JRAEE
Journal of Robotics, Automation, and Electronics Engineering
Vol. 4, No. 1, 2026, pp. 268-280
e-ISSN 3025-3780
https://journal.uny.ac.id/v3/jraee

Integrating Color Segmentation and Texture Analysis for Citrus Leaf Disease Identification through K-Means and Local Binary Pattern

Armin[a,1], Jayanti Yusmah Sari[b,2], Suharsono Bantun[c,3]

[a,b,c] Universitas Sembilanbelas November Kolaka, Indonesia
[1] armin@usn.ac.id; [2] jayanti@usn.ac.id; [3] suharsonob@usn.ac.id
[*] Corresponding Author

ARTICLE INFO

Article History
Received 30 Aug 2026
Revised 31 Aug 2026
Accepted 31 Aug 2026

ABSTRACT

Early identification of citrus leaf diseases is important for reducing crop losses and supporting effective disease management. This study proposes a citrus leaf disease identification framework combining K-means color segmentation and Local Binary Pattern (LBP) texture analysis. The method was evaluated using 90 citrus leaf images from three disease classes: Citrus Canker, Leaf Miner, and Sooty Mold. Images were preprocessed through cropping, resizing, and color space transformation. K-means clustering in the Lab color space was used to segment diseased regions, while LBP extracted texture features. Classification was performed using Euclidean Distance and evaluated with 10-fold cross-validation. The K-means-based method achieved an average accuracy of 76.65%, while LBP alone obtained 45.65%. Combining K-means and LBP produced the best performance, with an average accuracy of 80.02%. These results indicate that integrating color and texture features provides a more effective representation of citrus leaf disease symptoms and improves classification performance.

ABSTRAK

Identifikasi dini penyakit daun jeruk penting untuk mengurangi kerugian hasil panen dan mendukung pengelolaan penyakit yang efektif. Penelitian ini mengusulkan kerangka identifikasi penyakit daun jeruk dengan menggabungkan segmentasi warna K-means dan analisis tekstur Local Binary Pattern (LBP). Metode ini dievaluasi menggunakan 90 citra daun jeruk dari tiga kelas penyakit, yaitu Citrus Canker, Leaf Miner, dan Sooty Mold. Citra di praproses melalui cropping, resizing, dan transformasi ruang warna. K-means pada ruang warna Lab digunakan untuk melakukan segmentasi area yang terinfeksi, sedangkan LBP digunakan untuk mengekstraksi fitur tekstur. Klasifikasi dilakukan menggunakan Euclidean Distance dan dievaluasi dengan 10-fold cross-validation. Metode berbasis K-means menghasilkan akurasi rata-rata 76,65%, sedangkan LBP sebesar 45,65%. Kombinasi K-means dan LBP memberikan performa terbaik dengan akurasi rata-rata 80,02%. Hasil ini menunjukkan bahwa integrasi fitur warna dan tekstur dapat merepresentasikan gejala penyakit daun jeruk secara lebih efektif dan meningkatkan performa klasifikasi.

Keywords
Citrus leaf disease identification;
K-means clustering;
Local Binary Pattern;
texture analysis;
computer vision.

1. Introduction

Citrus is one of the most economically important fruit crops worldwide, contributing significantly to food security, agricultural trade, and rural livelihoods. The global citrus industry supports millions of farmers and is a major source of vitamins and nutritional compounds for human consumption [1]. However, citrus production faces numerous challenges caused by biotic stress, particularly plant diseases that affect vegetative growth, fruit quality, and overall productivity. Several diseases, including citrus canker, huanglongbing, citrus black spot, and citrus variegated chlorosis, are major threats to sustainable citrus cultivation worldwide [1]. Among these diseases, citrus canker is considered one of the most destructive bacterial diseases, causing lesions on leaves, stems, and fruits, which subsequently leads to reduced photosynthetic activity, premature fruit drop, deterioration of fruit quality, and significant yield losses [2].

Leaf diseases play a critical role in determining citrus productivity because leaves are directly involved in photosynthesis and metabolism. Disease symptoms commonly appear first on leaf surfaces, making leaves an important indicator of early disease diagnosis. Severe infections can result in defoliation, impaired nutrient transport, reduced fruit development, and substantial economic losses to farmers [2]. In addition, some diseases exhibit similar visual symptoms during their early stages, making manual diagnosis difficult and potentially leading to delayed treatment. Therefore, the development of reliable methods for early citrus leaf disease identification is essential to support disease management strategies and minimize production loss [1], [2].

Traditionally, disease diagnosis has been conducted through visual inspection by farmers or agricultural experts. Although this approach remains widely practiced, it is often subjective, time-consuming, and highly dependent on the expert knowledge. Variations in human perception and field conditions may also affect the consistency of the diagnosis. Recent advances in computer vision and image processing have created opportunities to develop automated disease identification systems capable of providing objective, rapid, and nondestructive assessments of plant health [3]. By analyzing digital images, computer vision systems can extract meaningful visual information, including color, texture, shape, and pattern characteristics, which can subsequently be used for disease recognition and classification [3], [4].

Image processing techniques have become increasingly important in agricultural applications because they enable automated monitoring and diagnosis without requiring direct physical contact with plants. In plant disease identification, image segmentation is commonly employed to isolate diseased regions from healthy tissues, whereas feature extraction techniques transform visual symptoms into quantitative representations suitable for classification [4]. Previous studies have demonstrated that disease symptoms frequently manifest as color abnormalities and texture alterations. Consequently, combining color- and texture-based information has the potential to improve disease representation and increase identification accuracy.

Several studies have investigated image-processing approaches for disease identification. Premana et al. [5] applied K-Means clustering for image segmentation using color and texture feature extraction on tomato fruit images. Their study demonstrated the feasibility of combining color and texture information; however, the classification performance and quantitative evaluation of the extracted features were not comprehensively reported. Consequently, the effectiveness of the integrated feature representation cannot be fully assessed.

Lestari et al. [6] developed a citrus leaf disease detection system based on RGB-HSV color segmentation and Fuzzy K-Nearest Neighbor (F-KNN) classification. This study focused on identifying citrus leaf diseases using color characteristics and achieved an accuracy of 69%. Although the proposed method demonstrated the usefulness of color information for disease recognition, the texture characteristics of disease symptoms were not explicitly incorporated into the identification process. As different diseases may exhibit similar color distributions, relying solely on color features may limit the classification performance.

Another study conducted by Ariesdianto et al. [7] employed K-Nearest Neighbor (K-NN) classification and gray-level co-occurrence matrix (GLCM) texture features for citrus leaf disease identification. Their approach analyzed two disease classes and one healthy leaf category, achieving an accuracy of 70%. Although texture analysis successfully represented the disease surface characteristics, color segmentation was not integrated into the framework. Consequently, complementary information associated with disease discoloration patterns has not been fully utilized.

The literature suggests that both color and texture features provide valuable information for identifying diseases. However, previous studies have generally focused on either color-based or texture-based approaches

ISSN 3025-3780 (online)
ISSN 3025-4590 (printed)

independently. Color-based methods may encounter difficulties when different diseases exhibit similar color characteristics, whereas texture-based methods may be affected by variations in image quality and environmental conditions. Furthermore, previous citrus disease identification studies have reported moderate classification performance, indicating opportunities for improving disease representation through the integration of complementary image features [5], [6], [7].

Based on these observations, a research gap has been identified. Existing studies have not sufficiently explored the integration of K-means color segmentation and Local Binary Pattern (LBP) texture analysis for citrus leaf disease identification. K-means clustering is widely recognized for its ability to separate image regions based on color similarity, whereas LBP is a computationally efficient texture descriptor that captures local texture patterns. Integration of these two techniques may provide a more comprehensive representation of disease symptoms by simultaneously considering chromatic and textural characteristics.

The novelty of this study lies in the integration of color segmentation and texture analysis through the combination of K-means clustering and Local Binary Pattern. Unlike previous studies that primarily emphasized either color or texture features separately, the proposed framework exploits both types of information within a unified disease identification process. Disease regions are first segmented using K-means clustering in the L*a*b color space and subsequently analyzed using LBP to extract texture descriptors. This integrated representation is expected to improve the discrimination capability of different citrus leaf diseases.

The proposed framework was designed to identify three common citrus leaf diseases: citrus canker, leaf miner infestation, and sooty mold disease. Disease identification was performed using Euclidean Distance-based similarity measurement, and the overall performance was evaluated using a 10-fold cross-validation procedure to ensure objective assessment and reliability.

This study makes three contributions. First, we developed a computer vision-based framework for citrus leaf disease identification using integrated color and texture information. Second, it combines K-means clustering and local binary patterns to exploit complementary disease characteristics extracted from citrus leaf images. Third, it evaluates the effectiveness of the proposed approach using a standardized cross-validation procedure, providing insights into the applicability of the integrated image features for agricultural disease identification. The findings are expected to contribute to the advancement of computer vision and image processing applications in agricultural monitoring and plant disease management.

2. Methods

2.1 Dataset Acquisition

The dataset used in this study consisted of 90 citrus leaf images collected from citrus plantations in Siompu District, South Buton Regency, Southeast Sulawesi, Indonesia, using an OPPO A12 smartphone camera under natural light conditions. The images were stored in JPEG format and processed using MATLAB R2015a software. The dataset comprised three disease categories, namely citrus canker, leaf miner infestation, and sooty mold disease, with 30 images in each category. These diseases were selected because they are among the most commonly observed citrus leaf diseases and exhibit distinct color and texture. The distribution of images across the disease categories is presented in Table 1. As shown in the table, the dataset was evenly distributed among the three classes, reducing class imbalance and providing a fair basis for evaluating the proposed disease identification framework. To assess the robustness and generalization capability of the proposed method, a 10-fold cross-validation strategy was employed, in which the dataset was divided into ten subsets, with one subset used for testing and the remaining subsets used for training in each iteration. The process was repeated until all subsets served as testing data once, and the average performance across all iterations was reported as the final evaluation result.

Table 1. Distribution of Citrus Leaf Disease Dataset

Disease Category	Number of Images	Percentage (%)
Citrus Canker	30	33.33
Leaf Miner	30	33.33
Sooty Mold	30	33.33
Total	**90**	**100**

2.2 Research Framework

Figure 1 illustrates the overall research framework proposed for citrus leaf disease identification by integrating color segmentation and texture analysis. The framework begins with the acquisition of citrus leaf images, followed by an image preprocessing stage consisting of cropping, resizing, and RGB-to-L*a*b color space conversion to enhance the image consistency and facilitate color-based analysis. Subsequently, K-means clustering was employed to segment disease regions from healthy leaf tissues based on color characteristics.

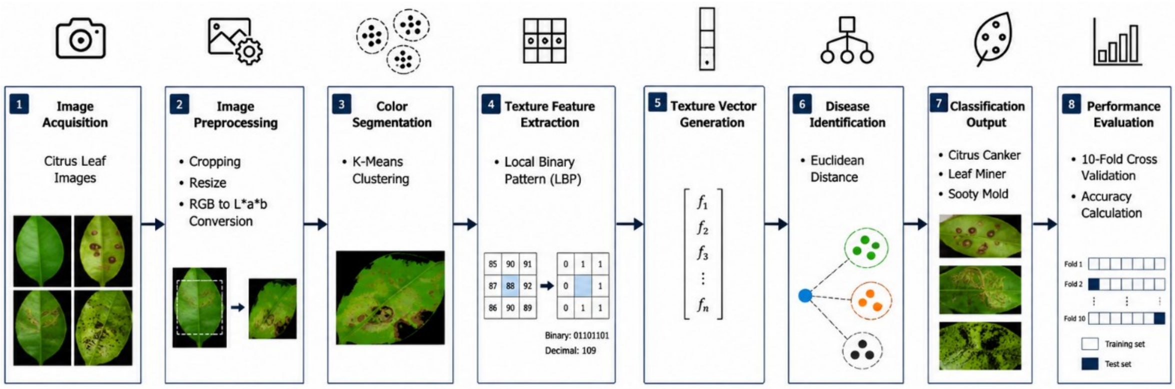

Fig. 1 Overall research framework for citrus leaf disease identification

The segmented regions were then processed using LBP to extract texture descriptors that represent disease-specific surface patterns. The resulting texture features were transformed into feature vectors and compared with stored feature vectors using the Euclidean Distance for disease identification. The classification output consisted of three disease categories: citrus canker, leaf miner infestation, and sooty mold disease. Finally, the performance of the proposed framework was evaluated using a 10-fold cross-validation scheme, and the classification accuracy was calculated to assess the effectiveness of integrating K-means color segmentation and Local Binary Pattern texture analysis for citrus leaf disease identification.

2.3 Image Preprocessing

Image preprocessing was performed to improve the image quality and ensure data consistency prior to feature extraction and disease identification. This stage involved cropping, resizing, and color space conversion, as illustrated in Figure 2. Cropping was applied to isolate the region of interest containing disease symptoms and eliminate irrelevant background information that could interfere with subsequent analyses. The cropped images were then resized to obtain uniform image dimensions, thereby reducing the computational complexity and ensuring consistent processing across all samples. Following resizing, the images were converted from the RGB color space to the L*a*b color space, which separates the luminance information from the chromatic components and provides a more effective representation of the color variations associated with disease symptoms. As shown in Figure 2, the preprocessing workflow transforms the original citrus leaf image into a standardized representation suitable for color-based analysis by sequentially performing region extraction, image normalization, and color-space transformation. These operations enhanced the visual distinction between diseased and healthy leaf tissues and provided a more reliable input for the subsequent K-means clustering stage, which segmented disease regions based on color characteristics.

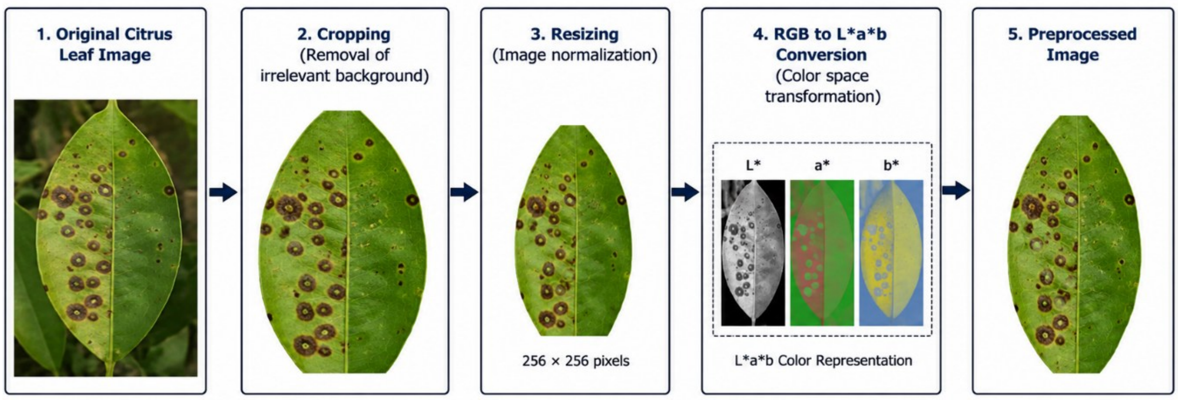

Fig. 2 Image preprocessing stages

2.4 Color Segmentation Using K-Means Clustering

Color segmentation was performed using the K-means clustering algorithm to isolate disease regions from

ISSN 3025-3780 (online)
ISSN 3025-4590 (printed)

healthy leaf tissues based on color characteristics. This stage plays a crucial role in citrus leaf disease identification because disease symptoms often appear as discoloration patterns that differ from the surrounding healthy tissues. Prior to clustering, the preprocessed images were represented in the L*a*b color space, which separates the luminance information from the chromatic components and facilitates more effective color discrimination. K-means clustering partitions image pixels into groups according to color similarity by iteratively assigning pixels to the nearest cluster centroid and updating the centroid positions until convergence is achieved. The distance calculation used during the cluster assignment follows the Euclidean Distance formulation presented in Equation (1).

$$D(x_i, c_j) = \sqrt{\sum_{k=1}^{n} (x_{ik}, c_{jk})^2} \tag{1}$$

where:

x_i= feature vector of pixel i

c_j= centroid of cluster j

n= number of features

As illustrated in Figure 3, the segmentation process transforms the pre-processed image into clustered regions, enabling the extraction of disease-specific areas while suppressing irrelevant background and healthy tissue information. The resulting segmented regions provide a clearer representation of visual disease symptoms, including lesions, spots, and discoloration patterns associated with citrus canker, leaf miner infestation, and sooty mold. By isolating these disease regions, K-means clustering enhances the quality of subsequent texture analysis and improves the effectiveness of the overall disease identification framework.

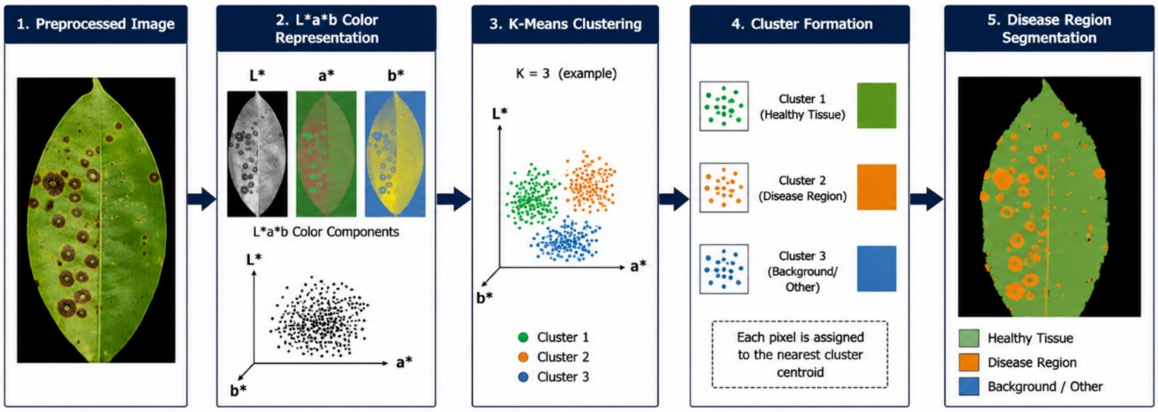

Fig. 3 K-Means color segmentation process

2.5 Texture Feature Extraction Using LBP

Following the segmentation stage, texture features were extracted using the LBP operator to characterize the surface patterns of the citrus leaf diseases. Texture analysis is particularly important in disease identification because many disease symptoms exhibit distinctive visual patterns, such as lesions, spots, and surface irregularities that may not be fully represented by color information alone. LBP is a computationally efficient texture descriptor that captures local texture variations by comparing the intensity values of each neighboring pixel with that of a center pixel within a predefined neighborhood. Neighboring pixels with intensity values greater than or equal to the center pixel are assigned a value of one, whereas lower values are assigned zero, resulting in a binary pattern that is subsequently converted into a decimal representation. The LBP computation follows the formulation presented in Equation (2), and the thresholding operation is defined in Equation (3).

$$LBP(x_c, y_c) = \sum_{P=0}^{P-1} s(i_p, i_c)2^p \tag{2}$$

where:

i_c = center pixel intensity

i_p = neighboring pixel intensity

P = number of neighboring pixels

$$s(x) = \{1, x \geq 0\ 0, x < 0 \tag{3}$$

As illustrated in Figure 4, the texture extraction process begins with a segmented disease region, followed by a pixel neighborhood comparison, binary pattern generation, and formation of texture descriptors. This procedure enables the extraction of disease-specific texture characteristics that effectively represent the visual appearance of citrus canker, leaf miner infestation, and sooty mold. By transforming local texture information into discriminative feature vectors, LBP enhances the ability of the proposed framework to distinguish between different disease categories and supports subsequent disease classification.

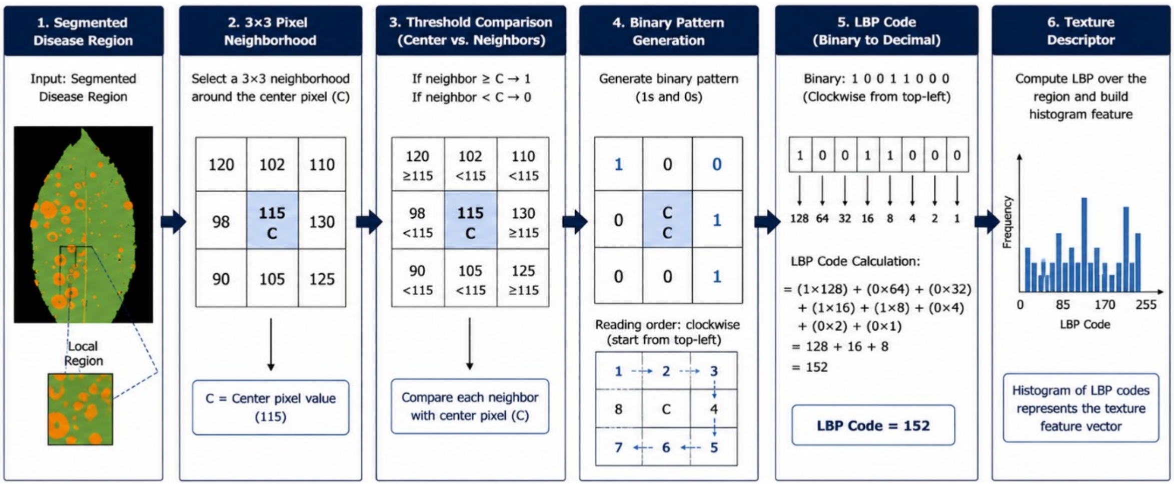

Fig. 4 Local binary pattern texture extraction process

2.6 Disease Classification

Disease classification was performed using the Euclidean Distance method to determine the similarity between feature vectors extracted from test images and those stored in the disease feature database. After the texture descriptors were generated using LBP, each feature vector was compared with the reference feature vectors representing the three disease classes, namely, Citrus Canker, Leaf Miner, and Sooty Mold. The similarity measurement was calculated using the Euclidean Distance formulation presented in Equation (4), where smaller distance values indicate a higher degree of similarity between the feature vectors.

$$d_{ij} = \sqrt{\sum_{k=1}^{n} (x_{ik}, x_{jk})^2} \tag{4}$$

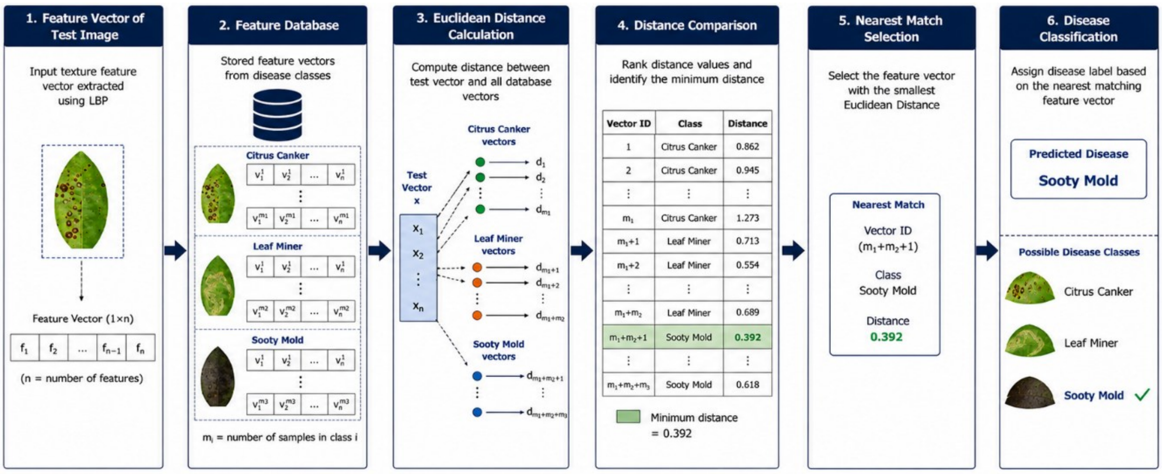

Fig. 5 Disease identification procedure using euclidean distance

The classification process follows the nearest-distance principle, in which a test image is assigned to the

ISSN 3025-3780 (online)
ISSN 3025-4590 (printed)

disease class associated with the minimum Euclidean Distance. As illustrated in Figure 5, the disease identification procedure begins with the extraction of feature vectors from the segmented disease region, followed by a similarity comparison with the stored feature database and the selection of the closest matching disease class. This approach enables the identification system to distinguish between different citrus leaf diseases based on their texture characteristics and provides a straightforward yet effective mechanism for disease recognition. The resulting classification output served as the basis for evaluating the effectiveness of the proposed framework in identifying citrus leaf diseases.

2.7 Performance Evaluation

The performance of the proposed citrus leaf disease identification framework was evaluated using a 10-fold cross-validation strategy to assess its robustness and generalization capability. In this approach, the dataset was divided into ten approximately equal subsets, where one subset was used as the testing data and the remaining nine subsets were used for training and feature database construction during each iteration. The process was repeated ten times until every subset had served as testing data once, ensuring that all images were included in both the training and testing phases. As illustrated in Figure 6, the cross-validation scheme systematically rotates the testing subset across all folds, thereby reducing the influence of data partitioning bias and providing a more reliable estimate of the model performance. This evaluation strategy is particularly suitable for relatively small datasets because it maximizes the utilization of the available data while improving the stability of the assessment results. The classification performance was measured using accuracy, which represents the proportion of correctly identified disease samples relative to the total number of testing samples and is calculated according to Equation (5).

$$Accuracy = \frac{N_{correct}}{N_{total}} x100\% \tag{5}$$

where:

$N_{correct}$ = number of correctly classified samples

N_{total} = total number of testing samples

The use of repeated training and testing across multiple folds enabled a more comprehensive evaluation of the proposed framework and minimized the likelihood of overestimating performance owing to favorable data splits. The average accuracy obtained from all ten folds was reported as the final performance indicator of the proposed citrus leaf disease identification system.

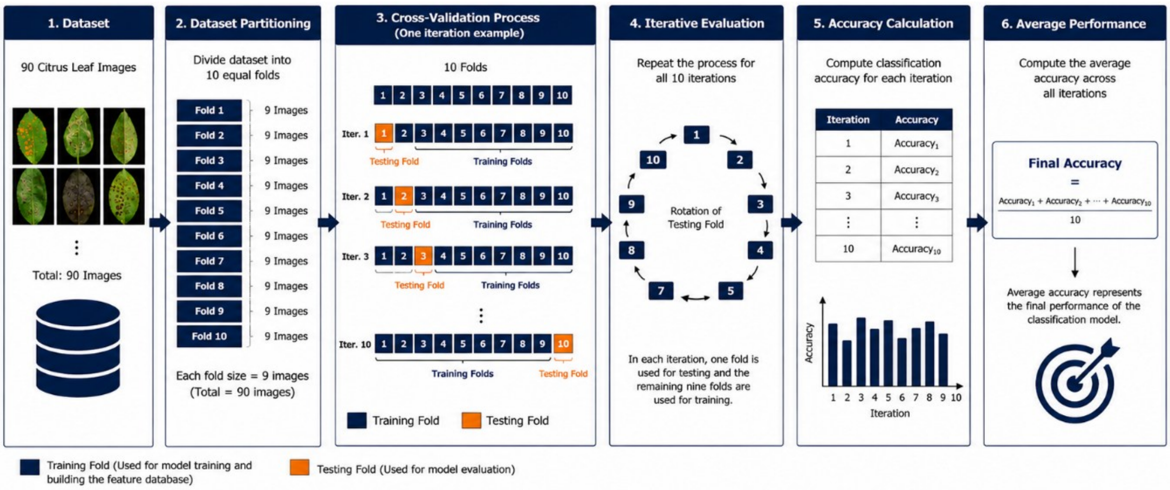

Fig. 6 Ten-fold cross validation scheme

3. Result and Discussion

3.1 Image Preprocessing Results

The image preprocessing stage transformed the acquired citrus leaf images into a standardized representation that was suitable for subsequent segmentation and feature extraction. As illustrated in Figure 7, cropping successfully isolated the region of interest by removing most irrelevant background information, allowing the analysis to focus on the leaf surface and visible disease symptoms. The resizing process further normalized the image dimensions across all samples, ensuring a consistent image representation throughout the dataset and reducing the variability associated with differences in image resolution. The resulting L*a*b images preserved the visual appearance of the disease lesions while providing a color representation in which the luminance information was separated from the chromatic components. This transformation enhanced the visibility of lesion-related color differences and reduced the influence of illumination variations in the original RGB images. As shown in Figure 7, the disease symptoms remained clearly distinguishable after preprocessing, indicating that important visual characteristics were retained throughout the transformation process. Consequently, the preprocessing stage produced a consistent and standardized dataset that facilitated subsequent K-means segmentation and supported the extraction of disease-related color and texture information in the proposed citrus leaf disease identification framework.

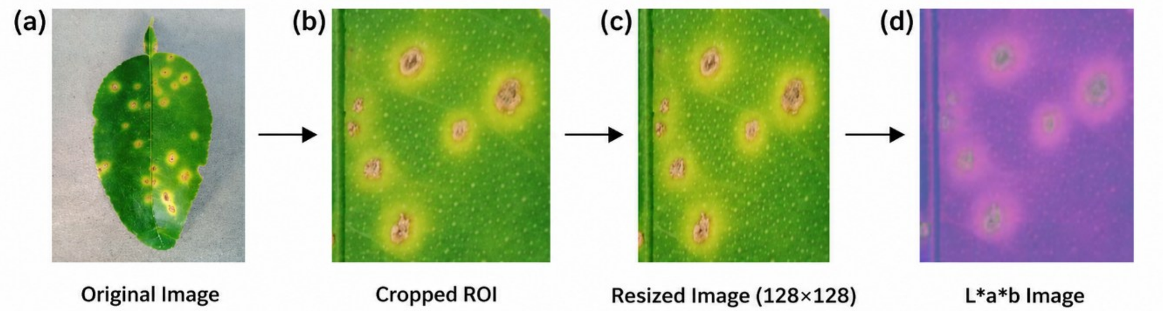

Fig. 7 Image preprocessing results used in the proposed citrus-leaf disease identification framework.

3.2K-Means Segmentation Results

The segmentation results presented in Figure 8 demonstrate that color-based clustering can separate disease-affected regions from healthy leaf tissues by grouping pixels with similar color characteristics. The segmented outputs revealed distinct visual patterns for each disease category. Citrus Canker images exhibit localized lesion regions with a clear color contrast relative to the surrounding healthy tissue, enabling the symptomatic areas to be isolated more distinctly. In the Leaf Miner samples, the segmented regions corresponded to the characteristic mining trails that appeared as elongated discoloration patterns across the leaf surface. Meanwhile, the Sooty Mold images show concentrated dark regions associated with fungal deposits, which form visually distinguishable clusters within the segmented output.

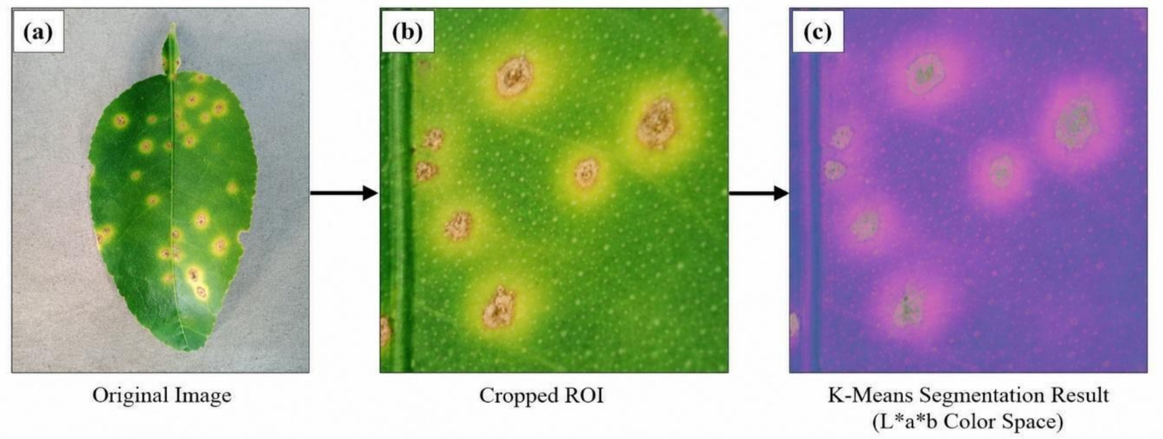

Fig. 8 Results of K-means-based color segmentation applied to citrus leaf images.

As illustrated in Figure 8, the segmentation process reduced the influence of irrelevant visual information and emphasized disease-related areas, resulting in a simplified representation of the leaves' surface. This outcome suggests that the color differences associated with disease symptoms provide sufficient information for clustering-based separation, allowing infected regions to be distinguished from healthy tissues. Similar observations have been reported in previous plant disease identification studies employing K-means segmentation, where color clustering effectively isolated symptomatic regions prior to feature extraction. The segmented images generated in this study also provided a more focused input for subsequent texture analysis, ensuring that the LBP features were extracted primarily from disease-related regions rather than from

ISSN 3025-3780 (online)
ISSN 3025-4590 (printed)

unaffected regions. Consequently, the segmentation results indicate that K-means clustering successfully highlighted the visual manifestations of citrus leaf diseases and established an appropriate foundation for the extraction of discriminative texture features in the next stage of the proposed framework.

3.3 Texture Analysis Results Using LBP

The texture analysis results shown in Figure 9 illustrate the transformation of the segmented citrus leaf image through several preprocessing stages before LBP feature extraction. Grayscale conversion simplified the image representation by retaining the intensity information while removing the color components, enabling the subsequent texture analysis to focus on the structural variations within the disease region. Median filtering further reduced the image noise while preserving important lesion boundaries, resulting in a smoother visual appearance. The histogram equalization stage enhanced the image contrast and increased the visibility of the local surface structures associated with the disease symptoms. As shown in Figure 9, these preprocessing operations produced a more distinct representation of the lesion area before applying the LBP. The final LBP image reveals detailed local texture patterns characterized by variations in pixel intensity relationships across the disease regions. Compared with the preceding images, the LBP representation emphasizes micro-texture structures that are less apparent in grayscale and filtered images. These results indicate that the LBP transformation successfully captured local texture information associated with citrus leaf disease symptoms and converted it into a representation suitable for subsequent feature extraction and Euclidean Distance-based classification. Therefore, the texture analysis process shown in Figure 9 provides an effective mechanism for preserving disease-related surface characteristics while enhancing the discriminative information required for identifying diseases.

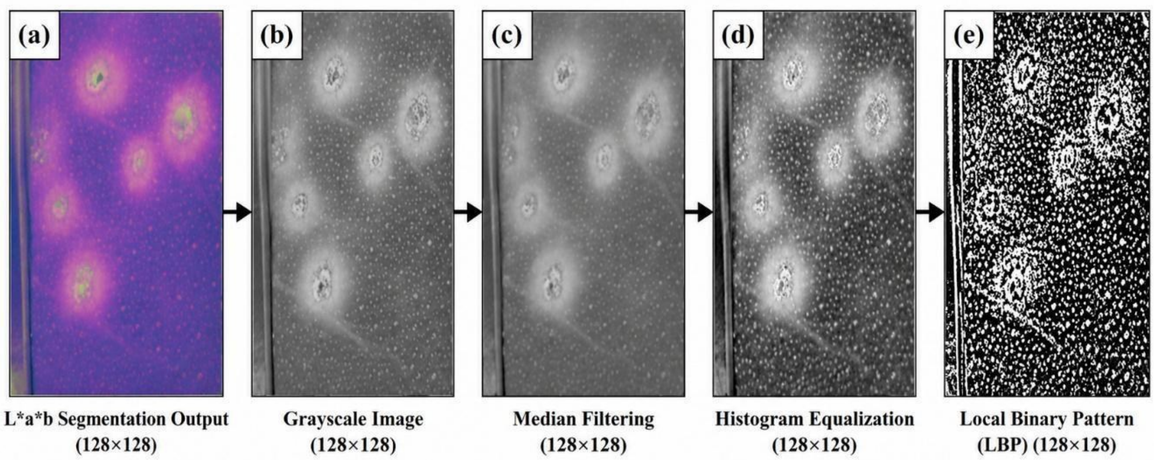

Fig. 9 LBP texture analysis results.

3.4Feature Extraction Results

The statistical characteristics of the features extracted from the K-means and LBP processing stages are summarized in Table 2. The results indicate that the three disease classes exhibit different distributions of mean feature values, whereas the maximum (Max) and minimum (Min) values remain constant at 255 and 0, respectively, across all samples. This observation suggests that the mean feature serves as the primary discriminative descriptor in the proposed framework. As shown in Table 2, leaf miners generally exhibited the highest mean feature values, ranging from 109.64 to 124.50, reflecting the presence of elongated mining trails that generated relatively brighter texture responses after segmentation and LBP transformation. Citrus Canker presented lower mean values, ranging from 102.50 to 117.21, which correspond to localized lesion regions and irregular surface damage. Meanwhile, Sooty Mold demonstrated the widest variation, with mean values ranging from 97.60 to 122.81, indicating substantial diversity in the distribution of fungal deposits across the affected leaf surface. Despite these differences, a noticeable overlap exists among the mean feature ranges of the three disease classes. For example, several Citrus Canker samples exhibited mean values comparable to those observed in Leaf Miner and Sooty Mold samples, resulting in similar numerical representations. Such overlap reduces the separability of the feature space and may contribute to classification ambiguity when the Euclidean Distance is used to identify the nearest matching sample. Similar observations have been reported in previous plant disease identification studies employing statistical image descriptors, where simple intensity-based features effectively capture disease characteristics but may experience reduced discriminative capability when different disease symptoms share comparable visual patterns. Nevertheless, the results presented in Table 2 demonstrate that the extracted features provide a compact representation of disease-specific texture information and constitute an appropriate basis for the subsequent classification.

Table 2. Statistical Summary of Features Extracted from K-Means and LBP Processing

Disease Class	n	Mean Feature Range	Range Mean ± SD	Max	Min
Citrus Canker	30	102.50–117.21	To be calculated from all samples	255	0
Leaf Miner	30	109.64–124.50	To be calculated from all samples	255	0
Sooty Mold	30	97.60–122.81	To be calculated from all samples	255	0

3.5 Classification Performance

The classification performance of the proposed framework is shown in Figure 10. The experimental results show that the integration of K-means color segmentation and LBP texture analysis achieved an average classification accuracy of 80.02%, indicating the effectiveness of combining color- and texture-based information for citrus leaf disease identification.

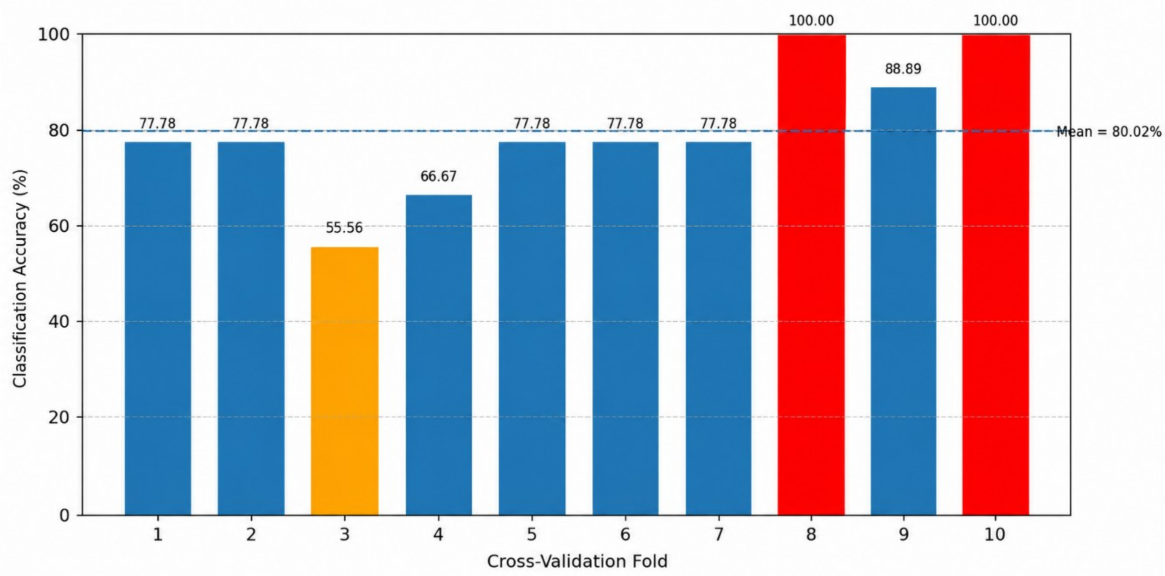

Fig. 10 Classification accuracy across 10-fold cross-validation

As shown in Figure 10, the classification accuracy varied across the ten cross-validation folds, ranging from 55.56% in Fold 3 to 100.00% in Folds 8 and 10. Despite this variation, most folds achieved accuracies above 75%, suggesting that the proposed framework maintained a relatively stable performance across different training and testing partitions. The occurrence of perfect classification performance in two folds demonstrates that the extracted color and texture features were highly discriminative for certain data partitions, whereas lower accuracies in other folds indicate the presence of samples with more challenging visual characteristics. The observed performance variation is consistent with the feature analysis presented in Section 3.4, where an overlap in mean feature values was observed among some disease categories. Such overlap may reduce class separability and increase the likelihood of misclassification when the Euclidean Distance is used as the similarity measure. Nevertheless, the overall trend shown in Figure 10 indicates that the proposed framework consistently achieved a satisfactory classification performance across different folds. Compared with previous citrus leaf disease identification studies, the obtained average accuracy of 80.02% is higher than the 69% reported by Lestari et al. using RGB-HSV color segmentation and Fuzzy K-Nearest Neighbor classification and the 70% reported by Ariesdianto et al. using GLCM texture features and K-Nearest Neighbor classification. These findings suggest that the integration of K-means segmentation and LBP texture descriptors provides a more informative representation of disease symptoms and improves the classification performance. Overall, the results demonstrate that the proposed framework effectively exploits complementary color and texture information to support reliable citrus-leaf disease identification.

3.6 Discussion

ISSN 3025-3780 (online)
ISSN 3025-4590 (printed)

The experimental findings demonstrate that the integration of color segmentation and texture analysis provides a more effective representation of citrus leaf disease symptoms than the use of individual visual descriptors. The preprocessing stage successfully standardized the input images through cropping, resizing, and color space transformation, ensuring that disease-related information was preserved while reducing irrelevant variations caused by the background elements and image acquisition conditions. This standardized representation subsequently enhanced the effectiveness of K-means segmentation, which isolated disease-affected regions by exploiting the color differences associated with lesion development, chlorosis, and fungal contamination. The importance of color information is evident from the classification results, where the K-means-based approach achieved an average accuracy of 76.65%, which is substantially higher than the 45.65% obtained using LBP-based texture analysis alone. These findings indicate that color characteristics are an important visual cue for differentiating citrus leaf diseases. However, the additional improvement to 80.02% achieved by integrating K-means and LBP demonstrates that color information alone is insufficient to fully characterize disease symptoms and that texture descriptors contribute complementary discriminatory information. From a computer vision perspective, K-means segmentation primarily improves disease localization by emphasizing disease-related color distributions, whereas LBP captures local structural variations associated with lesion boundaries, tissue degradation, mining trails, and fungal deposits. Consequently, segmentation quality directly influences texture feature quality because accurate disease region extraction reduces the inclusion of irrelevant healthy tissue and background pixels, enabling LBP to encode more representative disease patterns.

The feature extraction results further support this conclusion. While the Max and Min feature values remained constant at 255 and 0 across all disease categories, the Mean feature exhibited noticeable variation among Citrus Canker, Leaf Miner, and Sooty Mold samples, indicating that the Mean feature carried the primary discriminative information used during classification. However, partial overlap among the mean feature ranges was observed, suggesting that certain disease categories shared similar visual characteristics. This overlap is consistent with the pattern recognition theory, which states that the classification performance is strongly influenced by the class separability within the feature space. Because the Euclidean Distance classification relies on measuring the proximity among feature vectors, overlapping feature distributions reduce inter-class discrimination and increase the likelihood of misclassification. This phenomenon is reflected in the cross-validation results, where the classification accuracy varied from 55.56% to 100.00%, with an overall average of 80.02%. The occurrence of perfect classification performance in Folds 8 and 10 indicates that the extracted features were highly discriminative for specific testing partitions, whereas the lower accuracy observed in Fold 3 suggests the presence of samples with highly similar feature representation. Despite this variation, most folds achieved accuracies above 75%, indicating relatively stable performance and acceptable generalization capability across different training and testing partitions.

Compared with previous studies, the proposed framework demonstrated competitive performance. Lestari et al. reported an accuracy of 69% using RGB-HSV color segmentation combined with Fuzzy K-Nearest Neighbor (KNN) classification, while Ariesdianto et al. achieved 70% using GLCM texture features and K-Nearest Neighbor classification. The higher accuracy achieved in this study suggests that combining K-means segmentation and LBP texture descriptors provides a richer representation of disease symptoms by integrating complementary color and texture information. Nevertheless, this study has several limitations. The framework relies on handcrafted features and a simple Euclidean Distance classifier, which may not fully capture complex intra-class variability and inter-class similarity. Furthermore, overlapping feature distributions and visually similar disease manifestations, particularly among categories exhibiting comparable texture characteristics, continue to limit the classification performance. Future research may benefit from incorporating more discriminative feature representations and advanced classification models that can learn complex decision boundaries. Overall, the results confirm that the integration of color segmentation and texture analysis constitutes an effective and computationally efficient strategy for citrus leaf disease identification and provides a practical foundation for agricultural computer vision applications.

4. Conclusion

In this study, we propose a citrus leaf disease identification framework that integrates K-means color segmentation and LBP texture analysis to improve the disease recognition performance. The experimental results demonstrated that the preprocessing stage effectively standardized image representation through cropping, resizing, and color space transformation, thereby preserving the disease-related visual characteristics while reducing irrelevant image variations. K-means segmentation successfully isolated disease-affected regions and emphasized the color information associated with lesion development, whereas LBP captured local texture patterns related to tissue damage and disease symptoms. Feature analysis revealed that the mean feature provided greater discriminative information than the maximum and minimum features, which were constant across disease categories. The classification experiments further showed that K-means segmentation alone achieved an average accuracy of 76.65%, whereas LBP-based texture analysis alone achieved 45.65%. The integration of K-means and LBP produced the highest performance, achieving an average classification accuracy of 80.02% under 10-fold cross-validation, indicating that color and texture information provide complementary representations for citrus leaf disease identification.

The findings confirm that combining color-based segmentation and texture-based feature extraction improves disease representation and classification effectiveness compared with using either approach independently. Nevertheless, this study has several limitations. Partial overlap among feature distributions reduces class separability, particularly for disease categories exhibiting similar visual characteristics. In addition, the use of handcrafted features and Euclidean Distance classification may limit the ability of the framework to capture more complex disease variations. Therefore, future research should investigate more discriminative feature representations, larger and more diverse datasets, and advanced machine or deep learning classifiers to further improve the classification accuracy and generalization performance. Overall, the proposed framework provides an effective, computationally efficient, and practical approach for automated citrus leaf disease identification, with potential applications in agricultural monitoring and decision-support systems.

Author Contribution:

All authors contributed equally as the main contributors to this paper. All authors read and approved the final manuscript.

Funding:

This research received no external funding.

Acknowledgement:

The authors thank the Department of Electrical and Electronics Engineering, Vocational Faculty, Universitas Negeri Yogyakarta, for providing the facilities used in this research.

Conflicts of Interest:

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

[1] H. B. Barbieri, L. S. Fernandes, J. G. de M. Pontes, A. K. Pereira, and T. P. Fill, "An overview of the most threatening diseases that affect worldwide citriculture: Main features, diagnose, and current control strategies," *Front. Nat. Prod.*, vol. 2, p. 1045364, 2023.

[2] S. Ali *et al.*, "Citrus canker: a persistent threat to the worldwide citrus industry—an analysis," *Agronomy*, vol. 13, no. 4, p. 1112, 2023.

[3] J. R. Sykes, K. J. Denby, and D. W. Franks, "Computer vision for plant pathology: A review with examples from cocoa agriculture," *Appl. Plant Sci.*, vol. 12, no. 2, p. e11559, 2024.

[4] A. Upadhyay *et al.*, "Deep learning and computer vision in plant disease detection: a comprehensive review of techniques, models, and trends in precision agriculture," *Artif. Intell. Rev.*, vol. 58, no. 3, p. 92, 2025.

[5] A. Premana, R. M. H. Bhakti, and D. Prayogi, "Segmentasi K-Means Clustering Pada Citra Menggunakan Ekstrasi Fitur Warna dan Tekstur," *J. Ilm. Intech Inf. Technol. J. UMUS*, vol. 2, no. 01, pp. 89–97, 2020.

[6] F. Lestari, J. Sari, I. P. Sutardi, and N. Purnama, "Deteksi Penyakit Tanaman Jeruk Siam Berdasarkan Citra Daun Menggunakan Segmentasi Warna RGB-HSV," in *Seminar Nasional Teknologi Terapan Berbasis Kearifan Lokal (SNT2BKL)*, 2018, pp. 276–283.

[7] R. H. Ariesdianto, Z. E. Fitri, A. Madjid, and A. M. N. Imron, "Identifikasi Penyakit Daun Jeruk Siam Menggunakan K-Nearest Neighbor," *J. Ilmu Komput. dan Inform.*, vol. 1, no. 2, pp. 133–140, 2021.