ISSN: 2093-5374 / E-ISSN: 2093-5382

Enhancing Cloves Quality Classification Based on Ensemble Features Selection

Firman Tempola[1], Aina Musdholifah[2], and Wahyono[3*]

[1]Doctoral Student, Doctoral Program Department of Computer Science and Electronics, Faculty of Mathematics and Natural Science, Universitas Gadjah Mada, Yogyakarta, Indonesia; Department of Informatics, Khairun University, North Maluku, Indonesia. firman.tempola@mail.ugm.ac.id, https://orcid.org/0000-0002-3471-391X

[2]Assistant Professor, Department of Computer Science and Electronics, Faculty of Mathematics and Natural Science, Universitas Gadjah Mada, Yogyakarta, Indonesia. aina_m@ugm.ac.id, https://orcid.org/0000-0002-9076-6389

[3*]Associate Professor, Department of Computer Science and Electronics, Faculty of Mathematics and Natural Science, Universitas Gadjah Mada, Yogyakarta, Indonesia. wahyo@ugm.ac.id, https://orcid.org/0000-0002-2639-8411

Received: April 06, 2025; Revised: May 25, 2025; Accepted: June 10, 2025; Published: June 30, 2025

Abstract

Based on the Indonesian National Standard (SNI) 01-3392-1994, clove quality in Indonesia is categorized into three distinct classes (Quality 1, 2, and 3). However, quality determination presents significant challenges due to the similar morphological characteristics across classes and limited available data, necessitating manual feature extraction. This manual process often yields irrelevant features, highlighting the need for robust feature selection methods. This study demonstrates that ensemble feature selection significantly enhances the performance of conventional machine learning models (K-NN and Naïve Bayes) in SNI-compliant clove quality classification. The proposed approach employs five distinct feature selection methods to assign importance scores to features, with the final selected features representing those consistently identified across all methods. Experimental results indicate notable improvements, including a 2.66% accuracy increase in K-NN classification and reduced computational time across all tested models. These findings substantiate the effectiveness of ensemble feature selection for optimizing conventional machine learning applications in agricultural quality assessment.

Keywords: Ensemble Feature Selection, Machine Learning Conventional, Quality of Cloves.

1 Introduction

Processed cloves are available in various products, including food, spices, medicines, and cosmetics (Chaichi et al., 2021; Sharma et al., 2022). The clove qualities in Indonesia that influence product value are regulated by the Indonesian National Standard (SNI) number 01-3392-1994. The SNI is divided into three quality levels: 1, 2, and 3. Several tests should be carried out in the laboratory to determine the clove's qualities, such as testing for water content, shape, colors, smell, and influence of foreign

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), volume: 16, number: 2 (June), pp. 927-943. DOI: *10.58346/JOWUA.2025.I2.057*

*Corresponding author: Associate Professor, Department of Computer Science and Electronics, Faculty of Mathematics and Natural Science, Universitas Gadjah Mada, Yogyakarta, Indonesia.

materials. However, the conventional method requires a long time and depends on the laboratory assistants' capabilities. Meanwhile, in conventional market transactions, clove quality assessment predominantly relies on visual inspection by traders. However, this subjective approach often leads to inconsistent evaluations, resulting in financial losses for farmers during sales. To address this limitation, the implementation of digital image processing techniques offers a more objective and reliable method for quality determination.

Digital image processing can be done to help the clove quality determination process. Clove quality classification is the same as the classification of other agricultural commodities; it can be done using deep learning and machine learning (Vij & Prashant 2024). Both types of classification have their advantages and disadvantages (Juma & Mdodo 2024). For example, deep learning can automatically extract features (Polly & Devi, 2024) and handle very complex data such as text and video data (Ramesh et al., 2024). Performance is better when the data is extensive (Ojo & Zahid, 2022). However, the weakness is that when the data is small, it will be difficult to classify (Ojo & Zahid, 2022); the need for high computation and increasingly complex representation because each neuron is interconnected and requires several layers (Samanta et al., 2023), many experiments need to be done to find the correct parameters in classifying (Alzubaidi et al., 2021), and the training process is long (Wonggasem et al., 2024). In machine learning, the advantages are being able to classify with small amounts of data (Juwono et al., 2023), faster computing time (Adhinata et al., 2024), and fewer resources (Omaye et al., 2024). However, the manual feature extraction process could be improved. And the feature selection process takes work (Corrales et al., 2022). They do not directly conclude that the feature is optimal while the others are not. So, it is necessary to extract some features to complement the weaknesses of other features. They represent the characteristics contained in the dataset.

This study will conduct color and texture feature extraction. Color characteristics in clove classification are very influential. In the clove quality standards in Indonesia, one of the requirements is the color of dry cloves to distinguish the quality of cloves. However, color characteristics alone are not enough to classify clove quality because cloves of quality one and quality two have almost the same color; the only difference is the completeness of the cloves, the same as quality two and quality 3, so texture characteristics are needed to help with deficiencies in color characteristics. In addition, by extracting texture features, it can distinguish between dry and undried cloves because they can be distinguished from the texture of the cloves and then the completeness of the dry cloves because the best quality cloves are cloves that are still complete or the buds are still not detached (Patnaik et al., 2025). So combining these two types of features is very important to get quality cloves.

Several previous researchers have combined several features to classify agricultural commodities with machine learning (Huong & Dung 2023). Jitanan & Chimlek, (2019) extracted texture and color features for soybean quality classification. Santos et al. (Dos Santos et al., 2020) extracted shape and color features to get coffee quality. Hayit et al., (2024) extracted texture and color features for chickpea classification. However, the problem is that there are still irrelevant features, so a feature selection stage is needed to obtain relevant features and improve the machine learning model’s performance, even though the amount of data is limited (Naik et al., 2024 ; Michael & Jackson 2025).

Of course, many features will be produced based on the extracted features. Thus, some features must be relevant to the data to be classified, as stated by (Naik et al., 2024). so it is necessary to carry out a feature selection process. The feature selection process is used to produce features that have an influence or are relevant to the data so that they are easier to recognize and manage for the classification process (Diaz & Jiju, 2022; Feizi-Derakhsh & Kadhim, 2023; Nguyen et al., 2023).

Three primary feature selection paradigms exist in machine learning: wrapper, filter, and embedded methods, each employing distinct selection approaches (Chauhan & Desai 2022). Recent studies (Gill et al., 2022; Thakur & Muppavaram, 2024) demonstrate that implementing feature selection techniques in agricultural applications can significantly enhance algorithmic performance. Furthermore, research by (Lasso et al., 2020) reveals that different feature selection methods yield varying optimal feature subsets when applied to agricultural datasets, highlighting the domain-specific nature of feature importance.

Different feature selection results have different algorithm performances. Lasso et al., (2020) compared various feature selection methods for coffee leaf rust disease. His research also showed that the dataset also affects the type of feature selection that can improve the algorithm’s performance.

Each feature selection method produces different selected features; it is certainly a challenge to accommodate these features to be processed further at the classification stage. For this reason, this study proposes an ensemble feature selection model to answer this challenge. Ensemble feature selection is a feature selection model combining five feature selection methods. In the process of combining by giving a score to each selected feature, a feature score with a value of 5 is selected for classification. The features selected using ensemble feature selection are consistent because each time a feature selection method is applied, the feature is selected in all feature selection methods. This proposed study has three contributions: 1) Classification of cloves according to SNI standards. 2). Dataset produced independently. 3). By employing ensemble feature selection, conventional machine learning models can perform better. This is how the rest of the paper is structured:

The literature on cloves and feature selection methods is reviewed in Section 2. The experimental results from the test function assessments are presented in Section 3, which also describes the technique suggested in this work, including its novel contributions. Section 4 discusses pertinent previous research and compares the performance of the suggested ensemble feature selection methodology with traditional methods. Section 5 concludes by summarizing the main findings of the study and outlining potential directions for further research.

2 Related Works

Previous research on cloves was conducted by (Tempola et al., 2023), but the clove classification process is not by the standards in Indonesia; only the classification of clove types such as whole clove, mother clove, and headless clove; the training process also takes a long time. Then, in another study on cloves, only classification is based on the level of maturity of the clove (Rosihan et al., 2024), and the image acquisition process on cloves individually is not stacked. Chalik & Al Maki, (2021) classify cloves into four classes. In contrast to this study, the classification process will be based on quality standards in Indonesia, and data acquisition will be stacked.

The feature selection process is very important to obtain features relevant to the object’s class and improving system accuracy (Gill et al., 2022). In agricultural applications, (She et al., 2020) implemented the sequential forward selection (SFS) method, a wrapper-based technique whose performance depends on the classifier type. Contrastingly, (Iniyan & Jebakumar, 2022) employed mutual information a filter-based approach that relies exclusively on dataset characteristics rather than learning algorithms for soybean and corn yield prediction. This distinction highlights the methodological divide between wrapper methods (classifier-dependent) and filter methods (data-dependent).

Lasso et al., (2020) experimented by comparing three types of feature selection models, namely embedded, filter, and wrapper on four datasets using seven feature selection methods. The experimental results showed that the 4D and 7D datasets had lower Mean Absolute Error (MAE) values when using the Embedded feature selection method. In contrast, the 3D and 14D MAE values were lower when using the wrapper sequential forward selection feature selection method for the 14D and 3D datasets using Recursive Forward Selection (Nam et al., 2023). This experiment shows that each feature selection process has different feature selection results and has its advantages depending on the dataset used. Shows that each embedded filter, and wrapper feature selection method has advantages and disadvantages (Seijo-Pardo et al., 2019).

The feature selection process of an ensemble feature selection has been carried out (Faizin et al., 2024; Seijo-Pardo et al., 2019); the way it works in feature selection is by giving a score. Then, when calculating the probability value of each feature, this method certainly still has unstable features but is still accommodated. Meanwhile, according to (Singh & Singh, 2021), the ensembling process was carried out in stages, first selected using several filter feature selection methods, then giving a score to each selected feature, and its threshold was determined; the selected feature is the feature that passes the set threshold. Furthermore, the selection process is carried out again with several wrapper feature selection methods, then the threshold is determined again, and the selected feature is the feature that passes the set threshold; there are two threshold determination processes, of course, it is not practical in feature selection, besides the embedded feature selection method has not been applied, so there are still inconsistent features that are selected.

3 Method

This study proposes an ensemble feature selection (EFS) approach to enhance machine learning model performance for clove quality classification according to SNI 01-3392-1994 standards. The methodology follows a structured pipeline consisting of several key stages. Initially, image acquisition, preprocessing, segmentation, and feature extraction are performed to prepare the input data. Subsequently, the ensemble feature selection process is applied to identify optimal feature subsets. Following feature selection, the pipeline proceeds with data normalization, iterative model training and testing, and comprehensive performance evaluation. The complete workflow of the proposed model is illustrated in Figure 1.

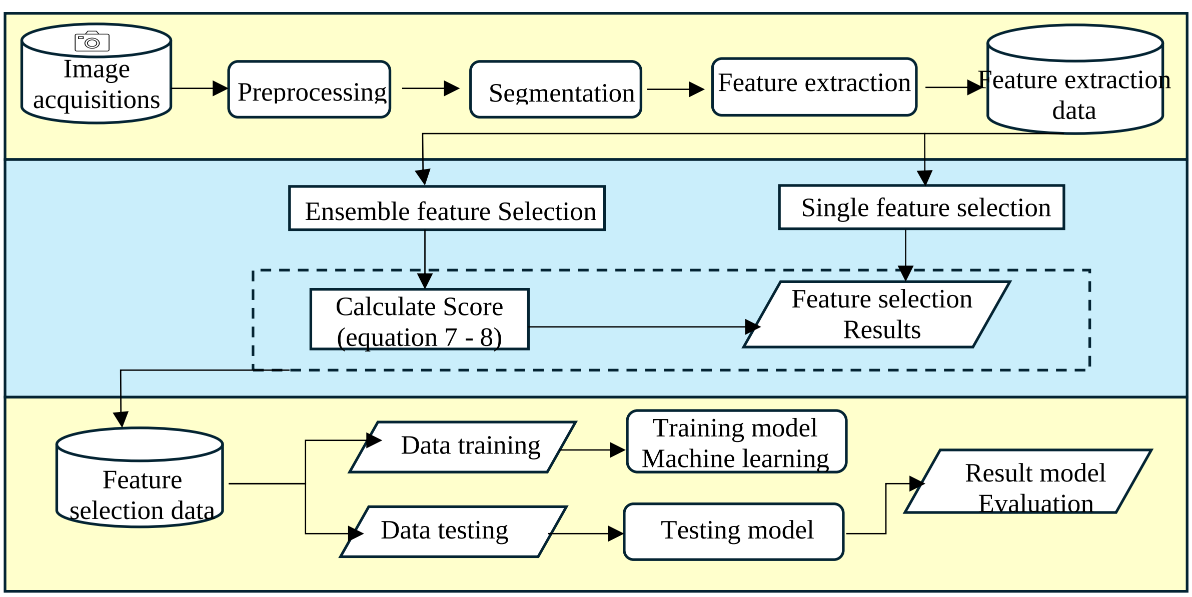

Figure 1: Clove Quality Classification Workflow

Image Acquisition

The study commenced with image acquisition of clove samples cultivated under specific environmental conditions: 60-80% relative humidity, temperatures ranging from 22-30°C, and altitudes between 200-600 meters above sea level. The sample distribution across quality classes was as follows: 34.17% for Quality 1, 31.85% for Quality 2, and 33.97% for Quality 3. The acquired clove image data conforms to Indonesian quality standards (SNI 01-3392-1994), ensuring that regional variability across Indonesian islands does not affect the quality assessment. Hardware used in the image acquisition procedure includes an OPPO A31 smartphone camera with rear camera specs of 12 MP + 2 MP + 2 MP, ISO-142, 4096 x 3072, and 72 dpi resolution. And a box that has LED lighting. Cloves are placed at the bottom of the box; then, the camera position is at the top. The camera was positioned at a distance of 24 cm from the clove image object. The clove image capture process is shown in Figure 2.

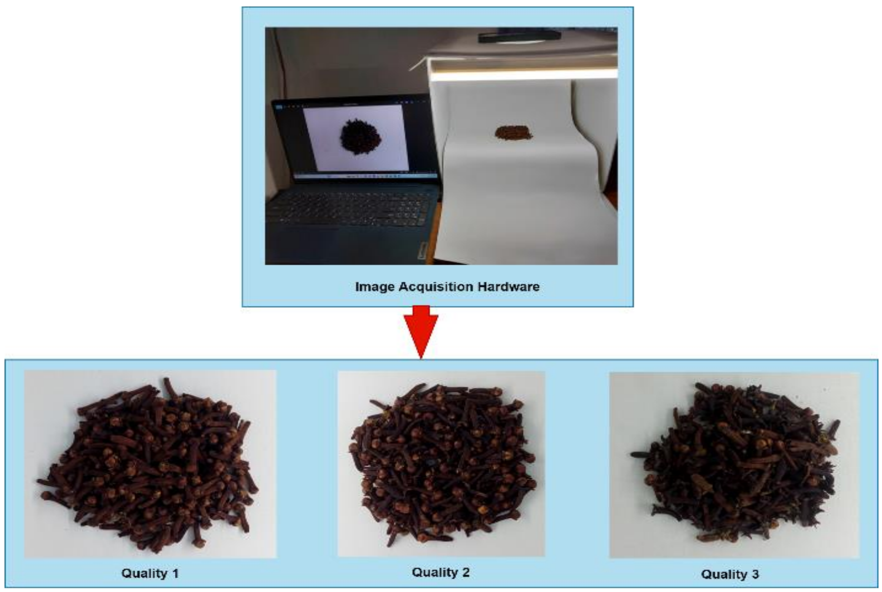

Figure 2: Image Acquisition Setup for Clove Classification

Preprocessing and Feature Extraction

Preprocessing is essential to speed up computing (Setyawan et al., 2022). After the preprocessing stage, the feature extraction stage continues. Feature extraction is carried out to facilitate the process of classifying clove quality. In this study, two main features were used, namely color features and texture features. The features extracted in the color feature section are RGB, HSV, and color moment. The use of these features is considered capable of recognizing objects well, as has been done by (Rahadiyan et al., 2023; Wang et al., 2021).

$$\text{Correlation} = \frac{\sum_i \sum_j (ij)\, P\,(i,j) \;-\; \mu_i\, \mu_j}{\sigma_i\, \sigma_j} \tag{1}$$

$$\text{ASM} = \sum_i \sum_j P\,(i,j\,)^2 \tag{2}$$

$$\text{Dissimilarity} = \sum_i \sum_j |i - j|\,.\,P\,(i,j) \tag{3}$$

$$\text{Homogeneity} = \frac{\sum_i \sum_j P(i,j)}{1 + |i - j|} \tag{4}$$

$$\text{Contrast} = \sum_i \sum_j |i - j|^2 P(i,j) \tag{5}$$

$$\text{Energy} = \sqrt{ASM} \tag{6}$$

The Gray Level Co-occurrence Matrix (GLCM) serves as the primary method for texture feature extraction in this study. As a fundamental technique in image processing, GLCM operates by computing probability values representing the spatial relationship between pixel pairs at specific distances and angular orientations. This matrix-based approach quantitatively characterizes texture by measuring the frequency of occurrence for pixel intensity pairs within defined spatial parameters of distance and direction (Chityala & Pudipeddi, 2021)—the values of the GLCM parameters (Allagwail et al., 2019) correlation, Angular Second Moment (ASM), Dissimilarity, Homogeneity, Contrast, and energy are obtained from angles 00, 450, 900, and 1350. Each value is obtained from Equation (1-6).

Ensemble Feature Selection

Feature selection techniques have demonstrated significant improvements in machine learning model performance, particularly regarding computational efficiency and predictive accuracy. This study introduces an ensemble feature selection approach, which combines multiple selection methods to yield more stable and consistent feature subsets compared to individual techniques. By reducing the feature space dimensionality, this method effectively decreases computational requirements while preserving - and potentially enhancing - model performance relative to conventional machine learning approaches.

The proposed ensemble feature selection (EFS) method uses three filter methods, namely Chi-Square, Mutual Information, and Variance Threshold. The wrapper and embedded methods used in the research are Recursive Feature Selection and Lasso Regression. These five feature selection methods are used because each has advantages and disadvantages, allowing them to complement one another. For example, the chi-square method is highly reliable for large datasets but performs poorly with non-linear data. Mutual information can accommodate non-linear relationships but may still select redundant features or those with high correlations. The variance threshold method is efficient for large datasets; however, it risks removing important features if their variance is low—even if they significantly impact the target variable. Recursive feature elimination (RFE) handles feature interactions well but heavily depends on the initial model. Lasso regression eliminates unimportant features by shrinking their coefficients to zero, but it becomes unstable with highly correlated features, potentially selecting them at random. Combining these methods can mitigate their limitations while leveraging their strengths (Effrosynidis & Arampatzis, 2021). To get the feature selection results with ensemble feature selection, give each feature a score in each feature selection method. Features that have a score = 5 are chosen. An illustration of ensemble feature selection is shown in Figure 3. This method is adopted from the concept of intersection. An intersection is the intersection of two sets, A and B, which contain all members of A that belong to B (or all members of B that belong to A). If connected to the ensemble feature selection, then there are five sets, namely Fa (Chi-Square), Fb (Mutual Information), Fc (Variance Threshold), Fd (Recursive Feature Elimination), and Fe (Lasso Regression). Equation (7-8) is used to calculate the score of each feature.

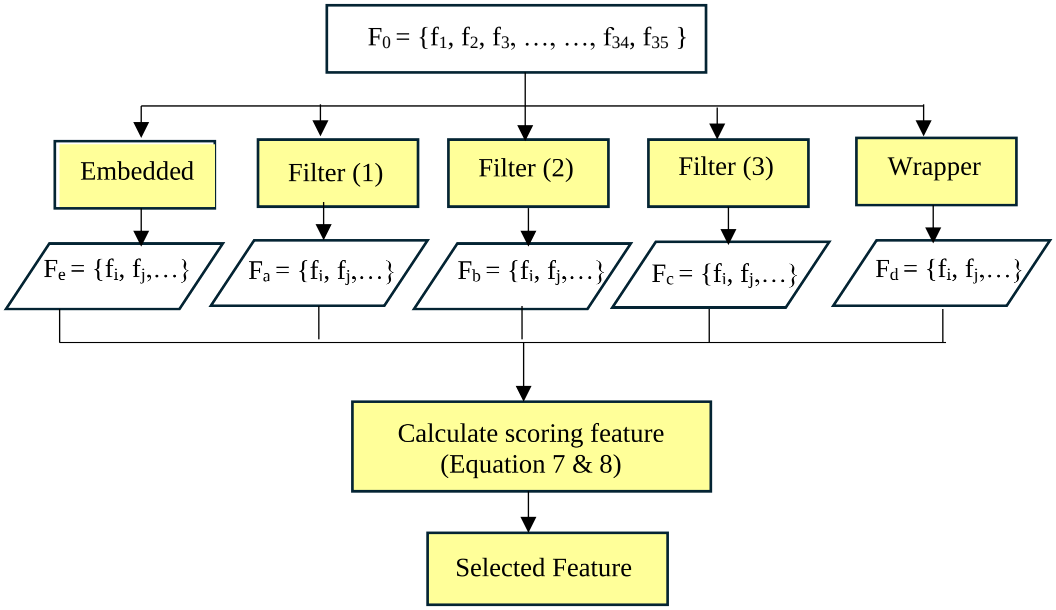

Figure 3. Model Ensemble Feature Selection

$$score\ (f_i) = \sum_{j=1}^{5} exist\ (F_j, f_1) \qquad (7)$$

$$exist\ (F_j, f_1) = \begin{cases} 1, if\ f_1\ in\ F_j \\ 0, \quad other \end{cases} \qquad (8)$$

where fi is the index of each ith feature, Fj is the index of the jth applied feature selection method. All clove quality features are denoted as F0 = {f1, f2, f3, …, …, …, f34, f35}. If we assume the set of each feature selection result is as follows:

$$F_a = \{f_1, f_3, f_6, f_8, f_9, f_{11}, f_{13}, f_{15}, f_{20}, f_{24}, f_{28}, f_{29}, f_{31}, f_{32}, f_{34}, f_{35}\}$$

$$F_b = \{f_1, f_2, f_3, f_4, f_5, f_9, f_{12}, f_{15}, f_{17}, f_{18}, f_{20}, f_{25}, f_{27}, f_{31}, f_{32}, f_{34}, f_{35}\}$$

$$F_c = \{f_1, f_3, f_7, f_{10}, f_{11}, f_{12}, f_{15}, f_{16}, f_{22}, f_{24}, f_{26}, f_{31}, f_{32}, f_{33}, f_{34}, f_{35}\}$$

$$F_d = \{f_1, f_2, f_3, f_5, f_8, f_9, f_{12}, f_{14}, f_{15}, f_{21}, f_{29}, f_{30}, f_{31}, f_{32}, f_{33}, f_{34}, f_{35}\}$$

$$F_e = \{f_1, f_3, f_4, f_5, f_9, f_{10}, f_{13}, f_{15}, f_{17}, f_{21}, f_{27}, f_{28}, f_{31}, f_{32}, f_{33}, f_{34}, f_{35}\}$$

So, to calculate the $score\ (f_i)$ of each feature, the selected feature is produced or forms a new set (X), namely {f1, f3, f15, f20, f31, f32, f34, f35}. If calculated, each member element in the new set has a score = 5. This means that the member element is always present in every five sets.

Machine Learning Evaluation Model

This study will compare several methods in machine learning, such as k-nearest neighbour, naïve bayes, multilayer perceptron, and support vector machines (Ahmed & Husien, 2024). The reason for using several machine learning methods is to get a technique that suits the data used. Each machine-learning method has advantages depending on the data used (Gope & Fukai, 2022). Furthermore, a confusion matrix measures the machine learning model's performance, and its accuracy value is calculated. In this study, clove quality is categorized into three classes, necessitating the use of a confusion matrix for multiple classes. The classification performance across the three target classes is quantitatively represented in the confusion matrix provided in Table 1.

Table 1: Confusion matrix 3 classes

		Predicted class (C_j)		
		C_1	C_2	C_3
actual class (C_i)	C_1	C_{11}	C_{12}	C_{13}
	C_2	C_{21}	C_{22}	C_{23}
	C_3	C_{31}	C_{31}	C_{31}

In the confusion matrix, True Positive (TP) values correspond to the diagonal elements (C_{11}, C_{22}, C_{33}). For any given class, the False Negative (FN) is computed as the sum of off-diagonal elements in its row, while the False Positive (FP) is derived from off-diagonal elements in its column. For instance, class 1's FN equals $C_{12} + C_{13}$, and its FP equals $C_{21} + C_{31}$. The True Negative (TN) for a class represents the sum of all elements excluding its row and column (e.g., TN for class 1 is $C_{22} + C_{23} + C_{32} + C_{33}$). These metrics, along with the overall True Positive (TP_{all}), can be formally expressed through Equations (9)-(13). Here, n represents the number of classes, and Cij represents the value in the confusion matrix at the intersection of the ith row and jth column. Subsequently, precision, recall, accuracy, and f1-score values can be calculated using Equation (14) – (17).

$$TP_i = C_{jj\ (j=i)} \tag{9}$$

$$TP_{all} = \sum_{j=1}^{n} C_{jj} \tag{10}$$

$$FN_i = \sum_{j=1, j\neq i}^{n} C_{ij} \tag{11}$$

$$FP_i = \sum_{j=1, j\neq i}^{n} C_{ji} \tag{12}$$

$$TN_i = \sum_{j=1, j\neq i}^{n} \sum_{k=1\ k\neq i}^{n} C_{jk} \tag{13}$$

$$Precision_i = \frac{TP_i}{TP_i + FP_i} \times 100\% \tag{14}$$

$$Recall_i = \frac{TP_i}{TP_i + FN_i}\ x\ 100\% \tag{15}$$

$$Accuracy = \frac{\sum_{i=1}^{n} C_{ii}}{\sum_{i=1}^{n}\sum_{j=1}^{n} C_{ij}}\ x\ 100\% \tag{16}$$

$$F1_i = \frac{2\,.\,Precision_i\,.\,Recall_i}{Precision_i + Recall_i}$$

$$Average\ F1\ Score = \frac{1}{n}\sum_{i=1}^{n} F1_i \tag{17}$$

where n is the number of classes, Cii represents the diagonal elements of the confusion matrix, which indicate the number of data points from class i correctly predicted as class i (True Positives for each class). Cij represents the elements in row i and column j, indicating the number of data points from class i that are incorrectly predicted as class j. The term $\sum_{i=1}^{n} C_{ii}$ represents the total number of correct predictions across all classes (i.e., the sum of the diagonal elements). Meanwhile, $\sum_{i=1}^{n}\sum_{j=1}^{n} C_{ij}$ represents the sum of all elements in the confusion matrix, corresponding to the total number of data points. Additionally, the model is evaluated using k-fold cross-validation, with a starting k value of 10. This procedure ensures that every dataset, specifically training and testing data, has the same privileges. Additionally, the model is evaluated using k-fold cross-validation, with a starting k value of 10. Every dataset has the same rights thanks to this procedure.

4 Result And Discussion

Initial Experimental Results

Stacked clove image acquisition: The total number of clove images successfully acquired was 2025, divided into three classes: quality 1, equal 692; quality 2, 645; and quality 3, 688. Then, preprocessing and feature extraction were carried out. The total number of features extracted was 35. After feature extraction, testing was carried out by applying four machine-learning models. The testing parameters are presented in Table 2, while the corresponding results are displayed in Table 3.

Table 2: Machine Learning Model Testing Parameters

Method	Parameters	Tuning Values	Optimal Parameter
k-NN	Distance k	Minkowski, Euclidean 3, 5, 7	Minkowski 7
NBC	Smoothing Priors	1e-09 None	1e-09 None
SVM	Kernel regularization	Linear C=1.0	Linear C=1.0
mlp	Hidden layer Neuron Learning rate Maximal iteration Activation	1 7 0.001 5000 Relu	1 7 0.001 5000 Relu

Table 3: Preliminary Experimental Results

Method	Accuracy	Precision	Recall	F1 Score
K-NN	90.92%	90.95%	90.83%	90.83%
NBC	81.04%	81.57%	80.95%	80.80%
MLP	95.43%	95.54%	95.44%	95.44%
SVM	94.93%	95.09%	94.96%	94.96%

The preliminary experimental results demonstrate significant performance variations across the evaluated machine learning models. The Multilayer Perceptron (MLP) achieved optimal performance when utilizing all 35 extracted features, attaining an accuracy of 95.43%, with precision=85.36%, recall=84.28%, and F1-score=83.99%. In contrast, the Naïve Bayes classifier exhibited the lowest performance metrics (accuracy=81.04%, precision=81.57%, recall=80.95%, F1-score=80.80%). Initial single-feature testing revealed that color features produced the highest predictive accuracy among individual features. In this configuration, three models k-NN, SVM, and MLP consistently exceeded 90% accuracy, while Naïve Bayes trailed at 84.75% accuracy. These results indicate that the stacked image acquisition process produces more consistent colors, and the globally dominant clove color appears more homogeneous, facilitating easier classification. In contrast, the texture feature resulted in accuracy below 80% across all tested machine-learning models. This is because overlapping clove surfaces introduce noise that disrupts the natural texture patterns during the stacked image acquisition process. Consequently, the extracted GLCM features become less effective and fail to produce representative results.

The color features demonstrated the most substantial contribution to clove quality classification, as evidenced by the performance of four machine learning models: k-NN (90.92% accuracy), Naïve Bayes (84.75%), Support Vector Machine (93.58%), and Multilayer Perceptron (92.83%). The superior performance of color features compared to texture features can be attributed to the strong correlation between clove color and quality. Extracted color features (RGB and HSV) effectively captured these quality distinctions. In contrast, texture characteristics (e.g., wrinkles, surface roughness) exhibited greater variability due to external factors such as drying processes, storage conditions, and plant varieties, resulting in weaker quality correlations. Furthermore, the stacked image acquisition method introduced additional texture variability, as overlapping cloves obscured individual surface features.

Experimental results with feature selection

The feature selection models used in clove quality classification include chi-square, mutual information, Variance Threshold, Recursive feature Elimination, Lasso Regression, and ensemble feature selection (Proposed Model). The feature selection results from each feature selection method applied were then tested with a machine learning model. The results of clove quality classification using feature selection are shown in Table 4.

Table 4: The Results of Machine Learning Model Performance After Feature Selection

Feature Selection	**Method Machine Learning**			
	KNN	**Naïve Bayes**	**MLP**	**SVM**
Chi-Square	91.79%	80.74%	93.88%	93.39%
Mutual Information	91.97%	80.55%	93.33%	93.39%
Variance Threshold	91.79%	80.74%	93.88%	93.39%
Recursive Feature Elimination	92.16%	81.04%	94.69%	94.62%
Lasso Regression	93.20%	84.01%	94.38%	94.93%
Ensemble feature selection (threshold)	91.79%	80.74%	93.88%	93.39%
Proposed Method: Ensemble feature selection (intersect)	93.58%	82.71%	93.02%	94.25%

The experiment results with single feature selection show that the k-NN method has increased in all feature selection models. The highest accuracy when using lasso feature selection with an accuracy of 93.20% with a total of 20 selected features. For the naïve Bayes method, the feature selection method that experienced an increase in accuracy was when lasso regression feature selection was applied; the resulting accuracy was 84.01%. Furthermore, there was no increase in accuracy for SVM and multilayer perceptron when a single feature selection was used. On the contrary, accuracy decreased.

In contrast to the application of single feature selection, when ensemble feature selection is applied, there are 10 selected features: red, grey, hue, saturation, homogeneity_135, contrast_90, ASM_135, energy_0, energy_90, and energy_135. These 10 features are highly consistent and relevant because they are consistently selected in every feature selection model. They consist of 4 colour features and six texture features. The results of this feature selection demonstrate that combining features in clove quality classification is crucial, as it can significantly impact the machine learning model's performance. Thist is evidenced by the increased accuracy of the k-Nearest Neighbors and naïveBayess models when applying ensemble feature selection. The accuracy of k-NN reached 93.58%, the highest achieved during the entire experimental process using the k-NN model. Similarly, the accuracy of Naïve Bayes improved to 82.71%. However, in contrast, the accuracy of the multilayer perceptron and support vector machine models either decreased or did not improve. This decline is likely due to the loss of other features that contain essential information, which can reduce the performance of these models.

The ensemble feature selection process revealed that selected color features constituted 36.36% of the total extracted color features (4 out of 11 features). In comparison, selected texture features accounted for 25% of the total texture features (6 out of 24 features). These results indicate that color features demonstrate greater suitability for the ensemble feature selection model than texture features. Notably, none of the color moment features were selected during the ensemble feature selection process, suggesting their limited contribution to feature selection. This lack of contribution stems from two key factors: (1) during single feature selection, none of the color moment features achieved a score of 5, and (2) color moment features only provide global color summarization of clove quality, lacking the detailed discriminative power of RGB and HSV features.

Ensemble feature selection is suitable for machine learning models such as k-NN and Naïve Bayes. Both models benefit substantially from this process, though through distinct mechanisms. For k-NN, ensemble feature selection improves distance calculation accuracy by eliminating redundant feature noise. In naïve Bayes, the method enhances model performance by reducing feature correlations that violate the algorithm's independence assumptions. Furthermore, ensemble feature selection simplifies the classification problem by identifying and retaining only the most relevant and consistent attributes within color and texture feature spaces. The k-NN model shows particular promise for automated clove quality sorting applications. Implementing this approach could significantly improve upon current manual assessment methods, which are inherently subjective due to their reliance on visual inspection and human experience.

In contrast to k-nearest neighbor and naïve bayes models, the support vector machine and multilayer perceptron exhibited reduced classification accuracy following feature selection. The SVM model demonstrated a 0.68% accuracy decrease, while MLP performance declined by 2.41%. This performance reduction can be attributed to fundamental differences in model architectures: (1) SVM relies on kernel transformations to map data into high-dimensional spaces, where removal of features critical for class separation may diminish effectiveness; and (2) MLP's performance degradation results from the loss of essential feature interactions when numerous features are eliminated, as observed across all feature selection processes applied to clove quality classification.

Furthermore, the 10 features retained after ensemble feature selection proved overly simplistic for these models, potentially leading to underfitting. This simplification particularly affected the MLP's ability to learn complex, non-linear relationships between features, while the SVM struggled with reduced dimensionality in the transformed feature space. In addition to testing performance in classification, the research also tested computing time. The results of the computing time test of the four applied machine learning models are shown in Table 5.

Table 5: The Computation Time Required for Each Experiment (in milliseconds)

Model	Without FS	Model Feature Selection					
		CS	**MI**	**VT**	**RFE**	**RL**	**EFS**
K-NN	217	206	207	204	212	499	**123**
NBC	110	40.7	40.4	39. 4	37.8	93.4	**32**
SVM	6020	3080	3010	3080	2630	2570	**1390**
MLP	84000	68000	64000	72000	75000	76000	**55100**

The test results in Table 5 show that applying ensemble feature selection can reduce the computation time of all machine learning models tested. The application of ensemble feature selection can reduce features by 25 features. Moreover, there are only ten selected features. These ten features are consistent features when applying a single feature selection. These ten features are features that are always selected. This is so that the EFS model can improve the machine learning model's performance, especially conventional machine learning models.

In contrast to the application of single feature selection, some models experience an increase in computation time, such as in the feature selection method with lasso regression. Previously, without applying feature selection, the resulting computation time for K-NN 217 milliseconds increased to 499 milliseconds.

Comparative Study

This study is also compared with several other studies using single-feature and ensemble feature selection. Traditional feature selection has been made by (Corrales et al., 2022; Iniyan & Jebakumar, 2022; Lasso et al., 2020; Olu-Ajayi et al., 2023) Where the outcomes demonstrate that the machine learning model's performance may be impacted by the use of feature selection. Then the application of ensembling feature selection there have also been several that have done it, such as (Faizin et al., 2024; Seijo-Pardo et al., 2019), the process of determining the threshold when applied to the quality of cloves, the accuracy of the cloves decreased as tested in Table 4. N. Singh & P. Singh (Singh & Singh, 2021). The method used was not applied to the case of clove quality because the process of providing thresholds was repeated. The results of the comparison of existing research and those carried out are shown in Table 6.

Based on comparing accuracy from several studies shown in Table 6 by applying feature selection, the machine learning method has advantages, such as in (Iniyan & Jebakumar, 2022) SVM has the best accuracy. (Olu-Ajayi et al., 2023) have the best accuracy when using decision trees, (Singh) naïve Bayes has the best performance, and (Faizin et al., 2024) K-NN has the best performance. Furthermore, compared with the proposed EFS method, its accuracy is better than other EFS methods, such as those used by (Seijo-Pardo et al., 2019). When using the EFS method (threshold) for clove quality classification, the accuracy is 91.79% for K-NN and Naïve Bayes 80.74%. Different from the proposed ensemble method for K-NN accuracy, reaching 93.58%, and naïve Bayes at 82.71%. This shows that the proposed EFS process effectively filters features that still have noise, in contrast to using thresholds that still accommodate these features even though they are not yet relevant and consistent because they have passed the set threshold. In addition, the proposed ensemble method can reduce the feature dimension so that the resulting feature subset is smaller and more informative. Then, the K-NN and naïve Bayes methods experienced increased accuracy performance because both methods are susceptible to high dimensionality. The comparison did not use the same data in this study because the clove quality image data was acquired independently. For this reason, the approach used has the potential to be applied in a broader context or on a more diverse dataset.

Table 6: Performance Comparison of the Proposed Method with Existing Ones.

References	Method Feature Selection	Method	Accuracy
S. Iniyan [29]	Mutual Information	SVM	92.43%
Olu-Ajayi, et al. [40]	Comparison of feature selection	DT	68%
Seijo-Pardo et. El [30]	Ensemble Feature Selection (Threshold)	SVM	82%
Singh [37]	Ensemble Feature Selection (Hybrid)	NB	88%
Faizin [36]	Ensemble Feature Selection (Threshold)	K-NN	85.87%
Proposed method	Ensemble Feature Selection	K-NN	93.58%

Finally, the findings from the ensemble feature selection process provide valuable insights for agricultural applications, particularly in the classification of clove quality. By identifying the most relevant color and texture features, this approach can improve the accuracy and efficiency of automated clove classification systems. In agriculture, traditional manual classification methods rely heavily on subjective visual inspection, which can lead to inconsistent and unreliable results. The EFS model prioritizes color features (36.36% of the total color features). It is particularly beneficial for machine learning models like k-NN and Naïve Bayes, as it helps eliminate redundant features and reduces noise. This leads to more accurate clove quality classification, making the sorting process faster and more objective. Furthermore, the reduced performance of models like SVM and MLP highlights the importance of careful feature selection, as simplifying the feature set too much can lead to underfitting, which may negatively impact the ability to classify complex agricultural data. Ensemble feature selection can significantly improve clove quality classification, streamlining the sorting process and ensuring more reliable outcomes in agricultural practices.

5 Conclusion

This study develops an ensemble feature selection model to enhance clove quality classification in accordance with SNI 01-3392-1994 standards. From an initial set of 35 extracted features, five distinct selection methods were employed: chi-square, mutual information, variance threshold, recursive feature elimination, and lasso regression. The proposed model operates by aggregating features that consistently achieve a perfect selection score (score=5) across all methods, thereby creating an optimized feature subset. Four machine learning algorithms—k-NN, SVM, Naïve Bayes, and Multilayer Perceptron—were evaluated. The analysis revealed significant variations in both selected features and model performance across different feature selection approaches. The ensemble method demonstrated notable improvements, increasing accuracy by 2.66% for k-NN and 1.66% for Naïve Bayes when utilizing the optimal 10-feature subset. However, performance degradation was observed in MLP and SVM models. These findings suggest that feature weighting adjustments could potentially enhance performance for these algorithms. Importantly, the ensemble approach consistently reduced computational time across all tested models, demonstrating its efficiency.

Acknowledgments

This work was partially supported by the Department of Computer Science and Electronics, Universitas Gadjah Mada, under the Publication Funding Year 2025.

References

[1] Adhinata, F. D., Wahyono, & Sumiharto, R. (2024). A comprehensive survey on weed and crop classification using machine learning and deep learning. *Artificial Intelligence in Agriculture, 13*, 45–63. https://doi.org/10.1016/j.aiia.2024.06.005

[2] Ahmed, M., & Husien, I. (2024). Hybrid Machine Learning Approach for Accurate Heart Disease Prediction. *International Journal of Intelligent Engineering and Systems, 17*(4), 728–737. https://doi.org/10.22266/IJIES2024.0831.55

[3] Allagwail, S., Gedik, O. S., & Rahebi, J. (2019). Face recognition with symmetrical face training samples based on local binary patterns and the gabor filter. *Symmetry, 11*(2). https://doi.org/10.3390/sym11020157

[4] Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. In Journal of Big Data. 8(1). *Springer International Publishing.* https://doi.org/10.1186/s40537-021-00444-8

[5] Chaichi, M., Mohammadi, A., Badii, F., & Hashemi, M. (2021). Triple synergistic essential oils prevent pathogenic and spoilage bacteria growth in the refrigerated chicken breast meat. *Biocatalysis and Agricultural Biotechnology, 32*(l). https://doi.org/10.1016/j.bcab.2021.101926

[6] Chalik, F. A., & Al Maki, W. F. (2021). Classification of Dried Clove Flower Quality using Convolutional Neural Network. *2021 International Conference on Data Science, Artificial Intelligence, and Business Analytics, DATABIA 2021 - Proceedings, 40–45.* https://doi.org/10.1109/DATABIA53375.2021.9650199

[7] Chauhan, S., & Desai, M. (2022). Machine Learning-Based Predictive Maintenance Scheduling in Industrial Settings. *International Academic Journal of Science and Engineering, 9*(4), 9–12. https://doi.org/10.71086/IAJSE/V9I4/IAJSE0929

[8] Chityala, R., & Pudipeddi, S. (2021). Image Processing and Acquisition using Python. *CRC Press.* https://lccn.loc.gov/2020015556

[9] Corrales, D. C., Schoving, C., Raynal, H., Debaeke, P., Journet, E. P., & Constantin, J. (2022). A surrogate model based on feature selection techniques and regression learners to improve soybean yield prediction in southern France. *Computers and Electronics in Agriculture, 192(November 2021).* https://doi.org/10.1016/j.compag.2021.106578

[10] Diaz, P. M., & Jiju, M. J. E. (2022). A comparative analysis of meta-heuristic optimization algorithms for feature selection and feature weighting in neural networks. *Evolutionary Intelligence, 15*(4), 2631–2650. https://doi.org/10.1007/s12065-021-00634-6

[11] Dos Santos, F. F. L., Rosas, J. T. F., Martins, R. N., Araújo, G. de M., Viana, L. de A., & Gonçalves, J. de P. (2020). Quality assessment of coffee beans through computer vision and machine learning algorithms. *Coffee Science, 15*(1), 1–9. https://doi.org/10.25186/.v15i.1752

[12] Effrosynidis, D., & Arampatzis, A. (2021). An evaluation of feature selection methods for environmental data. *Ecological Informatics, 61(January),* 101224. https://doi.org/10.1016/j.ecoinf.2021.101224

[13] Faizin, M. A., Kurniasari, D. T., Elqolby, N., Putra, M. A. R., & Ahmad, T. (2024). Optimizing Feature Selection Method in Intrusion Detection System Using Thresholding. *International Journal of Intelligent Engineering and Systems, 17*(3), 214–226.

[14] Feizi-Derakhsh, M. R., & Kadhim, E. A. (2023). An Improved Binary Cuckoo Search Algorithm For Feature Selection Using Filter Method And Chaotic Map. *Journal of Applied Science and Engineering, 26*(6), 897–903. https://doi.org/10.6180/jase.202306_26(6).0015

[15] Gill, M., Anderson, R., Hu, H., Bennamoun, M., Petereit, J., Valliyodan, B., Nguyen, H. T., Batley, J., Bayer, P. E., & Edwards, D. (2022). Machine learning models outperform deep learning models, provide interpretation and facilitate feature selection for soybean trait prediction. *BMC Plant Biology, 22*(1), 1–8. https://doi.org/10.1186/s12870-022-03559-z

[16] Gope, H. L., & Fukai, H. (2022). Peaberry and normal coffee bean classification using CNN,

SVM, and KNN: Their implementation in and the limitations of Raspberry Pi 3. *AIMS Agriculture and Food*, *7*(1), 149–167. https://doi.org/10.3934/AGRFOOD.2022010

[17] Hayit, T., Endes, A., & Hayit, F. (2024). KNN-based approach for the classification of fusarium wilt disease in chickpea based on color and texture features. *European Journal of Plant Pathology*, *168*(4), 665–681. https://doi.org/10.1007/s10658-023-02791-z

[18] Huong, N. T., & Dung, N. T. (2023). The Experience of the US and Japan in Agricultural Economic Development Policy and Lessons for Vietnam. *International Journal of Advances in Engineering and Emerging Technology*, *14*(1), 16–21

[19] Iniyan, S., & Jebakumar, R. (2022). Mutual Information Feature Selection (MIFS) Based Crop Yield Prediction on Corn and Soybean Crops Using Multilayer Stacked Ensemble Regression (MSER). *Wireless Personal Communications*, *126*(3), 1935–1964. https://doi.org/10.1007/s11277-021-08712-9

[20] Jitanan, S., & Chimlek, P. (2019). Quality grading of soybean seeds using image analysis. *International Journal of Electrical and Computer Engineering*, *9*(5), 3495–3503. https://doi.org/10.11591/ijece.v9i5.pp3495-3503

[21] Juma, J., & Mdodo, R. (2024). Implementing Machine Learning Algorithms for Predictive Maintenance in Manufacturing. *International Academic Journal of Innovative Research*, *11*(3), 21–26.

[22] Juwono, F. H., Wong, W. K., Verma, S., Shekhawat, N., Lease, B. A., & Apriono, C. (2023). Machine learning for weed–plant discrimination in agriculture 5.0: An in-depth review. *Artificial Intelligence in Agriculture*, *10*, 13–25. https://doi.org/10.1016/j.aiia.2023.09.002

[23] Lasso, E., Corrales, D. C., Avelino, J., de Melo Virginio Filho, E., & Corrales, J. C. (2020). Discovering weather periods and crop properties favorable for coffee rust incidence from feature selection approaches. *Computers and Electronics in Agriculture*, *176(August)*, 105640. https://doi.org/10.1016/j.compag.2020.105640

[24] Michael, P., & Jackson, K. (2025). Advancing scientific discovery: A high performance computing architecture for AI and machine learning. *Journal of Integrated VLSI, Embedded and Computing Technologies*, *2*(2), 18–26.

[25] Naik, N. K., Subbarao, M. V., Sethy, P. K., Behera, S. K., & Panigrahi, G. R. (2024). Machine learning with analysis-of-variance-based method for identifying rice varieties. *Journal of Agriculture and Food Research*, *18*(September), 101397. https://doi.org/10.1016/j.jafr.2024.101397

[26] Nam, H., Nunes, M. G. V., & Loukachevitch, N. (2023). 3D printing: Next-generation realization for future applications. *International Journal of Communication and Computer Technologies*, *11*(2), 19-24. https://doi.org/10.31838/IJCCTS/11.02.04

[27] Nguyen, T. H. T., Van Pham, N., Nguyen, V. N. N., Pham, H. M., & Phan, Q. B. (2022). Forecasting wind speed using a hybrid model of convolutional neural network and long-short term memory with Boruta Algorithm-based feature selection. *Journal of Applied Science and Engineering*, *26*(8), 1053-1060. https://doi.org/10.6180/jase.202308_26(8).0001

[28] Ojo, M. O., & Zahid, A. (2022). Deep learning in controlled environment agriculture: A review of recent advancements, challenges and prospects. *Sensors*, *22*(20), 7965. https://doi.org/10.3390/s22207965

[29] Olu-Ajayi, R., Alaka, H., Sulaimon, I., Balogun, H., Wusu, G., Yusuf, W., & Adegoke, M. (2023). Building energy performance prediction: A reliability analysis and evaluation of feature selection methods. *Expert Systems with Applications*, *225*, 120109. https://doi.org/10.1016/j.eswa.2023.120109

[30] Omaye, J. D., Ogbuju, E., Ataguba, G., Jaiyeoba, O., Aneke, J., & Oladipo, F. (2024). Cross-comparative review of Machine learning for plant disease detection: apple, cassava, cotton and potato plants. *Artificial Intelligence in Agriculture*, *12*, 127–151. https://doi.org/10.1016/j.aiia.2024.04.002

[31] Patnaik, D. I., Priyadarshini, P., Mishra, D., & Chaudhuri, S. D. (2025). Text to Tech: Preserving Ancient Hindu Texts in the Digital Realm. *Indian Journal of Information Sources and Services, 15*(1), 42–46. https://doi.org/10.51983/ijiss-2025.IJISS.15.1.07

[32] Polly, R., & Devi, E. A. (2024). Semantic segmentation for plant leaf disease classification and damage detection: A deep learning approach. *Smart Agricultural Technology, 9*(May), 100526. https://doi.org/10.1016/j.atech.2024.100526

[33] Rahadiyan, D., Hartati, S., Wahyono, & Nugroho, A. P. (2023). Feature aggregation for nutrient deficiency identification in chili based on machine learning. *Artificial Intelligence in Agriculture, 8*, 77–90. https://doi.org/10.1016/j.aiia.2023.04.001

[34] Ramesh, K., Mushtaq, F., Deshmukh, S., Ray, T., Parimi, C., Basem, A., & Elsheikh, A. (2024). An investigation of deep learning approaches for efficient assembly component identification. *Beni-Suef University Journal of Basic and Applied Sciences, 13*(1). https://doi.org/10.1186/s43088-024-00537-2

[35] Rosihan, Tempola, F., Nurtanzis Sutoyo, M., & Eri Gunawan, C. (2024). Improving System Accuracy by Modifying the Transfer Learning Architecture for Detecting Clove Maturity Levels. *Journal of Advances in Information Technology, 15*(3), 407–413. https://doi.org/10.12720/jait.15.3.407-413

[36] Samanta, I. S., Panda, S., Rout, P. K., Bajaj, M., Piecha, M., Blazek, V., & Prokop, L. (2023). A Comprehensive Review of Deep-Learning Applications to Power Quality Analysis. *Energies, 16*(11), 1–31. https://doi.org/10.3390/en16114406

[37] Seijo-Pardo, B., Bolón-Canedo, V., & Alonso-Betanzos, A. (2019). On developing an automatic threshold applied to feature selection ensembles. *Information Fusion*, *45*, 227-245. https://doi.org/10.1016/j.inffus.2018.02.007

[38] Setyawan, D., Wardoyo, R., Wibowo, M. E., & Murhandarwati, E. E. H. (2022). Classification of plasmodium falciparum based on textural and morphological features. *International Journal of Electrical and Computer Engineering, 12*(5), 5036–5048. https://doi.org/10.11591/ijece.v12i5.pp5036-5048

[39] Sharma, M., Mann, B., Pothuraju, R., Sharma, R., & Kumar, R. (2022). Physico-chemical characterization of ultrasound assisted clove oil-loaded nanoemulsion: as enhanced antimicrobial potential. *Biotechnology Reports*, *34*, e00720. https://doi.org/10.1016/j.btre.2022.e00720

[40] She, B., Yang, Y. Y., Zhao, Z. G., Huang, L. S., Liang, D., & Zhang, D. Y. (2020). Identification and mapping of soybean and maize crops based on Sentinel-2 data. *International Journal of Agricultural and Biological Engineering, 13*(6), 171–182. https://doi.org/10.25165/j.ijabe.20201306.6183

[41] Singh, N., & Singh, P. (2021). A hybrid ensemble-filter wrapper feature selection approach for medical data classification. *Chemometrics and Intelligent Laboratory Systems*, *217*, 104396. https://doi.org/10.1016/j.chemolab.2021.104396

[42] Tempola, F., Wardoyo, R., Musdholifah, A., & Nugroho, A. K. (2023, October). Implementation of Deep Learning With Resnet50 Modification for Clove classification. In *2023 IEEE 9th Information Technology International Seminar (ITIS)* (pp. 1-4). IEEE. https://doi.org/10.1109/ITIS59651.2023.10419960

[43] Thakur, A. S., & Muppavaram, K. (2024). Feature Selection Using Adaptive Weight Based Grey Wolf Optimization for Malware Detection in Android. *International Journal of Intelligent Engineering & Systems*, *17*(3). https://doi.org/10.22266/ijies2024.0630.49

[44] Vij, P., & Prashant, P. M. (2024). Analyzing Soil Pollution by Image Processing and Machine Learning at Contaminated Agricultural Field. *Natural and Engineering Sciences, 9*(2), 335-346. https://doi.org/10.28978/nesciences.1575484

[45] Wang, S., Xia, X., Ye, L., & Yang, B. (2021). Automatic detection and classification of steel surface defect using deep convolutional neural networks. *Metals, 11*(3), 1–23. https://doi.org/10.3390/met11030388

[46] Wonggasem, K., Chakranon, P., & Wongchaisuwat, P. (2024). Automated quality inspection of baby corn using image processing and deep learning. *Artificial Intelligence in Agriculture, 11*, 61–69. https://doi.org/10.1016/j.aiia.2024.01.001

Authors Biography

Firman Tempola obtained his Master of Computer Science in 2015 from Universitas Gadjah Mada. He is a senior lecturer in the Department of Informatics Engineering at Khairun University. His research interest is in machine learning and artificial intelligence.

Aina Musdholifah received bachelor's and master's degrees in computer science from Universitas Gadjah Mada (UGM), Yogyakarta, Indonesia, in 2003 and 2006, respectively, and the Ph.D. degree in computer science from Universiti Teknologi Malaysia (UTM). She is an Assistant Professor and a member of the Intelligent System Laboratory, Department of Computer Science and Instrumentation, UGM, Yogyakarta. Her research interests include genetics algorithms, machine learning, fuzzy logic, and bioinformatics.

Wahyono received a Bachelor of Computer Science from the Universitas Gadjah Mada, Indonesia, in 2010 and a Doctor of Electrical Engineering from the University of Ulsan, South Korea, in 2017. Since 2022, he has been an Associate Professor with the Department of Computer Science and Electronics, Universitas Gadjah Mada, Indonesia. His research interests include intelligent surveillance systems, computer vision, image processing, pattern recognition, machine learning, and software engineering. Dr. Wahyono was a recipient of the IES Student Travel Award for the International Conference on Industrial Technology, funded by IEEE Industrial Electronic Society, 2017. He is currently registered as a member of IEEE, IEEE Industrial Electronics Society, IEEE System Man and Cybernetics Society, International Association of Engineers (IAENG), Indonesian Association in Pattern Recognition, and IndoCEISS.