Received April 17th, 2025; accepted June 10th, 2025. Date of publication June 30th, 2025
Digital Object Identifier: https://doi.org/10.25047/jtit.v12i1.444

Color Feature Selection Optimized with Bio-Inspired Algorithms in Classify Purity of Luwak Coffee

SHINTA WIDYANINGTYAS[1], MUHAMMAD ARWANI[2]

[1]Politeknik Negeri Jember, Jember, Indonesia
[2]Universitas Nahdlatul Ulama Indonesia, Jakarta, Indonesia

CORESPONDING AUTHOR: SHINTA WIDYANINGTYAS (email:shinta_widya@polije.ac.id)

ABSTRACT Assessing the purity of Luwak Coffee is a complex challenge due to its unique production and limited availability, as visual inspection is unreliable. This study explores the use of image processing and feature selection to classify Luwak Coffee purity by analyzing 11 color features including RGB, HSV, HSL, and Lab color spaces. Two classification methods k-Nearest Neighbors (k-NN) and Random Forest (RF) were optimized using six Bio-Inspired Algorithms (Differential Evolution, Firefly Algorithm, Flower Pollination Algorithm, Harris Hawk Algorithm, Jaya Algorithm, and Particle Swarm Optimization) to identify the most important features for classifying the purity of Luwak Coffee. The results revealed that feature selection significantly improved accuracy, with the Jaya Algorithm paired with k-NN achieving the highest accuracy (0.918) using only three features (R_Mean, B_Mean, and H_Mean). For RF, the Flower Pollination Algorithm yielded the best performance (0.899) with three features. The study demonstrates a classification method coupled with Bio-Inspired Algorithms for classifying Luwak Coffee purity providing high accuracy as a non-destructive method. These findings contribute to the development of reliable tools for classifying purity of Luwak Coffee

KEYWORDS: Bio-Inspired Algorithm, Classification, Feature Selection, Image Processing, Luwak Coffee

I. INTRODUCTION

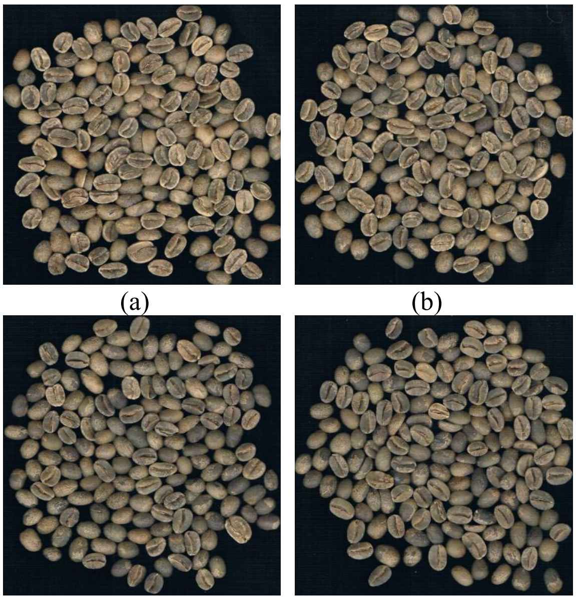
Purity is recognized as a key quality parameter of Luwak Coffee due to its unique production process and limited availability. However, the high market value and scarcity of authentic Luwak Coffee have led to widespread adulteration, where regular coffee beans are mixed with regular coffee. Currently, coffee farmers and traders rely on traditional methods, such as visual inspection and manual sorting, to assess the quality of Luwak Coffee. However, these methods are still highly subjective and inaccurate, as the visual traits that distinguish adulterated coffee beans cannot be seen by humans. In addition, destructive testing and chemical procedures, are laborious, expensive, and inefficient for local traders, small-scale farmers, and even larger manufacturers. Therefore, determining the accuracy of Luwak Coffee poses a significant challenge in designing a quick and inexpensive tool to classify its purity.

One approach that can overcome this challenge is to use image processing that can distinguish between Luwak Coffee and regular coffee. Image processing techniques are capable of extracting various color features according to the color space ranging from Red, Green, Blue (RGB); Hue, Saturation, Value (HSV); Hue, Saturation, Lightness (HSL). These color features have been successfully used in classifying the ripeness of Citrus limon L. [1], predicting total dissolved solids and pH of strawberries [2], and classifying fruit ripeness [3]. However, not all of these color features are relevant in determining the purity of Luwak Coffee [4]. Some features may have a strong correlation with purity, while other features are irrelevant and cause noise in the classification process.

To solve this problem, feature selection plays a crucial role in identifying the most relevant color features that have a strong relationship with the purity of Luwak Coffee. One of the feature selection methods that can be utilized is the wrapper method. The wrapper method comprehensively searches for the best combination of feature subsets. This is accomplished by using an optimized classification method with a learning algorithm to find the best features [5]. Common classification methods that can be employed include k-Nearest Neighbors (k-NN) [6], [7] and Random Forest (RF) [8]. Furthermore, the learning algorithm for optimization of wrapper methods can be conducted using bio-inspired algorithms, such as Different Evolution (DE), Firefly Algorithm (FA), Flower Pollination Algorithm (FPA), Harris Hawk Optimization (HHO), Jaya Algorithm (JA), and Particle Swarm Optimization (PSO). These algorithms are known for their efficiency in solving complex optimization problems and can help identify the optimal subset of features for accurate classification. Previous study [5] successfully developed a machine vision system using color and texture features for classifying Luwak Coffee purity using wrapper feature selection and classic machine learning algorithm including k-NN, RF and SVM. While, in this study focuses primarily on color features (RGB, HSV, HSL, Lab) to assess whether texture analysis is truly necessary for accurate purity classification. The system would become computationally lighter and reliable for real time applications. This is particularly important for build rapid and low cost tools rather than complex texture feature models. In addition, this study also explores other bio-inspired algorithm. The k-NN and RF classifiers were used in this study due to their reliability in classification problems. Currently, there is no study that used this sixth learning algorithm for classify purity of Luwak Coffee. In addition, there is previous research has explored the use of multiple learning algorithm for color feature selection. Through the use of bio-inspired learning algorithms and classification techniques, this study seeks to determine the most effective color features for classifying the purity of Luwak Coffee in order to increase the accuracy of the model. The findings of this study are expected to contribute to the development of reliable tools for detecting the purity of Luwak Coffee, supporting its authenticity.

II. METHOD

This study used Longan Arabica Luwak Coffee and regular Arabica coffee, both varieties of green beans that were purchased from PT. Perkebunan Nusantara XII in Banyuwangi, Indonesia. Wet processing was used to processed the standard green beans. The four purity levels depicted in Figure 1 were extremely low (0–25%), low (25–50%), medium (50–75%), and high (75–100%). A total of 160 green beans were used for each mixture. The analysis tools included a custom-built Visual Basic 6.0 program for color and texture analysis, and Python 3.11.5 with Visual Studio Code for feature selection. A computer with an Intel Core i7-1165G7 2.80 GHz processor and 8 GB of RAM was used for the experiments [5].



(a) (b) (c) (d)

FIGURE 1. Luwak Coffee Samples at several level of Purity (a) very low (0-25%), (b) low (25-50%), (c) medium (50-75%), and (d) high (75-100%)

The research focused on classifying the purity of Luwak Coffee through a five-step process: image acquisition, feature extraction, feature selection, model design, and evaluation. The input data consisted of co-occurrence matrices generated from feature extraction, which included color features, texture features, and a combination of both.

1. Image Acquisition

   Images of the green beans were captured using a Nikon Coolpix A10 digital camera (16 megapixels). The camera was placed 20 cm above a tray with an area of 256 cm², and consistent lighting to reduce variations in color analysis. Prior to analysis, the images were cropped to focus solely on the green beans, resulting in a dataset of 528 images.

2. Feature Extraction

   This step involved identifying unique characteristics from the image data to differentiate between samples. In the field of computer vision, an image sample is converted into a grid of pixels, where each pixel represents a point in three-dimensional color space defined by intensity values of the three fundamental colors: red, green, and blue (RGB color model). This RGB representation serves as the primary color space in digital imaging, where each color channel typically ranges from 0 to 255, creating a comprehensive color gamut suitable for digital analysis. The hue-saturation-value (HSV) and hue-saturation-lightness (HSL) color spaces were derived from the RGB color space using Equations (1) to (5). In this study, the output of feature extraction consisted of eleven color features: R_Mean, G_Mean, B_Mean, H_Mean, S_HSL_Mean, S_HSV_Mean, L_Mean, V_Mean, Lab_L_Mean, Lab_a_Mean, and Lab_b_Mean, which were kept in files in the CSV format. These features facilitated the differentiation of green beans based on their visual properties.

$$h = \begin{cases} 0, & if\ max = min \\ 60° x \dfrac{G-B}{max-min} + 0°, & if\ max = R \\ 60° x \dfrac{B-R}{max-min} + 120°, & if\ max = G \\ 60° x \dfrac{R-G}{max-min} + 240°, & if\ max = B \end{cases} \quad (1)$$

$$S_{HSV} = \begin{cases} 0, & if\ max = 0 \\ \dfrac{max-min}{max} = 1 - \dfrac{min}{max}, & if\ \max \neq 0 \end{cases} \quad (2)$$

$$V_{HSV} = max \quad (3)$$

$$S_{HSL} = \begin{cases} 0, & if\ max = min \\ \dfrac{max-min}{max+min} = \dfrac{max-min}{2l}, & if\ l \leq \dfrac{1}{2} \\ \dfrac{max-min}{2-(max-min)} = \dfrac{max-min}{2-2l}, & if\ l > \dfrac{1}{2} \end{cases} \quad (4)$$

$$V_{HSL} = \frac{1}{2}(max + \min) \quad (5)$$

3. Sensitivity Analysis
This study conducted sensitivity analysis to evaluate the optimal classification accuracy for determining Luwak Coffee purity. The classification methods implemented were k-NN and RF optimized using bio-inspired algorithms (DE, FA, FPA, HHO, JA, PSO). Sensitivity analysis systematically evaluates the performance variation by modifying the hyperparameters of the learning algorithm and classification method. The methodological framework for the classification model and optimization algorithm is detailed as follows:

A. Model Design

a. K-Nearest Neighbor (k-NN)
The k-Nearest Neighbors (k-NN) algorithm represents a fundamental supervised learning approach that is widely adopted in classification tasks. The underlying principle of the k-NN methodology is to predict the class label of an unclassified data point through majority voting, where the prediction is determined by the most dominant class among the k nearest training examples in the feature space. The parameter k indicates the number of nearest neighbors evaluated during the classification process. This study implements k values of 1, 3, 5, 7, and 9, based on [9].

b. Random Forest (RF)
Random Forest is an ensemble-based supervised learning algorithm distinguished by its computational robustness and predictive accuracy. An important hyperparameter in Random Forest optimization is the number of decision trees (ntree) in the ensemble. To determine the optimal configuration to maximize classification accuracy, this study systematically evaluated ntree values of 100, 200, 500, and 1000 [5].

B. Learning Algorithm

a. Differential Evolution (DE)
The DE introduced by Storn and Price (1997), represents a powerful metaheuristic optimization algorithm [10]. This methodology has emerged as a predominant approach for addressing global optimization challenges [11]. DE operates within the framework of evolutionary algorithms, employing biologically-inspired operators including mutation, crossover, and selection [10], [12].

b. Firefly Algorithm (FA)
The FA represents a nature-inspired metaheuristic optimization methodology that emulates the bioluminescent communication patterns of fireflies [13]. There are about two thousand firefly species, and most fireflies produce short and rhythmic flashes that are used as a communication means between the fireflies. Consequently, the primary objective of a firefly's flash is to attract other fireflies [14], [15].

c. Flower Pollination Algorithm (FPA)
The FPA introduced by Yang in 2012, represents a bio-inspired metaheuristic optimization methodology that emulates the biological pollination mechanisms observed in flowering plants [16]. The optimization process begins with a population of randomized solutions, utilizing an adaptive switching mechanism to alternate between local and global pollination operators. By systematically integrating the exploitation phase (local pollination) with the exploration phase (global pollination), the algorithm iteratively refines candidate solutions [17]. The optimization procedure continues until predetermined convergence criteria are satisfied, facilitating effective solution optimization across various problem domains.

d. Harris Hawk Optimization (HHO)
The HHO is a modern,population-based metaheuristic algorithm inspired by the natural hunting behaviors of Harris' hawks. Developed by Heidari et al. in 2019, this gradient-free optimization method mimics various predation strategies observed in nature, including tracking, flanking, and surprise attacks on prey. Similar to other nature-inspired algorithms, HHO operates through two fundamental phases: exploration and exploitation. The algorithm

mimics the cooperative hunting techniques of these birds of prey, where multiple hawks work together to capture their target. This innovative approach provides an effective optimization framework without the need for gradient information [18].

e. Jaya Algorithm (JA)

The JA is a contemporary metaheuristic method developed by Rao [19] to address both constrained and unconstrained optimization problems. As a population-based approach, it takes inspiration from the natural principle of "survival of the fittest", which guides solutions in the population to converge towards the best performing option while avoiding the worst. Essentially, the algorithm's search mechanism seeks to progress towards the optimal solution by approaching success and avoid failure by staying away from poorly performing alternatives. Among its main benefits, JA is very easy to implement and requires no special parameters.

e. Particle Swarm Optimization (PSO)

The PSO is a highly effective optimization method known for several key advantages. First, its implementation is very simple and requires minimal coding effort, so it can be used for a wide variety of applications. Additionally, PSO operates with only three adjustable parameters—inertia weight, cognitive coefficient, and social coefficient—offering an optimal balance between simplicity and controllability [20].

4. Evaluation

The primary evaluation metric employed in this study is accuracy, calculated based on the confusion matrix. Accuracy, recognized as the most straightforward performance measure, is defined as the proportion of correctly classified instances to the total number of instances evaluated, as computed using Equation 6.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \qquad (6)$$

where :

1. True Positive (TP) – prediction and actual both are yes
2. True Negatives (TN) – prediction is no and actual is yes
3. False Positives (FP) – prediction is yes and actual is no
4. False Negatives (FN) – prediction is no and actual is no.

The optimal model is determined based on two key factors: accuracy and the number of selected features. A higher accuracy signifies improved model performance, as it represents a larger ratio of correctly classified objects. Furthermore, the number of selected features is also considered important, as fewer features facilitate the development of automated systems or tools for detecting Luwak Coffee purity.

III. Result and Discussion

The image processing feature extraction process resulted in a total of 11 distinct color features, namely R_Mean, G_Mean, B_Mean, H_Mean, S_HSL_Mean, S_HSV_Mean, L_Mean, V_Mean, Lab_L_Mean, Lab_a_Mean, and Lab_b_Mean. The study assessed feature selection using two modeling approaches: k-NN and RF enhanced by six optimization algorithms: DE, FA, FPA, HHO, JA, and PSO. The results of this feature selection process, employing the k-NN method in conjunction with the six optimization algorithms, are presented in **Table 1**.

TABLE 1. Color Feature Selection Using k-NN

Learning Algorithm	k-Value	Number of Selected Features	Accuracy
All Feature	1	-	0.764
	3	-	0.811
	5	-	0.772
	7	-	0.641
	9	-	0.741
Differential Evolution	1	3	0.890
	3	**4**	**0.918**
	5	3	0.843
	7	3	0.899
	9	5	0.868
Firefly Algorithm	**1**	**5**	**0.899**
	3	5	0.912
	5	3	0.881
	7	7	0.893
	9	5	0.912
Flower Pollination Algorithm	**1**	**5**	**0.918**
	3	7	0.912
	5	5	0.893
	7	6	0.906
	9	4	0.843
Harris Hawk Optimization	1	3	0.893
	3	4	0.912
	5	6	0.868
	7	4	0.874
	9	3	0.849
Jaya Algorithm	1	4	0.893
	3	3	0.887
	5	5	0.899
	7	5	0.906
	9	**3**	**0.918**
Particle Swarm Optimization	1	4	0.868
	3	5	0.912
	5	5	0.906
	7	4	0.887
	9	4	0.868

Table 1 presents the performance of various learning presents the performance of various learning algorithms in feature selection, highlighting the relationship between the number of selected features and classification accuracy. Feature selection enhances the accuracy of the k-NN classifier compared to using all features, with a peak accuracy of 0.918 achieved by DE (k-value =

3, 4 selected features), the FPA (k-value = 1, 5 selected features), and the JA (k-value = 9, 3 selected features). This strategy aims to minimize the features, so the optimal model selection is based on the number of features and accuracy [5], [21]. For example, the JA achieves the best performance with 3 features selected, while DE and FPA require 4 and 5 features respectively to achieve the same accuracy. This efficiency reduces computational complexity and reduces overfitting, improving robustness and interpretability [22], [23]. The optimal k-NN model for classifying the purity of Luwak Coffee using JA with R_Mean, B_Mean, and H_Mean as the selected features, due to its simplicity and ability to optimize the feature subsets without extensive parameter tuning, achieving optimal results with fewer function evaluations and showing fast convergence [19]. The research result [24] showed that the proposed feature selection approach, optimized with JA, effectively eliminated irrelevant and redundant features using Naive Bayes, k-NN, Linear Discriminant Analysis, and Random Tree classifiers. In this study, the k-NN method, coupled with the JA, achieves an optimal balance between accuracy and simplicity, making it ideal for classify purity of Luwak Coffee.

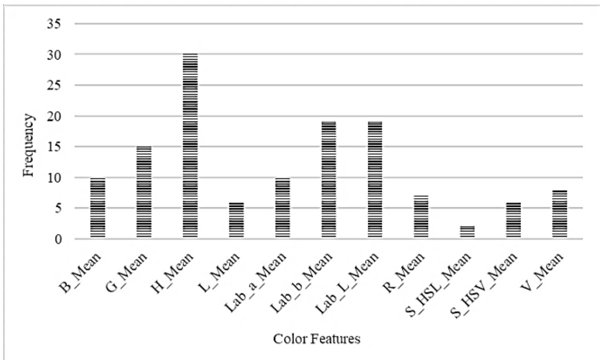


FIGURE 2. Distribution of Color Feature Frequency in k-NN Feature Selection

The feature selection process using the k-NN method across six different learning algorithms identified various combinations of color features to classify the purity of Luwak Coffee. **Figure 2** illustrates the frequency with which each color feature was selected across all algorithms using the k-NN method. H_Mean consistently selected across multiple algorithms, indicating its high importance in distinguishing the purity of Luwak Coffee. Hue refers to the dominant wavelength of color stimulus and is defined by angles [25]. Its frequent selection suggests that the hue component is a critical factor in the classification. The lightness component in the L*a*b (Lab_L_Mean) color space also appeared frequently, highlighting its relevance lightness variations in the coffee samples. The b component in the L*a*b (Lab_b_Mean) which represents the yellow-blue axis [26], was another frequently selected feature, indicating the ability of color space differentiating coffee samples at various purity levels. **Figure 2** also demonstrates that L_Mean, S_HSL_Mean, and S_HSV_Mean were selected less frequently compared to other features. This is attributed to their limited capacity to capture the specific color characteristics most effective in distinguishing the purity of Luwak Coffee.

TABLE 2. Color Feature Selection Using RF

Learning Algorithm	ntree	Number of Selected Features	Accuracy
All Feature	100	-	0.511
	200	-	0.684
	500	-	0.662
	1000	-	0.510
Differential Evolution	100	3	0.874
	200	3	0.881
	500	3	0.881
	1000	3	0.893
Firefly Algorithm	100	6	0.874
	200	5	0.862
	500	5	0.854
	1000	5	0.868
Flower Pollination Algorithm	100	4	0.874
	200	3	0.893
	500	4	0.887
	1000	**3**	**0.899**
Harris Hawk Optimization	100	5	0.843
	200	4	0.874
	500	3	0.881
	1000	3	0.868
Jaya Algorithm	100	2	0.562
	200	3	0.571
	500	2	0.552
	1000	2	0.642
Particle Swarm Optimization	100	4	0.875
	200	4	0.882
	500	3	0.861
	1000	3	0.763

Table 2 presents the performance of various learning algorithms in selecting color features using RF classification. The results indicate that feature selection enhances the accuracy of the RF classifier compared to utilizing all features. The FPA demonstrates strong performance, particularly with 1000 trees, achieving an accuracy of 0.899 by selecting B_Mean, H_Mean, and V_Mean. Furthermore, the adaptability of the FPA with varying ntree underscores its robustness in classifying the purity of Luwak Coffee. According to [27] the major advantages of the FPA are its flexibility and efficiency in searching for an optimal feature subset. When integrated with RF, the FPA provides an optimal balance between accuracy and model simplicity. Its ability to achieve higher performance with fewer features makes the FPA a preferred method for classifying the purity of Luwak Coffee.

Figure 3 shows that G_Mean, H_Mean, and R_Mean have the highest frequency of selection, indicating their strong relevance to the sample. The analysis of feature selection across various models highlights G_Mean, R_Mean, and H_Mean as the most frequently chosen features. G_Mean and R_Mean, representing the mean values of the green and red colors, respectively, are critical as they capture fundamental color information directly related to the visual characteristics of the sample.

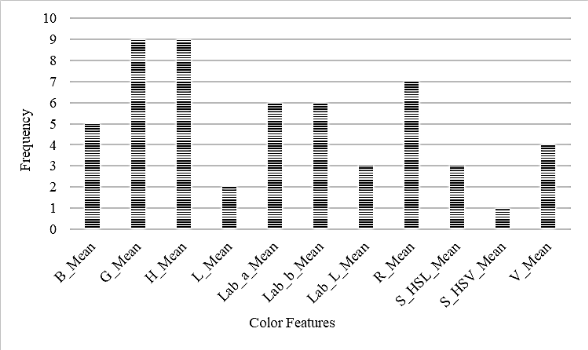


FIGURE 3. Distribution of Color Feature Frequency in Random Forest Feature Selection

H_Mean represents the mean value of the hue component, underscoring its importance. **Figure 3** also illustrates that Lab_L_Mean, S_HSL_Mean, L_Mean, and S_HSV_Mean were selected less frequently compared to other features. This is due to their limited ability to capture the specific color characteristics most effective in distinguishing the purity of Luwak Coffee.

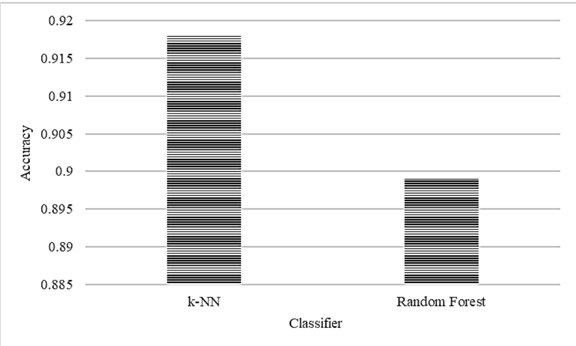


FIGURE 4. Accuracy of the Optimal Model for Each Classifier

Figure 4 illustrates the k-NN classifier, when combined with the JA, achieved the highest accuracy of 0.918 using only three selected features (R_Mean, B_Mean, and H_Mean), demonstrating superior feature optimization capabilities. The main advantages of k-NN include low time complexity and the ability to classify highly non-linear data at high speeds compared to other machine learning algorithms [9], [28]. Most notably, the k-NN optimized with the JA achieved higher accuracy while utilizing fewer features, indicating a more efficient use of discriminative color features. Previous studies [5] used several bio-inspired algorithms combined with k-NN, RF and SVM classifiers. The result was RF with Grey Wolf Optimizer (GWO) being the preferred method with the combination of colour and texture features (accuracy 0.981). Compared to the results of this study, it turns out that colour features are not giving good accuracy. Therefore, there is a need for further studies on the best combination of colour and texture features in classifying Luwak Coffee purity using the 6 learning algorithms that were used in this study.

IV. CONCLUSION

This study shows that effective feature selection significantly improves the classification accuracy of k-NN and RF models in assessing the purity of Luwak Coffee. Among the six optimization algorithms evaluated, JA combined with k-NN achieving the highest accuracy of 0.918 with three optimally selected features (R_Mean, B_Mean, and H_Mean). This combination outperformed other learning algorithm and classifier method. For the RF classifier, the FPA provided the best results, achieving an accuracy of 0.899 with three selected features. These findings highlight that k-NN optimized with the JA is compatible for classifying purity of Luwak Coffee. This study provides a non-destructive method for coffee purity assessment, which contributes to Kopi Luwak authentication.

REFERENCE

[1] A. C. P. Utomo, C. A. Sari, E. H. Rachmawanto, H. P. Hadi, S. A. Mostafa, and W. S. T. Al-Dayyeni, "Citrus Limon L. Ripeness Identification Using Feature Extraction of Mean RGB, HSV, LBP and K-Nearest Neighbor," in *2023 International Seminar on Application for Technology of Information and Communication (iSemantic)*, 2023, pp. 423–428.

[2] J. K. Basak, B. G. K. Madhavi, B. Paudel, N. E. Kim, and H. T. Kim, "Prediction of total soluble solids and pH of strawberry fruits using RGB, HSV and HSL colour spaces and machine learning models," *Foods*, vol. 11, no. 14, p. 2086, 2022.

[3] J. Pardede, M. G. Husada, A. N. Hermana, and S. A. Rumapea, "Fruit ripeness based on RGB, HSV, HSL, L* a* b* color feature using SVM," in *2019 international conference of computer science and information technology (ICoSNIKOM)*, 2019, pp. 1–5.

[4] Y. Hendrawan, S. Widyaningtyas, and S. Sucipto, "Computer vision for purity, phenol, and pH detection of Luwak Coffee green bean," *TELKOMNIKA (Telecommunication Comput. Electron. Control.*, vol. 17, no. 6, pp. 3073–3085, 2019.

[5] S. Widyaningtyas, S. Arwani, Muhammad, Sucipto, and Y. Hendrawan, "Improving the accuracy of green bean palm civet coffee purity classification using wrapper feature selection," *Coffee Sci. 1984-3909*, vol. 20, pp.

e202277–e202277, 2025.

[6] M. Effendi, M. Jannah, and U. Effendi, “Corn quality identification using image processing with k-nearest neighbor classifier based on color and texture features,” in *IOP Conference Series: Earth and Environmental Science*, 2019, vol. 230, no. 1, p. 12066.

[7] I. G. I. Sudipa, R. A. Azdy, I. Arfiani, and N. M. Setiohardjo, “Leveraging k-nearest neighbors for enhanced fruit classification and quality assessment,” *Indones. J. Data Sci.*, vol. 5, no. 1, pp. 30–36, 2024.

[8] J. Basavaiah and A. Arlene Anthony, “Tomato leaf disease classification using multiple feature extraction techniques,” *Wirel. Pers. Commun.*, vol. 115, no. 1, pp. 633–651, 2020.

[9] S. Uddin, I. Haque, H. Lu, M. A. Moni, and E. Gide, “Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction,” *Sci. Rep.*, vol. 12, no. 1, p. 6256, 2022.

[10] R. Storn and K. Price, “Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces,” *J. Glob. Optim.*, vol. 11, pp. 341–359, 1997.

[11] S. Hassan, A. M. Hemeida, S. Alkhalaf, A.-A. Mohamed, and T. Senjyu, “Multi-variant differential evolution algorithm for feature selection,” *Sci. Rep.*, vol. 10, no. 1, p. 17261, 2020.

[12] E. Hancer, B. Xue, and M. Zhang, “Differential evolution for filter feature selection based on information theory and feature ranking,” *Knowledge-Based Syst.*, vol. 140, pp. 103–119, 2018.

[13] X.-S. Yang, “Firefly algorithm, stochastic test functions and design optimisation,” *Int. J. bio-inspired Comput.*, vol. 2, no. 2, pp. 78–84, 2010.

[14] M. K. Sayadi, R. Ramezanian, and N. Ghaffari-Nasab, “A discrete firefly meta-heuristic with local search for makespan minimization in permutation flow shop scheduling problems,” *Int. J. Ind. Eng. Comput.*, vol. 1, no. 1, pp. 1–10, 2010.

[15] N. Chai-Ead, P. Aungkulanon, and P. Luangpaiboon, “Bees and firefly algorithms for noisy non-linear optimisation problems,” in *Proceedings of the international multi conference of engineering and computer scientists*, 2011, vol. 2.

[16] Z. A. A. Alyasseri, A. T. Khader, M. A. Al-Betar, M. A. Awadallah, and X.-S. Yang, “Variants of the flower pollination algorithm: a review,” *Nature-inspired algorithms Appl. Optim.*, pp. 91–118, 2018.

[17] H. M. Zawbaa and E. Emary, “Applications of flower pollination algorithm in feature selection and knapsack problems,” in *Nature-inspired algorithms and applied optimization*, Springer, 2017, pp. 217–243.

[18] R. Sihwail, K. Omar, K. A. Z. Ariffin, and M. Tubishat, “Improved harris hawks optimization using elite opposition-based learning and novel search mechanism for feature selection,” *Ieee Access*, vol. 8, pp. 121127–121145, 2020.

[19] R. Rao, “Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems,” *Int. J. Ind. Eng. Comput.*, vol. 7, no. 1, pp. 19–34, 2016.

[20] T. M. Shami, A. A. El-Saleh, M. Alswaitti, Q. Al-Tashi, M. A. Summakieh, and S. Mirjalili, “Particle swarm optimization: A comprehensive survey,” *Ieee Access*, vol. 10, pp. 10031–10061, 2022.

[21] S. Widyaningtyas, M. Arwani, S. Sucipto, and Y. Hendrawan, “Filter Feature Selection for Detecting Mixture, Total Phenol, and pH of Civet Coffee,” *Int. J. Artif. Intell. Robot.*, vol. 6, no. 2, pp. 75–82, 2024.

[22] X. Cheng, “A Comprehensive Study of Feature Selection Techniques in Machine Learning Models,” *Insights Comput. Signals Syst.*, vol. 1, no. 1, pp. 10–70088, 2024.

[23] T. Dokeroglu, A. Deniz, and H. E. Kiziloz, “A comprehensive survey on recent metaheuristics for feature selection,” *Neurocomputing*, vol. 494, pp. 269–296, 2022.

[24] H. Das, B. Naik, and H. S. Behera, “A Jaya algorithm based wrapper method for optimal feature selection in supervised classification,” *J. King Saud Univ. Inf. Sci.*, vol. 34, no. 6, pp. 3851–3863, 2022.

[25] L. Tao and G. Xu, “Color in machine vision and its application,” *Chinese Sci. Bull.*, vol. 46, pp. 1411–1421, 2001.

[26] V. S. Rathore, M. S. Kumar, and A. Verma, “Colour based image segmentation using L* a* b* colour space based on genetic algorithm,” *Int. J. Emerg. Technol. Adv. Eng.*, vol. 2, no. 6, pp. 156–162, 2012.

[27] C. Yan *et al.*, “A novel feature selection method based on MRMR and enhanced flower pollination algorithm for high dimensional biomedical data,” *Curr. Bioinform.*, vol. 17, no. 2, pp. 133–149, 2022.

[28] M. R. Romadhon and F. Kurniawan, “A comparison of naive Bayes methods, logistic regression and KNN for predicting healing of Covid-19 patients in Indonesia,” in *2021 3rd east Indonesia conference on computer and information technology (eiconcit)*, 2021, pp. 41–44.

SHINTA WIDYANINGTYAS received her Bachelor's Degree in Agroindustrial Technology Universitas Brawijaya in 2016 and Master's Degree in Agroindustrial Technology in 2018. Her research focus is non destructive testing for assesing quality of agricultural product using machine vision and artificial intelligence.

MUHAMMAD ARWANI, received his Bachelor's Degree in Agroindustrial Technology from Universitas Brawijaya in 2016 and his Master's Degree in the same field in 2018. His research focuses on the halal food industry and non-destructive testing for agricultural products.