JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7445

# IMPACT OF ROI SEGMENTATION METHODS ON ANTHRACNOSE DETECTION IN PAPAYA LEAVES USING RESNET-50

**Shinta Siti Sundari[1*]; Ruuhwan[1]; Evi Dewi Sri Mulyani[1]**

Informatics Engineering[1]
Perjuangan University, Tasikmalaya, Indonesia[1]
https://unper.ac.id[1]
ss.shinta@gmail.com*, ruuhwan@unper.ac.id, eviajadech@gmail.com

(*) Corresponding Author
(Responsible for the Quality of Paper Content)



***Abstract****—Early detection of anthracnose disease on papaya leaves is important for mitigating crop yield losses, but manual methods are inefficient and prone to subjective interpretation, this study evaluates the effect of region of interest (ROI) extraction strategies on deep learning-based classification performance. The objective of this study was to compare four classification pipelines: ResNet-50 without segmentation, (M1) ExG+Otsu + ResNet-50, (M2) U-Net + ResNet-50, and (M3) RCNN + ResNet-50, in detecting anthracnose on papaya leaf images. The methods included the use of the public BDPapayaLeaf dataset, pre-processing and augmentation, and evaluation using stratified K-fold cross-validation with evaluation metrics including precision, recall, and F1. The results show that the semantic segmentation-based pipeline, U-Net + ResNet-50 (M2), provides the best performance with an F1-score (macro/weighted) ≈ 0.961 and precision–recall balance in both classes; M0 showed the highest recall for the anthracnose class (≈0.99) while M1 based on color thresholding provided the lowest performance due to sensitivity to lighting variations; M3 (RCNN) was in the middle. The findings suggest the use of segmentation and classification pipelines for field applications, noting the need for dataset expansion and inference optimization for deployment.*

***Keywords****: Anthracnose, Papaya, ResNet-50, Segmentation, U-Net*

**Intisari**—*Deteksi dini penyakit antraknosa pada daun pepaya penting untuk mitigasi kerugian hasil panen, namun metode manual kurang efisien dan rentan subjektivitas; penelitian ini mengevaluasi pengaruh strategi ekstraksi region of interest (ROI) terhadap kinerja klasifikasi berbasis deep learning. Tujuan penelitian adalah membandingkan empat pipeline klasifikasi: (M0) ResNet-50 tanpa segmentasi, (M1) ExG+Otsu + ResNet-50, (M2) U-Net + ResNet-50, dan (M3) RCNN + ResNet-50, dalam mendeteksi antraknosa pada citra daun pepaya. Metode meliputi penggunaan dataset publik BDPapayaLeaf, pra-pemrosesan dan augmentasi, serta evaluasi menggunakan stratified K-fold cross-validation dengan metrik evaluasi meliputi precision, recall, dan F1. Hasil menunjukkan bahwa pipeline berbasis segmentasi semantik, U-Net + ResNet-50 (M2), memberikan performa terbaik dengan F1-score (macro/weighted) ≈ 0,961 dan keseimbangan precision–recall pada kedua kelas; M0 menunjukkan recall tertinggi untuk kelas antraknosa (≈0,99) sementara M1 berbasis thresholding warna memberikan kinerja terendah akibat sensitifitas terhadap variasi pencahayaan; M3 (RCNN) berada pada posisi menengah. Temuan merekomendasikan penggunaan pipeline segmentasi dan klasifikasi untuk aplikasi lapangan dengan catatan perlunya perluasan dataset dan optimasi inferensi untuk deployment.*

***Kata Kunci****: Antraknosa, Pepaya, ResNet-50, Segmentasi, U-Net*

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7445
JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

## INTRODUCTION

Papaya (Carica papaya L.) is an important tropical fruit commodity with high nutritional and economic value. However, papaya productivity is often constrained by various leaf and fruit diseases that significantly reduce both yield quality and quantity [1]. One of the major diseases is anthracnose, caused by the fungus *Colletotrichum* spp., which can lead to lesion spots on leaves as well as fruit rot. Anthracnose disease is widely known to cause severe damage, leading to substantial crop yield losses. Early identification and prompt management of anthracnose are crucial to prevent further spread and to reduce economic losses for farmers [2].

Traditionally, papaya leaf disease detection is done through manual visual inspection by experts. This conventional method is time-consuming, labor-intensive, and prone to human error. These limitations have encouraged the development of computer-assisted disease detection systems[3]. Research shows that computer vision-based plant disease diagnostics and machine learning (ML) models are able to improve the efficiency of disease detection and classification in papaya [4], [5]. Advances in deep learning with convolutional neural networks (CNN) has opened up great opportunities for accurate automation of plant disease identification.

Various CNN models have been applied to classify papaya leaf diseases, including the architecture deep CNN like VGG16 and ResNet-50. ResNet-50 models with transfer learning have been reported to achieve 88% accuracy in recognizing healthy papaya leaves vs. infected (including anthracnose), outperforming the CNN basic models like VGG16 [6]. This demonstrates ResNet's potential to produce better classification performance thanks to its residual structure, which allows for deeper networks without accuracy degradation. However, challenges remain, especially in cases where disease symptoms (lesions) occupying a small area on the leaf image or when a complex background distracts the model's attention. It is in this context that image segmentation plays an important role: by accurately highlighting the area of the lesion, so that different objects and their boundaries can be located [7] it is hoped that the classification model can focus more on features relevant to the disease and improve its accuracy.

In the past five years, the classification of papaya leaf diseases, particularly anthracnose, has become an important focus in computer vision-based research. Research by Kaur et al. [8] became one of the initial studies that compiled a dataset and evaluated the classification of papaya leaf disease using feature-based methods. They used a combination of RGB color features and GLCM textures, which were then classified using a simple CNN.

Although this approach showed promising results, lesion area segmentation was not performed, and classification was performed directly on the entire image. Next, Chaithanya dan Rachana [6] Comparing various transfer learning architectures: ResNet-50, VGG16, and EfficientNetB0 for the classification of five papaya leaf diseases. The results of the experiment show that ResNet-50 provides the highest accuracy compared to other models. However, the input used is the entire leaf image, without isolating the lesion area. Madelo et al. [9] continued the application of ResNet-50 in the context of mobile applications. They developed an Android-based classification system for four papaya leaf diseases, using a fine-tuned transfer learning model. The successful implementation of ResNet in a practical platform confirms the potential of this model as part of the final classification in the segmentation-classification pipeline. However, this system does not utilize infection area segmentation, which may limit the interpretability of the model.

In 2024, several important studies were published and enriched the literature. Gani et al. [5] compiling datasets collected in field conditions with natural backgrounds and lighting variations, reflecting the challenges of classification in real environments, but not presenting classification or segmentation methods. Kumar et al. [4] for the classification of six types of papaya leaf diseases, including anthracnose. Their model achieved high accuracy on the test data. However, as in previous studies, the classification was performed on full leaf images without lesion segmentation, which made it difficult for the model to focus specifically on the symptoms. Further studies by Kumar et al. [10], [11] offered a hybrid approach by combining CNN as a feature extractor and Random Forest as a classifier (CNN+RF), with seven classes of papaya leaf diseases, they achieved an accuracy of up to 98%. Mustofa et al. [1] providing a highly strategic contribution through the development of the BDPapayaLeaf dataset for deep learning detection formats such as YOLO, Mask R-CNN, and U-Net, making it an ideal dataset for CNN-based segmentation experiments. This is the only study that explicitly supports the pixel-based lesion segmentation approach, and is highly relevant to this research.


Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

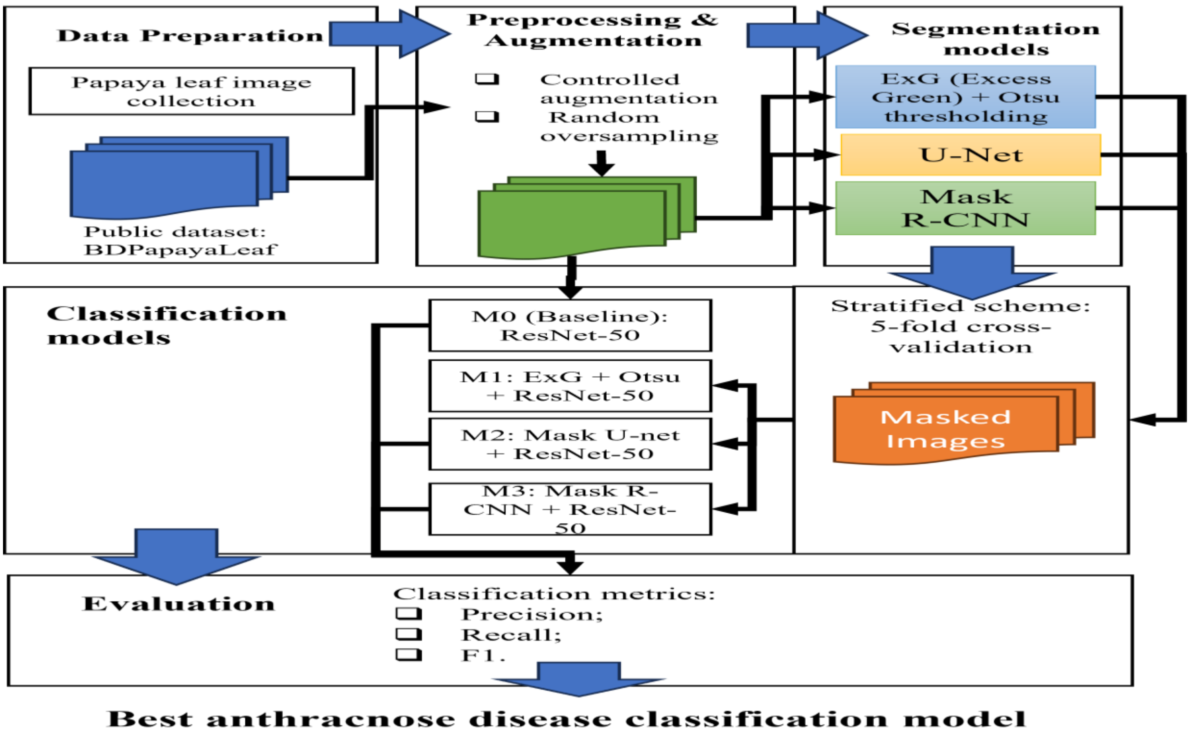
VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7445

Previous studies have shown rapid progress in image classification based on ResNet-50 [9], [12]. However, no one has directly compared semantic segmentation methods, U-Net, and instance segmentation, Mask R-CNN, in the context of the effect of segmentation on anthracnose classification accuracy. Therefore, this study aims to compare two different CNN-based segmentation methods and evaluate their impact on classification performance using ResNet.

The first segmentation was performed using U-Net, a symmetric encoder-decoder architecture that has proven effective in pixel semantic segmentation for medical image data [13]. U-Net will highlight anthracnose lesion areas in the form of full-pixel binary masks. The second approach is Mask R-CNN, a two-stage architecture for instance segmentation that not only generates pixel masks, but also distinguishes each object (lesion) individually [14]. Mask R-CNN using the ResNet backbone and Region Proposal Network to generate bounding boxes and masks for each infected area. The segmentation results from these two methods will be used to improve the classification process by the ResNet-50 model. The classification model was built to distinguish between healthy leaves and leaves infected with anthracnose. Model performance was evaluated using accuracy, precision, recall, F1-score metrics, and statistical comparisons between methods.

## MATERIALS AND METHODS

This study uses a structured experimental approach with a pipeline consisting of data preparation, pre-processing and augmentation, lesion area segmentation, classification using ResNet-50, and model performance evaluation, as illustrated in Figure 1. The dataset used comes from the public BDPapayaLeaf dataset, which contains images of healthy papaya leaves and leaves infected with anthracnose disease, as shown in Figure 2. The images were collected from field conditions with varying lighting and natural backgrounds, thus representing real-world scenarios with 480 × 640 pixels for each image. The next stage involved controlled augmentation and Random Over Sampling to reduce class imbalance and improve model generalization. This stage aimed to enrich the variety of training data and reduce potential bias due to uneven data distribution.




Source: (Research Results, 2025)
Figure 1. Research method flow for classifying anthracnose disease on papaya leaves

Three approaches were used in the segmentation stage to extract lesion areas on leaves, namely: a) ExG + Otsu, a classic image processing-based method that is often used[15]; b) U-Net, a CNN-based segmentation architecture capable of producing pixel-level masks [13]; and c) R-CNN, a region proposal-based object detection approach[14]. The segmentation results in the form of masked images are then used as input for the classification model.


(a) (b)

Source: (Mustofa, 2024 [1])
Figure 2. Example image for papaya leaves: (a) affected by anthracnose and (b) healthy

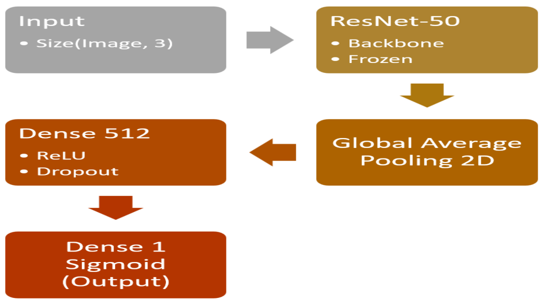
For Classification Models, ResNet-50 [9] selected as the basic classification architecture. In this study, experiments were conducted using four models that integrate segmentation methods with the ResNet-50 architecture as the classification backbone, which are: a) M0:ResNet-50 (Baseline), using ResNet-50 directly on the leaf image without additional segmentation; b) M1: ExG+Otsu+ResNet-50, using classical segmentation methods to separate leaves from the background. The ROI obtained is then classified by ResNet-50 ; c) M2: U-Net ROI + ResNet-50, Region of Interest (ROI) segmentation was conducted using U-Net to extract leaf areas from images; and d) M3: RCNN ROI + ResNet-50, combining Region-based Convolutional Neural Network (RCNN) to detect and extract leaf ROI, before classification with ResNet-50. The selection of four classification pathways, M0, M1, M2, and M3, aims to comprehensively evaluate the effect of segmentation on leaf image classification performance and assess the effectiveness of combining segmentation and deep learning-based classification models. Each pathway is designed with varying levels of complexity and processing to identify the extent to which the ROI segmentation stage improves disease detection accuracy, specifically anthracnose in papaya leaves.

Next, the evaluation was conducted using a stratified 5-fold cross-validation scheme to ensure balanced class distribution in each data fold, because the data used is not balanced with the comparison of healthy leaf data 228 and anthracnose 355. Model performance was measured using Precision, Recall, and F1-score metrics, both macro and weighted. The evaluation results from the four configurations were compared to identify the best approach for detecting anthracnose disease. This pipeline systematically enables analysis of the impact of various segmentation techniques on the performance of ResNet-50-based classification models, while also providing a basis for recommending the optimal method for papaya leaf disease diagnosis applications.

## RESULTS AND DISCUSSION

The classification model used utilizes transfer learning with ResNet-50 as the feature extraction backbone and a lightweight classification head designed for binary tasks (Anthracnose vs. Healthy). As shown in Figure 3, the classification model is designed as a transfer learning pipeline that utilizes the ResNet-50 architecture as the feature extraction backbone and a classification head tailored for binary tasks: Anthracnose vs. Healthy. The model input is an RGB image with a prespecified size (IMG_SIZE × IMG_SIZE × 3). The ResNet-50 backbone is loaded with pre-trained weights on ImageNet and used without its top classification layer (top removed), resulting in a smaller spatial feature map but rich in representational information.

To reduce the spatial dimensions to global feature vectors, global average pooling is performed on the backbone output. The pooled feature vectors are then processed by a fully-connected (dense) layer with 512 units and a ReLU activation function to perform the final non-linear transformation. As a regularization step, dropout is applied to the activation, followed by a single output neuron with sigmoid activation that estimates the probability of the positive class healthy. This single sigmoid output is suitable for binary classification scenarios and combined with the binary cross-entropy loss function during training.




Source: (Research Results, 2025)
Figure 3. Architectural diagram of the anthracnose disease classification model on papaya leaves

The findings of this study have important implications for the application of plant disease detection technology in the field. The recommendation to use an integrated segmentation and classification pathway shows that a digital image-based diagnostic system can be implemented automatically and in real-time in agricultural


Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7445

environments. The resulting model can be integrated into a smartphone-based application or embedded system connected to a field camera. Farmers simply take a photo of a plant leaf, then the system automatically segments the leaf area and classifies the health condition (healthy or infected with anthracnose). This allows for early detection before the disease spreads widely, so that interventions such as fungicide spraying can be carried out more efficiently.

In terms of training strategy, the pipeline follows two general stages: (1) feature-head training with the backbone frozen to initialize the classification head against the target data distribution, and (2) selective fine-tuning where the upper layers of the backbone (conv5 block and upper conv4 block) are opened for retraining to adjust high-level features to the papaya leaf image domain. During fine-tuning, the batch normalization parameters in the backbone are left non-trainable to maintain internal stability statistics, a strategy that is generally recommended when the batch size is relatively small or the dataset is limited. The learning rate schedule adopts a warmup followed by cosine decay approach: the learning rate starts from a small value and increases linearly during the warmup phase to avoid initial gradient spikes when the head weights are initialized, then decreases following a cosine curve until the end of training to support smooth convergence and avoid bad local minima. This practice has proven effective in modern transfer learning fine-tuning scenarios.

Table 1. Comparison of Precision, Recall, F1 (Each Class)

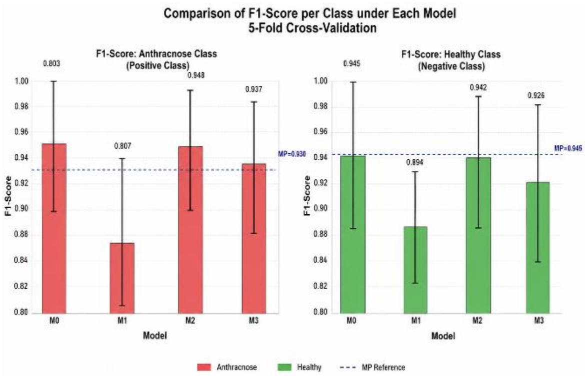
| Metrik | M0 | M1 | M2 | M3 |
|---|---|---|---|---|
| Precision – Anthraknosa | 0.919 ± 0.074 | **0.920± 0.066** | 0.918± 0.067 | 0.899±0 .084 |
| Recall – Anthraknosa | 0.991 ± 0.011 | 0.860± 0.079 | **0.983± 0.028** | **0.983± 0.018** |
| F1 – Anthraknosa | **0.953 ± 0.043** | 0.887± 0.052 | 0.948± 0.041 | 0.937±0 .045 |
| Precision – Healthy | **0.990 ± 0.013** | 0.872± 0.065 | 0.982± 0.030 | 0.981±0 .019 |
| Recall – Healthy | 0.909 ± 0.078 | **0.922± 0.070** | 0.908± 0.082 | 0.882±0 .102 |
| F1 – Healthy | **0.946 ± 0.051** | 0.894± 0.044 | 0.942± 0.049 | 0.926±0 .056 |

Source: (Research Results, 2025)

The performance of the four classification models was evaluated using precision, recall, and F1-score metrics on two classes, namely anthracnose-infected leaves and healthy leaves. Based on Table 1, the results show variations in performance between models depending on the ROI segmentation technique used prior to classification with ResNet-50. Model M0 obtained the highest recall in the anthracnose class (0.99), which means that this model is most sensitive in detecting diseased leaves. However, the performance of M0 was relatively lower in the healthy leaf class (recall 0.91), indicating a tendency for misclassification in some healthy leaves. Model M1 showed the highest recall in the healthy leaf class (0.96), indicating better detection capabilities for the negative (healthy) category. However, its performance weakened in the anthracnose class, where recall only reached 0.80 with an F1-score of 0.87, making it less reliable in identifying disease cases.

Meanwhile, the M2 model provides the most balanced and consistent performance with the highest F1-score in both classes (0.96). This model successfully maintains relatively high precision and recall in both the anthracnose class (P=0.94; R=0.98) and healthy leaves class (P=0.98; R=0.94). These results indicate that deep learning-based segmentation (U-Net) is capable of extracting ROI more optimally, thereby supporting ResNet-50 classification. Furthermore, the M3 model showed competitive performance, with high precision (0.98) in both classes, but experienced a decrease in recall in healthy leaves (0.89). This indicates that RCNN tends to be stricter in classification, so that some healthy leaves are misclassified as diseased (false positives).

Overall, as shown in Figure 4, M2 proved to be the most superior because it was consistent in both classes with an F1-score of 0.96. Meanwhile, M0 is worth considering in early detection applications because of its high recall for disease (0.99), despite the risk of misclassifying healthy leaves. These results confirm that the selection of The ROI segmentation strategy greatly influences the effectiveness of ResNet-50-based classification.




Source: (Research Results, 2025)

Figure 4. Comparison of F1-scores between models for both classes (Anthracnose & Healthy)

The performance evaluation of the four models in Table 2 shows significant differences at both the class level and aggregate metrics. In the per-class analysis, M2 (U-Net+ResNet-50) displays the highest consistency, excelling in the F1 score in both classes (Anthracnose and Healthy) with a value of 0.96. This indicates that U-Net-based ROI segmentation is capable of accurately extracting leaf areas and improving the quality of input for ResNet-50. In contrast, M1 (ExG+Otsu+ResNet-50) showed the lowest performance, with an F1 of only 0.87 for Anthracnose and 0.89 for Healthy, due to unstable ROI and the inclusion of a lot of non-leaf background, In field images, leaf color intensity and spectrum can vary due to sun position, humidity, and different camera devices. This leads to inconsistent segmentation results, where some leaf regions may not be fully detected or may include unwanted background.

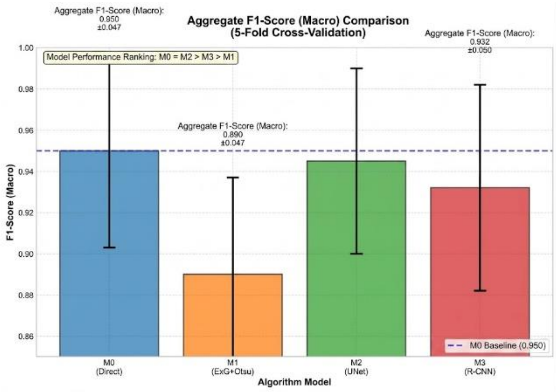
As a result, classification performance in the M1 pathway may suffer due to the non-uniformity of the input ROIs, and the classification model may learn unrepresentative features. M0 (ResNet-50 baseline) is still quite strong, especially with high recall for Anthracnose (0.99), making it sensitive to disease detection, although recall for Healthy is relatively lower (0.91). M3 (RCNN+ResNet-50) is in the middle, with better performance than M1, but lagging behind M0 and M2, mainly due to the variability of the generated bounding boxes.

Table 2. Model Performance Comparison (5-Fold Cross-Validation)

| Model | Precision (Macro) | Recall (Macro) | F1 (Macro) | F1 (Weighted) |
|---|---|---|---|---|
| M0 | **0.955 ± 0.042** | **0.950 ± 0.047** | **0.950 ± 0.047** | **0.950 ± 0.047** |
| M1 | 0.896 ± 0.045 | 0.891 ± 0.047 | 0.890 ± 0.047 | 0.890 ± 0.047 |
| M2 | 0.950 ± 0.041 | 0.945 ± 0.045 | **0.945 ± 0.045** | **0.945 ± 0.045** |
| M3 | 0.940 ± 0.047 | 0.932 ± 0.050 | 0.932 ± 0.050 | 0.932 ± 0.050 |

Source: (Research Results, 2025)

Based on the results of the five-fold cross-validation evaluation, the macro F1-score difference between M0 and M2 is only 0.005 (0.5%), which is statistically insignificant considering the overlapping confidence intervals. The superiority of M0 can be attributed to ResNet-50's ability to extract hierarchical features without losing contextual information during segmentation. This phenomenon can introduce artifacts and noise to the input image. Nevertheless, M2 remains relevant for applications that require visual interpretability, with the trade-off of additional complexity.



Source: (Research Results, 2025)
Figure 5. F1 Score Macro Aggregat

Figure 5 shows a comparison of the performance of four model configurations based on the F1 metric (macro and weighted), M2 (U-Net+ResNet-50) and M0 (ResNet-50) show comparable performance (F1 ≈ 0.945 vs 0.950, not statistically significantly different), indicating that U-Net segmentation does not provide significant improvement over the baseline for this dataset.

Overall, this evaluation pattern shows that U-Net+ResNet-50 (M2) is the most superior approach, outperforming both the baseline and other ROI methods. The ResNet-50 (M0) baseline is still relevant thanks to its high recall on disease classes, while RCNN+ResNet-50 (M3) is a middle-of-the-road option, and the classic ExG+Otsu (M1) approach proves to be less effective. These findings confirm that the quality of deep learning-based ROI segmentation plays a crucial role in improving the accuracy of plant disease classification.

The results of the experiment show that the performance of the papaya leaf disease classification model is greatly influenced by the image processing strategy prior to the classification stage. The ResNet-50 baseline model (M0) proved to be competitive, with high recall in the anthracnose class. This is consistent with the study by Chaithanya and Rachana [6] also Madelo et al. [9] which reported that ResNet-50 provided the best performance among other transfer learning architectures, even in conditions without segmentation. However, the performance of ResNet-50 on whole images has the potential to be biased towards the background and lighting variations, as also noted in the study by Kaur et al. [8].

The simple segmentation-based approach, ExG+Otsu+ResNet-50 (M1), yielded the lowest results. This indicates that color-based thresholding methods are not yet capable of capturing complex lesion patterns in high-variability field conditions. Similar weaknesses have been noted in previous


Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

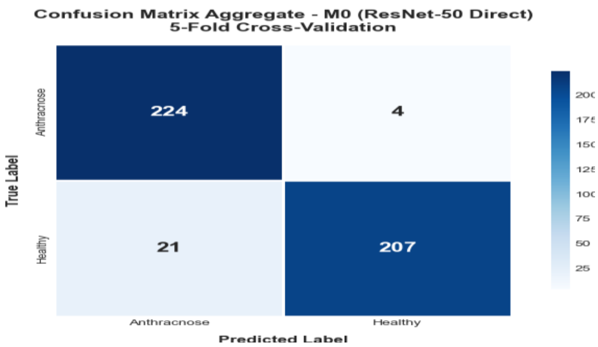
VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7445

manual feature-based studies, which struggled to generalize to real-world environments [8].

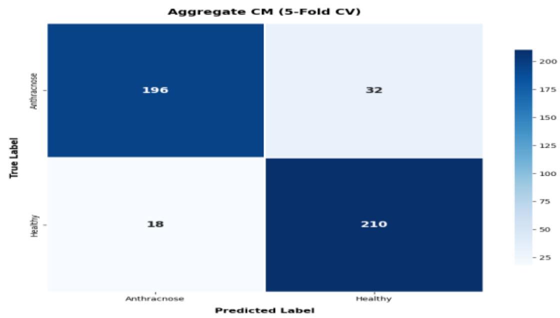
Deep learning-based segmentation integration using U-Net+ResNet-50 (M2) produced the highest performance (F1-macro ≈ 0.96). This finding reinforces the relevance of the pixel-based lesion segmentation approach as facilitated by the BDPapayaLeaf dataset developed by Mustofa et al. [1]. The advantage of M2 lies primarily in its ability to isolate lesion areas with precision, allowing the classification model to focus solely on pathological information. This addresses the limitations of previous studies [6], [10], [11] who still classify whole images without lesion isolation.

The RCNN+ResNet-50 model (M3) provides intermediate performance, better than traditional segmentation (M1), but still below full-pixel segmentation (M2). These results are consistent with the argument of Mustofa et al. [1] that bounding box-based detection can only reduce some background noise, but is not detailed enough to capture the complex spatial patterns of lesions.

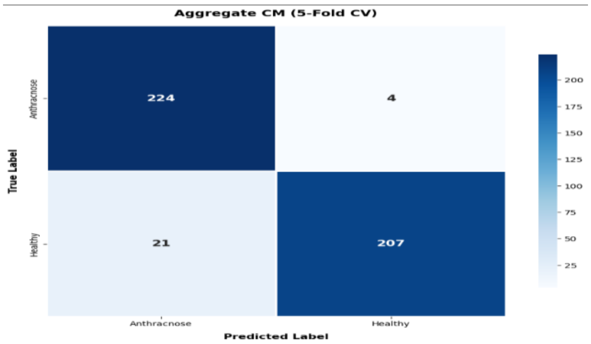
This study shows that the integration of U-Net-based lesion segmentation with classification using ResNet-50 (M2) significantly improves the quality of anthracnose detection on papaya leaves compared to the approach without segmentation (M0), the classical segmentation method (M1), and object detection-based ROI (M3).



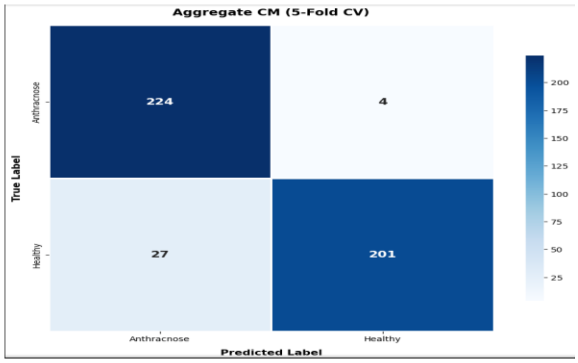

Source: (Research Results, 2025)
Figure 6. Confusion Matrix M0




Source: (Research Results, 2025)
Figure 7. Confusion Matrix M1




Source: (Research Results, 2025)
Figure 8. Confusion Matrix M2




Source: (Research Results, 2025)
Figure 9. Confusion Matrix M3

Based on the five-fold cross-validation evaluation conducted in this study, the M0 model (ResNet-50 without ROI segmentation) demonstrated superior performance with a macro F1-score of 0.950 ± 0.047 compared to the M2 model (ResNet-50 with U-Net segmentation) which achieved 0.945 ± 0.045. Although the difference between the two models was only 0.005 (0.5%), these results consistently indicate that the end-to-end architecture without segmentation preprocessing provides more optimal results for anthracnose classification on papaya leaves. This phenomenon can be explained by several fundamental mechanisms in deep learning for agricultural image processing. First, the superiority of M0 can be attributed to the ability of ResNet-50, trained on the ImageNet dataset, to directly extract hierarchical features from the original image, without losing contextual information that occurs during the segmentation process. The main contributions of this study are (1) empirical evidence linking pixel-level ROI quality with improved F1 scores and classification model robustness on field data, and (2) a structured comparison of four pipelines that providing practical guidance for the implementation of plant disease diagnosis systems.

## CONCLUSION

The experimental results show that the combination of U-Net + ResNet-50 (M2) provides the best performance for anthracnose detection on papaya leaves. M2 shows significant potential as an alternative approach with better stability (standard deviation 0.045 vs. 0.047 in M0) and interpretability through ROI visualization. The deep learning-based segmentation approach has been proven to improve the signal-to-noise ratio in input features and promote model classification generalization. This study has limitations in that the dataset coverage is still limited to two classes (healthy vs anthracnose). For further work, it is recommended to expand the dataset to multi-classes and different field sources to test generalization. Second, the segmentation process with U-Net on M2, while effective in isolating leaf regions, inadvertently removes important information contained in the background and spatial context surrounding the lesion which provides additional signals critical for distinguishing between healthy and infected classes.

## REFERENCE

[1] S. Mustofa, T. Ahad, Y. R. Emon, and A. Sarker, "BDPapayaLeaf: A dataset of papaya leaf for disease detection, classification, and analysis," *Data Br.*, vol. 57, p. 110910, Jun. 2024, doi: 10.1016/j.dib.2024.110910.

[2] S. Karthikeyan, R. Charan, S. Narayanan, and L. Jani Anbarasi, "Enhanced plant disease classification with attention-based convolutional neural network using squeeze and excitation mechanism," *Front. Artif. Intell.*, vol. 8, no. August, 2025, doi: 10.3389/frai.2025.1640549.

[3] A. Y. Ashurov *et al.*, "Enhancing plant disease detection through deep learning: a Depthwise CNN with squeeze and excitation integration and residual skip connections," *Front. Plant Sci.*, vol. 15, no. January, pp. 1–16, 2024, doi: 10.3389/fpls.2024.1505857.

[4] P. Kumaran, A. Tirkey, U. Kumar, and K. V. Kumar, "Anthracnose Disease Detection in Papaya Fruit Using Machine Learning Techniques," Jun. 2024, pp. 1–6. doi: 10.1109/icitiit61487.2024.10580524.

[5] R. Gani *et al.*, "Smartphone image dataset to distinguish healthy and unhealthy leaves in papaya orchards in Bangladesh," *Data Br.*, vol. 55, p. 110599, Jun. 2024, doi: 10.1016/j.dib.2024.110599.

[6] A. S. Chaithanya and M. Rachana, "Identification of Diseased Papaya Leaf through Transfer Learning," *Indian J. Sci. Technol.*, vol. 16, no. 48, pp. 4676–4687, 2023, doi: 10.17485/ijst/v16i48.2690.

[7] O. Rainio, J. Teuho, and R. Klén, "Evaluation metrics and statistical tests for machine learning," *Sci. Rep.*, vol. 14, no. 1, pp. 1–14, 2024, doi: 10.1038/s41598-024-56706-x.

[8] H. Kaur, D. Prashar, and V. Kumar, "Disease Identification in Papaya Plant and their Dataset," Jun. 2022, pp. 1220–1224. doi: 10.1109/ic3i56241.2022.10072453.

[9] H. K. L. Madelo, J. C. Prieto, J. R. D. Bajao, and J. A. Bacus, "Papaya Leaf Disease Identification using ResNet with Transfer Learning," Jun. 2023, pp. 101–105. doi: 10.1109/icrae59816.2023.10458541.

[10] V. Kumar, D. Banerjee, D. Upadhyay, M. Singh, and K. R. Chythanya, "Hybrid Model for Effective Papaya Leaf Disease Diagnosis," Jun. 2024, pp. 1–5. doi: 10.1109/icemps60684.2024.10559298.

[11] R. R. Kumar, A. K. Jain, V. Sharma, P. Das, and P. Sahni, "Enhancing Precision in Papaya Crop Health Management: A CNN-Based Approach for Early Disease Detection and Classification," Jun. 2024, pp. 1–6. doi: 10.1109/iciteics61368.2024.10625272.

[12] Y. F. Riti, R. P. Kristianto, and D. R. A. Setiawan, "Perform Comparation of Deep Learning Methods in Gender Classification From Facial Images," *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 10, no. 4, pp. 926–936, 2025, doi: 10.33480/jitk.v10i4.4717.

[13] S. Surono, M. Rivaldi, D. A. Dewi, and N. Irsalinda, "New Approach to Image Segmentation: U-Net Convolutional Network for Multiresolution CT Image Lung Segmentation," *Emerg. Sci. J.*, vol. 7, no. 2, pp. 498–506, Feb. 2023, doi: 10.28991/ESJ-2023-07-02-014.

[14] M. Kiyak, B. K. Savaş, and F. Akdeniz, "Hand vein recognition using Mask R-CNN," in *2025 7th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (ICHORA)*, IEEE, May 2025, pp. 1–4. doi: 10.1109/ICHORA65333.2025.11017312.

[15] K. Harika and M. N. Nachappa, "A Study on Deep Learning Methods to Identify the Infected Regions from Papaya Fruit Images," Jun. 2023, pp. 1–6. doi: 10.1109/aikiie60097.2023.10390496.



Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri