A Hybrid Deep Feature-Based VGG19 and Support Vector Machine Approach for Durian Leaf Classification

[1*]Jennifer Velensia Santoti, [2]Siska Devella
[1,2]Department of Informatics, Facculty of Computer Science and Engineering,
Universitas Multi Data Palembang, Indonesia
Email: [1]jennifervelensiasantoti_2226250014@mhs.mdp.ac.id, [2]siskadevella@mdp.ac.id

Article Info
Article history:
Received Dec 22th, 2025
Revised Jan 13th, 2026
Accepted Feb 24th, 2026

Keyword:
Durian Leaf Classification
Image Classification
Support Vector Machine
Transfer Learning
VGG-19

ABSTRACT
Durian leaf classification has remained challenging due to high visual similarity among superior durian varieties and the limited robustness of conventional convolutional neural network models that rely on Softmax classifiers. This study aimed to address this limitation by investigating a deep feature-based classification framework that combined VGG19 as a feature extractor with a Support Vector Machine classifier. The experiments were conducted on a dataset of 1,530 durian leaf images representing four varieties: Bawor, Duri Hitam, Musang King, and Super Tembaga. Four experimental scenarios were designed to evaluate classification performance using Support Vector Machine and Softmax classifiers under both imbalanced and balanced data conditions through the application of Synthetic Minority Over-sampling Technique. The research gap addressed in this study lay in the absence of prior investigations that systematically evaluated the integration of VGG19 and Support Vector Machine for durian leaf variety classification under varying data distributions. Experimental results showed that the proposed VGG19–Support Vector Machine framework consistently achieved higher accuracy and more stable performance than Softmax-based models. This study demonstrated that replacing the conventional Softmax classifier with a Support Vector Machine significantly improved classification robustness compared to previous approaches that employed end-to-end convolutional neural network architectures.

Corresponding Author:
Jennifer Velensia Santoti,
Department of Informatics,
Universitas Multi Data Palembang,
14 Rajawali Street, 9 Ilir, East Ilir II District, Palembang City, South Sumatra, 30113, Indonesia.
Email: jennifervelensiasantoti_2226250014@mhs.mdp.ac.id
DOI: http://dx.doi.org/10.24014/ijaidm.v9i1.38831

1. INTRODUCTION
1.2. Research Background

Durian (Durio zibethinus L.) is an important horticultural commodity in Indonesia, particularly in regions such as Borneo, which are recognized as centers of diversity for the Durio genus [1]. Accurate identification of superior durian varieties is essential to support seedling certification, plantation management, and productivity improvement in modern agricultural practices. However, varietal identification remains challenging due to the high visual similarity among durian varieties, especially when relying on manual observation. Leaf images provide an effective visual reference for classification because their morphological characteristics, including shape and texture, remain relatively stable throughout the year and can be captured consistently under various conditions [2].

Although Convolutional Neural Network (CNN)–based approaches have demonstrated high accuracy in durian leaf classification [2], [3], most existing studies rely on conventional Softmax classifiers. These classifiers are often sensitive to feature overlap and class imbalance, which frequently results in performance degradation when handling visually similar varieties or uneven data distributions [4]. Such limitations highlight the need for more robust classification strategies that maintain stable performance under complex visual conditions. This necessity motivates the investigation of hybrid deep feature–based frameworks that integrate powerful feature extraction with more effective classifiers. Without addressing these limitations, automated durian variety identification systems may remain insufficiently reliable for practical deployment in real-world agricultural settings.

Recent advances in computer vision and deep learning have enabled the development of automated solutions for plant-based image classification. Several studies have demonstrated that convolutional neural networks can extract discriminative visual features from durian leaf images. CNN-based approaches using architectures such as MobileNet, AlexNet, InceptionV3, and MobileNetV2 have reported high classification accuracy for durian variety identification based on leaf images [2], [3], [5]. These results indicate that deep learning models are effective in capturing subtle differences in leaf texture and structure. Nevertheless, most of these studies rely on conventional Softmax classifiers, which may experience performance degradation when visual similarities between classes are high.

Beyond varietal classification, leaf images have also been widely utilized in plant disease detection tasks. Several international studies have applied deep learning architectures such as ResNet-50, Vision Transformer, VGG-16, and VGG-19 to classify diseases in durian leaves and other agricultural crops, including tomato, grapevine, and medicinal plants [6], [7], [8], [9], [10], [11], [12]. These studies consistently reported strong classification performance, confirming that leaf morphology contains rich visual information suitable for deep feature extraction. However, their primary focus was disease identification rather than varietal classification, and most did not investigate alternative classification strategies beyond standard CNN–Softmax configurations.

Despite continuous architectural improvements, convolutional neural network models that rely solely on Softmax classifiers remain sensitive to feature overlap and overfitting when handling complex and visually similar classes. Advanced CNN variants and hybrid feature representations have been proposed to mitigate these issues; however, classification instability is still observed under challenging visual conditions, indicating that classifier selection plays a critical role in determining model robustness [13].

To address these limitations, hybrid classification models that combine convolutional neural networks with support vector machines have gained increasing attention in various image classification domains. International studies have shown that CNN–support vector machine models can improve classification stability, reduce overfitting, and achieve higher accuracy compared to conventional CNN–Softmax approaches [4], [14]. Support vector machines are particularly effective in handling high-dimensional feature spaces through appropriate kernel functions, enabling non-linear class separation that is difficult to achieve with linear classifiers [15]. Furthermore, the robustness and flexibility of support vector machines have been demonstrated in optimized computational implementations, highlighting their reliability for complex classification tasks [16].

Among various convolutional neural network architectures, VGG-19 is widely recognized for its deep convolutional structure and strong ability to capture hierarchical spatial features [15]. Through transfer learning, VGG-19 models pretrained on large-scale datasets have demonstrated strong adaptability across various agricultural image classification tasks, even under limited training data conditions [13], [16]. Despite its proven effectiveness, existing studies on durian leaf classification have not explored the integration of VGG-19 as a feature extractor with a support vector machine classifier, nor have they systematically evaluated such a hybrid approach under different data distribution scenarios.

Based on this research landscape, it can be observed that although convolutional neural networks have been extensively applied for leaf-based image classification, a comprehensive investigation of a VGG-19 and support vector machine hybrid model for durian leaf variety classification remains unexplored. Therefore, this study proposes a transfer learning–based framework that integrates VGG-19 feature extraction with a support vector machine classifier and evaluates its performance under multiple experimental scenarios, including balanced and imbalanced datasets. This approach aims to improve classification accuracy and stability, thereby advancing more reliable durian variety identification systems to support precision agriculture in Indonesia.

1.2. Related Work On CNN–SVM and SMOTE

Various studies have demonstrated that integrating Convolutional Neural Networks (CNN) with Support Vector Machines (SVM) can enhance classification performance compared to using CNNs with conventional classifiers like Softmax. Yohannes et al. [4] reported that features extracted using VGG-19 and ResNet-50, subsequently classified with an SVM, achieved an accuracy of 93.28% in identifying Van Gogh paintings, thereby confirming SVM's ability to provide more stable separation of high-dimensional features. In skin cancer classification [14], the CNN–SVM approach also showed competitive performance; specifically, the VGG-19 architecture utilizing a linear kernel and patch-based preprocessing achieved a peak accuracy of 65.33%, outperforming ResNet-50 configurations under the same testing scheme. These findings suggest that the combination of a CNN architecture and an SVM kernel significantly improves the accuracy of skin lesion identification. Furthermore, the issue of data imbalance also impacts model performance. Research by Ningrum et al. [17] proved that the application of the Synthetic Minority Over-sampling Technique (SMOTE) improved classification accuracy—notably increasing MobileNetV2's performance from 91.11% to 92.42%—by creating a more balanced class distribution. Collectively, these results reinforce that the integration of CNN–SVM is effective in enhancing classification accuracy and stability, and that addressing data imbalance is a critical factor in optimizing model performance.

2. RESEARCH METHOD

2.1. Research Dataset

In this study, the public Durian Leaf dataset was used, which was developed from a previous study titled, "Durian Seedling Variety Identification Using MobileNetV2 Based on Leaf Images." [3]. The dataset was collected through direct observation, with leaves from four superior durian varieties (Bawor, Duri Hitam, Musang King, and Super Tembaga) photographed on durian farms in Indonesia. The dataset consisted of a total of 1,530 durian leaf images distributed across four classes: Bawor (400 images), Super Tembaga (400 images), Musang King (400 images), and Duri Hitam (330 images). The dataset was divided using a random splitting strategy, where 80% of the images from each class were allocated for training and the remaining 20% were used for testing. No separate validation set was employed in this study, as model evaluation was conducted exclusively on the testing dataset to ensure consistent performance comparison across all experimental scenarios. Details of the dataset division are shown in Table 1, and examples of durian leaves from each class are displayed in Figure 1.

Table 1. Durian Leaf Dataset Splitting

No	Durian Variety	Total Images	Train	Test
1	Bawor	400	320	80
2	Super Tembaga	400	320	80
3	Musang King	400	320	80
4.	Duri Hitam	330	264	66

Figure 1. Durian Leaf Image Samples for Classification

2.2. Proposed Architecture

As shown in Figure 2, this research was conducted through four main scenarios. In the first scenario, the VGG-19–SVM model was trained on a class-imbalanced dataset without SMOTE oversampling, so training was performed directly on the original data distribution. The second scenario used the same setup but included SMOTE to balance the class distributions before training. The third and fourth scenarios involved using the VGG19 architecture as the primary model. In the third scenario, VGG19 was trained without data balancing, while in the fourth, the same model was combined with SMOTE. Comparing these four scenarios provides a basis for understanding the effects of data imbalance, the utility of SMOTE, and how architectural differences influence the overall performance of the classification system.

The stages begin with the Durian Leaf dataset from Kaggle, which contains leaf images of four leading durian varieties: Bawor, Duri Hitam, Musang King, and Super Tembaga. Next, data splitting was performed across all scenarios, with 80% of the dataset used for training and 20% for testing to ensure consistent model evaluation. Once the dataset was prepared, preprocessing and image augmentation were performed via a data generator to ensure consistent input representation and enhance model generalization. During the training phase, images were resized, normalized, and subjected to various augmentation techniques to increase data diversity without altering class labels, as summarized in Table 2.

Conversely, the testing phase involved only resizing and normalization to ensure that the evaluation accurately reflects the model's performance on unseen data. Additionally, the data generator applied one-hot encoding to the class labels, enabling the classification model to process them directly.

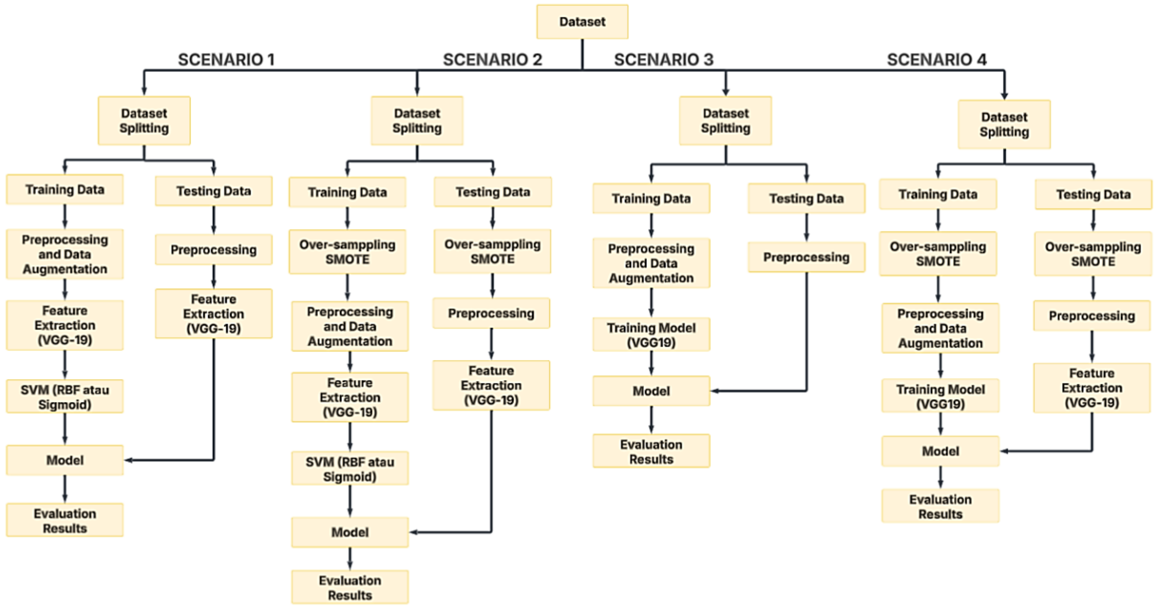


Figure 2. Proposed stages

Table 2. Image Pre-processing and Augmentation

Process	Parameter
Resize	target_size = (224, 224)
Rescaling	rescale = 1./255
Brightness	brightness_range = [0.8, 1.2]
Zooming	zoom_range = 0.2
Width Shift	width_shift_range = 0.1
Height_Shift	height_shift_range = 0.1
Horizontal flip	horizontal_flip = True

2.2.1. Scenario 1 and 2

The first scenario involves progressing to the feature extraction stage and developing a classification model using transfer learning with the VGG-19 architecture. The pre-trained VGG-19 model from ImageNet is loaded without the fully connected layers (include_top=False), effectively functioning as a feature extractor. Most of the convolutional layers are frozen, while the final layers remain trainable to adapt to the specific features of the durian leaf images. The backbone output is condensed using Global Average Pooling to produce a feature vector for each image.

Deep features were extracted sequentially from both the training and testing datasets using the VGG-19 network and subsequently combined into a unified feature matrix. The corresponding labels, originally encoded using a one-hot representation, were converted back into class indices using the argmax operation to ensure compatibility with the Support Vector Machine (SVM) classifier. These deep feature vectors served as the input for the SVM model.

The SVM classifier was configured using a radial basis function (RBF) kernel due to its effectiveness in modeling non-linear decision boundaries in high-dimensional feature spaces. For comparative analysis, a sigmoid kernel was also evaluated under the same parameter settings. The regularization parameter C was set to 10, while the gamma parameter was defined as scale to allow adaptive kernel scaling based on feature variance. Multi-class classification was implemented using a one-vs-one strategy, enabling effective class separation among all durian varieties. This configuration was selected to balance classification accuracy and generalization performance when handling complex feature representations generated by the VGG-19 architecture.

In the second scenario, the process stages follow the same flow as the first scenario, but with the added use of the Synthetic Minority Over-sampling Technique (SMOTE). All images were first standardized to 224×224 pixels and converted into feature vectors to enable oversampling of minority classes until the data distribution became balanced. SMOTE creates synthetic samples through interpolation by selecting one minority sample, finding several of its nearest neighbors, and then generating new points between them based on the differences in feature values. This process increases data variation without straying from the minority class distribution pattern. Therefore, it not only boosts the amount of data but also improves its representation.

Oversampling was performed after data splitting because, if done before splitting, synthetic samples could end up in the test data, compromising the model's evaluation accuracy.

2.2.2. Scenario 3 and 4

In the third scenario, the VGG-19 architecture was fully implemented, using the SoftMax activation function in the final classification layer and pretrained weights from ImageNet. The include_top=True setting was kept to maintain the original structure of the model's top section, including VGG-19's built-in series of fully connected layers. At this point, most convolutional layers were frozen to prevent weight updates, while a few of the final layers remained unfrozen to adapt to the specific visual features of the durian leaf images.

The original VGG-19 output layer, which has 1000 neurons, was not used because it does not match the number of classes in this research dataset. Instead, a new Dense layer was added with a number of neurons specifically adjusted to match the total number of durian classes. The SoftMax activation function was applied to this layer to produce a probability distribution across each class, effectively supporting the requirements of multi-class classification.

In the fourth scenario, the VGG-19 modeling pipeline remained consistent with the third scenario, but SMOTE (Synthetic Minority Over-sampling Technique) was added to address data imbalance. All images were first resized to 224 x 224 pixels and converted into feature vectors, allowing the oversampling process to be performed numerically. SMOTE then created synthetic samples for the minority classes by generating new data points based on the proximity between existing feature vectors.

2.3. Model Evaluation

The model evaluation results are computed to assess the effectiveness of the method used. Evaluation metrics are derived from a Confusion Matrix to obtain precision, recall, accuracy, and F1-Score values, as described in Equations (1), (2), (3), and (4) [4], [18].

$$Precision = \frac{TP}{TP+FP} \times 100\% \tag{1}$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \tag{2}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \tag{3}$$

$$f1 - score = 2 \times \frac{Precision\ x\ Recall}{Precision + Recall} \tag{4}$$

2.4. Hyperparameter Tuning

Effectively setting hyperparameters is crucial for optimizing a deep learning model's performance and ensuring its robust generalization [19]. In this study, hyperparameter tuning was performed across all scenarios. The process, from preprocessing to evaluation, was repeated by testing various hyperparameter values, including batch size, learning rate, and optimizer. This phase aimed to find the configuration that provides the best classification performance on the durian leaf varieties. The hyperparameters used during training are listed in Table 3.

Table 3. List of Hyperparameters Used

Hyperparameter	Hyperparameter Configuration
Batch Size	32, 64, 128
Learning Rate	0.001, 0.0001, 0.00001
Optimizer	Adam, SGD, RMSprop

2.5. Literature Review

Fitriani and Litanianda [2]applied an end-to-end CNN-based classification approach using AlexNet, InceptionV3, and MobileNet architectures for durian leaf classification without separating feature extraction and classification stages or employing data balancing techniques. The dataset consisted of five durian leaf classes, with images preprocessed by normalization, cropping, resizing to 150×150 pixels, and basic augmentation. The results showed that InceptionV3 and AlexNet achieved near-perfect classification accuracy, outperforming MobileNet, which exhibited only minor misclassifications, indicating the effectiveness of deeper CNN architectures for durian leaf classification.

Kurniawan and Ariatmanto [3] employed a transfer learning approach by fine-tuning MobileNetV2 as an end-to-end classification model using basic preprocessing techniques, including image resizing and augmentation, without class balancing. The model retained the original MobileNetV2 classifier, integrating feature extraction and classification within a single architecture. The study reported a maximum accuracy of 90% achieved with a learning rate of 0.0001 and a batch size of 32, with performance evaluation focused on accuracy and loss metrics.

Ramadhan et al. [7] employed an end-to-end CNN-based approach to classify three durian leaf conditions, namely healthy leaves, leaf spot disease, and downy mildew, using a single network for both feature extraction and classification without applying data balancing or alternative classification strategies. The study reported that the ResNet-50 architecture achieved excellent performance for both binary and multiclass classification tasks, reaching a maximum accuracy of 99.6% and a macro F1-score of 96.9%, demonstrating the effectiveness of deep CNN models for durian leaf disease classification despite limited and imbalanced datasets.

Yohannes et al. [4] developed a CNN–SVM–based classification system to distinguish Van Gogh paintings from non–Van Gogh artworks, where convolutional neural networks served as feature extractors and SVM replaced the Softmax classifier. The study compared VGG-19 and ResNet-50 architectures using linear SVM kernels optimized through random and grid search strategies on a dataset of 124 Van Gogh paintings and 207 non–Van Gogh paintings. The results showed that VGG-19 achieved superior performance, reaching an accuracy of 93.28% with grid optimization, outperforming ResNet-50, which attained a maximum accuracy of 90.28%, indicating the stronger feature representation capability of VGG-19 for artwork classification.

Yohannes and Ezar Al Rivan [14] applied a CNN–SVM framework using VGG-19 and ResNet-50 as feature extractors combined with an SVM classifier to classify skin cancer images from the HAM10000 dataset. The study evaluated multiple preprocessing scenarios, including image resizing and patch-based techniques, across five skin cancer classes. The results demonstrated that VGG-19 consistently outperformed ResNet-50 in classification performance, confirming the effectiveness of CNN–SVM models for medical image classification.

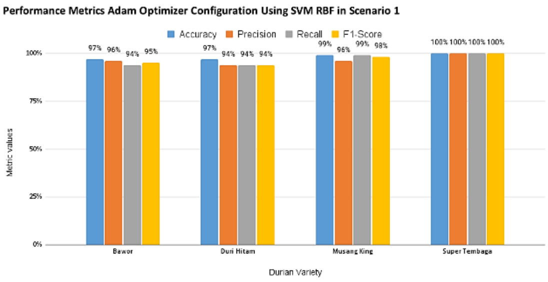
3. RESULTS AND ANALYSIS

3.1. Scenario 1: VGG19–SVM Performance without SMOTE

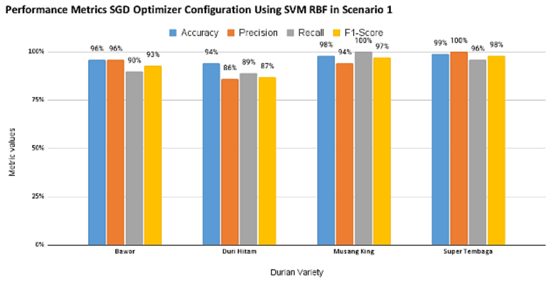
Based on Figures 3(a), 3(b), and 3(c), the RBF kernel consistently achieved the highest performance across all classes, indicating its effectiveness in handling non-linear separability in deep feature representations. Accuracy remained stable between 96% and 100%, with high F1-scores across classes. In Figure 3(a), Super Tembaga achieved 100% on all metrics, while Musang King reached 99% accuracy with an F1-score of 98%, reflecting effective feature separation by VGG19 combined with the RBF decision boundary. Similar trends were observed in Figure 3(c), where both classes maintained F1-scores above 98%, demonstrating stability under different optimizer configurations. The Bawor class showed consistent performance with F1-scores of 95% and 96%, whereas Duri Hitam exhibited the lowest performance, with F1-scores of 94% and 93%, which can be attributed to visual similarity with Bawor and increased feature overlap. Overall, these results indicate that the RBF kernel can construct robust non-linear decision boundaries that enhance the discriminative capability of VGG19 features.

In contrast, Figures 4(a), 4(b), and 4(c) show that the Sigmoid kernel produced unstable results and substantial performance degradation. In Figure 4(a), Super Tembaga achieved only 68% accuracy with an F1-score of 83%, while Duri Hitam experienced further declines, with F1-scores of 74% and 79% in Figures 4(a) and 4(b), respectively. These findings suggest that the Sigmoid kernel is less capable of modeling the complex, high-dimensional feature distributions generated by VGG19, leading to reduced classification reliability. Consequently, the RBF kernel remains the more suitable choice for this classification task.

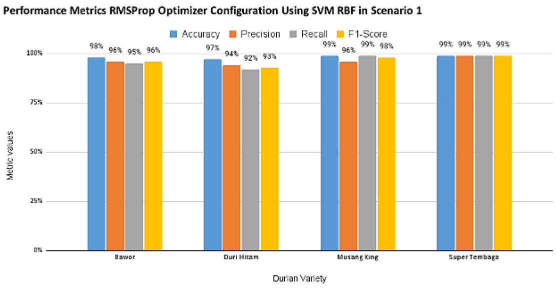
The optimizer influenced feature stability, with adaptive methods such as Adam and RMSProp achieving consistently high F1-scores above 94% across all classes, while SGD showed greater variability, particularly for the Duri Hitam class, which dropped to an F1-score of 87%. However, results in Figures 3 and 4 indicate that kernel selection has a more substantial impact on overall performance than optimizer choice. The RBF kernel maintained stable accuracy between 96% and 100% with F1-scores ranging from 93% to 99%, whereas the Sigmoid kernel suffered marked degradation. Within the controlled experimental setting of Scenario 1, these results demonstrate that the VGG19–SVM model with the RBF kernel, especially when paired with adaptive optimizers, is more robust to feature overlap and class similarity without implying universal generalization beyond the evaluated conditions.



(a)

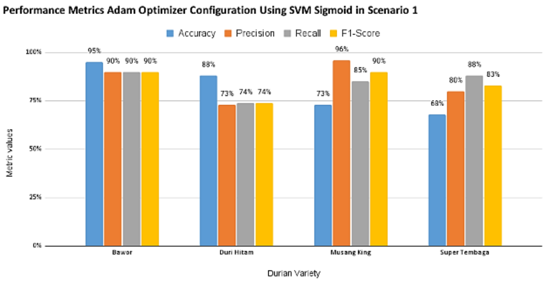


(b)

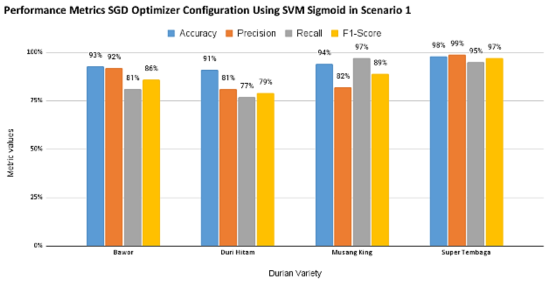


(c)

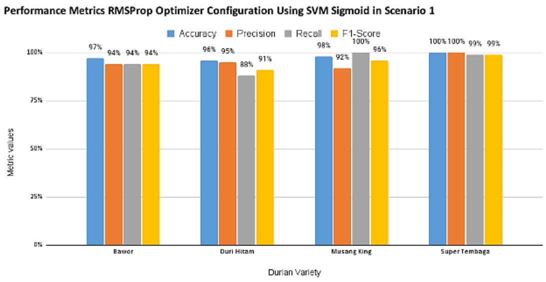
Figure 3. Performance Metrics Optimizer Configuration Using SVM RBF in Scenario 1



(a)



(b)

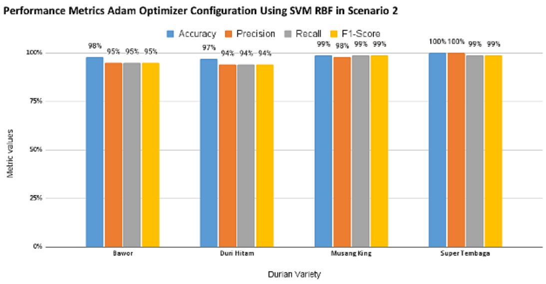


(c)

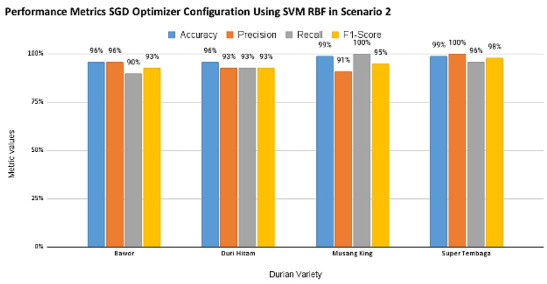
Figure 4. Performance Metrics Optimizer Configuration Using SVM Sigmoid in Scenario 1

3.2. Scenario 2 : VGG19–SVM Performance with SMOTE

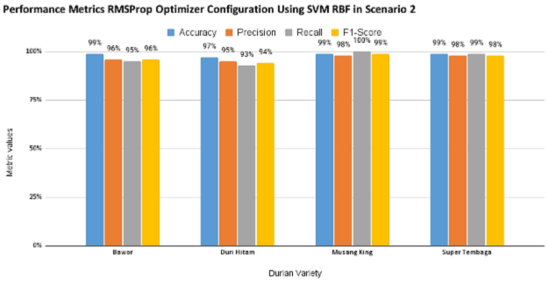
Scenario 2 shows that applying SMOTE improved performance stability by reducing class imbalance effects, as illustrated in Figures 5(a), 5(b), and 5(c). The most notable improvement occurred in the Duri Hitam class, which achieved an F1-score of 94% in Figure 5(a) and remained stable between 93% and 94% across Figures 5(b) and 5(c). This indicates that interpolation-based oversampling helped the classifier learn minority class decision regions more effectively without distorting the feature space. The Bawor class also showed improved consistency, achieving an F1-score of 95%, suggesting that balanced class representation contributed to more reliable decision boundaries. Although the overall accuracy gain was limited due to the already strong VGG19 features, SMOTE enhanced robustness across experimental configurations rather than merely increasing peak performance.



(a)



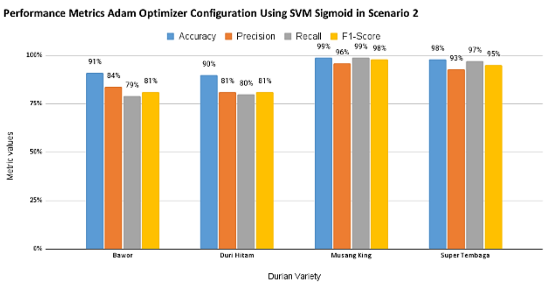
(b)



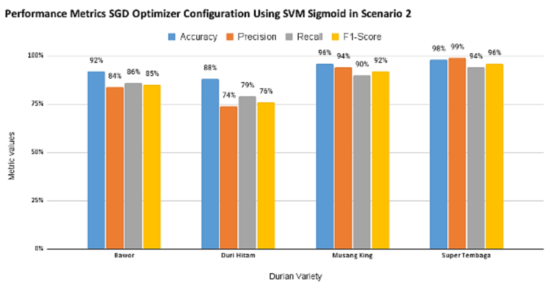
(c)

Figure 5. Performance Metrics Optimizer Configuration Using SVM RBF in Scenario 2

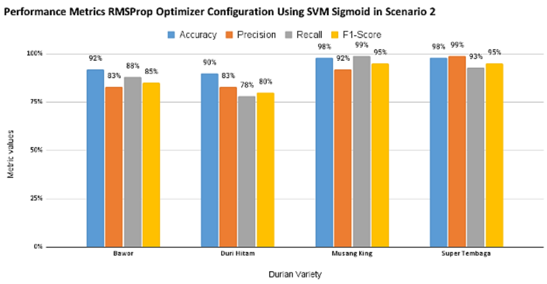
In contrast, Figures 6(a), 6(b), and 6(c) indicate that the Sigmoid kernel remained suboptimal despite data balancing. The Duri Hitam class achieved only 81% and 76% F1-scores in Figures 6(a) and 6(b), respectively, indicating that the kernel's limitations arise from its inability to model complex non-linear decision boundaries in high-dimensional feature spaces. These results confirm that correcting class distribution alone is insufficient when the decision function lacks adequate representational capacity, reinforcing the greater influence of kernel selection over oversampling.



(a)



(b)



(c)

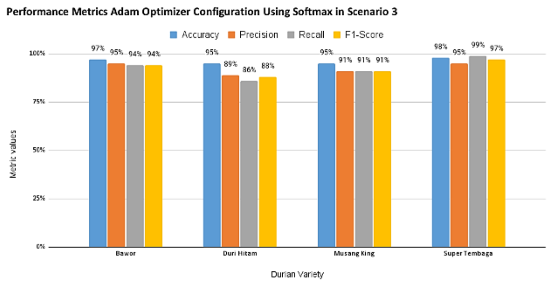
Figure 6. Performance Metrics Optimizer Configuration Using SVM Sigmoid in Scenario 2

Optimizer behavior in Scenario 2 was consistent with Scenario 1, with adaptive methods such as Adam and RMSProp producing stable F1-scores between 94% and 99%, while SGD remained less consistent despite slight improvement, as reflected by the Bawor class reaching an F1-score of 93%. Although SMOTE improved minority class stability under the RBF kernel, particularly for Duri Hitam, performance under the

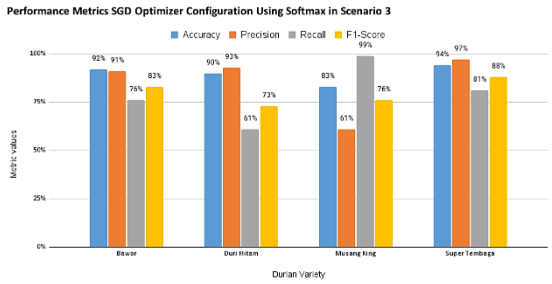
Sigmoid kernel remained below 81%, demonstrating that data balancing enhances robustness only when paired with an appropriate kernel. Overall, these findings confirm that kernel selection has a greater impact on classification robustness than optimizer choice, and that the VGG19–SVM model with the RBF kernel and adaptive optimizers provides stronger resilience to class imbalance within the evaluated experimental conditions.

3.3. Scenario 3 : VGG19–Softmax Performance without SMOTE

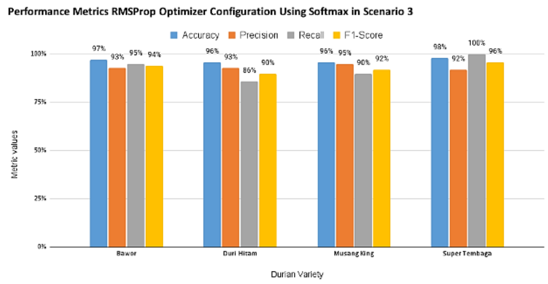
Figures 7(a)–7(c) indicate that using Softmax as the classifier increases sensitivity to feature variation, resulting in less stable performance compared to the SVM-based approach. This behavior reflects Softmax's reliance on linear separability, which becomes problematic when class features overlap. In Figure 7(a), the Adam optimizer maintained relatively stable accuracy and F1-scores across most classes, with Bawor achieving 97% accuracy and a 94% F1-score, and Super Tembaga reaching 98% accuracy; however, the Duri Hitam class showed a notable decline with an F1-score of 88%, indicating difficulty in separating visually similar classes. Figure 7(b) further emphasizes this limitation, where the SGD optimizer produced pronounced precision–recall imbalance, exemplified by Musang King attaining 61% precision and 99% recall, suggesting overprediction, with a similar pattern observed for Duri Hitam. The RMSProp optimizer in Figure 7(c) yielded more consistent results than SGD, with Duri Hitam achieving an F1-score of approximately 90%, indicating that adaptive optimization can partially mitigate classifier sensitivity. Nevertheless, this performance remains inferior to that of the SVM-based model, demonstrating that optimization alone cannot compensate for Softmax limitations under overlapping feature distributions.



(a)



(b)



(c)

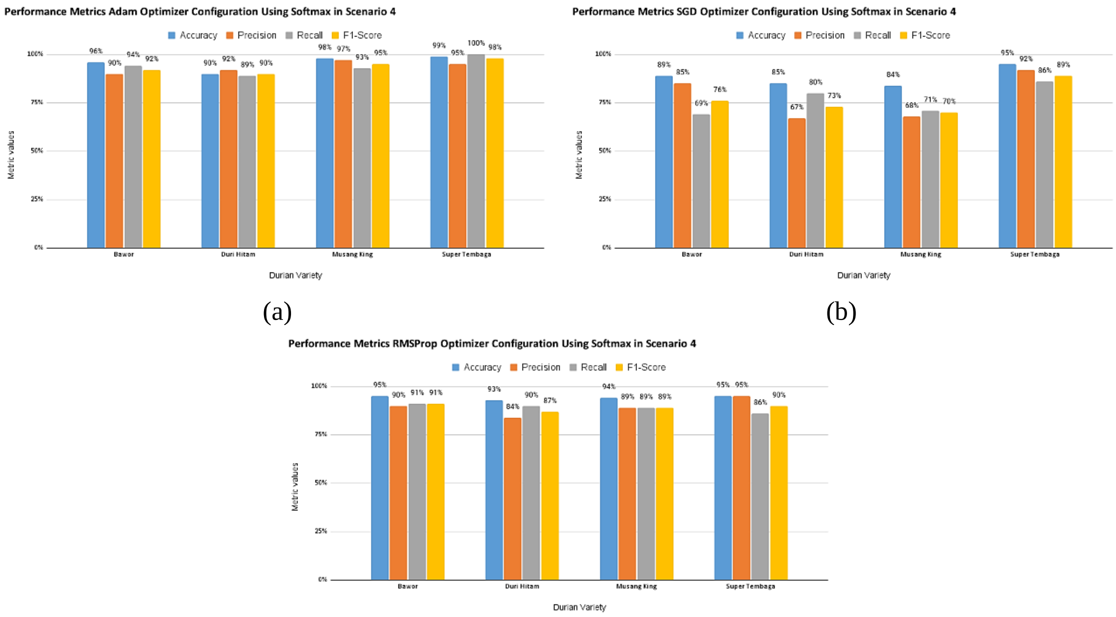
Figure 7. Performance Metrics Optimizer Configuration Using Softmax in Scenario 3

In Scenario 3, Softmax performance was therefore strongly dependent on optimizer choice. Adam and RMSProp produced relatively stable results, whereas SGD caused substantial instability, particularly for visually similar classes such as Musang King and Duri Hitam. Although RMSProp improved consistency, overall performance remained below that of the SVM-based approach. These findings indicate that Softmax lacks intrinsic robustness in complex feature spaces and relies heavily on adaptive optimization, reinforcing the advantage of SVM-based classification for deep feature representations.

3.4. Scenario 4 : VGG19–Softmax Performance with SMOTE

In Scenario 4, applying SMOTE resulted in a moderate improvement in Softmax performance, but it was insufficient to fully stabilize the model, as illustrated in Figures 8(a)–8(c). In Figure 8(a), the Adam optimizer showed noticeable gains in several classes, particularly Musang King, which achieved an F1-score of 95%, surpassing its performance in Scenario 3. This improvement suggests that data balancing helped Softmax better capture minority class patterns when combined with adaptive optimization. However, certain classes, such as Bawor, continued to exhibit precision–recall imbalance, with precision remaining around 90%, indicating that feature overlap still affected class separation. Figure 8(b) further demonstrates that although

SMOTE corrected class distribution, the SGD optimizer continued to perform poorly, with Bawor achieving only a 76% F1-score and Musang King declining to 70%, confirming that SGD is ineffective for Softmax under complex feature conditions. In contrast, Figure 8(c) shows that RMSProp produced the most consistent results after SMOTE, with the Duri Hitam class improving to an F1-score of 87%, indicating that adaptive parameter updates can better accommodate data distribution shifts.



(a) (b)

(c)

Figure 8. Performance Metrics Optimizer Configuration Using Softmax in Scenario 4

Overall, Scenario 4 confirms that Softmax performance remains highly dependent on optimizer choice, even after data balancing. Quantitatively, Adam and RMSProp achieved moderate improvements, whereas SGD consistently resulted in degraded performance across multiple classes. Although SMOTE enhanced minority class representation and improved stability under adaptive optimizers, Softmax remained constrained by overlapping feature distributions and limited decision boundary flexibility. Consequently, its performance did not reach that of the SVM–RBF models observed in Scenarios 1 and 2, indicating that classifier capability plays a more decisive role than data balancing alone in achieving robust classification for deep feature representations.

3.5. Discussions

The results of this study demonstrate that the hybrid VGG-19–SVM framework with an RBF kernel provides more stable and consistent performance than CNN–Softmax approaches for durian leaf variety classification. Across all scenarios, the proposed method achieved accuracy between 97% and 100% with F1-scores ranging from 94% to 100%, particularly maintaining robustness for visually similar classes such as Bawor and Duri Hitam. These findings are more consistent than the end-to-end CNN approaches reported in previous research, where InceptionNetV3 and AlexNet reached nearly 100% accuracy without misclassification, while MobileNet still exhibited classification errors [2], and MobileNetV2 only achieved a maximum accuracy of 90% while being highly dependent on hyperparameter tuning [3]. The findings confirm that separating feature extraction and classification through a CNN–SVM architecture reduces sensitivity to feature overlap and optimization instability compared to end-to-end CNN–Softmax models.

The superiority of the RBF kernel aligns with previous CNN–SVM studies, which reported accuracy improvements up to 93.28% due to its ability to form nonlinear decision boundaries in high-dimensional feature spaces [4], [14]. Nevertheless, the RBF kernel possesses limitations, such as sensitivity to gamma and C parameters, which potentially lead to overfitting on datasets of limited size; meanwhile, the Sigmoid kernel tends to produce less stable performance when handling complex visual features [4]. While the application of SMOTE improved performance consistency for minority classes, the overall increase in global accuracy remained relatively limited because synthesis occurs in a numerical feature space that may not fully represent real-world visual variations [17]. Furthermore, constraints such as limited dataset size, potential lighting biases, and the lack of testing under real-field conditions remain challenges in generalizing the model for real-world applications [6], [7]. Overall, these results indicate that classifier capability, especially non-linear kernel-based modeling, plays a more critical role than optimizer choice or oversampling alone in achieving robust agricultural image classification.

4. CONCLUSION

This study demonstrates that integrating VGG19 as a deep feature extractor with an SVM classifier utilizing the RBF kernel offers a robust and stable solution for classifying four durian leaf varieties, particularly under conditions of high visual similarity and class imbalance. In Scenarios 1 and 2, the VGG19–SVM (RBF) configuration consistently yielded superior performance, achieving accuracy ranging from 97% to 100% and F1-scores between 94% and 100%, both on the original and SMOTE-augmented datasets. Conversely, Scenarios 3 and 4 revealed that Softmax-based classification resulted in instability, particularly for visually similar classes, even when supported by adaptive optimizers and data balancing. These findings validate the research objective by confirming that decoupling feature extraction from classification via a CNN–SVM framework yields greater stability than end-to-end CNN–Softmax models.

The primary contribution of this research lies in the empirical demonstration that the CNN–SVM approach maintains stability across diverse experimental conditions, including varying optimizers and data distributions. The RBF kernel proved effective in modeling non-linear decision boundaries within high-dimensional feature spaces. Furthermore, while SMOTE enhanced minority class representation without compromising stability when paired with a suitable classifier, the benefits of data balancing were negligible when classifier capacity was insufficient. This underscores that kernel selection plays a more decisive role than optimization strategies or oversampling alone.

Notwithstanding these strengths, several limitations warrant mention. The experiments were restricted to a single dataset acquired under controlled conditions, which may constrain generalizability to unstructured agricultural environments. Model performance also remains sensitive to feature distribution consistency and SVM hyperparameter selection, particularly for the RBF kernel. Additionally, SMOTE-based oversampling operates within the numerical feature space and may not fully encapsulate the natural visual variability inherent in field-acquired imagery.

Consequently, future research should prioritize evaluating the proposed framework on larger, heterogeneous datasets that encompass variations in lighting, viewing angles, and leaf conditions. Investigating advanced feature extractors, such as EfficientNet or Vision Transformers, alongside alternative data augmentation and regularization strategies, could further enhance generalization. These advancements are essential for supporting the development of reliable and deployable automated systems for durian variety classification in precision agriculture.

ACKNOWLEDGMENT

Gratitude is expressed to everyone involved in this research, including colleagues who supported the academic process, family members whose moral support facilitated smooth progress, and Universitas Multi Data Palembang for providing the guidance and academic environment that enabled the successful completion of both the research and the manuscript.

REFERENCES

[1] G. Khaksar, S. Kasemcholathan, and S. Sirikantaramas, “Durian (Durio zibethinus L.): Nutritional Composition, Pharmacological Implications, Value-Added Products, and Omics-Based Investigations,” *Horticulturae*, vol. 10, no. 4, Mar. 2024, doi: 10.3390/horticulturae10040342.

[2] L. M. M. Fitriani and Y. Litanianda, “Classification of Durian Leaf Images using CNN (Convolutional Neural Network) Algorithm,” *JIKO (Jurnal Informatika dan Komputer)*, vol. 7, no. 2, pp. 150–157, Aug. 2024, doi: 10.33387/jiko.v7i2.8576.

[3] D. Kurniawan and D. Ariatmanto, “Identifikasi Varietas Bibit Durian Menggunakan Mobilenetv2 Berdasarkan Gambar Daun,” *Jurnal Informatika & Rekayasa Elektronika)*, vol. 7, no. 2, Nov. 2024, doi: 10.36595/jire.v7i2.1236.

[4] S. J. Elroy, F. Nurdiyansyah, G. Priyandoko, and K. Kunci, “Klasifikasi Daun Durian Pada Citra Dalam Menentukan Jenis Menggunakan Convolutional Neutral Network,” *3RD MDP Student Conference (MSC) 2024*, vol. 3, no. 1, Apr. 2024, doi: 10.35957/mdp-sc.v3i1.7199.

[5] M. Mat Daud, A. Abualqumssan, ‘Atyka Nor Rashid Fadilla, M. Hanif Md Saad, W. M. D. W. Zaki, and N. S. M. Satar, “Durian Disease Classification using Vision Transformer for Cutting-Edge Disease Control,” *IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 14, no. 12, pp. 446–452, 2023, doi: 10.14569/IJACSA.2023.0141246.

[6] A. S. Ramadhan, Y. Rahmawati, I. ratna Indra Astutik, and Sumarno, “Identifikasi Penyakit Daun Durian Menggunakan Penerapan Algoritma Residual Network (RESNET-50),” *Infotek: Jurnal Informatika dan Teknologi*, vol. 8, no. 2, pp. 435–446, Jul. 2025, doi: 10.29408/jit.v8i2.30293.

[7] A. B. L. Sabarre, A. N. S. Navidad, D. S. Torbela, and J. J. Adtoon, “Development of Durian Leaf Disease Detection on Android Device,” *International Journal of Electrical and Computer Engineering*, vol. 11, no. 6, pp. 4962–4971, Dec. 2021, doi: 10.11591/ijece.v11i6.pp4962-4971.

[8] V. Ayumi *et al.*, "Transfer Learning for Medicinal Plant Leaves Recognition: A Comparison with and without a Fine-Tuning Strategy," *IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 13, no. 9, p. 2022, 2022, doi: 10.14569/IJACSA.2022.0130916.
[9] T. H. Nguyen, T. N. Nguyen, and B. V. Ngo, "A VGG-19 Model with Transfer Learning and Image Segmentation for Classification of Tomato Leaf Disease," *AgriEngineering*, vol. 4, no. 4, pp. 871–887, Dec. 2022, doi: 10.3390/agriengineering4040056.
[10] M. A. Rajab, F. A. Abdullatif, and T. Sutikno, "Classification of Grapevine Leaves Images using VGG-16 and VGG-19 Deep Learning Nets," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 22, no. 2, pp. 445–453, Apr. 2024, doi: 10.12928/TELKOMNIKA.v22i2.25840.
[11] K. Tyagi, S. Vats, and Vashisht; Vasudha, "Implementing Inception v3, VGG-16 and VGG-19 Architectures of CNN for Medicinal Plant leaves Identification and Disease Detection," *Journal of Electrical Systems*, vol. 20, no. 7, 2024, doi: 10.52783/jes.3989.
[12] M. P. Mathew, S. Elayidom, J. V Raj P, and A. K. M, "Enhanced Bell Pepper and Grape Leaf Disease Classification Using a Depthwise Separable VGG19-Capsule Network," *Electronic Letters on Computer Vision and Image Analysis*, vol. 24, no. 1, pp. 179–194, 2025, doi: 10.5565/rev/elcvia.1964.
[13] Yohannes, D. Udjulawa, and F. Febbiola, "Klasifikasi Lukisan Karya Van Gogh Menggunakan Convolutional Neural Network-Support Vector Machine," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 7, no. 1, Apr. 2021, doi: 10.28932/jutisi.v7i1.3399.
[14] R. Yohannes and M. Ezar Al Rivan, "Klasifikasi Jenis Kanker Kulit Menggunakan CNN-SVM," *Jurnal Algoritme*, vol. 2, no. 2, pp. 133–144, Apr. 2022, doi: 10.35957/algoritme.v2i2.2363.
[15] W. Sholihah and A. Silvia Handayani, "Revolutionizing Healthcare: Comprehensive Evaluation and Optimization of SVM Kernels for Precise General Health Diagnosis," *Scientific Journal of Informatics*, vol. 10, no. 4, pp. 445–454, Nov. 2023, doi: 10.15294/sji.v10i4.46430.
[16] M. H. Yacoub, S. M. Ismail, L. A. Said, A. H. Madian, and A. G. Radwan, "Support Vector Machine Reconfigurable Hardware Implementation on FPGA," *Franklin Open*, vol. 7, Jun. 2024, doi: 10.1016/J.FRAOPE.2024.100115.
[17] A. P. Ningrum, S. Winarno, and V. Praskatama, "Klasifikasi Kualitas Biji Kedelai Menggunakan Transfer Learning Convolutional Neural Network Dan SMOTE," *Journal of Applied Computer Science and Technology*, vol. 5, no. 2, pp. 155–164, Dec. 2024, doi: 10.52158/jacost.v5i2.1002.
[18] A. B. Ballo, A. Coulibaly, and M. Diaby, "Optimization of Mango Leaf Disease Detection Using a Hybrid Machine Learning Approach," *International Journal of Computer Theory and Engineering*, vol. 16, no. 3, pp. 94–103, 2024, doi: 10.7763/IJCTE.2024.V16.1358.
[19] T. M. Arkan, A. Sugiharto, and H. A. Wibawa, "Optimizing VGG16 Architecture with Bayesian Hyperparameter Tuning for Tomato Leaf Disease Classification," *Jurnal Masyarakat Informatika*, vol. 16, no. 2, pp. 162–173, Jun. 2025, doi: 10.14710/jmasif.16.2.73168.

BIBLIOGRAPHY OF AUTHORS

Jennifer Velensia Santoti is an undergraduate student majoring in Informatics at Universitas Multi Data Palembang, Indonesia. The author's academic interests focus on computer vision, image classification, and the application of deep learning techniques in agricultural and visual data analysis. She can be contacted via email at jennifervelensiasantoti_2226250014@mhs.mdp.ac.id.

Siska Devella is a lecturer at Universitas Multi Data Palembang. Her research interests include artificial intelligence, machine learning, and digital image processing. She can be contacted via email at siskadevella@mdp.ac.id