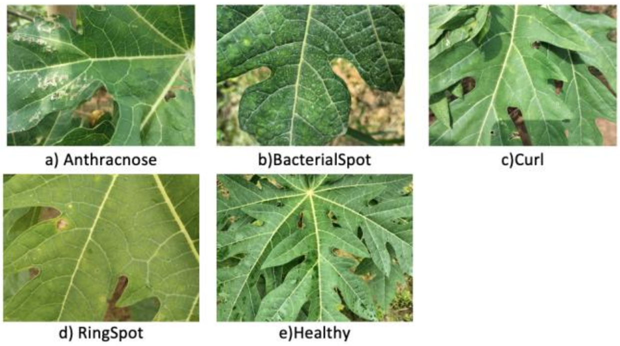

Figure 2 Sample images from the dataset

The images were taken under natural lighting conditions and contain various variations in leaf orientation, background details, disease severity, and texture characteristics. To ensure an objective evaluation, the dataset was randomly divided into training, validation, and test subsets for each class. While the target ratio during the splitting process was approximately 70% training, 15% validation, and 15% test, these ratios fluctuated slightly due to imbalances in the number of class samples. The full numerical distribution of images across the five classes is detailed in Table 1.

Table 1 Class-wise distribution of papaya leaf images

Class Name	Total Images	Train	Validation	Test
Anthracnose	353	248	52	53
Bacterial Spot	458	320	70	68
Curl	585	409	89	87
Healthy	228	159	35	34
Ring Spot	533	373	80	80
Total	2,159	1,509	326	322

2.2. Improved Darknet-19 Architecture

Darknet-19, a lightweight convolutional neural network architecture that formed the basis of the first YOLO models, consists of sequential convolution and max pooling layers. In this study, the model was adapted to the classification of papaya leaves and enhanced with some additional refinements. Specifically, batch normalization layers, the LeakyReLU activation function, and a dropout layer in the final fully connected block were added to reduce overfitting and increase model stability. Furthermore, label smoothing was applied within the loss function to improve decision boundaries between classes. The model processes images through a total of 19 convolutional layers and progressively extracts texture and shape features associated with disease symptoms. Thanks to its relatively limited depth, both training and inference processes are quite fast. In these respects, Darknet-19 offers a suitable and efficient architecture for both binary classification and multi-class leaf disease recognition tasks.

2.3. Improved Darknet-53 Architecture

Darknet-53 is a deeper neural network architecture consisting of 53 convolutional layers, organized into residual blocks. The skip connections in this structure strengthen gradient flow, allowing the model to more effectively learn complex patterns, particularly those seen in papaya leaf diseases, such as vein color changes, tissue distortions, and lesion boundaries. In this study, the Darknet-53 architecture was further enhanced with additional regularization methods. The integration of dropout, label smoothing, and a cosine-based learning rate reduction strategy both increased the stability of the training process and contributed to the model's more robust convergence. Furthermore, depending on the task type, the classification layer was restructured to include two output neurons (binary classification) or five output neurons (multi-class classification). Thanks to the model's deep structure and residual learning capabilities, Darknet-53 achieved the highest accuracy values in both classification scenarios, making it the most successful architecture in the study.

2.4. Evaluation Metrics

To quantitatively assess the performance of the proposed classification models, several standard evaluation metrics commonly used in machine learning and computer vision were employed. These include accuracy, precision, recall (sensitivity), and the F1-score [14]. All metrics were calculated using the confusion matrix obtained during the evaluation phase. TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively.

Accuracy;

Accuracy represents the overall proportion of correctly classified samples and serves as a global performance indicator. It is defined as:

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \tag{1}$$

Precision;

Precision measures the reliability of the model's positive predictions by quantifying the percentage of predicted positive samples that truly belong to the target class:

$$Precision = TP / (TP + FP) \tag{2}$$

Recall (Sensitivity);

Recall evaluates the model's ability to correctly identify all samples that belong to the positive class:

$$Recall = TP / (TP + FN) \quad (3)$$

F1-Score;

The F1-score is the harmonic mean of precision and recall, combining both aspects into a single balanced measure:

$$F1 = 2 * \frac{(Precision * Recall)}{(Precision + Recall)} \quad (4)$$

Multi-Class Evaluation;

For multi-class classification, the metrics above were computed using a one-vs-all strategy, treating each class independently as the positive class. Per-class performance was reported to highlight variations in recognition difficulty across disease categories. This evaluation framework ensures a rigorous and transparent comparison of the Darknet19 and Darknet53 architectures, providing deeper insight into their ability to capture discriminative patterns associated with papaya leaf diseases.

3. Results

This section presents experimental findings obtained using the optimized Darknet-19 and Darknet-53 architectures within the proposed classification framework. The models were evaluated under two different scenarios: 3.1. Five-class papaya leaf disease classification, and 3.2. Binary classification based on healthy/unhealthy discrimination. Accuracy was used as the primary evaluation metric, and both models were evaluated on completely independent and previously unseen test samples.

3.1. Multi-Class Classification Performance

In the five-class papaya leaf disease recognition problem, the proposed method demonstrated strong discrimination ability across both architectures. The Darknet-19 model achieved 79.81% accuracy, demonstrating that despite its limited depth, it was able to adequately capture key texture and color features associated with the disease. In contrast, the Darknet-53 architecture, which has deeper and redundant connections, achieved a higher accuracy of 84.78%. This increase can be attributed to the model's ability to learn finer details and small morphological differences between classes more effectively.

Confusion matrices are shown in Figure 3. An examination of the confusion matrices reveals that Darknet-53 makes fewer misclassifications, particularly in the formally similar Curl and Ring Spot categories. This demonstrates the model's improved ability to resolve complex visual patterns. Training and validation curves also reveal more stable convergence with the use of advanced regularization techniques (data augmentation, label smoothing, dropout, and cosine-based learning rate scheduling). These findings confirm that the applied strategies significantly improve the model's generalization ability.

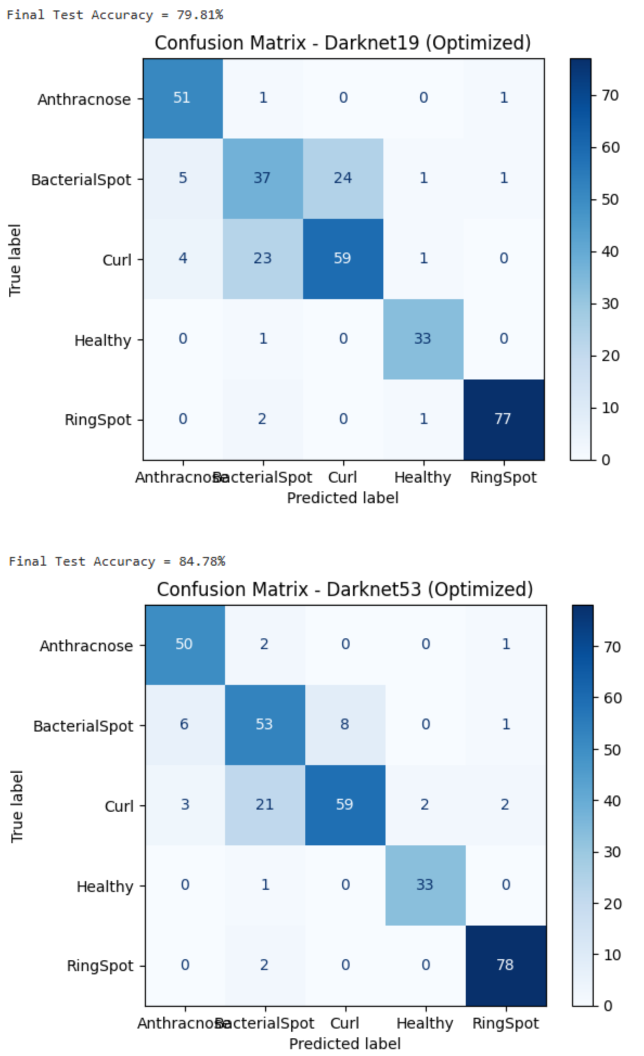

Figure 3 Confusion matrix for five classes

Table 2 summarizes the class-wise performance of the Darknet19 model on the five-category papaya leaf disease classification task.

Table 2 Performance Metrics for Darknet19 (Multi-Class Classification)

Class	Precision	Recall	F1-Score
Anthracnose	0.85	0.98	0.91
Bacterial Spot	0.58	0.54	0.56
Curl	0.71	0.67	0.69
Healthy	0.92	0.97	0.94
RingSpot	0.97	0.96	0.97

The results indicate that the model performs exceptionally well for the Healthy and RingSpot classes, achieving both high precision and recall. Anthracnose and Curl classes also demonstrate moderately strong performance, reflecting the model's ability to capture distinctive patterns in these categories. However, the Bacterial Spot class exhibits comparatively lower scores across all metrics, suggesting significant confusion with visually similar classes—particularly Curl. Overall, the findings highlight that while Darknet19 is effective in distinguishing several leaf diseases types, further optimization or feature enhancements may be required to improve classification accuracy for the more challenging Bacterial Spot category.

Table 3 reports the precision, recall, and F1-score values for each class in the multi-class evaluation of the Darknet53 model.

Table 3 Performance Metrics for Darknet53 (Multi-Class Classification)

Class	**Precision**	**Recall**	**F1-Score**
Anthracnose	0.847	0.943	0.892
Bacterial Spot	0.671	0.779	0.721
Curl	0.880	0.678	0.765
Healthy	0.943	0.971	0.957
RingSpot	0.951	0.975	0.963

The results demonstrate that Darknet53 delivers strong and stable performance across all classes, with particularly high scores for the Healthy and RingSpot categories. These findings indicate that the model is highly effective in learning distinguishing visual patterns for well-defined disease categories. Although the Bacterial Spot class shows relatively lower performance—likely due to its visual similarity to other disease types—the model still achieves moderate precision and recall in that category. Overall, the performance metrics confirm that Darknet53 provides robust classification capabilities and outperforms its shallower counterpart in complex multi-class papaya leaf disease recognition tasks.

3.2. Binary Classification Performance

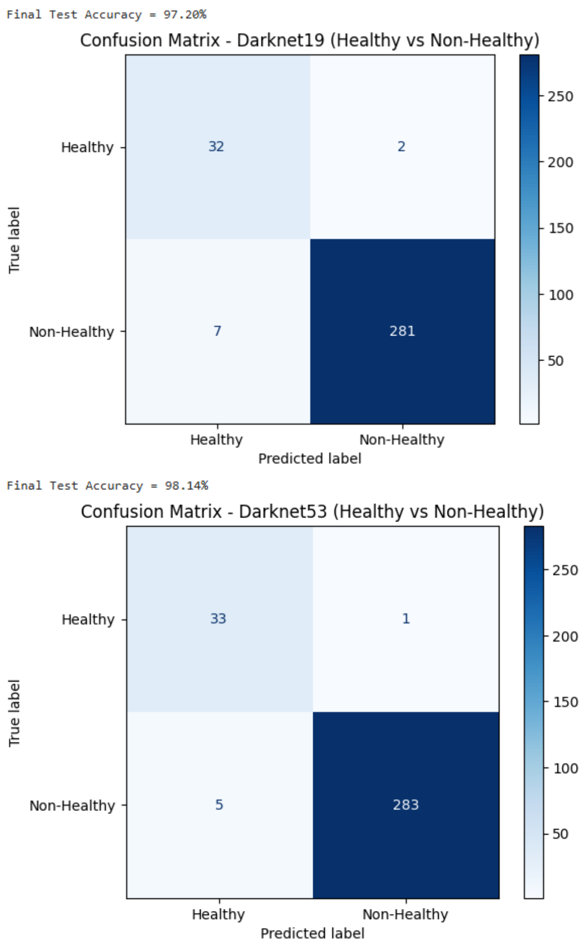

Figure 4 Confusion matrix for two classes

A significant increase in accuracy was observed in the Healthy–Unhealthy binary classification scenario; this is due to the lower complexity of the task compared to the multi-class version. When all disease types were combined into a single “Unhealthy” class, the Darknet-19 model achieved 97.20% accuracy, demonstrating its ability to distinguish healthy leaves from those showing any signs of disease. Darknet-53 took this performance a step further, achieving 98.14% accuracy, a nearly perfect distinction. Figure 4 shows confusion matrix.

An examination of the confusion matrices for this scenario reveals that class boundaries are clearly separated and misclassifications are extremely low. This finding confirms that the optimized training process makes the models’ decision-making mechanisms highly robust. This significant performance improvement observed in binary classification demonstrates that both architectures can effectively exploit the general structural differences between healthy and diseased leaves. Conversely, learning finer distinctions in the five-class scenario benefits particularly from the enhanced feature extraction capabilities offered by Darknet-53’s deep residual connections.

Table 4 Class-wise distribution of papaya leaf images

Metric	**Darknet19 Architecture**	**Darknet53 Architecture**
Accuracy	97.20%	98.14%
Precision (Healthy)	82.05%	86.84%
Recall (Healthy)	94.11%	97.05%

Table 4 presents a comparative evaluation of Darknet19 and Darknet53 using the Healthy class as the positive class in binary classification. The results show that Darknet53 consistently outperforms Darknet19 across all evaluated metrics. Darknet53 achieves higher precision and recall, indicating that it more accurately identifies healthy leaves while minimizing false classifications. The improved F1-score further confirms its balanced performance between sensitivity and precision. Additionally, Darknet53 attains the highest overall accuracy, demonstrating superior generalization capability. These findings highlight the effectiveness of the deeper residual architecture in Darknet53 and suggest that it provides a more robust solution for plant health assessment tasks compared to Darknet19.

3.2. Comparative Analysis

A combined evaluation of the results from the multi-class and binary classification experiments reveals that the deep-structured Darknet-53 architecture consistently outperforms Darknet-19. While both models demonstrate good general ability to recognize disease symptoms, Darknet-53’s residual learning structure and broader representational capacity allow it to extract more discriminative and fine-grained features. This superiority was particularly evident in the five-class scenario; when visual similarities between classes were high, Darknet-53’s more expressive feature hierarchy significantly improved accuracy. In contrast, the computationally lighter Darknet-19 occasionally failed to distinguish visually similar disease categories, resulting in more misclassifications. The performance difference between the two models became even clearer when examining the binary classification results (Figure 5).

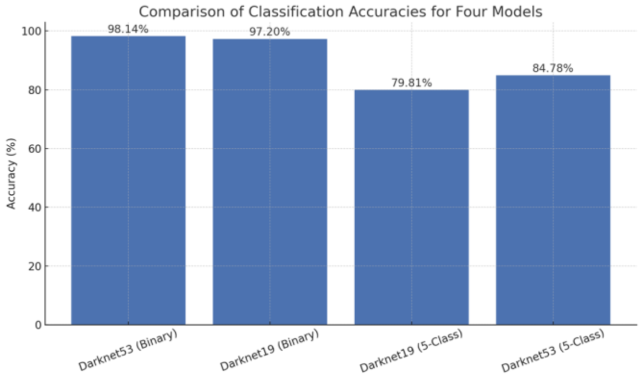

Figure 5 Comparison of accuracies

When the disease classes were combined into a single "Unhealthy" class, both models showed a significant increase in accuracy, indicating that it was easier to learn general structural distinctions between healthy and diseased leaves. However, Darknet-53 achieved near-perfect accuracy, demonstrating that it was able to capture subtle changes in vein structure, color uniformity, and lesion boundaries more effectively than Darknet-19. Overall, while Darknet-19 provides a stronger baseline model with lower computational cost, Darknet-53 provides higher accuracy and more robust generalization performance in both multi-class and binary classification. These results clearly demonstrate that Darknet-53 is the preferred architecture for more complex agricultural image analysis.

4. Conclusion

This study examined the performance of the optimized Darknet-19 and Darknet-53 architectures for the evaluation of papaya leaves in both multi-class and binary classification scenarios. The combined use of data augmentation methods, label smoothing, dropout, and cosine-based learning rate planning significantly improved the generalization success of the models. The findings revealed that Darknet-53 achieved higher accuracy compared to Darknet-19 in both tasks, demonstrating a significant superiority in five-class classification and providing nearly error-free results in binary classification. These results demonstrate that optimized Darknet-based approaches offer a robust solution for reliable and efficient plant health assessment and provide a solid foundation for future use in smart agricultural monitoring systems.

Author contributions

RK: Conceptualization; Methodology; Software; Validation; Formal analysis; Investigation; Data curation; Writing – original draft; Visualization.

HKS: Conceptualization; Methodology; Validation; Resources; Writing – review & editing; Supervision; Project administration.

Conflict of Interest

The authors have no conflicts of interest to declare.

References

[1] Gulzar Y. Papaya leaf disease classification using pre-trained deep learning models: A comparative study. *Applied Fruit Science* (2025) **67**(4):287.

[2] Mustofa S, Ahad MT, Emon YR, Sarker A. BDPapayaLeaf: A dataset of Papaya leaf for disease detection, classification, and analysis. *Data in Brief* (2024) **57**:110910.

[3] Hossen MS, Haque I, Islam MS, Ahmed MT, Nime MJ, Islam MA. Deep learning based classification of papaya disease recognition. In: *2020 3rd International Conference on Intelligent Sustainable Systems (ICISS*: IEEE (2020). p. 945–951.

[4] Moraes JL, Oliveira Neto J, Correia-Silva JR, Paixão TM, Badue C, Oliveira-Santos T, et al. Sisfrutos Papaya: A Dataset for Detection and Classification of Diseases in Papaya. In: *International Conference on Artificial Neural Networks*. Cham: Springer International Publishing (2021). p. 28–39.

[5] Gurunathan K, Bharathkumar V, Meeran M, Hariprasath K, Jidendiran R. Classification of Cultivars Employing the Alexnet Technique Using Deep Learning. In: *2023 International Conference on Bio Signals, Images, and Instrumentation (ICBSII*: IEEE (2023). p. 1–7.

[6] Pawar SE, Surana AV, Sharma P, Pujeri R. Fruit disease detection and classification using machine learning and deep learning techniques. *Int J Intell Syst Appl Eng* (2024) **12**(4s):440–453.

[7] Madelo H, Prieto JC, Bajao J, Bacus JA. Papaya Leaf Disease Identification using ResNet with Transfer Learning. In: *2023 8th International Conference on Robotics and Automation Engineering (ICRAE*: IEEE (2023). p. 101–105.

[8] Moraes JL, Oliveira Neto J, Badue C, Oliveira-Santos T, Souza AF. Yolo-Papaya: A papaya fruit disease detector and classifier using CNNs and convolutional block attention modules. *Electronics* (2023) **12**(10):2202.

[9] Yashodharan S. Neural network for papaya leaf disease detection. *Acta graphica: znanstveni časopis za tiskarstvo i grafičke komunikacije* (2019) **30**(3):11–24.

[10] Habib MT, Majumder A, Nandi RN, Ahmed F, Uddin MS. A comparative study of classifiers in the context of papaya disease recognition. In: *Proceedings of International Joint Conference on Computational Intelligence: IJCCI 2018*. Singapore: Springer Nature Singapore (2019). p. 417–429.

[11] Kumar RR, Jain AK, Sharma V, Das P, Sahni P. Enhancing precision in papaya crop health management: A cnn-based approach for early disease detection and classification. In: *2024 IEEE International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS*: IEEE (2024). p. 1–6.

[12] Gani R, Isty MN, Rashid M, Ahmed J, Islam T, Hasan M, et al. Attention guided convolutional neural network with explainable AI for papaya leaf disease detection in edge and drone agricultural systems. *Scientific Reports* (2025) **15**(1):41489.

[13] Calgan H, Gokyildirim A, Ilten E, Demirtas M. Classification of fractional-order chaotic systems using deep learning methods. *The European Physical Journal Special Topics* (2025):1–19.

[14] Powers DM. *Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation* (2020).