**Journal of Applied Informatics and Computing (JAIC)**
Vol.10, No.4, August 2026, pp. 3514~3522
e-ISSN: 2548-6861


# Support Vector Machine and Multilayer Perceptron Optimization for Banana Leaf Disease Classification Using Hue Saturation Value and Gray Level Co-occurrence Matrix

**Haris Pujianto [1], Christy Atika Sari [2]*, Musab Iqtait [3]**
[1,2]Department of Informatics Engineering, Universitas Dian Nuswantoro, Indonesia
[3]Department of Data Science and Artificial Intelligence, Zarqa University, Jordan
111202315196@mhs.dinus.ac.id [1], christy.atika.sari@dsn.dinus.ac.id [2], migtait@zu.edu.jo [3]

**Article Info**

***Article history:***

Received 2026-07-07
Revised 2026-07-27
Accepted 2026-08-08

***Keyword:***

*Banana Leaf Disease,*
*GLCM,*
*HSV,*
*Multilayer Perceptron,*
*Support Vector Machine.*

**ABSTRACT**

Banana leaf disease poses a significant threat to the quality and productivity of banana plants. Conventional disease identification relies heavily on expert knowledge and is time-consuming, highlighting the need for an automated and efficient solution. This study presents a classification system for banana leaf diseases by comparing the performance of Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel and Multilayer Perceptron (MLP). The dataset employed was the Banana Leaf Spot Diseases (BananaLSD) dataset, comprising four disease categories: healthy, cordana, pestalotiopsis, and sigatoka. Preprocessing steps included image resizing, data cleaning, feature extraction using Hue Saturation Value (HSV) color features and Gray Level Co-occurrence Matrix (GLCM) texture features, and data normalization via StandardScaler. Experimental results demonstrate that SVM achieved an accuracy of 94.34%, outperforming MLP which reached 93.40%. These findings confirm that the integration of HSV and GLCM features with SVM constitutes an effective approach for automated banana leaf disease classification, offering a promising foundation for intelligent plant health monitoring systems.



## I. INTRODUCTION

Banana plants are one of Indonesia's leading agricultural commodities, possessing high economic value. According to data from the Central Statistics Agency (BPS), banana production in Indonesia in 2024 reached approximately 9.26 million tons, making it the fruit with the highest production volume compared to other fruits. This indicates that bananas have high economic value and serve as a primary food source for the population [1]. Disease outbreaks on banana leaves are a major challenge in cultivation. Common diseases affecting banana leaves include Sigatoka, characterized by yellow and brown spots; Cardona Mosae, featuring oval-shaped yellow and brown lesions; Cladosporium musae, causing small spots; and Fusarium oxysporum, resulting in yellowing leaves that gradually wither from the edges. These diseases can lead to errors in disease management, causing banana growth to slow or fail, resulting in losses for farmers due to failed harvests [2].

The management of banana diseases using digital technology can be addressed through digital image processing. One of the image processing methods that can be easily implemented is computer vision. Computer vision-based identification relies on machine learning algorithms, deep learning, and other algorithms [3]. Some of the machine learning algorithms frequently used by researchers, particularly for detecting leaves resembling those of tomatoes [4], mangroves [5], [6], mangoes [7], corn [8], [9], sugarcane [10], and rice [11] include CNN [12], SVM [13], BPNN, MLP [14].

Previous research on the classification of banana leaf diseases has been conducted using machine learning methods. One study employed the Support Vector Machine (SVM) method with color feature extraction, such as the HSV histogram. The results showed that the SVM method with HSV histogram features achieved the highest accuracy of 83.33% [3] Additionally, the Multilayer Perceptron (MLP), as a feedforward artificial neural network approach, consists of an input layer, one or more hidden layers, and an output layer, and relies on non-linear activation functions to identify complex patterns in the dataset [15]. Previous studies have tended to focus on the use of color features without combining them with texture features, so the representation of image information has not been fully optimized. Furthermore, studies that directly compare the performance of SVM and MLP methods in classifying banana leaf diseases remain limited.

Existing studies have primarily investigated individual machine learning classifiers or emphasized a single type of image feature, making it difficult to determine the contribution of different feature representations under comparable experimental settings. A comprehensive comparison between SVM and MLP using identical preprocessing procedures and the same combination of HSV color features and GLCM texture features has rarely been discussed for banana leaf disease classification. As a result, it remains unclear whether the combination of HSV color features and GLCM texture features provides consistent performance improvements for different machine learning classifiers when evaluated under the same experimental conditions.

Addressing this limitation requires a feature representation capable of capturing both color and texture characteristics while allowing different classifiers to be evaluated under identical experimental conditions. The HSV color feature was selected for this study because it provides a more stable representation of color across varying lighting conditions. Meanwhile, the Gray Level Co-occurrence Matrix (GLCM) texture feature was used to capture the textural characteristics of leaf surfaces, which serve as important indicators for disease identification. The SVM method is effective at handling high-dimensional data and is efficient for classification with limited data, whereas the MLP is capable of learning complex non-linear relationships through its multilayer neural network architecture [13]. Several studies indicate that both methods perform well in classifying images of plant leaf diseases. This study presents a comparative evaluation of Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel and Multilayer Perceptron (MLP) using identical preprocessing procedures and combined HSV color and GLCM texture features to identify the best-performing classifier for banana leaf disease classification.

## II. METHOD

### A. Literature Review

Based on previous research studies in Table I, it can be concluded that the SVM and MLP methods both have advantages in classifying plant leaf images [16], [17]. SVM excels in handling stable high-dimensional data on limited datasets, especially with the use of RBF, while MLP has better capabilities in modeling complex non-linear patterns. Most studies only focus on one method or one type of feature, and have not studied the comparison between SVM and MLP with a combination of color and texture features on banana leaf objects. This study is important to conduct to analyze the effect of using HSV color features and GLCM texture on the performance of the SVM and MLP methods, as well as to determine the most optimal method in classifying.

TABLE I
RELATED RESEARCH

| ***Research by*** | ***Metthods*** | ***Summary*** | ***Relevant*** |
|---|---|---|---|
| Ahmed et al. (2020)[18] | Texture Feature, SVM | Uses leaf image texture analysis and SVM for plant disease classification. | Relevant because it applies image-based plant disease classification using SVM. |
| He & Chen (2021)[19] | MLP | Develops an MLP architecture for hyperspectral image classification. | Relevant because it compares MLP and SVM for image classification. |
| Hossin & Sulaiman (2021)[20] | Accuracy, Precision, Recall, F1-Score, Confusion Matrix | Evaluates machine learning models using confusion matrix, accuracy, precision, recall, and F1-score. | Relevant because it uses the same evaluation metrics. |
| Kiran & Chandrappa (2023) [21] | GLCM, SVM | Applies image preprocessing, GLCM feature extraction, and machine learning for leaf disease classification. | Relevant because it employs GLCM and SVM. |
| Jha et al. (2023) [22] | SVM, Feature Extraction | Uses SVM with color feature extraction for plant disease classification. | Relevant because it combines SVM with image feature extraction. |
| Ahmad et al. (2021)[23] | HSV, GLCM, SVM | Combines HSV color and GLCM texture features for leaf disease identification. | Relevant because it utilizes both HSV and GLCM features. |
| Singh et al. (2021)[24] | Feature Extraction, Image Processing | Investigates interpretable image features for plant leaf classification using machine learning. | Relevant because it focuses on image feature extraction for leaf classification. |
| Thaiyalnayaki & Joseph (2021)[25] | SVM, MLP | Compares SVM and MLP for plant disease classification using leaf images and attributes. | Relevant because it compares SVM and MLP methods. |

| | | | |
|---|---|---|---|
| Garg & Biswas (2021)[26] | Neural Network, SVM | Employs neural networks and SVM for plant disease identification with hidden layer optimization. | Relevant because it uses neural networks and SVM for leaf disease classification. |
| Kurmi et al. (2022) [27] | Image Processing, Feature Extraction, SVM, Neural Network | Utilizes image preprocessing, feature extraction, SVM, and neural networks for plant disease classification. | Relevant because it combines image processing, feature extraction, SVM, and neural networks. |

### *B. Dataset and Preprossesing*

The data source for this study was obtained from the Kaggle platform, specifically the Banana Leaf Spot Diseases (BananaLSD) Dataset. This dataset consists of 400 images of banana leaves classified into four categories: cordana, healthy, pestalotiopsis, and sigatoka. The dataset used consists of two parts: the Originalset and the Augmentedset. The Originalset comprises the original banana leaf images, while the Augmentedset consists of augmented images such as those rotated, flipped, or with altered lighting designed to increase the volume and diversity of the data. Both datasets were then combined and used as input data in the training and testing of the banana leaf disease classification model.

TABLE II

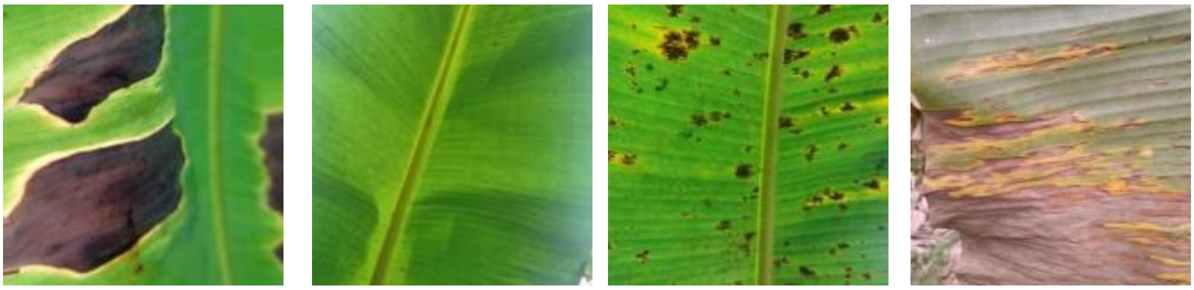
DATASET DISTRIBUTION

| Class | Number of Images | Percentage |
|---|---|---|
| Healthy | 392 | 24.69% |
| Cordana | 396 | 24.94% |
| Pestalotiopsis | 400 | 25.19% |
| Sigatoka | 400 | 25.19% |
| Total | 1588 | 100% |

As presented in Table II, the final dataset consists of 1,588 banana leaf images distributed across four classes. The number of images in each class is relatively balanced, thereby reducing the risk of classification bias toward a particular class during model training and evaluation.

All images are in JPEG format and are used as input data in the training and testing of the banana leaf disease classification model, as shown in Figure 1. Data preprocessing is the initial step performed to prepare the data before it is used in the model training process. This step aims to improve data quality so that it can yield more optimal model performance [28]. Preprocessing includes data cleaning to remove duplicate and invalid data. Next, a resizing process is performed to standardize the image size to 128 x 128 pixels. Additionally, data normalization is performed using StandardScaler to stabilize the distribution of feature values. This step is crucial to ensure that each feature has a comparable scale so as not to affect the model training process. After preprocessing, the data is ready for use in the feature extraction and classification stages using the SVM and MLP methods.

After image preprocessing, the dataset was divided into 80% training data and 20% testing data using stratified train–test splitting to preserve the proportion of each disease class. Feature normalization using StandardScaler was then fitted only on the training data and subsequently applied to both the training and testing sets to prevent data leakage. Hyperparameter optimization was performed exclusively on the training set using five-fold Stratified Cross Validation, while the testing set was reserved only for the final performance evaluation.



Cordana Healthy Pesalotipsis Sigatoka

Figure 1 Banana Leaf Disease Image Dataset

### *C. Feature Extraction*

Feature extraction is a crucial step in digital image processing aimed at obtaining characteristic information from images so that it can be used in the classification process. In this study, a combination of color and texture features was used, namely Hue Saturation Value (HSV) and Gray Level Co-occurrence Matrix (GLCM). This combination of methods was chosen because it can represent color information and image texture patterns more optimally than using only one type of feature [29]. Color features in this study were extracted using the HSV (Hue Saturation Value) color space. The HSV color space was chosen because it can separate color components and light intensity, making it more stable against changes in lighting compared to the RGB color space.

In the HSV color space, the hue component indicates the type of color, saturation indicates the level of color saturation, while value indicates the image's brightness level. These characteristics make HSV more effective for use in digital image analysis, particularly in the classification of plant leaf diseases. Feature extraction is performed by converting RGB images into the HSV color space using the following equation [30], where $H$ = Hue, $S$ = Saturation, $V$= Value, $R$= Red, $G$= Green, $B$= Blue. The resulting HSV values are then used to represent the color of banana leaf images in the classification process.

$$H = \arctan(G - B, R - B) \quad (1)$$
$$S = \frac{max-min}{max} \quad (2)$$
$$V = \max(R, G, B) \quad (3)$$

In addition to color features, this study also employs texture features using the Gray Level Co-occurrence Matrix (GLCM) method. GLCM is a second-order statistical method used to analyze spatial relationships between pixels in grayscale images. This method is capable of representing image texture characteristics based on the distribution of pixel intensities at specific distances and angles [31]. The use of texture features is crucial in leaf disease classification because disease spots on the leaf surface produce specific texture patterns that can be distinguished based on pixel intensity distribution. In this study, four GLCM parameters were used: contrast, energy, homogeneity, and correlation. The contrast value is used to measure the degree of intensity difference between pixels in an image. Energy is used to indicate the level of uniformity in the texture pattern. Homogeneity is used to determine the degree of similarity in the distribution of elements within the GLCM matrix, while correlation is used to measure the linear relationship between pixels in the image. The mathematical equations for the GLCM features used are given in formulas (4) through (7), where $P(i, j)$ = the probability of a pixel's occurrence, $i, j$ = the pixel's position in the matrix, $\mu$ = the mean pixel intensity, and $\sigma$ = the standard deviation. The combination of HSV and GLCM features is considered effective in improving the system's ability to recognize image characteristics based on color and texture. The use of these two features has been widely applied in image classification research, particularly in the identification of plant diseases based on digital image processing [32].

$$Contrast = \sum_{i,j} |i - j|^2 . P(i, j) \quad (4)$$
$$Energy = \sum_{i,j} P(i, j)^2 \quad (5)$$
$$Homogeneity = \sum_{i,j} \frac{P(i,j)}{1+|i-j|} \quad (6)$$
$$Correlation = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)P(i,j)}{\sigma_i \sigma_j} \quad (7)$$

## D. Support Vector Machine (SVM)

Support Vector Machines (SVM) are a classification method that works by determining the optimal hyperplane to separate data across classes. This method can handle both linear and non-linear data through the use of kernel functions such as linear, polynomial, sigmoid, and Radial Basis Function (RBF). In this study, the kernel function used is the Radial Basis Function (RBF). The RBF kernel function is defined in Equation (8), Where $x$ and $y$ are two data points in the feature space representing test data and training data, respectively, while $||x - y||^2$ is the Euclidean distance squared between the two points, and $\gamma$ is the parameter that controls the influence range of each training data point.

The performance of RBF-based methods is highly dependent on the selection of parameter values, making the determination of the appropriate values crucial. Specifically, the accuracy of the γ parameter in the RBF kernel is the primary determining factor for SVM classification performance [33]. Given the crucial role of the $\gamma$ parameter, optimizing its value is an important aspect in the application of SVM across various domains, including the field of image processing. In image processing, SVMs are widely used because they have good generalization capabilities and are effective in processing high-dimensional data. The application of SVMs to the classification of banana leaf diseases has proven capable of producing high accuracy by utilizing image color and texture features [31].

$$K(x, y) = \exp(-\gamma ||x - y||^2 \quad (8)$$

## E. Multilayer Perceptron (MLP)

A Multilayer Perceptron (MLP) is a type of artificial neural network consisting of an input layer, hidden layers, and an output layer, and it uses nonlinear activation functions to learn complex patterns in data. MLPs excel at capturing nonlinear relationships between features, making them widely used in image classification. Compared to other methods, the MLP is capable of modeling more complex patterns, although its performance is heavily influenced by the amount of data and the parameters used [34]. Technically, each neuron in the hidden layer receives a signal from the previous layer, then calculates the weighted sum of all received inputs, and subsequently passes it through an activation function to produce a non-linear output. This process allows the network to represent relationships that cannot be captured by ordinary linear models [35].

The activation function used in the hidden layer in this study is ReLU (Rectified Linear Unit) as shown in Equation (9). This function works by passing the input value directly if it is positive, and outputting a value of zero if the input is negative. $ReLU$ was chosen for its ability to address the vanishing gradient problem commonly encountered in conventional activation functions such as sigmoid and tanh, thereby making the network training process more stable and efficient [35]. Meanwhile, in the output layer, the Softmax function is used to convert the network's raw output values into a probability distribution across classes [36], as shown in Equation (10). This function ensures that all output values fall within the range [0,1] and sum to one, allowing them to be directly interpreted as the probability that a data point belongs to a specific class. The combination of ReLU in the hidden layer and Softmax in the output layer is a configuration commonly used in multi-class classification because it has been empirically proven to deliver good performance.

$$ReLU(x) = \max(0, x) \quad (9)$$

$$Softmax(z_i) = \frac{e^{zi}}{\sum_j e^{zj}} \quad (10)$$

### *F. Hyperparameter Optimization and Fair Comparison*

To obtain the optimal classifier configuration, hyperparameter optimization was performed using GridSearchCV with five-fold Stratified Cross Validation. For SVM, the regularization parameter (C) and kernel coefficient (γ) of the RBF kernel were optimized. For MLP, the hidden layer architecture, activation function, learning rate, and regularization parameter were optimized. The classifier achieving the highest mean cross-validation accuracy was selected as the final model for performance evaluation. To ensure a fair comparison, both SVM and MLP were implemented using the same image preprocessing pipeline, identical HSV and GLCM feature representations, the same training and testing data partitions, and identical evaluation metrics. This experimental design allows both classifiers to be evaluated under consistent conditions so that any observed performance differences can be attributed primarily to the learning characteristics of each classifier rather than differences in data preparation or feature extraction.

### *G. Research Stages*

This research was conducted systematically to classify banana leaf diseases using the Support Vector Machine (SVM) and Multilayer Perceptron (MLP) methods. The research flowchart is shown in Figure 3.

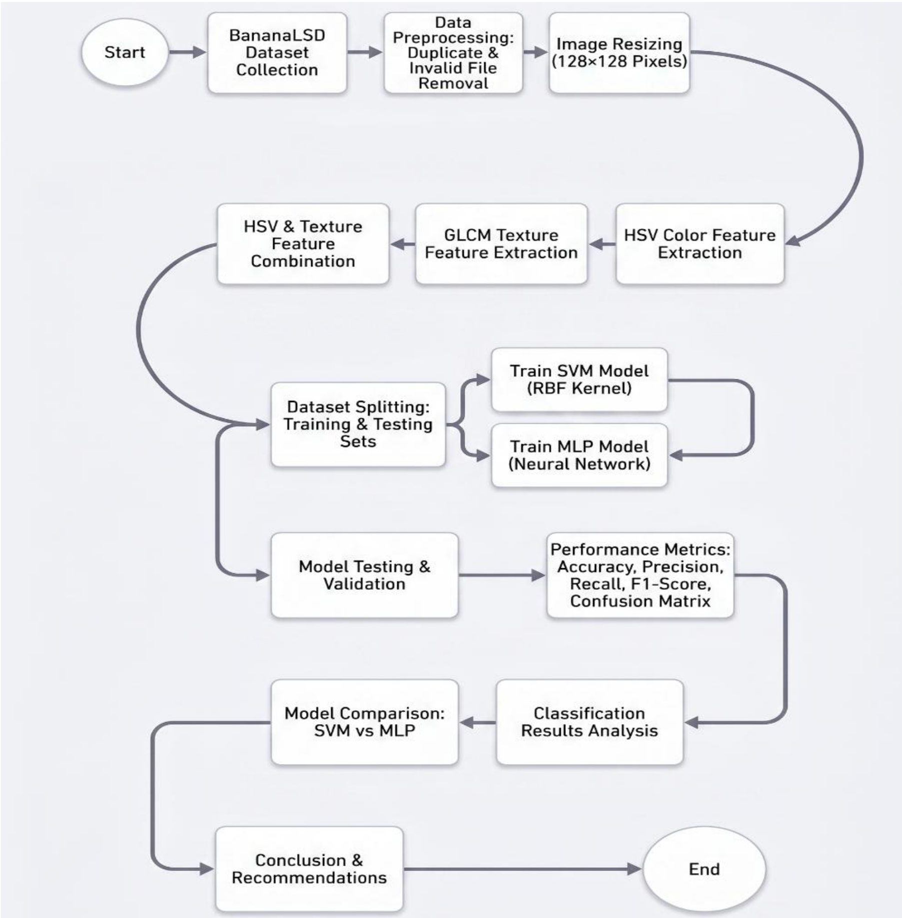



Figure 3 Proposed Method

### *H. Evaluation*

The performance of the classification model was measured using a confusion matrix, which is a table used to evaluate prediction results by comparing actual data with the model's predictions. The confusion matrix consists of four categories: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), each of which indicates the number of correct and incorrect predictions made by the model [37].

Based on the confusion matrix, model performance is evaluated using several evaluation metrics: accuracy, precision, recall, and F1-score. The use of multiple evaluation metrics aims to provide a more accurate and comprehensive analysis of the model's performance. Accuracy is used to measure the model's precision in classifying the entire dataset. Precision is used to determine the accuracy of the positive predictions generated by the model. Recall is used to measure the model's ability to correctly identify positive data, while the F1-score is used to balance the values of precision and recall. The mathematical equations used in the evaluation process are shown in formulas (11) through (14), where $TP$ = True Positive, $TN$ = True Negative, $FP$ = False Positive, dan $FN$ = False Negative. The use of the confusion matrix and these evaluation metrics is considered effective in measuring the performance of image classification models, particularly in machine learning-based research on plant disease classification.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{11}$$

$$Precision = \frac{TP}{TP+FP} \tag{12}$$

$$Recall = \frac{TP}{TP+FN} \tag{13}$$

$$F1Score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \tag{14}$$

## III. HASIL DAN PEMBAHASAN

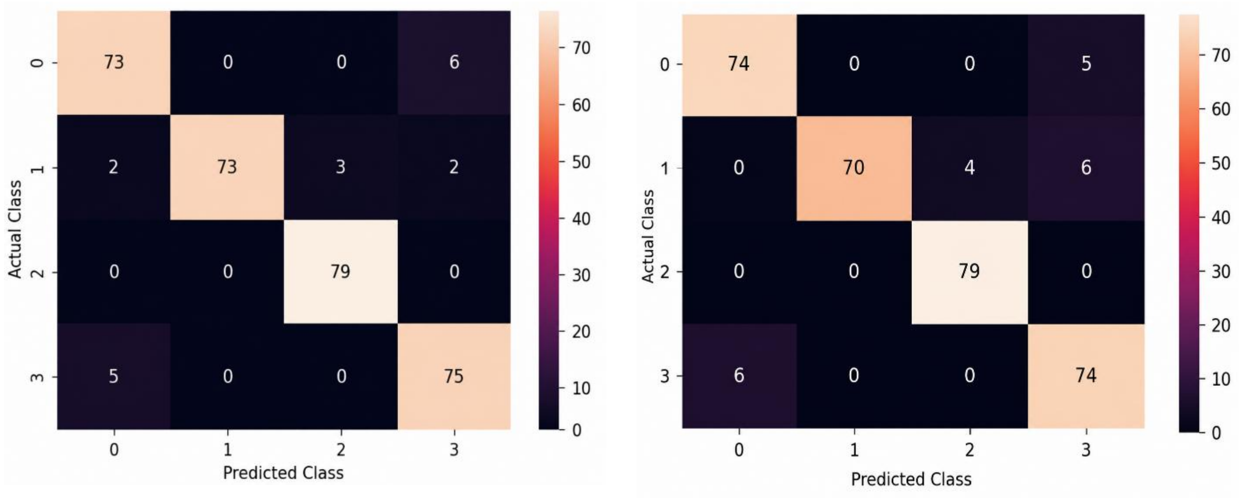
The testing aimed to evaluate the performance of SVM and MLP in classifying banana leaf diseases based on HSV and GLCM features. The dataset included four classes: healthy, cordana, pestalotiopsis, and sigatoka. The test results show that SVM outperforms MLP and compares favorably with other studies, as shown in Table I. From the test results, SVM proved to be better than MLP at classifying banana leaf diseases. SVM's superiority is primarily supported by the RBF kernel's ability to optimally separate non-linear data. The HSV and GLCM features also produce structured data representations, making the class separation process by the SVM hyperplane easier. Specifically, the HSV feature is capable of consistently representing leaf color even when lighting conditions change, while the GLCM feature effectively captures the texture patterns of disease spots on the leaf surface. On the other hand, while MLP is capable of learning complex non-linear patterns, its performance is highly dependent on data availability and the appropriate network configuration. The limited dataset size in this study resulted in MLP performance failing to surpass that of SVM. Based on the experiments conducted, SVM is deemed more suitable for the banana leaf image dataset utilizing a combination of color and texture features.

A confusion matrix is used to evaluate the model's success in classifying each class of banana leaf diseases. With SVM, nearly all data were classified correctly. Classification errors only occurred in a small portion of images that shared similar visual characteristics across classes. In the MLP, the classification error rate was higher than that of the SVM. A number of disease images were still misclassified because they had relatively similar color and texture patterns. Overall, the confusion matrix indicates that SVM has better generalization capabilities in distinguishing types of banana leaf diseases. SVM with an RBF kernel achieved the highest accuracy of 94.34%, while the MLP achieved an accuracy of 93.34%. A comparison of the accuracy of the two methods is presented in Figure 4.

The results of the evaluation using a confusion matrix show that SVM outperforms MLP in terms of classification accuracy. The SVM model was able to accurately identify most images within each category of banana leaf disease. The "healthy," "cordana," and "pestalotiopsis" categories were found to have satisfactory prediction accuracy with relatively low error rates. The SVM model is not entirely free from errors, particularly in the "sigatoka" and "pestalotiopsis" categories. These two diseases share similarities in both spot color and texture patterns, causing the model to fail to accurately distinguish some images. These shared visual characteristics are the primary cause of prediction errors in several test samples. The MLP demonstrated relatively lower performance compared to the SVM. Data from the confusion matrix indicate that the MLP produced more errors, particularly in the "healthy" and "Pestalotiopsis" categories. This suggests that the MLP requires a larger volume of training data to capture complex patterns more effectively. HSV color features and GLCM texture features in this study has been shown to significantly contribute to the model's ability to identify types of banana leaf diseases. HSV helps distinguish color variations in disease spots, while GLCM captures texture patterns on the leaf surface. The combination of these two features yields fairly optimal classification performance, particularly when applied to an SVM with an RBF kernel.

Based on the comparison presented in Table 2, the Support Vector Machine (SVM) method demonstrates better classification performance compared to the Multilayer Perceptron (MLP) method in several studies related to leaf disease classification using digital images. Research conducted by Nur Sholehah Mat Said et al.[3] reported an accuracy of 83.33% using the SVM method for the classification of banana leaf diseases. In addition, the study by Thaiyalnayaki and Joseph [2]achieved an accuracy of 94.14% using SVM and 88.73% using MLP for the classification of plant diseases based on leaf images.

Table III demonstrates that the combination of HSV and GLCM features consistently yields better classification performance than using either feature independently. HSV provides discriminative color information, whereas GLCM contributes complementary texture characteristics. The integration of both feature representations produces a richer image description, resulting in improved classification accuracy and F1-score for banana leaf disease classification. These findings indicate that both color and texture information are important for distinguishing visually similar banana leaf diseases.




SVM confusion matrix MLP confusion matrix

Figure 4 Comparison of confusion matrix

TABLE III
ABLATION STUDY

| Feature Representation | Accuracy | F1-score |
|---|---|---|
| HSV | 93.40 | 0.9343 |
| GLCM | 69.18 | 0.6909 |
| HSV + GLCM | 94.34 | 0.9435 |

TABLE IV
COMPUTATIONAL EFFICIENCY OF SVM AND MLP

| Method | Feature Extraction (s) | Training Time (s) | Testing Time (s) |
|---|---|---|---|
| SVM | 13.33 | 10.58 | 0.0182 |
| MLP | 13.33 | 43.05 | 0.0017 |

Table IV presents the computational efficiency of both classifiers in terms of feature extraction, training, and testing time. Since both models use the same HSV and GLCM feature extraction pipeline, the feature extraction time is identical. However, the SVM required substantially less training time than the MLP while achieving higher classification accuracy. Although the MLP exhibited a slightly shorter testing time, the difference was negligible in practical applications. Overall, SVM provides a more favorable balance between predictive performance and computational cost, making it more suitable for banana leaf disease classification on moderate-sized datasets.

Based on Table II, the proposed method in this research achieved an accuracy of 94.34% using the SVM method with the RBF kernel, while the MLP method obtained an accuracy of 93.40%. These results indicate that the SVM method still performs better than MLP in classifying banana leaf diseases. The use of the RBF kernel in SVM is considered more effective in handling non-linear data patterns found in digital leaf images, allowing the classification model to separate disease characteristics more accurately. The high classification performance achieved in this research is also influenced by the combination of HSV color features and GLCM texture features. The HSV feature extraction method is capable of representing leaf color characteristics such as green, yellow, and brown variations, which are important indicators in identifying banana leaf diseases. Meanwhile, the GLCM method captures texture information through parameters such as contrast, energy, homogeneity, and correlation. The combination of these two features provides a more comprehensive image representation and contributes significantly to improving the classification accuracy of both SVM and MLP models.

Recent deep learning approaches, including ResNet, EfficientNet, MobileNet, and Vision Transformer, have demonstrated excellent performance in plant leaf disease classification, particularly when trained using large-scale image datasets [38], [39], [9]. However, these approaches generally require substantially greater computational resources, longer training times, and larger annotated datasets than classical machine learning methods [38], [40]. The objective of this study was not to outperform deep learning models, but rather to provide a fair comparison between optimized SVM and MLP classifiers using identical HSV and GLCM feature representations. Considering the moderate dataset size and the computational efficiency achieved in this study, classical machine learning remains a practical and effective alternative for banana leaf disease classification.

The use of HSV color features contributed to the robustness of the proposed approach under moderate illumination variations because the HSV color space separates chromatic information from brightness, making color representation less sensitive to lighting changes than RGB [41]. In addition, the incorporation of GLCM texture features provides complementary information regarding disease spot patterns on the leaf surface. However, images acquired under extreme lighting conditions, complex natural backgrounds, or poor image quality may still affect the consistency of feature extraction and reduce classification performance [42]. Therefore, further evaluation using field images collected under more diverse environmental conditions is recommended to assess the robustness and generalization capability of the proposed method.

This study evaluated model generalization using five-fold Stratified Cross Validation, which provides a more reliable estimate of classification performance than a single train–test split by reducing the influence of data partitioning [6]. However, the evaluation was conducted using a single publicly available dataset. Consequently, the robustness of the proposed approach under different acquisition conditions, such as images captured using different devices, illumination levels, and environmental backgrounds, has not yet been fully validated. Future research should evaluate the proposed method using external datasets collected under more diverse field conditions to further assess its generalization capability.

TABLE V
COMPARISON RESULT

| Research by | Method | Subject | Result |
|---|---|---|---|
| [3] | SVM | Classification of banana leaf diseases | SVM = 83,33% |
| [2] | SVM and MLP | Classification of plant diseases based on leaf images | SVM = 94,14% MLP = 88,73% |
| Our | SVM Kernel RBF and MLP | Classification of banana leaf diseases | SVM = 94,34% MLP = 93,40% |

## IV. CONCLUSION

This study successfully developed a classification system for banana leaf diseases using the Support Vector Machine (SVM) method based on the Radial Basis Function (RBF) kernel and the Multilayer Perceptron (MLP). The system development process involved a series of steps, ranging from the collection of the BananaLSD dataset, data preprocessing, and the extraction of HSV color and GLCM texture features, to model training and comprehensive performance evaluation. Based on the evaluation conducted, the SVM method achieved an accuracy of 94.34%, while the MLP reached an accuracy of 93.40%. This difference in performance indicates that SVM is superior in classifying banana leaf diseases on the dataset used in this study. The advantages of SVM are supported by its ability to determine the optimal hyperplane for separating data classes, particularly with the help of the RBF kernel, which can handle non-linear data distributions. On the other hand, MLP continues to demonstrate good performance, but its performance is highly dependent on the availability of a larger dataset so that complex patterns can be learned more effectively. Confusion matrix analysis indicates that most images can be classified accurately by both methods. Nevertheless, a small number of classification errors are still found in disease classes that share visual similarities in terms of color and texture. This underscores the importance of feature representation quality in determining the success of the classification process. This study demonstrates that the integration of HSV color features and GLCM texture features with the RBF kernel SVM method is an effective approach for classifying banana leaf diseases. These findings are expected to serve as a scientific foundation for the development of artificial intelligence-based plant disease detection systems in the future.

## REFERENCES

[1] J. D. Thiagarajan *et al.*, “Analysis of banana plant health using machine learning techniques,” *Sci. Rep.*, vol. 14, no. 1, p. 15041, Jul. 2024, doi: 10.1038/s41598-024-63930-y.

[2] V. Chaudhari and M. P. Patil, “Detection and Classification of Banana Leaf Disease Using Novel Segmentation and Ensemble Machine Learning Approach,” *Applied Computer Systems*, vol. 28, no. 1, pp. 92–99, Jun. 2023, doi: 10.2478/acss-2023-0009.

[3] N. S. M. Said, H. Madzin, S. K. Ali, and N. S. Beng, “Comparison of color-based feature extraction methods in banana leaf diseases classification using SVM and K-NN,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 24, no. 3, p. 1523, Dec. 2021, doi: 10.11591/ijeecs.v24.i3.pp1523-1533.

[4] A. Anton, S. Rustad, G. F. Shidik, and A. Syukur, “Comparative Analysis Of Cnn Architectures For Tomato Leaf Disease Classification Using Transfer Learning,” *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, vol. 11, no. 1, pp. 125–135, Aug. 2025, doi: 10.33480/jitk.v11i1.7014.

[5] A. Sudrajat, R. D. Apnena, A. H. Rahayu, and M. Iqtait, “Multi-Class Mangrove Classification Using Transfer Learning with MobileNet-V3 on Multi-Organ Images,” *Jurnal Teknik Informatika (Jutif)*, vol. 6, no. 3, pp. 1631–1642, Jul. 2025, doi: 10.52436/1.jutif.2025.6.3.4683.

[6] F. Farhan, C. A. Sari, E. H. Rachmawanto, and N. R. D. Cahyo, “Mangrove Tree Species Classification Based on Leaf, Stem, and Seed Characteristics Using Convolutional Neural Networks with K-Folds Cross Validation Optimalization,” *Advance Sustainable Science Engineering and Technology*, vol. 5, no. 3, p. 02303011, Oct. 2023, doi: 10.26877/asset.v5i3.17188.

[7] Md. Fahim-Ul-Islam, A. Chakrabarty, R. Rahman, H. Moon, and Md. J. Piran, “Advancing mango leaf variant identification with a robust multi-layer perceptron model,” *Sci. Rep.*, vol. 14, no. 1, p. 27406, Nov. 2024, doi: 10.1038/s41598-024-74612-0.

[8] Y. Sari, M. Maulida, R. Maulana, J. Wahyudi, and A. Shalludin, “Detection of Corn Leaves Nutrient Deficiency Using Support Vector Machine (SVM),” in *2021 4th International Conference of Computer and Informatics Engineering (IC2IE)*, IEEE, Sep. 2021, pp. 396–400. doi: 10.1109/IC2IE53219.2021.9649375.

[9] R. Ariska, C. A. Sari, and E. H. Rachmawanto, “Classification of Corn Leaf Disease Using Convolutional Neural Network,” *Advance Sustainable Science Engineering and Technology*, vol. 6, no. 4, p. 0240408, Aug. 2024, doi: 10.26877/asset.v6i4.772.

[10] S. Srinivasan, S. M. Prabin, S. K. Mathivanan, H. Rajadurai, S. Kulandaivelu, and M. A. Shah, “Sugarcane leaf disease classification using deep neural network approach,” *BMC Plant Biol.*, vol. 25, no. 1, p. 282, Mar. 2025, doi: 10.1186/s12870-025-06289-0.

[11] R. Setiawan, H. Zein, R. A. Azdy, and S. Sulistyowati, “Rice Leaf Disease Classification with Machine Learning: An Approach Using Nu-SVM,” *Indonesian Journal of Data and Science*, vol. 4, no. 3, pp. 136–144, Dec. 2023, doi: 10.56705/ijodas.v4i3.114.

[12] Y. CAMGÖZLÜ and Y. KUTLU, “Leaf Image Classification Based on Pre-trained Convolutional Neural Network Models,” *Nat. Eng. Sci.*, vol. 8, no. 3, pp. 214–232, Dec. 2023, doi: 10.28978/nesciences.1405175.

[13] P. R. and V. A., “Disease diagnosis in banana leaves: a review on AI powered techniques,” *PeerJ Comput. Sci.*, vol. 11, p. e3310, Dec. 2025, doi: 10.7717/peerj-cs.3310.

[14] S. Bagheri, S. Taridashti, H. Farahani, P. Watson, and E. Rezvani, “Multilayer perceptron modeling for social dysfunction prediction based on general health factors in an Iranian women sample,” *Front. Psychiatry*, vol. 14, Dec. 2023, doi: 10.3389/fpsyt.2023.1283095.

[15] D.-C. Rodríguez-Lira, D.-M. Córdova-Esparza, J. M. Álvarez-Alvarado, J. Terven, J.-A. Romero-González, and J. Rodríguez-Reséndiz, “Trends in Machine and Deep Learning Techniques for Plant Disease Identification: A Systematic Review,” *Agriculture*, vol. 14, no. 12, p. 2188, Nov. 2024, doi: 10.3390/agriculture14122188.

[16] N. R. D. Cahyo, C. A. Sari, E. H. Rachmawanto, C. Jatmoko, R. R. A. Al-Jawry, and M. A. Alkhafaji, “A Comparison of Multi Class Support Vector Machine vs Deep Convolutional Neural Network for Brain Tumor Classification,” in *2023 International Seminar on Application for Technology of Information and Communication (iSemantic)*, IEEE, Sep. 2023, pp. 358–363. doi: 10.1109/iSemantic59612.2023.10295336.

[17] P. Pungki, C. Atika Sari, D. R. Ignatius Moses Setiadi, and E. Hari Rachmawanto, “Classification of Plant Types based on Leaf Image using the Artificial Neural Network Method,” in *2020 International Seminar on Application for Technology of Information and Communication (iSemantic)*, IEEE, Sep. 2020, pp. 67–72. doi: 10.1109/iSemantic50169.2020.9234196.

[18] Anjna, M. Sood, and P. K. Singh, “Hybrid System for Detection and Classification of Plant Disease Using Qualitative Texture Features Analysis,” *Procedia Comput. Sci.*, vol. 167, pp. 1056–1065, 2020, doi: 10.1016/j.procs.2020.03.404.

[19] X. He and Y. Chen, “Modifications of the Multi-Layer Perceptron for Hyperspectral Image Classification,” *Remote Sens. (Basel).*, vol. 13, no. 17, p. 3547, Sep. 2021, doi: 10.3390/rs13173547.

[20] M. A. Al-Hashem, A. M. Alqudah, and Q. Qananwah, “Performance Evaluation of Different Machine Learning Classification Algorithms for Disease Diagnosis,” *International Journal of E-Health and Medical Communications*, vol. 12, no. 6, 2021, doi: 10.4018/IJEHMC.20211101.oa5.

[21] S. M. Kiran and D. N. Chandrappa, “Plant Leaf Disease Detection Using Efficient Image Processing and Machine Learning Algorithms,” *Journal of Robotics and Control (JRC)*, vol. 4, no. 6, pp. 840–848, 2023, doi: 10.18196/jrc.v4i6.20342.

[22] S. Jha, V. Luhach, G. S. Gupta, and B. Singh, “Crop Disease Classification using Support Vector Machines with Green Chromatic Coordinate (GCC) and Attention based feature extraction for IoT based Smart Agricultural Applications,” Nov. 2023, doi: 10.48550/arXiv.2311.00429.

[23] N. Ahmad, H. M. S. Asif, G. Saleem, M. U. Younus, S. Anwar, and M. R. Anjum, “Leaf Image-Based Plant Disease Identification Using Color and Texture Features,” *Wirel. Pers. Commun.*, vol. 121, no. 2, pp. 1139–1168, Nov. 2021, doi: 10.1007/s11277-021-09054-2.

[24] J. P. G. Lakshika and T. S. Talagala, “Computer-aided Interpretable Features for Leaf Image Classification,” Aug. 2021, doi: 10.48550/arXiv.2106.08077.

[25] K. Thaiyalnayaki and C. Joseph, “Classification of plant disease using SVM and deep learning,” *Mater. Today Proc.*, vol. 47, pp. 468–470, 2021, doi: 10.1016/j.matpr.2021.05.029.

[26] G. Garg and M. Biswas, “Improved Neural Network-Based Plant Diseases Identification,” 2021, pp. 79–89. doi: 10.1007/978-981-15-5341-7_6.

[27] Y. Kurmi and S. Gangwar, “A leaf image localization based algorithm for different crops disease classification,” *Information Processing in Agriculture*, vol. 9, no. 3, pp. 456–474, Sep. 2022, doi: 10.1016/j.inpa.2021.03.001.

[28] H.-J. Park, Y.-S. Koo, H.-Y. Yang, Y.-S. Han, and C.-S. Nam, “Study on Data Preprocessing for Machine Learning Based on Semiconductor Manufacturing Processes,” *Sensors*, vol. 24, no. 17, p. 5461, Aug. 2024, doi: 10.3390/s24175461.

[29] K. V. Gokula, D. J., R. Pinagadi Venkateswara, D. V., and K. S., “An automated segmentation and classification model for banana leaf disease detection,” *J. Appl. Biol. Biotechnol.*, vol. 10, no. 1, pp. 213–220, Oct. 2022, doi: 10.7324/JABB.2021.100126.

[30] K.-P. Lee, J. Yip, K.-L. Yick, C. Lu, L. Lu, and Q.-W. E. Lei, “A Novel Force-Sensing Smart Textile: Inserting Silicone-Embedded FBG Sensors into a Knitted Undergarment,” *Sensors*, vol. 23, no. 11, p. 5145, May 2023, doi: 10.3390/s23115145.

[31] S. M. Javidan, A. Banakar, K. Rahnama, K. A. Vakilian, and Y. Ampatzidis, “Feature engineering to identify plant diseases using image processing and artificial intelligence: A comprehensive review,” *Smart Agricultural Technology*, vol. 8, p. 100480, Aug. 2024, doi: 10.1016/j.atech.2024.100480.

[32] Siti Saniah and Mhd. Furqan, “Classification Of Rice Plant Diseases Using K-Nearest Neighbor Algorithm Based On Hue Saturation Value Color Extraction And Gray Level Co-Occurrence Matrix Features,” *Jurnal Teknologi Dan Open Source*, vol. 7, no. 2, pp. 212–223, Dec. 2024, doi: 10.36378/jtos.v7i2.3972.

[33] Y. R. Denny, E. Permata, and L. D. Assaat, “Classification of Diseases of Banana plant Fusarium Wilted Banana Leaf Using Support Vector Machine,” *Gravity : Jurnal Ilmiah Penelitian dan Pembelajaran Fisika*, vol. 8, no. 1, Jun. 2022, doi: 10.30870/gravity.v8i1.15893.

[34] S. Wang *et al.*, “Advances in Deep Learning Applications for Plant Disease and Pest Detection: A Review,” *Remote Sens. (Basel).*, vol. 17, no. 4, p. 698, Feb. 2025, doi: 10.3390/rs17040698.

[35] T. Apleni, F. O. Isinkaye, and M. O. Olusanya, “Ensemble-based feature fusion for accurate plant disease classification using pre-trained models,” *Sci. Rep.*, vol. 15, no. 1, p. 41925, Nov. 2025, doi: 10.1038/s41598-025-25927-z.

[36] B. Asadi and H. Jiang, “On Approximation Capabilities of ReLU Activation and Softmax Output Layer in Neural Networks,” Feb. 2020, [Online]. Available: http://arxiv.org/abs/2002.04060

[37] N. Rendevski, A. Bocevska, Z. Kotevski, and T. Dimovski, “Immersive Technologies in 5G-Enabled Applications: Some Technical Challenges of the Novel Usage Models,” *International Journal of Computer Science and Information Technology*, vol. 13, no. 6, pp. 31–37, Dec. 2021, doi: 10.5121/ijcsit.2021.13603.

[38] M. K. Kasambya and K. M. Lauraine, “<p>Explainable and Robust Deep Learning for Banana Leaf Disease Classification: A Comparative Study of MobileNet, ResNet50, and EfficientNetB0 With Gradient-weighted Class Activation Mapping (Grad-CAM) and Statistical Validation</p>,” *Cureus Journal of Computer Science*, Jun. 2026, doi: 10.7759/s44389-026-00114-x.

[39] M. Syifa Aryanta, C. A. Sari, and H. Rachmawanto, “A Banana Disease Detection Using MobileNetV2 Model Based on Adam Optimizer,” 2025. [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC

[40] Ü. Atila, M. Uçar, K. Akyol, and E. Uçar, “Plant leaf disease classification using EfficientNet deep learning model,” *Ecol. Inform.*, vol. 61, p. 101182, Mar. 2021, doi: 10.1016/j.ecoinf.2020.101182.

[41] D. P. Sönmez and Ş. Kılıç, “PlantPathNet: a novel network based on cross-layer feature integration for plant disease classification,” *Front. Plant Sci.*, vol. 17, Apr. 2026, doi: 10.3389/fpls.2026.1783576.

[42] D. P. Sönmez and Ş. Kılıç, “PlantPathNet: a novel network based on cross-layer feature integration for plant disease classification,” *Front. Plant Sci.*, vol. 17, Apr. 2026, doi: 10.3389/fpls.2026.1783576.