Integrating Self-Organizing Map and Principal Component Analysis for Enhanced Herbal Leaf Image Classification

Imam Ahmad[1], *Yessi* Jusman[2*], *Yuri* Rahmanto[1], *Ferry* Putrawansyah[3], and *Rohmat Indra* Borman[1]

[1]Faculty of Engineering and Computer Science, Universitas Teknokrat Indonesia, Bandar Lampung, Indonesia
[2]Department of Electrical Engineering, Faculty of Engineering, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia
[3]Informatics Engineering Study Program, Institut Teknologi Pagar Alam, Pagar Alam, Indonesia

Abstract. This study presents a classification model for herbal leaf images by combining Principal Component Analysis (PCA) for dimensionality reduction with Self-Organizing Map (SOM) as an unsupervised classification algorithm. The model was developed using a dataset of five herbal leaf types: Betel, Papaya, Moringa, Katuk, and Turmeric. Shape-based morphological features were extracted from segmented leaf images, including area, perimeter, eccentricity, major axis length, and minor axis length. PCA was applied to reduce the five-dimensional feature vectors into two principal components, enhancing data representation and reducing redundancy. SOM was then used to classify the PCA-transformed data. The proposed model achieved an accuracy of 94.44%, outperforming the SOM-only configuration, which attained 85.56%. The improvement demonstrates that PCA effectively enhances SOM performance by providing more informative inputs. These results confirm the potential of integrating dimensionality reduction with unsupervised learning for accurate and efficient herbal plant classification.

1 Introduction

In recent years, the use of herbal plants has gained increasing attention as a natural alternative for healthcare and disease prevention. Indonesia, as one of the most biodiverse countries in the world, is home to over 30,000 plant species, of which approximately 9,600 are classified as having medicinal potential [1]. This rich biodiversity offers considerable opportunities for the development of traditional and modern medicine. However, accurately identifying herbal plant species remains a significant challenge due to the high degree of visual similarity among leaf shapes and textures. Manual identification methods require expert knowledge and are often prone to error, which limits their practicality for widespread use.

* Corresponding author: yjusman@umy.ac.id

Advancements in artificial intelligence (AI), particularly in image processing and classification, have opened new pathways for addressing this challenge. Machine learning-based classification has become a cornerstone in various domains such as autonomous driving, medical imaging, and precision agriculture. Neural networks, in particular, have demonstrated strong performance in image classification tasks. Deep learning architectures deliver high accuracy but require large annotated datasets and substantial computational resources, whereas shallow neural networks provide a more efficient alternative for applications with smaller datasets and limited processing capacity [2].

Several studies have explored supervised learning methods for herbal plant classification. For example, a Backpropagation Neural Network (BPNN) achieved an accuracy of 92% [3], Learning Vector Quantization (LVQ) reached 92.50% [4], and a Radial Basis Function (RBF) neural network obtained 92.08% [5]. Although these models demonstrated promising results, their dependency on labeled data and sensitivity to training configurations may restrict their flexibility in real-world settings [6]. To address these limitations, this study employs an unsupervised learning approach using the Self-Organizing Map (SOM) algorithm. SOM is capable of organizing high-dimensional data into a lower-dimensional space for visualization and classification, without requiring labeled input [7]. However, when applied to complex datasets with many correlated features, SOM may struggle with efficiency and representation quality.

To overcome this, Principal Component Analysis (PCA) is applied as a dimensionality reduction technique prior to SOM classification. PCA transforms the original feature set into a smaller number of uncorrelated principal components while preserving the majority of the data's variance [8]. This integration reduces data complexity, enhances computational efficiency, and improves the quality of clustering in the SOM phase. Although PCA–SOM integration has been explored in earlier works, its application to the herbal plant domain remains limited. Alternative dimensionality reduction techniques such as Multivariate Direction Scoring (MDS) have also been proposed [9], but PCA offers a simpler and computationally efficient solution particularly suitable for relatively small datasets such as herbal leaf images.

This study contributes by demonstrating how PCA–SOM integration can be effectively adapted for herbal leaf classification, a domain with significant agricultural and medicinal importance yet underexplored in AI-based classification research. The proposed lightweight and unsupervised framework is intended to support automated plant species recognition, especially in environments with limited labeled data and constrained computational resources. By combining effective feature reduction with unsupervised learning, this work provides a practical and scalable solution for smart agriculture and herbal medicine applications.

2 Methodology

This study adopts a sequential image processing and classification approach that integrates Principal Component Analysis (PCA) for dimensionality reduction and Self-Organizing Maps (SOM) for unsupervised classification of herbal leaf images. The overall workflow is illustrated in Fig. 1.

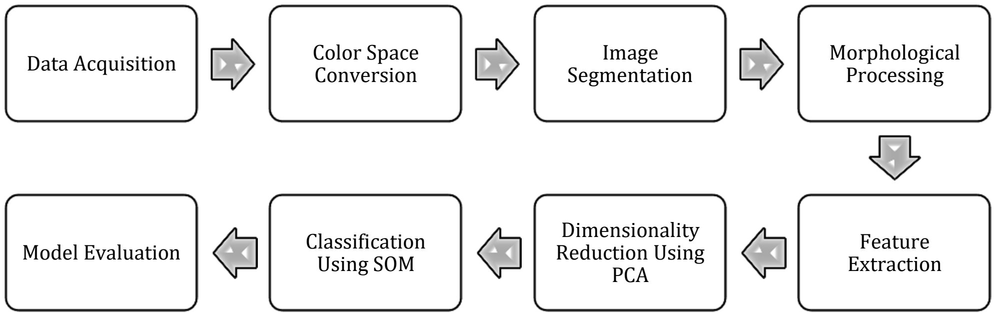


Fig. 1. Workflow of the Proposed Classification Framework.

2.1 Data Acquisition

The dataset used in this study consists of 600 herbal leaf images representing five plant species: *Betel*, *Papaya*, *Moringa*, *Katuk*, and *Turmeric*. These images were obtained from a publicly available source on the Kaggle platform. To support effective model training and evaluation, the dataset was divided based on a 70:30 ratio, a commonly adopted standard that balances the need for sufficient training data and a meaningful evaluation set. Accordingly, 420 images were allocated for training and the remaining 180 images for testing.

2.2 Color Space Conversion

To improve the accuracy of color-based image analysis, each image was converted from the RGB color space to the HSV (Hue, Saturation, Value) color space. Unlike RGB, which combines chromatic and luminance information into three interdependent channels, HSV separates these components, making it more aligned with human visual perception [10]. In this space, Hue represents the dominant color (e.g., red, green, blue), Saturation indicates the purity or intensity of the color, and Value reflects its brightness level. This separation facilitates more intuitive and effective color segmentation, allowing targeted enhancement of relevant features without altering other visual aspects.

2.3 Image Segmentation

Image segmentation is a fundamental process in image analysis, used to partition an image into distinct regions based on specific visual characteristics [11]. In this study, segmentation is applied to isolate the leaf object from the background, enabling further analysis of the target region. A thresholding technique is employed, wherein segmentation is guided by pixel intensity values [12]. This method is particularly effective when there is a strong contrast between the object and the background. The grayscale image is converted into a binary representation by applying a predefined threshold: pixels with intensity values above the threshold are assigned as foreground (typically white), while those below are assigned as background (typically black). This step ensures a clear separation between the herbal leaf and its surroundings, supporting reliable feature extraction in subsequent stages.

2.4 Morphological Processing

Following segmentation, morphological operations were applied to enhance the structural integrity of the extracted leaf regions. These operations manipulate object shapes within the binary image using small structuring elements to refine boundaries and eliminate noise [13]. In this study, three key morphological techniques were utilized: hole filling, opening, and closing. Hole filling was used to correct enclosed voids within the segmented object that may result from noise or imperfect segmentation. Opening, which consists of erosion followed by dilation, effectively removes small background artifacts while preserving the main structure of the leaf. Conversely, closing, composed of dilation followed by erosion, seals small internal gaps and connects adjacent components. Together, these operations improve the clarity and continuity of the leaf shapes, ensuring more reliable feature extraction in subsequent stages.

2.5 Feature Extraction

Following morphological processing, a set of shape-based features was extracted from each segmented leaf image to quantify its structural characteristics in numerical form. Morphological features are widely used in plant image analysis due to their effectiveness in capturing species-specific traits such as size, contour, and elongation patterns [14]. In this study, five morphological features were selected: area, perimeter, eccentricity, major axis length, and minor axis length. These features provide a concise yet informative description of the leaf's geometry and are suitable for subsequent computational processing. The output of this step is a structured dataset in which each image is represented as a five-dimensional feature vector, serving as the input for the dimensionality reduction phase.

2.6 Dimensionality Reduction Using PCA

To enhance computational efficiency and reduce feature redundancy, this study employs Principal Component Analysis (PCA) as a dimensionality reduction technique. PCA transforms the original correlated features into a set of uncorrelated principal components that capture the directions of maximum variance in the data [8]. This transformation allows the model to operate on a lower-dimensional space while preserving the most significant patterns [15].

The process begins with data standardization to ensure that all features contribute equally to the analysis. Each feature is normalized to have a mean of zero and a variance of unit using Equation 1.

$$x' = \frac{x - \bar{x}}{\sigma} \quad (1)$$

where x is the original feature value, $\bar{x}$ is the feature mean, and σ is the standard deviation.

The covariance matrix C is then calculated to evaluate the relationship between features using Equation (2).

$$C_{ij} = \frac{1}{n-1}\sum_{k=1}^{n}(x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j) \quad (2)$$

where n is the number of samples, and x_{ik}, x_{jk} are the values of features i and j for sample k.

$$Cv = \lambda v \quad (3)$$

where v is the eigenvector and λ is the corresponding eigenvalue.

Eigenvectors corresponding to the largest eigenvalues are selected as principal components, forming a reduced feature space. This transformation produces a compact representation of each image while preserving informative structure. By reducing feature dimensionality prior to SOM, PCA decreases noise, improves clustering quality, and ensures that SOM training focuses on the most discriminative attributes.

2.7 Classification Using Self-Organizing Map (SOM)

The Self-Organizing Map (SOM) is an unsupervised learning algorithm that maps multidimensional data into a simpler, structured form. Developed by Teuvo Kohonen in the 1980s, SOM mimics the cortical organization of the human brain, where neurons form spatial representations of input data [7]. It organizes high-dimensional data into a lower-dimensional space, usually a two-dimensional map, preserving the data's topology so that similar patterns or clusters remain close on the map [7].

In this study, SOM is applied to classify the feature vectors produced from the PCA-reduced dataset. The SOM architecture consists of neurons arranged in a two-dimensional grid, with each neuron associated with a weight vector that matches the dimensionality of the input data. During training, the network identifies the neuron whose weight vector is closest to the input vector. This neuron is referred to as the Best Matching Unit (BMU), and the proximity is determined using the Euclidean distance, as defined in Equation (4).

$$d(x, w_i) = \sqrt{\sum_{j=1}^{n} \left(x_j - w_{ij}\right)^2} \tag{4}$$

where $d(x, w_i)$ is the distance between the input vector x and the weight vector w_i of neuron i, and n is the number of dimensions of the input vector.

Once the BMU is identified, its weight vector and those of its neighboring neurons are updated to move closer to the input vector. The magnitude of the adjustment is influenced by the distance from the BMU and the learning rate, as described in Equation (5).

$$w_i(t+1) = w_i(t) + \theta(u, i, t)\alpha(t)\big(x(t) - w_i(t)\big) \tag{5}$$

where $w_i(t)$ is the weight of neuron i at iteration t, $\theta(u, i, t)$ is the neighborhood function that defines the influence of the BMU on neuron i, and $\alpha(t)$ is the learning rate, which gradually decreases over time.

This iterative process of competition, cooperation, and adaptation allows the map to self-organize. Neurons that respond to similar input patterns become spatially adjacent on the map, resulting in an organized representation that facilitates effective unsupervised classification of herbal leaf images.

2.8 Model Evaluation

The performance of the proposed classification model is evaluated using standard performance metrics derived from the confusion matrix, including accuracy, precision, and recall. The confusion matrix provides a comprehensive view of classification outcomes by comparing predicted class labels with actual ground truth. Accuracy indicates the proportion of correct predictions among all predictions made. Precision reflects the proportion of correctly identified positive instances among all predicted positives, while recall measures the ability of the model to detect all actual positive instances.

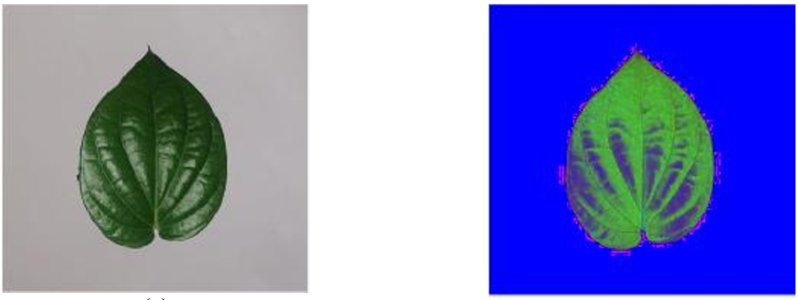
To assess the contribution of dimensionality reduction to the overall model performance, a comparative evaluation is conducted between two configurations: SOM without PCA and SOM with PCA. This comparison aims to determine the impact of Principal Component Analysis on both classification accuracy and computational efficiency. Evaluation is based on the aforementioned metrics as well as execution time, offering insights into how PCA enhances the robustness and scalability of the SOM model.

This comparative analysis is particularly relevant in smart agricultural applications, where reliable classification and efficient computation are critical for practical implementation in resource-constrained environments.

3 Results and Discussion

To develop a robust classification model for herbal leaf images using Principal Component Analysis (PCA) and Self-Organizing Map (SOM), a series of processing steps were implemented and evaluated. The dataset used in this study consisted of 600 images representing five types of leaves: *Betel*, *Papaya*, *Moringa*, *Katuk*, and *Turmeric*, obtained from the Kaggle platform. The data were divided into 70% for training (420 images) and 30% for testing (180 images), following standard machine learning practice to ensure sufficient data for model learning and performance evaluation.

The model construction involved several processing stages, beginning with the conversion of RGB images to the HSV color space. This transformation improves the representation of color and brightness, allowing for more effective image processing and segmentation. Fig. 2 illustrates the conversion from RGB to HSV color format.



(a) (b)

Fig. 2. (a) RGB Image of Herbal Leaf; (b) HSV Image for Enhanced Segmentation.

The result of the color space conversion is shown in Fig. 2, where the original RGB image is transformed into its HSV representation. This transformation improves the contrast between the leaf and the background, facilitating more effective segmentation. The next step was image segmentation using a global thresholding technique, which converts grayscale images into binary form by applying an intensity threshold. Pixels above the threshold are considered foreground (white), while those below are treated as background (black). The results of the segmentation process are illustrated in Fig. 3.

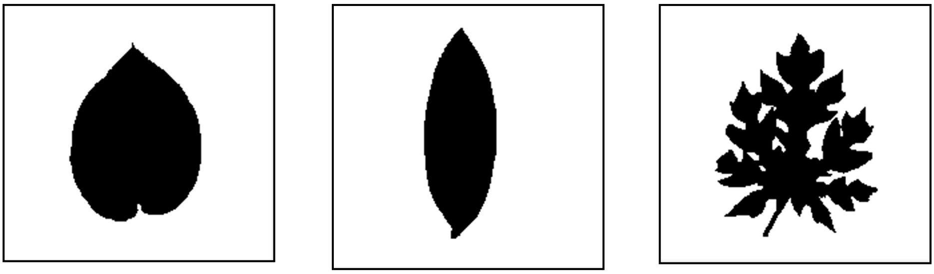


Fig. 3. Sample Images Segmented Using Thresholding.

Fig. 3 shows examples of binary segmentation results, where the leaf objects are clearly separated from the background. The foreground (leaf region) appears in white, while the background is rendered in black, indicating a successful thresholding process for isolating the leaf structure. To improve the quality of the segmented images, morphological operations were applied, including hole filling, opening, and closing. These operations refine the object's shape by removing noise and connecting fragmented regions, enhancing the clarity of the leaf structure. Fig. 4 shows the visual difference before and after the morphological enhancements.

(a) (b)

Fig. 4. (a) Segmented Image Before Morphological Enhancement; (b) After Morphological.

Fig. 4 shows the visual differences before and after morphological enhancement. The image in Fig. 4(a) still contains small gaps and irregularities, while Fig. 4(b) demonstrates a cleaner and more defined leaf shape after applying hole filling, opening, and closing operations. These improvements result in more accurate object boundaries, which are essential for the next stage of feature extraction.

Once the images were refined, shape-based features were extracted to numerically represent the structure of each leaf. The selected features included area, perimeter, eccentricity, major axis length, and minor axis length. Table 1 provides an example of the feature extraction results for a segmented image.

Table 1. Morphological Feature Extraction Results of a Segmented Leaf Image.

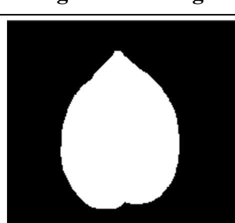
Segmented Image	Parameters	Value
	Area	38142
	Perimeter	721.8440
	Eccentricity	0.6046
	Major Axis Length	248.2113
	Minor Axis Length	197.7127

Table 1 shows the numerical results of morphological feature extraction from a single segmented leaf image. These features quantitatively describe the shape and structure of the object, serving as the foundation for the classification process. Each image in the dataset was represented as a five-dimensional feature vector consisting of area, perimeter, eccentricity, major axis length, and minor axis length.

Before classification, dimensionality reduction was performed using PCA to simplify data representation and remove feature redundancy. PCA transformed the original five-dimensional feature vectors into two principal components (PC1 and PC2), preserving the most significant variance in the data. The resulting distribution of image samples in the transformed space is presented in Fig. 5.

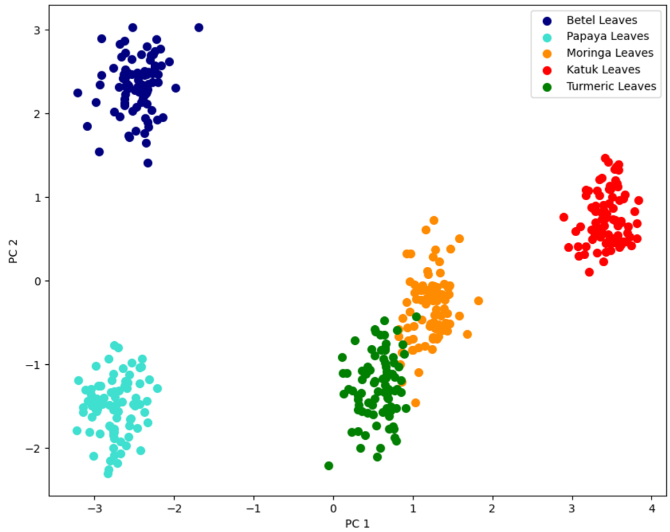


Fig. 5. Distribution of Data for Each Class in the PCA-Transformed Space.

Figure 5 visualizes the distribution of image data after dimensionality reduction using PCA. Each data point represents a sample projected onto two principal components, PC1 and PC2. The separation among the clusters indicates that PCA successfully preserved class-distinguishing characteristics within a lower-dimensional space, providing a compact yet informative input for the classification process.

Following PCA, classification was conducted using SOM. This algorithm projected the PCA-transformed data into a two-dimensional grid while preserving topological relationships. The weight adaptation during training resulted in organized neuron clusters that reflect the input data structure. Fig. 6 presents the neuron weight map generated by the SOM.

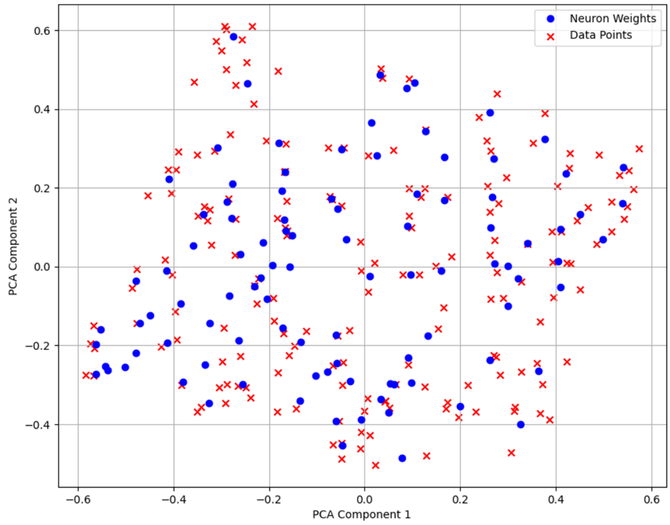


Fig. 6. SOM Neuron Weight Map Showing Clustering of PCA-Reduced Data Points.

Fig. 6 presents the neuron weight map generated by the Self-Organizing Map (SOM) after training on the PCA-reduced dataset. In the figure, each red 'x' represents a data point, while the blue dots indicate the positions of neurons in the SOM grid. The spatial proximity between data points and nearby neurons suggests that the SOM successfully captured the underlying structure of the dataset. Neurons that are closely surrounded by similar data points reflect effective clustering, indicating that the model has learned meaningful patterns from the input data.

The model's classification output was evaluated using a confusion matrix, which provides a detailed breakdown of the predicted and actual class distributions. The confusion matrix for the SOM + PCA model is shown in Fig. 7.

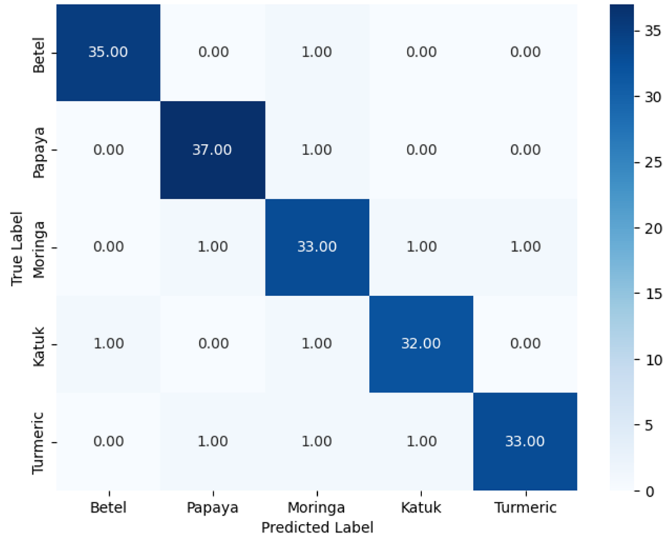


Fig. 7. Confusion matrix of the SOM + PCA model.

Fig. 7 shows that most predictions closely align with the ground truth labels, indicating effective classification performance across classes. *Healthy Leaf* and *Anthracnose* show the highest number of correctly classified instances (30 and 27, respectively). However, several misclassifications are observed between classes with visually similar features, such as *Cercospora Leaf Spot* and *Anthracnose*. This reflects the inherent challenge of shape-based classification, particularly when disease symptoms exhibit similar morphological appearances.

Evaluation metrics including precision, recall, and accuracy were derived from the confusion matrix to provide a comprehensive assessment of model performance. To further examine the impact of dimensionality reduction, a comparative evaluation was conducted between two configurations: SOM without PCA and SOM with PCA. Table 2 summarizes the precision, recall, and overall accuracy for each class under both models.

Table 2. Comparison of SOM and SOM + PCA performances.

Class Name	Model SOM			Model SOM + PCA		
	Precision	Recall	Accuracy	Precision	Recall	Accuracy
Betel Leaves	94.29	89.19	85.56	97.22	97.22	94.44
Papaya Leaves	87.50	89.74		94.87	97.37	
Moringa Leaves	81.82	77.14		89.19	91.67	
Katuk Leaves	82.35	80.00		94.12	94.12	
Turmeric Leaves	81.58	91.18		97.06	91.67	

The results in Table 2 demonstrate that the SOM with PCA configuration consistently outperformed the SOM-only model. The overall accuracy improved from 85.56% to 94.44%, indicating an increase of 8.88%. The most significant improvements were observed in *Moringa Leaves* and *Katuk Leaves*, suggesting that PCA helped the classifier better distinguish between classes with overlapping morphological traits.

When compared with previous studies, the proposed model achieved superior results. Backpropagation Neural Networks reached 92.00%, Learning Vector Quantization (LVQ) achieved 92.50% [9], and Radial Basis Function (RBF) networks recorded 92.08% [10]. In contrast, our model achieved 94.44%, demonstrating that the integration of dimensionality reduction with an unsupervised neural network can yield performance improvements beyond supervised models. This comparative analysis contextualizes our contribution within the literature and highlights the value of combining PCA with SOM for herbal leaf classification.

The improved performance can be attributed to the synergy between PCA and SOM. PCA reduces redundancy and retains informative features, enabling SOM to converge faster and form clearer clusters. SOM's topology-preserving property further supports visual interpretation, strengthening its role as an interpretable unsupervised classifier. Compared to more advanced dimensionality reduction methods [9], PCA was chosen for its simplicity and efficiency, which makes it particularly suitable for small herbal datasets.

Despite its promising performance, the model has several limitations. Misclassifications may still occur due to random initialization of neuron weights, the visual similarity among certain leaf types, or limited diversity in extracted features. Additionally, the relatively small dataset may restrict generalization to unseen data. Future work may explore improved weight initialization strategies, expand the dataset with more diverse samples, and incorporate additional features such as color and texture descriptors.

Practical implications of this research are significant. The proposed model can be integrated into smart agriculture systems for automated plant recognition, reducing reliance on expert knowledge and minimizing human error. In the healthcare and pharmaceutical sectors, accurate identification of herbal leaves can assist in ensuring quality control and authenticity in herbal medicine production. Thus, the framework contributes not only to methodological advancements but also to real-world applications in agriculture and healthcare.

4 Conclusion

This study presented a classification model for herbal leaf images by integrating Principal Component Analysis (PCA) for dimensionality reduction with a Self-Organizing Map (SOM) for unsupervised classification. Using a dataset of five herbal leaf types with shape-based features, the proposed SOM with PCA achieved an accuracy of 94.44%, representing an improvement of 8.88% over the SOM-only model, which reached 85.56%. Furthermore, the proposed approach outperformed several supervised benchmarks reported in previous studies, including Backpropagation Neural Networks with 92.00% accuracy, Learning Vector Quantization with 92.50%, and Radial Basis Function networks with 92.08%. The performance gain is primarily due to PCA's ability to reduce redundancy and highlight the most informative components, thereby enhancing SOM's clustering effectiveness and interpretability. These findings confirm that dimensionality reduction combined with unsupervised learning offers a lightweight yet effective framework for plant morphology classification, with practical implications for smart agriculture and herbal medicine, particularly in supporting species authenticity, quality control, and decision-making. Despite these promising results, limitations remain in the form of occasional misclassifications caused by morphological similarities and the relatively small dataset size, which may restrict generalization. Future studies should therefore focus on expanding the dataset with more diverse samples, incorporating complementary features such as color and texture, and exploring hybrid frameworks that integrate SOM with supervised techniques to further improve accuracy and robustness.

References

1. J. Magos and B. Nigel, Setting the priority medicinal plants for conservation in Indonesia. Genet. Resour. Crop Evol. **68**, 2019 (2021). https://doi.org/https://doi.org/10.1007/s10722-021-01115-6
2. M. A. Bansal, D. R. Sharma, and D. M. Kathuria, A Systematic Review on Data Scarcity Problem in Deep Learning: Solution and Applications. ACM Comput. Surv. **54**, 1 (2022). https://doi.org/10.1145/3502287
3. Elbasi, E., Al-Bahadili, H., & Al-Rawashdeh, A., "Enhanced Plant Leaf Classification over a Large Number of Classes Using Machine Learning," *Applied Sciences*, vol. 14, art. 10507, 2024. https://doi.org/10.3390/app142210507
4. Andrian Sah, Mulyadi Mulyadi, Allan Desi Alexander & Adam M. Tanniewa, "Pengembangan Model Klasifikasi Citra Penyakit Daun Lada Menggunakan Jaringan Syaraf Tiruan Learning Vector Quantization (LVQ)," *Jurnal Ilmiah Informatika dan Ilmu Komputer (JIMA-ILKOM)*, vol. 4, no. 1, pp. 34–44, 2025. https://doi.org/10.58602/jima-ilkom.v4i1.53
5. R. I. Borman, F. Rossi, D. Alamsyah, R. Nuraini, Y. Jusman, and F. Fauziyah, Classification of Medicinal Wild Plants Using Radial Basis Function Neural Network with Least Mean Square, in International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS), (IEEE, 2022), pp. 141–146, (2022)

   I. H. Sarker, AI-Based Modeling: Techniques, Applications and Research Issues Towards Automation, Intelligent and Smart Systems. SN Comput. Sci. **3**, 158 (2022). https://doi.org/10.1007/s42979-022-01043-x
6. I. Ahmad, Y. Rahmanto, R. I. Borman, F. Rossi, Y. Jusman, and A. D. Alexander, Identification of Pineapple Disease Based on Image Using Neural Network Self-Organizing Map (SOM) Model, in International Conference on Electronic and

Electrical Engineering and Intelligent System (ICE3IS), (IEEE, 2022), pp. 12–17, (2022)

7. Y. Yuan and X. Chen, Vegetable and fruit freshness detection based on deep features and principal component analysis. Curr. Res. Food Sci. **8**, 100656 (2024)
8. G. Biagetti, P. Crippa, L. Falaschetti, S. Orcioni, and C. Turchetti, Multivariate Direction Scoring for Dimensionality Reduction in Classification Problems, in Intelligent Decision Technologies, edited by I. Czarnowski, A. M. Caballero, R. J. Howlett, and L. C. Jain, (Springer International Publishing, Cham, 2016), pp. 413–423, (2016)
9. R. Napianto, A. Tri Prastowo, A. Deni Wahyudi, Wajiran, Y. E. P. Yudoutomo, and A. A. Suparto, Identification of Beef Freshness Using Linear Discriminant Analysis (LDA) Algorithm and Extraction of HSV Values, in International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS), (2022), pp. 114–119, (2022)
10. A. Mulyanto, W. Jatmiko, P. Mursanto, P. Prasetyawan, and R. I. Borman, A New Indonesian Traffic Obstacle Dataset and Performance Evaluation of YOLOv4 for ADAS. J. ICT Res. Appl. **14**, 286 (2021). https://doi.org/10.5614/itbj.ict.res.appl.2021.14.3.6
11. Ganbayar Batchuluun, Seung G. Kim, Jung S. Kim, Tahir Mahmood & Kang R. Park, "Artificial Intelligence-Based Segmentation and Classification of Plant Images with Missing Parts and Fractal Dimension Estimation," *Fractal and Fractional*, 8(11): 633, 2024. https://doi.org/10.3390/fractalfract8110633
12. A. Das, K. G. Dhal, S. Ray, and J. Gálvez, Histogram-based fast and robust image clustering using stochastic fractal search and morphological reconstruction. Neural Comput. Appl. **34**, 4531 (2022). https://doi.org/10.1007/s00521-021-06610-6
13. Burger W., Burge M. J. Digital Image Processing: An Algorithmic Introduction. 3rd Edition. Springer, 2022. https://doi.org/10.1007/978-3-031-05744-1
14. A. Hasan, M. Taufik, L. O. S. Bande, A. Khaeruni, R. Mallarangeng, Gusnawaty H. S., Asniah, Syair & A. Rahman, "Morphometric Analysis of Chili Leaves with Yellow Curly Symptom Using Digital Image Processing and Data Mining Algorithms," *Jurnal Fitopatologi Indonesia*, vol. 19, no. 6, pp. 231–237, 2023. https://doi.org/10.14692/jfi.19.6.231-237