International Journal of Informatics and Data Science
ISSN 3026-7315 (media online)
Volume 3, No 2, June 2026, Page: 100−114
DOI: doi.org/10.64366/ijids.v3i2.549
https://journals.adaresearch.or.id/ijids/index

Comparative Study of Background Removal for Mango Fruit Disease Classification Using ResNet50-SVM

Perani Rosyani*, Ines Heidiani Ikasari, Saprudin, Shita Nurul Ayasha, Alya Salsabila Az Zahra

Faculty of Computer Science, Informatics Engineering Study Program, Universitas Pamulang, South Tangerang, Indonesia
Email: [1,*]dosen00837@unpam.ac.id, [2]dosen01374@gmail.com, [3]dosen00845@unpam.ac.id, [4]shitanurulayasha@gmail.com, [5]Alyasalsabila@gmail.com
Correspondence Author Email: dosen00837@unpam.ac.id
Received: 03/05/2026, Revised: 20/05/2026, Accepted: 10/06/2026, Published: 30/06/2026

Abstract−Background removal is commonly applied in image classification to suppress irrelevant visual information and emphasize the target object. However, its contribution to disease classification may depend on whether background elimination also modifies contextual, boundary, and texture information used by deep feature extractors. This study evaluates the effect of background removal on five-class mango fruit disease classification using a hybrid ResNet50–Support Vector Machine (SVM) framework. A paired experimental design was constructed from 834 matched MangoFruitDDS image pairs representing Alternaria, Anthracnose, Black Mould Rot, Healthy, and Stem-End Rot. The paired data were stratified into 667 training and 167 test samples for each image condition. ImageNet-pretrained ResNet50 was used as a frozen feature extractor, producing 2,048-dimensional feature vectors that were standardized and classified using an optimized SVM. Background-removed images achieved a higher cross-validation macro F1-score (0.7559) than original images (0.7340). However, evaluation on the identical test set showed higher performance for original images, with 76.05% accuracy, 0.7687 macro F1-score, and 0.9541 macro ROC-AUC, compared with 74.25%, 0.7445, and 0.9425, respectively, after background removal. The effect was class-dependent: Healthy improved after background removal, whereas several disease classes declined, suggesting that removing the background may alter contextual, boundary, or texture-related information contributing to ResNet50 feature representations. Exact McNemar testing showed no statistically significant difference between the paired predictions ($p = 0.783846$). These findings indicate that developers should not assume hard background removal universally improves plant disease classification; class-specific validation and feature-preserving or attention-based alternatives should be considered.

Keywords: Background Removal; Mango Fruit Disease; ResNet50; Support Vector Machine; Transfer Learning

1. INTRODUCTION

Mango (*Mangifera indica L.*) is one of the most economically and nutritionally important fruit commodities. However, the quality and productivity of mango fruit can be adversely affected by various diseases occurring during both the growing and postharvest stages. Several diseases commonly found in mango fruit include Alternaria, Anthracnose, Black Mould Rot, and Stem-End Rot. Anthracnose may cause irregular black lesions on the fruit surface, whereas Black Mould Rot can lead to the development of black fungal growth and tissue softening. Stem-End Rot typically develops from the stem end of the fruit and may result in discoloration, tissue softening, and postharvest decay. These conditions highlight the importance of accurate disease identification to maintain fruit quality and reduce postharvest losses [1].

Manual disease identification based on visual observation has several limitations, as different diseases may exhibit similar visual characteristics. Advances in computer vision and deep learning offer an alternative approach to classifying diseases using visual features extracted from images. Convolutional Neural Networks (CNNs) are capable of automatically learning patterns related to color, texture, shape, and other visual characteristics. In addition to training CNNs from scratch, transfer learning enables the use of models pre-trained on large-scale datasets to address classification problems with relatively limited data. Prabhu et al. demonstrated the effectiveness of this approach for mango disease classification by comparing various CNN and Transformer architectures. ResNet50 achieved a testing accuracy of 97.34%, while the Swin Transformer achieved 98.28% [2].

ResNet50 is a CNN architecture that can be employed not only as an end-to-end classifier but also as a deep feature extractor. Feature representations generated by CNNs can be combined with machine learning algorithms to construct hybrid classification models. One algorithm suitable for this purpose is the Support Vector Machine (SVM), which determines an optimal decision boundary between classes in the feature space. The use of SVM for plant disease classification has demonstrated competitive performance. Mookkandi et al., for example, combined Local Gabor XOR Pattern (LGXP) feature extraction with several SVM kernels. On a mango disease dataset, SVM performance was strongly influenced by the kernel employed; the linear kernel achieved a testing accuracy of 99.88%, whereas the coarse Gaussian kernel achieved 83.50%. These findings indicate that, in addition to feature quality, classifier configuration plays an important role in distinguishing diseases with closely related visual characteristics.

In addition to model selection, image preprocessing can substantially influence the visual representations generated by a classification model. Segmentation and background removal are commonly employed to suppress visual information outside the target object, with the expectation that reducing irrelevant background content allows the model to focus more effectively on disease-related regions. Bezabh et al. [3] ,for example, incorporated K-Means and Mask R-CNN-based segmentation into a mango disease classification pipeline and reported an accuracy of 99.21% using an ensemble of GoogLeNet and VGG16. Such findings demonstrate the potential value of

ISSN 3026-7315 (media online)
Volume 3, No 2, June 2026, Page: 100−114
DOI: doi.org/10.64366/ijids.v3i2.549
https://journals.adaresearch.or.id/ijids/index

segmentation as part of a computer vision pipeline; however, they do not establish that removing the background itself is universally beneficial.

This assumption presents an important methodological issue. In object-centered image classification, background pixels are often treated as irrelevant information that should be removed before feature extraction. However, convolutional neural networks generate feature representations from spatial patterns of pixel intensities, local gradients, edges, textures, and their hierarchical combinations. Consequently, hard foreground-background separation may change not only the amount of background information but also the visual distribution around object boundaries. Contextual information associated with object appearance, illumination, contrast, and boundary structure may therefore be altered during background removal. For disease classes whose discriminative characteristics depend on subtle surface textures or lesion boundaries, such transformations may affect the resulting deep feature representation rather than simply eliminate noise.

A study particularly relevant to this issue was conducted by Yehulu et al. [1] using MangoFruitDDS, the same dataset employed in the present study. Their pipeline applied background removal using *rembg* together with additional preprocessing procedures, including resizing, normalization, denoising, and augmentation, followed by feature extraction using MobileNetV3 and classification using several machine-learning algorithms. The MobileNetV3-Large and XGBoost combination achieved high classification performance. However, because background removal was incorporated together with several other preprocessing and modeling procedures, the independent contribution of background removal could not be isolated from the overall pipeline.

Therefore, a specific research gap remains regarding whether background removal itself provides a consistent advantage for mango fruit disease classification when sample identity, data partitioning, feature extraction, classifier configuration, and evaluation procedures are controlled. This issue is particularly relevant because the intuitive assumption that a cleaner object representation should improve classification may not hold when background removal simultaneously changes contextual or boundary-related visual information. Furthermore, comparisons based on different image samples may confound the effect of preprocessing with differences in sample composition. A controlled paired comparison is therefore required to determine whether the removal of background information produces a measurable improvement in classification performance [4], [5].

The primary contribution of this study is a controlled evaluation of background removal in mango fruit disease classification using a paired-image design. Unlike pipelines in which background removal is combined with multiple preprocessing operations, the present comparison maintains identical sample identities, data partitions, feature extraction architecture, classifier family, and evaluation procedures across the two image conditions. In addition to aggregate classification metrics, class-wise performance and confusion patterns are analyzed to determine whether the effect of background removal is consistent across disease categories. Finally, the Exact McNemar Test is applied to the paired test predictions to determine whether differences between the two image conditions are statistically significant. Through this design, the study investigates whether removing the background consistently improves ResNet50–SVM classification or whether its effect depends on the visual characteristics of individual disease classes.

2. RESEARCH METHODOLOGY

This study employed a controlled comparative experimental design to evaluate the effect of background removal on mango fruit disease classification using a hybrid deep learning–machine learning framework [7]. ImageNet-pretrained ResNet50 was employed as a frozen deep feature extractor, while Support Vector Machine (SVM) was used as the classifier. Two corresponding image conditions were evaluated: original images and background-removed images [8]. To minimize confounding effects arising from differences in sample identity, both conditions were constructed from matched image pairs and processed using identical data partitions, ResNet50 preprocessing, feature extraction procedures, classifier optimization strategies, and evaluation metrics. Thus, the principal experimental factor being compared was the image condition associated with the presence or removal of background information.

2.1 Research Stages

The research procedure consisted of dataset validation and image pairing, stratified data splitting, image preprocessing, ResNet50-based deep feature extraction, feature standardization, SVM hyperparameter optimization, test-set evaluation, and paired statistical analysis. The MangoFruitDDS dataset contained five mango fruit condition classes: Alternaria, Anthracnose, Black Mould Rot, Healthy, and Stem-End Rot. Two image conditions were considered: original images and their corresponding background-removed versions.

Initially, 862 original images and 838 background-removed images were identified. Because the number of images differed between the two conditions, a pairing procedure based on class labels and image identities was performed before data splitting. Only images with valid counterparts under both conditions were retained. This procedure resulted in 834 matched image pairs, ensuring that the comparison between the two conditions was based on the same mango fruit samples.

International Journal of Informatics and Data Science
ISSN 3026-7315 (media online)
Volume 3, No 2, June 2026, Page: 100−114
DOI: doi.org/10.64366/ijids.v3i2.549
https://journals.adaresearch.or.id/ijids/index

The 834 matched pairs were subsequently divided into training and testing subsets using a stratified 80:20 split with a random state of 42. This procedure produced 667 training samples and 167 testing samples for each image condition. Identical split indices were applied to the original and background-removed images so that corresponding images remained in the same subset. An identity-based overlap check between the training and testing subsets yielded an overlap of zero, indicating that no paired image identity was shared between the two subsets.

Each image was resized to 224 × 224 pixels and represented using three RGB channels. ResNet50-specific preprocessing was then applied before feature extraction. ImageNet-pretrained ResNet50 was used without its final classification layer, and its parameters were kept frozen throughout feature extraction. Global average pooling produced a 2,048-dimensional feature vector for each image. Consequently, each condition generated a training feature matrix of 667 × 2,048 and a testing feature matrix of 167 × 2,048.

The extracted features were standardized using parameters estimated from the training data and subsequently classified using SVM. Hyperparameter optimization was performed using GridSearchCV with 5-fold stratified cross-validation and macro F1-score as the selection criterion. The optimized models were then evaluated on the identical paired test samples using accuracy, macro precision, macro recall, macro F1-score, confusion matrix, and ROC-AUC. Finally, the Exact McNemar Test was applied to the paired predictions to determine whether the difference between the two image conditions was statistically significant [11].

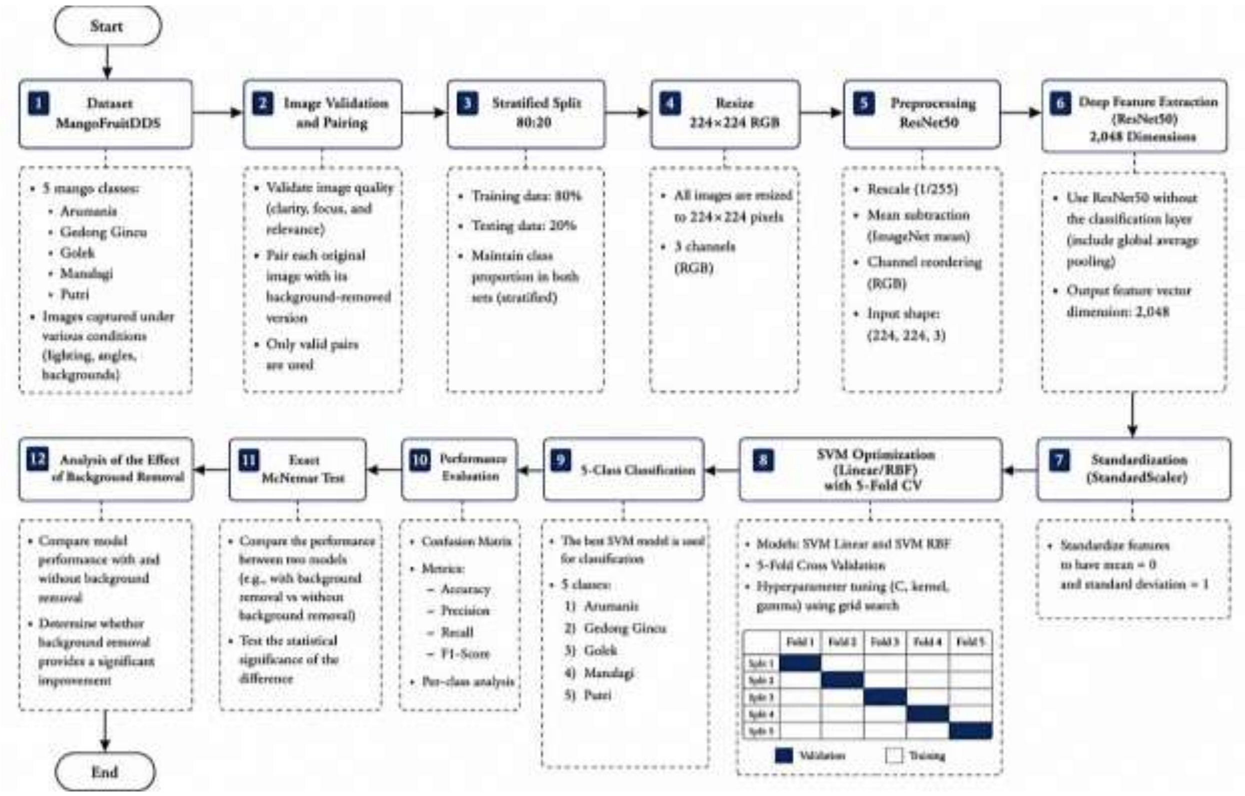

Figure 1. Research stages of mango fruit disease classification using ResNet50–SVM

Figure 1 illustrates the controlled paired experimental design employed in this study.. A single paired dataset is divided into two image conditions, namely Original Images and Background-Removed Images. Both conditions are then processed through an identical ResNet50–SVM pipeline and subsequently brought together at the evaluation stage for performance comparison and statistical significance testing using the Exact McNemar Test [12].

2.2 *Dataset* and *Preprocessing*

The dataset used in this study was MangoFruitDDS, which contains mango fruit images representing five conditions: Alternaria, Anthracnose, Black Mould Rot, Healthy, and Stem-End Rot. Two image representations were considered in the experiment: SenMangoFruitDDS_original, containing the original images, and SenMangoFruitDDS_bgremoved, containing their background-removed counterparts. According to the description of MangoFruitDDS reported by Yehulu et al. [1], background removal was performed using *rembg* as part of the dataset preprocessing procedure. In the present study, the available background-removed images were used as provided; therefore, the background-removal algorithm itself was not retrained or modified.

The original directory contained 862 images distributed across the five classes, whereas the background-removed directory contained 838 images. Because not every original image had a corresponding background-removed counterpart, the two directories were not directly used in their entirety. Instead, image validation and

International Journal of Informatics and Data Science
ISSN 3026-7315 (media online)
Volume 3, No 2, June 2026, Page: 100−114
DOI: doi.org/10.64366/ijids.v3i2.549
https://journals.adaresearch.or.id/ijids/index

pairing were performed using class labels and image identities. After this procedure, 834 valid image pairs were obtained. The final paired distribution consisted of 165 Alternaria, 127 Anthracnose, 182 Black Mould Rot, 205 Healthy, and 155 Stem-End Rot images under each condition, as presented in Table 1.

Table 1. Dataset distribution after the pairing process

Class	Original	Background Removed
Alternaria	165	165
Anthracnose	127	127
Black Mould Rot	182	182
Healthy	205	205
Stem End Rot	155	155
Total	834	834

Table 1 presents the distribution of paired images used in this experiment. To establish comparable experimental conditions, an image validation and matching process was performed based on class labels and the identity of each image. This stage also included an examination of file names that could potentially result in duplicate identities. After the validation and matching processes were completed, 834 valid image pairs available under both conditions were obtained.

Original Images

Background-Removed Images

Figure 2. Examples of original and background-removed images used in the study

Figure 2 presents representative images from the five classes under two conditions: the first row shows the original images, while the second row shows the corresponding background-removed images. The use of paired images was intended to control differences in sample composition. Each original image and its corresponding background-removed version represented the same underlying mango fruit sample. Therefore, identical sample identities and identical train-test assignments were maintained across the two experimental conditions.

An important consideration in this comparison is that background removal represents a transformation of the image rather than merely the deletion of unrelated pixels. Hard foreground-background separation can potentially modify pixel intensities around the fruit boundary and create sharper transitions between the foreground object and the replacement background. Such changes are relevant to CNN-based feature extraction because convolutional filters respond to local intensity variations, gradients, edges, and textures. Consequently, possible boundary alterations introduced by the background-removal process may contribute to differences between the feature representations extracted from the original and background-removed images.

However, the present study did not directly quantify boundary artifacts produced by *rembg*. Therefore, potential edge or boundary alterations are treated as a methodological consideration and a possible explanation for subsequent feature and classification differences rather than as an experimentally established artifact. This distinction is important because the primary objective of the present experiment was to evaluate the downstream classification effect of the two available image conditions, rather than to benchmark or optimize the background-removal algorithm itself.

The 834 paired images were subsequently partitioned using a stratified 80:20 split with a random state of 42, resulting in 667 training and 167 testing images for each condition. The same split indices were applied to both image representations. Before feature extraction, all images were resized to 224 × 224 pixels with three RGB channels and processed using the preprocessing function associated with ImageNet-pretrained ResNet50. Apart from the existing difference between the original and background-removed image representations, all subsequent preprocessing, feature extraction, classification, and evaluation procedures were kept identical between the two experimental conditions [13].

International Journal of Informatics and Data Science
ISSN 3026-7315 (media online)
Volume 3, No 2, June 2026, Page: 100−114
DOI: doi.org/10.64366/ijids.v3i2.549
https://journals.adaresearch.or.id/ijids/index

In this study, the original and background-removed images were not combined into a single training dataset. Instead, they were treated as two separate experimental conditions with identical image pairs and data partitions. This strategy was adopted to allow the effect of background removal on feature representations and classification performance to be directly evaluated without being confounded by differences in the number or identity of the samples.

2.3 ResNet50 Feature Extraction and SVM Classification

A hybrid deep learning–machine learning framework was employed by using ImageNet-pretrained ResNet50 as a frozen deep feature extractor and SVM as the final classifier. This configuration separates visual representation learning from classification: ResNet50 transforms each input image into a high-dimensional feature representation, whereas SVM determines the decision boundaries among the five mango fruit condition classes. [14] [15] .[16]

ResNet50 employs residual connections that allow information to propagate through deep convolutional layers. A residual block can be generally expressed as:

$$y = F(x, W) + x \quad (1)$$

where $x$denotes the input to the residual block, $F(x, W)$represents the residual transformation parameterized by W, and $y$denotes the block output. In this study, ResNet50 was initialized with ImageNet-pretrained weights and configured without its original classification layer. All ResNet50 parameters remained frozen, and no fine-tuning was performed. Each 224 × 224 × 3 input image was propagated through the convolutional backbone, and Global Average Pooling was applied to the final convolutional feature maps to obtain a 2,048-dimensional feature vector.

Feature extraction was performed separately on the training and testing sets for both image conditions. For the original images, the extraction process generated a feature matrix of 667 × 2,048 for the training set and 167 × 2,048 for the testing set. The same dimensions were obtained for the background-removed images. An examination of the resulting feature matrices confirmed the absence of Not a Number (NaN) and infinite (Inf) values, indicating that all extracted feature representations were valid for the subsequent classification stage.

Prior to SVM classification, the extracted feature vectors were standardized. For each feature dimension, standardization can be expressed as:

$$z_j = \frac{x_j - \mu_j^{train}}{\sigma_j^{train}} \quad (2)$$

where $x_j$is the original value of feature j, while $\mu_j^{train}$and $\sigma_j^{train}$are the mean and standard deviation estimated from the training data. The same training-derived parameters were subsequently applied to the test features to prevent information from the test set from influencing the transformation. The standardized feature vectors were classified using SVM. Both linear and Radial Basis Function (RBF) kernels were evaluated. For the nonlinear case, the RBF kernel is defined as:

$$K(x_i, x_j) = \exp(-\gamma \parallel x_i - x_j \parallel^2) \quad (3)$$

where γcontrols the influence of individual training samples in the feature space. Hyperparameter optimization was conducted using GridSearchCV with 5-fold stratified cross-validation. The evaluated values of Cwere 0.1, 1, 10, and 100, while the RBF gamma configurations included scale, auto, 0.001, and 0.01. Macro F1-score was used as the model-selection criterion to provide equal contribution from each class despite differences in class frequency.

2.4 Model Performance Evaluation

Model performance was evaluated using the testing data, which were not involved in the hyperparameter selection process. The evaluation metrics included accuracy, precision, recall, F1-score, confusion matrix, and ROC-AUC. For multiclass evaluation, precision, recall, and F1-score were calculated using macro averaging, ensuring that each class contributed equally to the final metric. Accuracy represents the proportion of correctly classified samples among all testing samples and is expressed as:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

Precision measures the model's ability to generate correct positive predictions:

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

Recall measures the model's ability to correctly identify actual positive samples:

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

Meanwhile, the F1-score represents the harmonic mean of precision and recall:

International Journal of Informatics and Data Science
ISSN 3026-7315 (media online)
Volume 3, No 2, June 2026, Page: 100−114
DOI: doi.org/10.64366/ijids.v3i2.549
https://journals.adaresearch.or.id/ijids/index

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

In addition to these metrics, a confusion matrix was employed to analyze classification error patterns for each class. This analysis is important for identifying disease classes with similar visual characteristics that are more likely to be confused by the model. The model's discriminative ability was also evaluated using the Receiver Operating Characteristic–Area Under the Curve (ROC-AUC). For the five-class classification problem, ROC-AUC was calculated using the One-vs-Rest (OvR) approach, allowing the model's ability to distinguish each individual class from the remaining classes to be analyzed separately.

Because the original and background-removed datasets used the same 167 paired test images, the evaluation was further complemented by the Exact McNemar Test [9]. This test was employed to determine whether differences in the patterns of correct and incorrect predictions between the two image conditions were statistically significant. The null hypothesis (H_0) stated that there was no significant difference in classification performance between the two conditions, whereas the alternative hypothesis (H_1) stated that a significant difference existed. The significance level was set at $\alpha = 0.05$. Therefore, H_0 was rejected when the p-value was < 0.05.

3. RESULT AND DISCUSSION

This study evaluated the effect of **background removal** on the classification of five mango fruit conditions using a hybrid ResNet50–SVM approach. The experiment was conducted under two image conditions, namely original images and background-removed images, using identical pairs of training and testing data. The evaluation was performed sequentially, starting from dataset validation, feature extraction using ResNet50, SVM hyperparameter optimization, performance testing on the test set, classification error analysis, and statistical evaluation using the Exact McNemar Test..

3.1 Dataset Analysis and Feature Extraction Results

The initial dataset under the original image condition consisted of 862 images covering five classes: Alternaria, Anthracnose, Black Mould Rot, Healthy, and Stem-End Rot. Meanwhile, the background-removed dataset contained fewer images. This difference in dataset size prevented the two datasets from being directly compared in their entirety because not all original images had corresponding counterparts in the background-removed condition. Therefore, a matching process based on class labels and image identities was conducted to obtain equivalent experimental samples.

The matching process resulted in 834 valid image pairs. The final distribution consisted of 165 Alternaria images, 127 Anthracnose images, 182 Black Mould Rot images, 205 Healthy images, and 155 Stem-End Rot images under each condition. By using identical image pairs, the effect of background removal could be analyzed without bias caused by differences in the objects included in the original and background-removed datasets.

Table 2. Distribution of images used in the experiment

Class	Original	Background Removed
Alternaria	165	165
Anthracnose	127	127
Black Mould Rot	182	182
Healthy	205	205
Stem End Rot	155	155
Total	834	834

As shown in Table 2, the 834 image pairs were subsequently divided using an 80:20 ratio, resulting in 667 training samples and 167 testing samples under each condition. Examination of the image identities showed an overlap value of 0 between the training and testing sets. This result indicated that there was no evidence of data leakage during the dataset splitting process.

Following preprocessing, ResNet50 was used to extract visual representations from each image. Each image produced a feature vector containing 2,048 features, resulting in a matrix of 667 × 2,048 for the training set and 167 × 2,048 for the testing set. The same dimensions were obtained for both the original and background-removed images. Examination of the extracted features also showed that all feature matrices contained 0 NaN values and 0 Inf values, indicating that all extracted features were valid for use in the classification process.

Table 3. Feature extraction results using ResNet50

Condition	Data	Number of Samples	Feature Dimension	NaN	Inf
Original	Training	667	2,048	0	0
Original	Testing	167	2,048	0	0
Background Removed	Training	667	2,048	0	0
Background Removed	Testing	167	2,048	0	0

International Journal of Informatics and Data Science
ISSN 3026-7315 (media online)
Volume 3, No 2, June 2026, Page: 100−114
DOI: doi.org/10.64366/ijids.v3i2.549
https://journals.adaresearch.or.id/ijids/index

The identical sample sizes and feature dimensions under both conditions provide a consistent basis for comparison. Because the same paired image identities, data partitions, ResNet50 architecture, and feature dimensions were maintained, differences observed at the classification stage can be examined in association with the two image representations. However, identical feature dimensionality does not imply identical feature distributions. Removing background pixels and altering foreground-background boundaries may change the activation patterns produced by convolutional filters, potentially shifting the location of samples within the 2,048-dimensional ResNet50 feature space. The present experiment does not directly quantify this representational shift; therefore, this mechanism is considered a plausible interpretation that is further examined through class-wise performance and error patterns.

3.2 SVM Hyperparameter Optimization Results

The features extracted by ResNet50 were subsequently standardized using StandardScaler[17] and classified using SVM. Hyperparameter optimization [18] was performed using GridSearchCV with 5-fold stratified cross-validation. The kernels compared were linear and RBF, with several combinations of C and gamma values. The macro F1-score was used as the model selection criterion because the evaluation involved five classes with sample sizes that were not perfectly balanced.

The optimization results showed that the RBF kernel was the best-performing kernel under both image conditions. For the original images, the optimal configuration was obtained using C = 10 and gamma = scale, achieving a cross-validation macro F1-score of 0.7340. Meanwhile, the background-removed images achieved their best configuration using C = 10 and gamma = auto, with a macro F1-score of 0.7559.

Table 4. SVM hyperparameter optimization results

Dataset	Best Kernel	C	Gamma	CV Macro F1
Original	RBF	10	scale	0,7340
Background Removed	RBF	10	auto	0,7559

Table 4 During cross-validation, the background-removed condition achieved a macro F1-score of 0.7559, which was 0.0219 higher than the 0.7340 obtained from the original images. This result suggests that, within the training folds, reducing background variation may have produced feature distributions that were more favorable to the SVM decision boundaries. By suppressing variations outside the fruit region, background removal may reduce part of the within-class variability unrelated to the fruit surface.

However, this cross-validation advantage should not be interpreted as evidence that background removal generally produces superior feature representations. Background removal simultaneously changes the pixel distribution of the input image, particularly around the foreground-background transition. Because ResNet50 extracts hierarchical representations from local gradients, textures, shapes, and their spatial combinations, such changes may affect the resulting feature vectors differently across samples and classes. Therefore, evaluation on an independent paired test set remains necessary to determine whether the cross-validation advantage generalizes to unseen images.

3.3 Model Performance Evaluation on the Test Set

Testing was conducted using the same 167 images under each condition. The results showed a different pattern from that observed during cross-validation. Although the background-removed condition achieved a higher macro F1-score during cross-validation, the model trained using original images produced numerically higher performance across all aggregate evaluation metrics on the test set.

Table 5. Comparison of ResNet50–SVM performance on the test set

Metric	Original	Background Removed	Difference (BG − Original)
Accuracy	0,7605	0,7425	−0,0180
Macro Precision	0,7784	0,7461	−0,0323
Macro Recall	0,7624	0,7441	−0,0183
Macro F1-score	0,7687	0,7445	−0,0243
Macro ROC-AUC	0,9541	0,9425	−0,0116

Table 5 shows that the model using original images achieved an accuracy of 76.05%, whereas the background-removed condition achieved 74.25%, representing a decrease of approximately 1.80 percentage points. A larger difference was observed in macro precision, which reached 77.84% for the original images and 74.61% for the background-removed images. The macro F1-score also decreased from 76.87% to 74.45%, while the macro ROC-AUC decreased from 95.41% to 94.25%.

The reversal between cross-validation and test-set performance provides evidence that the effect of background removal was not consistently beneficial. Original images achieved a test accuracy of 0.7605 and macro F1-score of 0.7687, whereas the background-removed images achieved 0.7425 and 0.7445, respectively. Macro

International Journal of Informatics and Data Science
ISSN 3026-7315 (media online)
Volume 3, No 2, June 2026, Page: 100−114
DOI: doi.org/10.64366/ijids.v3i2.549
https://journals.adaresearch.or.id/ijids/index

ROC-AUC also decreased from 0.9541 to 0.9425. Thus, the feature representation that appeared more favorable during cross-validation did not produce a corresponding generalization advantage on the independent test set.

One plausible explanation involves a change in the statistical and spatial characteristics of the input after background removal. In the original images, fruit boundaries are formed naturally by the contrast between the fruit and its surrounding scene. After hard background removal, pixels outside the object are suppressed or replaced, producing a substantially different intensity distribution and potentially sharper foreground-background transitions. Because convolutional filters respond to local intensity differences, gradients, edges, and texture patterns, this transformation may modify activations not only in background regions but also around the contour of the fruit. These changes can propagate through successive ResNet50 layers and result in different high-level feature vectors.

Importantly, this interpretation does not imply that background information itself is necessarily disease-specific. Rather, it suggests that removing the background changes the complete visual input from which ResNet50 constructs its representation. Consequently, the resulting feature-space geometry available to the SVM may differ between the two conditions. Depending on the visual characteristics of a class, this representational change may either increase or decrease class separability.

These results indicate that background removal did not improve overall generalization performance under the ResNet50–SVM configuration used in this study. The findings also demonstrate that successful preprocessing performance on training data or during cross-validation does not necessarily translate into improved performance on an independent test set. In other words, the increase in cross-validation macro F1-score for the background-removed condition was not maintained when the model was evaluated on unseen data. A class-wise analysis provides a more detailed view of these changes.

Table 6. Precision, recall, and F1-score for each class

Class	Prec. Original	Recall Original	F1 Original	Prec. BG	Recall BG	F1 BG
Alternaria	0,4857	0,5152	0,5000	0,5152	0,5152	0,5152
Anthracnose	0,9583	0,9200	0,9388	0,8148	0,8800	0,8462
Black Mould Rot	0,6410	0,6757	0,6579	0,5385	0,5676	0,5526
Healthy	0,8837	0,9268	0,9048	1,0000	0,9512	0,9750
Stem End Rot	0,9231	0,7742	0,8421	0,8621	0,8065	0,8333

Table 6 demonstrates that the effect of background removal was not uniform across all classes. The class-wise results demonstrate that background removal did not affect all categories in the same direction. The most notable improvement occurred for the Healthy class, whose F1-score increased from 0.9048 to 0.9750, while precision increased from 0.8837 to 1.0000. Alternaria also showed a small F1-score increase from 0.5000 to 0.5152. In contrast, Anthracnose decreased from 0.9388 to 0.8462, Black Mould Rot decreased from 0.6579 to 0.5526, and Stem-End Rot slightly decreased from 0.8421 to 0.8333.

The improvement observed for Healthy may be associated with greater visual consistency after background removal. Healthy fruit images are expected to be distinguished primarily through the appearance of the fruit surface rather than the presence of disease lesions. Suppressing heterogeneous background regions may therefore reduce visual variability unrelated to the fruit itself and allow the extracted representation to become more concentrated on fruit color, shape, and surface characteristics. This interpretation is consistent with the increase in Healthy precision to 1.0000 and the improvement in its ROC-AUC from 0.9861 to 0.9957.

The disease categories present a different problem because their discrimination depends more strongly on localized visual patterns. Lesion color, texture, shape, distribution, and contrast relative to surrounding fruit tissue may contribute to the feature representation. Background removal does not explicitly preserve or enhance these disease-specific characteristics; instead, it modifies the image globally outside the estimated foreground and potentially around the fruit contour. Therefore, a cleaner background does not necessarily imply improved separation between disease classes whose internal surface patterns remain visually similar.

This effect is particularly evident for Alternaria and Black Mould Rot. Their F1-scores under the background-removed condition were only 0.5152 and 0.5526, respectively. The limited improvement for Alternaria and substantial decline for Black Mould Rot indicate that eliminating background information did not resolve the visual overlap between these categories. Instead, the transformation may have altered the balance between global object-level information and localized disease-related texture information represented by ResNet50. Consequently, samples from these visually similar classes may remain close or become less separable within the extracted feature space.

These class-dependent changes support the interpretation of a possible feature representational shift rather than a uniform preprocessing benefit. In this context, feature representational shift refers to the possibility that the same mango fruit, when presented with different background conditions, produces a different position in the ResNet50 feature space because the input pixel distribution, boundary contrast, and contextual information have changed. Since the present study did not directly measure distances between paired feature vectors or visualize their distributions, this mechanism should be interpreted as a hypothesis consistent with the observed classification patterns rather than as direct evidence of a causal feature-space shift.

International Journal of Informatics and Data Science
ISSN 3026-7315 (media online)
Volume 3, No 2, June 2026, Page: 100−114
DOI: doi.org/10.64366/ijids.v3i2.549
https://journals.adaresearch.or.id/ijids/index

3.4 Confusion Matrix Analysis

The model's error patterns were further analyzed using the confusion matrix. For the original images, the numbers of correctly classified samples for Alternaria, Anthracnose, Black Mould Rot, Healthy, and Stem-End Rot were 17, 23, 25, 38, and 24 images, respectively. Therefore, 127 out of 167 images were correctly classified.

For the background-removed condition (Figure 4), the numbers of correctly classified samples for Alternaria, Anthracnose, Black Mould Rot, Healthy, and Stem-End Rot were 17, 22, 21, 39, and 25, respectively, resulting in 124 correctly classified images out of 167 test sample.

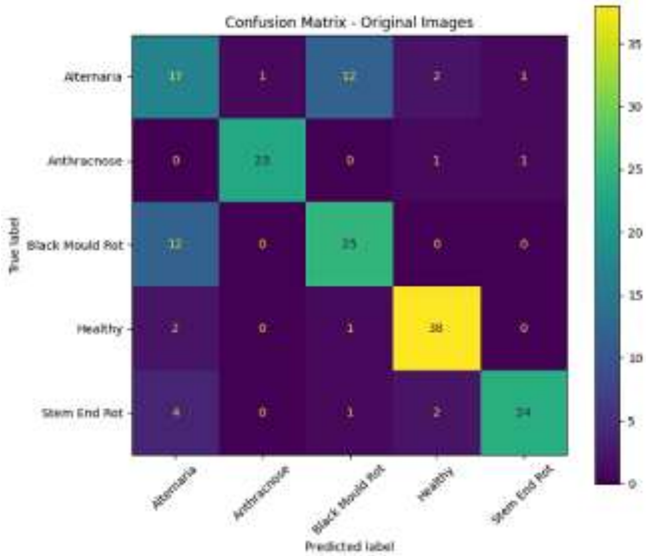

Figure 3. Confusion matrix for classification using original images

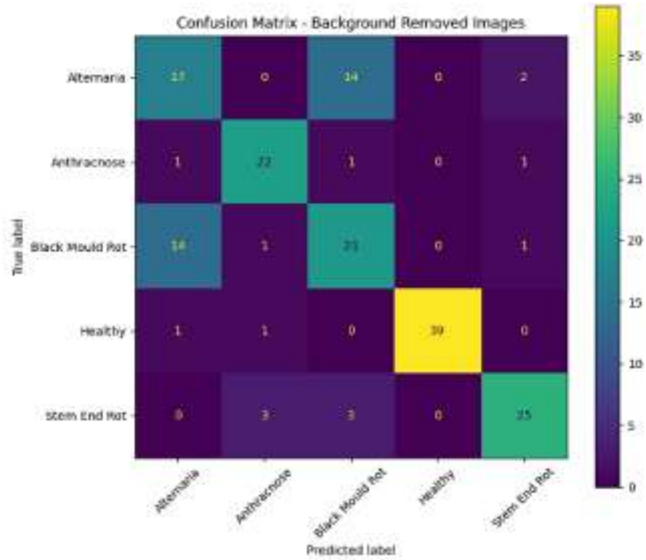

Figure 4. Confusion matrix for classification using background-removed images

Figure 4 provides further evidence that the effect of background removal was class-dependent. Under the original-image condition, 12 Alternaria samples were misclassified as Black Mould Rot and 12 Black Mould Rot samples were misclassified as Alternaria. After background removal, these bidirectional errors increased to 14 samples in each direction. Therefore, background removal did not improve the distinction between the two most frequently confused disease categories; instead, their observed confusion increased.

This pattern is important because removing the background would be expected to improve classification if background complexity were the primary source of confusion. The persistence and increase of Alternaria–Black Mould Rot errors suggest that the main classification difficulty is more likely associated with overlapping fruit-surface characteristics than with background variation alone. Furthermore, if background removal changes boundary contrast or the relative contribution of global and local image patterns to ResNet50 activations, the resulting representation may not preserve the same class-separation structure as the original images.

In contrast, the number of correctly classified Healthy samples increased from 38 to 39 out of 41, while Stem-End Rot increased from 24 to 25 out of 31. These modest improvements demonstrate that suppressing

International Journal of Informatics and Data Science
ISSN 3026-7315 (media online)
Volume 3, No 2, June 2026, Page: 100−114
DOI: doi.org/10.64366/ijids.v3i2.549
https://journals.adaresearch.or.id/ijids/index

background variation may benefit some image categories. However, they were insufficient to compensate for the additional errors in other disease classes. Taken together, the confusion matrices indicate that the usefulness of background removal depends on the discriminative visual information required for each class rather than simply on the amount of background present in the image.

3.5 ROC-AUC Analysis

The model's discriminative ability for each class was analyzed using ROC-AUC with the One-vs-Rest approach. Overall, both image conditions produced high AUC values, although the effect of background removal varied across classes.

Table 7. Comparison of ROC-AUC for each class

Class	Original	Background Removed	Difference
Alternaria	0,8912	0,8917	+0,0005
Anthracnose	0,9989	0,9744	−0,0245
Black Mould Rot	0,9345	0,8931	−0,0414
Healthy	0,9861	0,9957	+0,0097
Stem End Rot	0,9599	0,9578	−0,0021

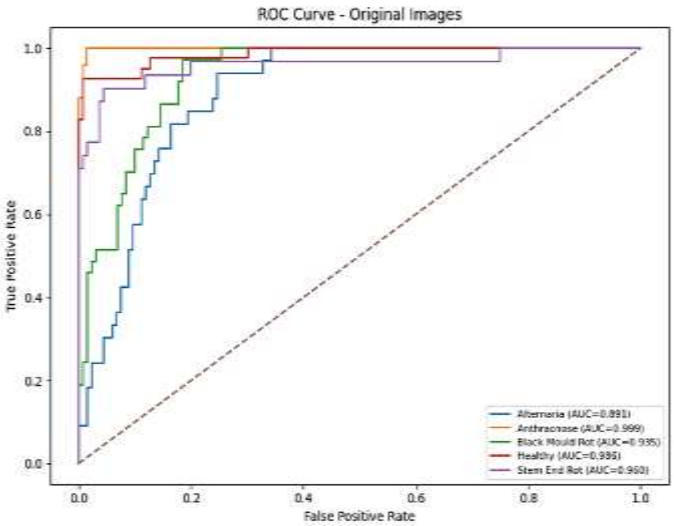

Figure 5. Multiclass ROC curves for original images

As shown in Figure 5, Anthracnose achieved the highest AUC value of 0.9989, followed by Healthy at 0.9861, Stem-End Rot at 0.9599, Black Mould Rot at 0.9345, and Alternaria at 0.8912. These values indicate that the model had very high discriminative ability, particularly for Anthracnose and Healthy.

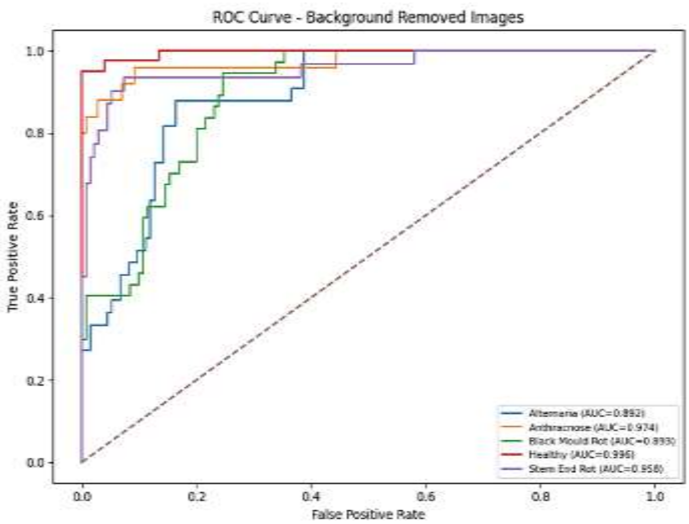

Figure 6. Multiclass ROC curves for background-removed images

As shown in Figure 6, after background removal, the AUC for Healthy increased to 0.9957, while Alternaria showed a very slight increase to 0.8917. In contrast, the AUC for Anthracnose decreased to 0.9744, Black Mould

ISSN 3026-7315 (media online)
Volume 3, No 2, June 2026, Page: 100−114
DOI: doi.org/10.64366/ijids.v3i2.549
https://journals.adaresearch.or.id/ijids/index

Rot decreased to 0.8931, and Stem-End Rot decreased to 0.9578. The largest decrease occurred in Black Mould Rot, with a reduction of 0.0414, which is consistent with the increased misclassification of this class observed in the confusion matrix.

At the aggregate level, the macro ROC-AUC for the original images reached 0.9541, whereas the background-removed condition achieved 0.9425. Although both values indicate high discriminative ability, these results again demonstrate that background removal did not improve overall performance under the experimental configuration used in this study

The class-wise AUC changes are consistent with the confusion-matrix and F1-score patterns. Healthy showed an AUC increase from 0.9861 to 0.9957, whereas Black Mould Rot exhibited the largest decrease, from 0.9345 to 0.8931. Anthracnose also decreased from 0.9989 to 0.9744. This consistency across different evaluation measures suggests that the observed class-dependent effects are not limited to a single classification metric.

Nevertheless, ROC-AUC describes discriminative performance rather than the internal mechanism responsible for the change. Therefore, the AUC differences should not be interpreted as direct evidence that particular background pixels or boundary artifacts caused the performance shift. Instead, when considered together with the F1-scores and confusion matrices, the results support the broader interpretation that the transformation from original to background-removed images altered the information available to the ResNet50–SVM pipeline in a class-dependent manner

3.6 Analysis of the Effect of Background Removal and Statistical Testing

Taken together, the cross-validation, test-set, class-wise, confusion-matrix, and ROC-AUC results show that background removal produced a non-uniform effect on mango fruit disease classification. During model selection, the background-removed condition achieved a higher cross-validation macro F1-score (0.7559) than the original condition (0.7340). However, this advantage was not retained on the independent paired test set, where the original images achieved higher accuracy, macro F1-score, and macro ROC-AUC.

The class-wise results provide additional insight into this reversal. Background removal improved the Healthy class but reduced performance for Anthracnose and Black Mould Rot, with only a minor change for Alternaria and Stem-End Rot. This pattern suggests that background removal should not be viewed merely as noise reduction. The operation changes the pixel distribution of the input image and may modify foreground-background transitions, local gradients, and contextual relationships. Because ResNet50 transforms these visual patterns hierarchically into a 2,048-dimensional representation, changes introduced at the input level may propagate to the feature representation subsequently used by the SVM.

From this perspective, the improved Healthy performance may reflect a situation in which suppressing background variability makes the fruit representation more visually consistent. Conversely, disease categories requiring discrimination based on subtle lesion texture and color patterns may not benefit from such simplification. The increased bidirectional confusion between Alternaria and Black Mould Rot after background removal further suggests that their classification difficulty originates primarily from overlapping disease-related visual characteristics rather than background complexity alone.

However, these mechanisms should be interpreted cautiously. The present study did not directly quantify changes in boundary pixels, convolutional activation maps, or distances between paired 2,048-dimensional feature vectors. Therefore, the observed class-dependent performance differences are consistent with a possible shift in feature representation, but they do not establish a causal relationship between specific background pixels, segmentation artifacts, and classification errors. Future feature-space analysis using paired feature distances, dimensionality-reduction visualization, or activation-based interpretation could test this hypothesis directly.

To determine whether the observed differences in paired classification outcomes were statistically significant, the Exact McNemar Test was applied to the 167 corresponding test samples. There were 28 cases in which the original-image model was correct and the background-removed model was incorrect, compared with 25 cases showing the opposite outcome, resulting in 53 discordant pairs. The Exact McNemar Test yielded a p-value of 0.783846, which exceeded the significance threshold of $\alpha = 0.05$. Therefore, the null hypothesis was not rejected.

The statistical result is particularly important for interpreting the numerical advantage of the original images. Although original images achieved 1.80 percentage points higher accuracy and 2.43 percentage points higher macro F1-score, these differences should not be interpreted as evidence that original images are statistically superior. Instead, the results indicate that background removal failed to provide a consistent and statistically significant improvement under the ResNet50–SVM configuration evaluated in this study. The practical implication is therefore not that background information should always be retained, but that background removal should be validated empirically for the target dataset and disease classes rather than adopted automatically as a universal preprocessing step

Table 8. Exact McNemar Test results

Parameter	Result
Original correct, BG incorrect	28
Original incorrect, BG correct	25

ISSN 3026-7315 (media online)
Volume 3, No 2, June 2026, Page: 100−114
DOI: doi.org/10.64366/ijids.v3i2.549
https://journals.adaresearch.or.id/ijids/index

Parameter	Result
Discordant pairs	53
p-value	0,783846
α	0,05
Decision	Not significant

The Exact McNemar Test produced a p-value of 0.783846, which was greater than the significance level of 0.05. Therefore, the null hypothesis was not rejected. This result indicates that there was no statistically significant difference in classification performance between the original and background-removed image conditions.

This finding is important when interpreting the numerical differences of 1.80 percentage points in accuracy and 2.43 percentage points in macro F1-score in favor of the original images. These numerical differences are insufficient to conclude that the original-image condition is statistically superior. Instead, the results indicate that, for the dataset and ResNet50–SVM configuration used in this study, background removal did not provide a consistent or statistically significant improvement in classification performance.

Overall, the findings suggest that background removal should not be regarded as a preprocessing step that will always improve mango fruit disease classification. Its effect depends on the class being analyzed. The Healthy class benefited relatively clearly from background removal, whereas several disease classes experienced decreased performance. Therefore, the selection of preprocessing techniques should consider whether the process preserves visual characteristics that are relevant for distinguishing each disease class, rather than merely reducing background complexity.

3.7 Discussion

Previous studies have demonstrated the effectiveness of deep learning, transfer learning, segmentation, and machine-learning classifiers for mango disease recognition. Yehulu et al. [1] investigated five mango fruit conditions in the MangoFruitDDS dataset using MobileNetV3-Large combined with several machine-learning classifiers. Their best-performing configuration, MobileNetV3-Large combined with XGBoost, achieved an accuracy of 98.95%. Background removal was incorporated as part of their preprocessing pipeline; however, the study did not perform a controlled paired comparison between original and background-removed representations of the same images. Consequently, although the study demonstrated the effectiveness of the overall preprocessing and classification pipeline, the independent contribution of background removal could not be determined.

Mookkandi et al. [3] investigated plant disease classification, including mango diseases, using Local Gabor XOR Pattern (LGXP) features combined with SVM. Their findings demonstrated that SVM can provide strong classification performance when supported by discriminative visual features. This finding supports the use of SVM for plant disease recognition, although the feature extraction strategy differs from that employed in the present study. Rather than relying on handcrafted LGXP descriptors, the present framework uses 2,048-dimensional deep feature representations extracted using ImageNet-pretrained ResNet50 and subsequently classified using an optimized SVM.

Segmentation-based approaches have also demonstrated strong performance in mango disease recognition. Bezabh et al. [19] employed K-Means and Mask R-CNN-based segmentation together with an ensemble of GoogLeNet and VGG16, achieving an accuracy of 99.21%. Their results demonstrate that segmentation can contribute to a high-performing disease classification pipeline. However, segmentation was incorporated as part of the overall framework rather than evaluated as an independent experimental factor. In contrast, the present study specifically investigates the effect of background removal while maintaining matched image identities, identical train-test partitions, the same ResNet50 feature extraction procedure, and the same SVM-based classification framework across the two image conditions.

Transfer-learning studies further demonstrate the effectiveness of pretrained feature representations for mango disease recognition. Prabhu et al. [2]evaluated ResNet50 and several other transfer-learning architectures and showed that pretrained deep models can effectively capture discriminative characteristics of mango diseases. Similarly, Hossain et al. [20] investigated CNN and Vision Transformer architectures using MangoLeafBD and reported an accuracy of 99.75% with an optimized DeiT model. These studies primarily focused on improving classification performance through model selection and optimization. In contrast, the objective of the present study is not to identify the highest-performing deep learning architecture, but to determine whether background removal itself provides a measurable advantage when the feature extraction and classification procedures are controlled.

In the present study, 834 paired MangoFruitDDS images were evaluated under original and background-removed conditions using ImageNet-pretrained ResNet50 as a frozen feature extractor and an optimized RBF-SVM classifier. The original-image condition achieved an accuracy of 76.05%, whereas the background-removed condition achieved 74.25%. However, these numerical differences should not be interpreted as evidence that the original-image condition is statistically superior. The Exact McNemar Test yielded a p-value of 0.783846, which was greater than the significance level of $\alpha = 0.05$, indicating that the difference between the paired classification outcomes was not statistically significant.

More importantly, the class-wise analysis demonstrated that the effect of background removal was not uniform. The Healthy class benefited from background removal, whereas performance decreased for several disease

ISSN 3026-7315 (media online)
Volume 3, No 2, June 2026, Page: 100−114
DOI: doi.org/10.64366/ijids.v3i2.549
https://journals.adaresearch.or.id/ijids/index

classes, particularly Anthracnose and Black Mould Rot. The increased bidirectional confusion between Alternaria and Black Mould Rot after background removal further suggests that eliminating background information does not necessarily resolve similarities in disease-related surface characteristics. These results complement previous segmentation-based studies by demonstrating that a preprocessing operation that is beneficial as part of a complete classification pipeline may not necessarily provide an independent advantage when evaluated under a controlled paired design.

Therefore, the contribution of this study should not be interpreted primarily in terms of achieving higher classification accuracy than previous mango disease studies, because direct comparison of accuracy values across different datasets, disease targets, preprocessing procedures, and model architectures would be inappropriate. Instead, the contribution lies in demonstrating through a controlled paired comparison, class-wise analysis, and statistical validation that hard background removal cannot be assumed to provide a universal classification advantage. The findings suggest that preprocessing effectiveness depends not only on removing visually irrelevant information but also on preserving contextual, boundary, and texture-related patterns that may contribute to deep feature representations. Accordingly, background removal should be empirically evaluated according to the characteristics of the target dataset and disease classes rather than automatically adopted as a standard preprocessing step.

4. CONCLUSION

This study evaluated the effect of background removal on five-class mango fruit disease classification using a hybrid ResNet50–SVM framework through a controlled paired experimental design. A total of 834 matched image pairs from MangoFruitDDS were divided using an identical stratified 80:20 split, resulting in 667 training and 167 test samples for each image condition. Although the background-removed condition achieved a higher cross-validation macro F1-score of 0.7559 compared with 0.7340 for the original images, this advantage was not maintained on the independent paired test set. Original images achieved an accuracy of 76.05%, macro F1-score of 0.7687, and macro ROC-AUC of 0.9541, whereas background-removed images achieved 74.25%, 0.7445, and 0.9425, respectively. The effect of background removal was class-dependent: the Healthy class improved, while performance declined for several disease classes, particularly Anthracnose and Black Mould Rot. The increased confusion between Alternaria and Black Mould Rot after background removal further indicates that eliminating background information does not necessarily improve the separability of visually similar disease categories. These findings are consistent with the possibility that hard background removal changes contextual information, foreground-background transitions, and texture- or edge-related visual patterns contributing to ResNet50 feature representations; however, such representational changes were not directly quantified in this study and should therefore not be interpreted as a demonstrated causal mechanism. The Exact McNemar Test produced a p-value of 0.783846, indicating that the numerical performance differences between the two image conditions were not statistically significant at $\alpha = 0.05$. Consequently, background removal should not be assumed to universally improve plant disease classification and should instead be validated according to the target dataset, disease class, and feature extraction architecture. Future research should investigate attention-based mechanisms, including spatial and channel attention, as alternatives to hard background segmentation. Rather than completely discarding contextual information, attention mechanisms may enable the model to emphasize disease-relevant fruit regions while retaining potentially useful contextual and boundary information. Further studies should also compare original images, hard-segmented images, and attention-guided representations while incorporating feature-space analysis, boundary-artifact assessment, explainable AI techniques, and fine-tuning to better determine how preprocessing strategies influence disease-specific visual representations and classification performance.

5. DECLARATIONS

5.1 Author Contributions

Conceptualization: P.R, I.H.I, and Saprudin;
Methodology: P.R, I.H.I, and Saprudin;
Software: P.R;
Validation: P.R, I.H.I, and S;
Formal Analysis: P.R, I.H.I, and S;
Investigation: P.R, I.H.I, and S;
Resources: P.R, I.H.I, and S;
Data Curation: P.R;
Writing Original Draft Preparation: P.R;
Writing Review and Editing: P.R, I.H.I, and S;
Visualization: P.R.
All authors have read and agreed to the published version of the manuscript.

ISSN 3026-7315 (media online)
Volume 3, No 2, June 2026, Page: 100−114
DOI: doi.org/10.64366/ijids.v3i2.549
https://journals.adaresearch.or.id/ijids/index

5.2 Data Availability

The dataset used in this study is publicly available through the MangoFruitDDS dataset repository on Kaggle. The processed data and experimental results generated during this study are available from the corresponding author upon reasonable request.

5.3 Funding

This research received no external funding. The publication and manuscript processing fees were independently covered by the authors.

5.4 Institutional Review Board Statement

Not applicable.

5.5 Informed Consent Statement

Not applicable.

5.6 Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

5.7 Generative AI and AI-assisted Technologies in the Writing Process

The authors used generative AI and AI-assisted technologies to improve the clarity, grammar, and readability of the manuscript. All AI-assisted content was subsequently reviewed, verified, and edited by the authors. The authors take full responsibility for the accuracy, integrity, and final content of the published manuscript.

REFERENCES

[1] G. T. Yehulu *et al.*, “Mango fruit disease detection and classification using MobileNetV3_large model,” *Sci. African*, vol. 30, no. October, p. e03061, 2025, doi: 10.1016/j.sciaf.2025.e03061.
[2] O. Prabhu, M. T, and S. Shetty, “An investigation on advances in transfer learning and explainable AI for mango leaf disease detection,” *Smart Agric. Technol.*, vol. 12, no. May, p. 101348, 2025, doi: 10.1016/j.atech.2025.101348.
[3] K. Mookkandi, M. K. Nath, and K. Balakrishnan, “Robust method for classification of agricultural crops diseases using LGXP descriptor,” *Results Eng.*, vol. 29, no. November 2025, p. 108826, 2026, doi: 10.1016/j.rineng.2025.108826.
[4] R. A. Rizvee *et al.*, “LeafNet: A proficient convolutional neural network for detecting seven prominent mango leaf diseases,” *J. Agric. Food Res.*, vol. 14, no. August, p. 100787, 2023, doi: 10.1016/j.jafr.2023.100787.
[5] R. Amalia, A. F. Zaidan, S. Ramadhan, F. Septian, A. M. Aqsha, and P. Rosyani, “Classification of Autoimmune Diseases Using the K-Nearest Neighbors Algorithm,” *Formosa J. Sci. Technol.*, vol. 4, no. 1, pp. 337–348, Jan. 2025, doi: 10.55927/fjst.v4i1.13443.
[6] M. Q. R. Pembury Smith and G. D. Ruxton, “Effective use of the McNemar test,” *Behav. Ecol. Sociobiol.*, vol. 74, no. 11, 2020, doi: 10.1007/s00265-020-02916-y.
[7] U. Sanath Rao *et al.*, “Deep Learning Precision Farming: Grapes and Mango Leaf Disease Detection by Transfer Learning,” *Glob. Transitions Proc.*, vol. 2, no. 2, pp. 535–544, 2021, doi: 10.1016/j.gltp.2021.08.002.
[8] M. T. Rahman, S. D. Dipto, I. J. June, A. Momin, and M. R. Al Mamun, “Machine Learning-based Disease Classification in Tomato (Solanum lycopersicum) Plants,” *J. Keteknikan Pertan. Trop. dan Biosist.*, vol. 12, no. 3, pp. 151–160, Dec. 2024, doi: 10.21776/ub.jkptb.2024.012.03.01.
[9] A. Shahcheraghian, A. Ilinca, and N. Sommerfeldt, “K-means and agglomerative clustering for source-load mapping in distributed district heating planning,” *Energy Convers. Manag. X*, vol. 25, no. October 2024, p. 100860, 2025, doi: 10.1016/j.ecmx.2024.100860.
[10] K. N. Rahman, S. C. Banik, R. Islam, and A. Al Fahim, “A real time monitoring system for accurate plant leaves disease detection using deep learning,” *Crop Des.*, vol. 4, no. 1, Feb. 2025, doi: 10.1016/j.cropd.2024.100092.
[11] G. Psaltakis, K. Rogdakis, M. Loizos, and E. Kymakis, “One-vs-One, One-vs-Rest, and a novel Outcome-Driven One-vs-One binary classifiers enabled by optoelectronic memristors towards overcoming hardware limitations in multiclass classification,” *Discov. Mater.* , vol. 4, no. 1, 2024, doi: 10.1007/s43939-024-00077-7.
[12] M. M. Islam *et al.*, “A two-stage model for enhanced mango leaf disease detection using an innovative handcrafted spatial feature extraction method and knowledge distillation process,” *Ecol. Inform.*, vol. 91, no. July, p. 103365, 2025, doi: 10.1016/j.ecoinf.2025.103365.
[13] F. Mahrus Fathoni, “Klasifikasi Penyakit Daun Tomat Menggunakan Algoritma K-NN Berdasarkan Ekstraksi Fitur GLCM dan LBP,” *J. Tek. Inform. dan Teknol. Inf.*, vol. 4, no. 1, pp. 39–50, Jan. 2024, doi: 10.55606/jutiti.v4i1.3417.
[14] T. Hu, M. Khishe, M. Mohammadi, G. R. Parvizi, S. H. Taher Karim, and T. A. Rashid, “Real-time COVID-19 diagnosis from X-Ray images using deep CNN and extreme learning machines stabilized by chimp optimization algorithm,” *Biomed. Signal Process. Control*, vol. 68, Jul. 2021, doi: 10.1016/j.bspc.2021.102764.
[15] A. Venkataramana, K. Suresh Kumar, N. Suganthi, and R. Rajeswari, “Prediction of Brinjal Plant Disease Using Support Vector Machine and Convolutional Neural Network Algorithm Based on Deep Learning,” *J. Mob. Multimed.*, vol. 18, no. 3, pp. 771–788, 2022, doi: 10.13052/jmm1550-4646.18315.
[16] M. Damar, A. Putra, and S. Rakasiwi, “Penerapan Model SVM dengan Ekstraksi Fitur ResNet50 untuk Identifikasi Sel Darah Terinfeksi Malaria,” vol. 7, no. 3, pp. 1866–1874, 2026, doi: 10.47065/bits.v7i3.8750.
[17] C. Wang, “ResNet50 Feature Extraction with SVM and kNN for Three-Class Dog Heart Size Classification”.

ISSN 3026-7315 (media online)
Volume 3, No 2, June 2026, Page: 100−114
DOI: doi.org/10.64366/ijids.v3i2.549
https://journals.adaresearch.or.id/ijids/index

[18] A. O. Anim-Ayeko, C. Schillaci, and A. Lipani, "Automatic blight disease detection in potato (Solanum tuberosum L.) and tomato (Solanum lycopersicum, L. 1753) plants using deep learning," *Smart Agric. Technol.*, vol. 4, Aug. 2023, doi: 10.1016/j.atech.2023.100178.

[19] Y. A. Bezabh, A. M. Ayalew, B. M. Abuhayi, T. N. Demlie, E. A. Awoke, and T. E. Mengistu, "Classification of mango disease using ensemble convolutional neural network," *Smart Agric. Technol.*, vol. 8, no. April, p. 100476, 2024, doi: 10.1016/j.atech.2024.100476.

[20] M. A. Hossain, S. Sakib, H. M. Abdullah, and S. E. Arman, "Deep learning for mango leaf disease identification: A vision transformer perspective," *Heliyon*, vol. 10, no. 17, p. e36361, 2024, doi: 10.1016/j.heliyon.2024.e36361.

