VOL. 10. NO. 3 FEBRUARY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480 /jitk.v10i3.6013

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

COMBINATION OF LEARNING VECTOR QUANTIZATION AND LINEAR DISCRIMINANT ANALYSIS FOR TEA LEAF DISEASE CLASSIFICATION

Mutasar[1*]; Chaeroen Niesa[2]

Informatics Study Program[1, 2]
Universitas Islam Kebangsaan Indonesia, Bireuen, Aceh, Indonesia[1, 2]
https://uniki.ac.id[1, 2]
mutasarstmik@gmail.com[1*], jeumalaniesa@gmail.com[2]

(*) Corresponding Author
(Responsible for the Quality of Paper Content)

Abstract*—Tea farming, one of the key pillars of Indonesia's economy, faces productivity challenges due to diseases affecting tea leaves. Manual identification of tea leaf diseases requires significant time and cost, making an automated solution necessary. This research develops an innovative model for classifying tea leaf diseases by synergizing Learning Vector Quantization (LVQ) and Linear Discriminant Analysis (LDA). By leveraging LVQ's prototype-based classification and LDA's dimensionality reduction, the model ensures accurate and efficient disease identification. During preprocessing, tea leaf images were converted to the CIELAB color space to enhance segmentation using Otsu's Thresholding. Features such as Mean Color and texture attributes based on Gray Level Co-occurrence Matrix (GLCM) were extracted, reduced via LDA, and classified using LVQ. Tested on five tea leaf disease classes, the model achieved 94.1% accuracy. This performance underscores its potential to significantly assist farmers in early detection and management of tea leaf diseases, while also providing researchers with a robust tool for advancing agricultural technology.*

Keywords*: GLCM texture features, linear discriminant analysis, learning vector quantization, mean color, tea leaf disease classification.*

Intisari— *Pertanian teh, salah satu pendukung utama ekonomi Indonesia, menghadapi tantangan produktivitas akibat penyakit yang menyerang daun teh. Identifikasi penyakit daun teh secara manual memerlukan waktu dan biaya yang signifikan, sehingga diperlukan solusi otomatis. Penelitian ini mengembangkan model inovatif untuk mengklasifikasikan penyakit daun teh melalui sinergi antara Learning Vector Quantization (LVQ) dan Linear Discriminant Analysis (LDA). Dengan memanfaatkan klasifikasi berbasis prototipe LVQ dan reduksi dimensi oleh LDA, model ini memastikan identifikasi penyakit yang akurat dan efisien. Pada tahap pra-pemrosesan, citra daun teh dikonversi ke ruang warna CIELAB untuk meningkatkan segmentasi menggunakan Otsu's Thresholding. Fitur seperti Mean Color dan atribut tekstur berdasarkan Gray Level Co-occurrence Matrix (GLCM) diekstraksi, direduksi melalui LDA, dan diklasifikasikan menggunakan LVQ. Model ini diuji pada lima kelas penyakit daun teh dan mencapai akurasi sebesar 94,1%. Kinerja ini menunjukkan potensinya untuk membantu petani dalam deteksi dini dan pengelolaan penyakit daun teh, serta memberikan peneliti alat yang kuat untuk memajukan teknologi pertanian.*

Kata Kunci*: fitur tekstur GLCM, linear discriminant analysis, learning vector quantization, mean color, klasifikasi penyakit daun teh.*

INTRODUCTION

Tea farming plays a vital role in Indonesia's economy, particularly as a leading export product and for domestic consumption. According to data from the Central Bureau of Statistics (BPS), Indonesia's tea production in 2022 reached over 130,000 tons, with more than 50% of it exported to various countries [1]. The main tea-producing regions in Indonesia include West Java, North

VOL. 10. NO. 3 FEBRUARY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i3.6013

Sumatra, and Central Java. However, one of the major challenges faced by tea farmers in Indonesia is the spread of diseases affecting tea plants, particularly the leaves, which are a critical part of production [2]. Various tea leaf diseases can reduce the quality of the leaves and ultimately affect tea production [3]. If not promptly identified and treated, these diseases can cause significant harm to the plants. Tea leaf disease identification is primarily performed manually by specialists, a process that is both time-intensive and expensive. Therefore, a more efficient and automated solution for identifying tea leaf diseases is needed to improve agricultural productivity and minimize production losses.

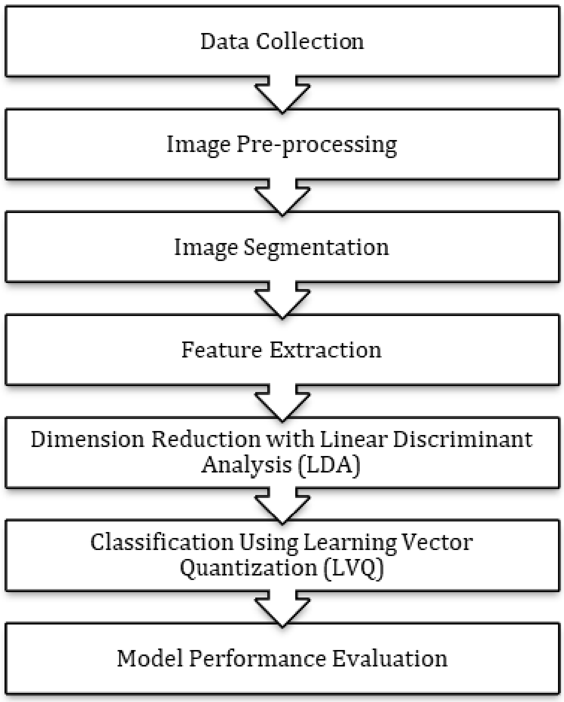
A study using Support Vector Machine (SVM) for mango leaf disease classification reported 80% accuracy [4], but SVM's effectiveness heavily depends on the selection of hyperparameters, such as the kernel function, which can be challenging to optimize. Furthermore, SVM struggles with imbalanced datasets, where one class significantly outnumbers others, potentially leading to biased predictions. Research using the Extreme Learning Machine (ELM) neural network identified tomato leaf diseases with 84.667% accuracy [5], though its reliance on randomly assigned input weights can lead to inconsistent training outcomes. This lack of control over weight initialization makes the model sensitive to variations in the dataset and prone to underfitting or overfitting, especially in scenarios with limited training data. The Backpropagation neural network also identified rice leaf spot disease with 85.8% accuracy [6], but it often encounters convergence issues, which can result in prolonged training times or the risk of getting trapped in local minima. This drawback is particularly problematic for high-dimensional data, where the optimization process becomes more complex. Additionally, Backpropagation requires careful tuning of learning rates to ensure stability and avoid overshooting during training.

Previous studies show that neural networks excel at learning and recognizing complex patterns from training data. However, given the complexity of leaf images and the constraints of limited labeled data, a method is needed that not only maps inputs for each class effectively but also provides interpretability for classification decisions. Learning Vector Quantization (LVQ) addresses these needs through its prototype-based classification approach. By representing each class with a prototype vector, LVQ offers a clear and interpretable way to understand how classification decisions are made [7]. This is particularly advantageous in plant disease classification, where transparent decision-making can guide agricultural interventions. Furthermore, LVQ’s computational efficiency and adaptability enable it to learn classification patterns quickly, making it highly suitable for multi-class problems like tea leaf disease classification [8]. Despite its strengths, one of the challenges in implementing LVQ is managing the high dimensionality of feature data, which can lead to overfitting and slower model performance [9]. To overcome this, Linear Discriminant Analysis (LDA) is integrated into the model for dimensionality reduction. LDA enhances the model by maximizing the separation between classes while minimizing variance within each class, resulting in more discriminative features [10], [11]. This ensures that the LVQ model processes only the most relevant features, reducing noise and computational burden. By combining LDA’s feature optimization capabilities with LVQ’s prototype-based classification, this study achieves a robust and efficient model for identifying tea leaf diseases.

The objective of this research is to develop a model for classifying tea leaf diseases by integrating Learning Vector Quantization (LVQ) and Linear Discriminant Analysis (LDA). In the preprocessing stage, tea leaf images are converted to the CIELAB color space to separate color and intensity components, enhancing subsequent analysis. Image segmentation is then performed using Otsu’s Thresholding method, which automatically isolates the infected leaf area from the background based on pixel intensity distribution. The next step involves feature extraction, focusing on two primary types of features: color features and texture features. Color features, such as Mean Color, provide insights into the intensity distribution of the infected area, while texture features are extracted using the Gray Level Co-occurrence Matrix (GLCM) to analyze texture patterns resulting from pixel intensity variations. These features are subsequently processed using LDA to reduce dimensionality, ensuring that only the most relevant data are used in the classification process by LVQ. The main contribution of this study is the development of a novel methodology for classifying tea leaf diseases through the integration of LVQ and LDA, an approach that remains underexplored in the agricultural sector.

MATERIALS AND METHODS

The research methodology serves as a critical systematic guide that directs the researcher throughout the process, ensuring a structured and logical approach [12]. This study aims to develop an image-based classification model for detecting tea leaf diseases, with the research steps outlined in Figure 1.



Source : (Research Results, 2024)
Figure 1. Research Steps Applied

Figure 1 presents a visual representation of the research methodology implemented. Below is a comprehensive description of each step, providing detailed explanations of the processes and actions taken in this research.

Data Collection

The initial stage of this research involves the collection of a dataset of tea leaf images, which includes samples of healthy leaves and leaves infected by various diseases. The dataset used is “Tea Leaf Disease Classification,” a publicly available dataset that can be accessed on Kaggle (https://www.kaggle.com/datasets/mamun009/tea-leaf-disease-clasification) [13]. This study focuses on five classes of tea leaf conditions: healthy leaves, algal spot, brown blight, gray blight, and red spot. The dataset consists of 5000 images, each with a resolution of 256 × 256 pixels, ensuring consistent quality and adequate detail for analysis. The images were captured under controlled lighting conditions to minimize noise and shadow artifacts. For model development, the dataset was split into training and testing sets with an 80:20 ratio, resulting in 4000 training images (800 per class) and 1000 testing images (200 per class). This balanced distribution ensures sufficient representation for each class during both the training and evaluation phases.

Image Pre-processing

After acquiring the tea leaf image dataset, the next step is image pre-processing. The primary goal of pre-processing is to enhance the quality of the images and prepare them for segmentation and feature extraction [14]. One of the key steps in this process is converting the images from the RGB color space to the CIELAB color space (commonly referred to as LAB). This transformation is critical because the CIELAB color space is designed to approximate human vision, making it more suitable for analyzing color-based patterns in images [15]. Unlike RGB, which is highly sensitive to lighting variations, CIELAB separates the image into three independent components: L (luminance or lightness), a (green to red), and b (blue to yellow). This separation ensures that color analysis is stable across varying lighting conditions, allowing algorithms to detect subtle color differences in infected leaves more effectively. For example, discoloration caused by diseases is more apparent in the a and b channels, enabling precise differentiation between healthy and infected areas.

Image Segmentation

At this stage, the pre-processed images are segmented to separate the main object from the surrounding area. Segmentation removes unnecessary information, such as background noise or irrelevant elements, aiding the classification process [16]. Otsu's Thresholding is used to automatically divide the image into the foreground and background. This technique was chosen for its ability to find the optimal segmentation threshold by analyzing pixel intensity distribution, enabling clear separation of image elements [17]. By using Otsu's Thresholding, the best threshold value is found to distinguish the leaf from its background, optimizing the image for feature extraction.

Feature Extraction

The objective of this feature extraction step is to gather pertinent data from the visual content, which will subsequently facilitate the classification task [18]. This study uses two types of features: color and texture. The Mean Color method is particularly suited for analyzing tea leaf diseases because discoloration is one of the primary symptoms of many leaf diseases. By calculating the average intensity of the L*, a*, and b* components in the CIELAB color space, Mean Color effectively captures subtle variations in hue and brightness that indicate disease progression or severity. For texture features, the Gray Level Co-occurrence Matrix (GLCM) is employed due to its ability to analyze spatial relationships between pixel pairs

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 10. NO. 3 FEBRUARY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i3.6013

with specific intensity levels. Texture plays a significant role in identifying disease-related patterns, such as lesions, spots, or irregular surface textures caused by infections. GLCM is ideal for this task as it provides quantitative metrics, including contrast, dissimilarity, homogeneity, energy, and correlation, which describe the distribution and organization of intensity variations in the leaf surface [19]. These features are particularly effective in differentiating between healthy and diseased areas, where diseases often create distinct textural anomalies. The combination of color and texture features ensures a comprehensive representation of the visual cues necessary for accurate disease classification.

Dimension Reduction with Linear Discriminant Analysis (LDA)

After extracting color and texture features, Linear Discriminant Analysis (LDA) is applied for dimensionality reduction. LDA reduces the number of features while retaining the most important ones for classification [20]. It enhances class separation by maximizing inter-class distance and minimizing within-class variance, leading to more discriminative features [21]. LDA identifies projections that optimize class separation and reduces within-class variance, assessed using the within-class scatter matrix, as shown in Equation (1).

$$S_W = \sum_{k=1}^{K} \sum_{x_i \in C_k} (x_i - \mu_k)(x_i - \mu_k)^T \quad (1)$$

where S_W denotes the within-class scatter matrix, K is the number of classes, C_k refers to the k-th class, x_i represents the feature vector of the i-th sample in class C_k, and μ_k is the mean feature vector for class C_k.

Next, the between-class scatter matrix is calculated to measure the variance between the means of each class relative to the overall mean. Equation (2) is used to calculate the between-class scatter matrix.

$$S_B = \sum_{k=1}^{K} N_k (\mu_k - \mu)(\mu_k - \mu)^T \quad (2)$$

where S_b refers to the between-class scatter matrix, N_k represents the number of samples in the k-th class, μ_k denotes the mean feature vector for class C_k, and μ indicates the global mean feature vector.

LDA optimizes the ratio of between-class scatter to within-class scatter. This ratio is expressed in Equation (3).

$$J(w) = {w^T S_B w} \big/ {w^T S_W w} \quad (3)$$

where $J(w)$ represents the scatter ratio to be maximized, w indicates the projection vector (LDA component) to be found, S_B refers to the between-class scatter matrix, and S_W denotes the within-class scatter matrix

LDA maximizes $J(w)$ by finding the projection vector w that optimizes class separation relative to within-class variance. This ensures that the data in the lower-dimensional space achieves optimal class separation, selecting the most relevant features for classification.

Classification Using Learning Vector Quantization (LVQ)

The features reduced by LDA are used to train the Learning Vector Quantization (LVQ) model, an artificial neural network that compares input features with learned prototype vectors [22]. LVQ's goal is to learn a prototype vector for each class to classify new data [23]. The core of LVQ is updating prototypes based on training data. To find the closest prototype to input vector x, LVQ calculates the Euclidean distance between x and each prototype w_j, as shown in Equation (4).

$$d(x, w_j) = \sqrt{\sum_{i=1}^{n} (x_i - w_{ji})^2} \quad (4)$$

where $d(x, w_j)$ is the Euclidean distance between the input vector x and the prototype w_j, x_i refers to the i-th component of the input vector x, w_{ji} is the i-th component of the prototype w_j, while n is the number of features in the vector x.

The prototype is then updated based on whether the classification is correct or incorrect, moving closer to correct data and farther from incorrect data to improve accuracy. The prototype vector gradually adapts to better represent each class's data pattern. The update for correct classifications is shown in Equation (5), and for incorrect classifications in Equation (6).

$$w_j^{new} = w_j^{old} + \alpha(x - w_j^{old}) \quad (5)$$

$$w_j^{new} = w_j^{old} - \alpha(x - w_j^{old}) \quad (6)$$

LVQ uses the Euclidean distance to select the nearest prototype and updates the prototype based on whether the classification was correct or incorrect.

Model Performance Evaluation

Once the LVQ model has been trained, the final step involves evaluating its performance using metrics such as accuracy, precision, recall, and the F1-score. Accuracy provides an overall measure of the model's ability to correctly classify tea leaf diseases, making it a straightforward indicator of performance [24]. Precision, which measures the proportion of correctly classified positive instances out of all predicted positives, is particularly relevant in this context to minimize false positives that could lead to unnecessary interventions [25]. Recall, or sensitivity, assesses the model's ability to identify all true positive instances, which is critical for ensuring that no diseased leaves are overlooked [26]. The F1-score, as the harmonic mean of precision and recall, is especially important when dealing with imbalanced datasets, as it balances the trade-off between false positives and false negatives [26]. Together, these metrics provide a comprehensive evaluation of the model's performance, ensuring its reliability in accurately identifying and classifying tea leaf diseases under real-world conditions.

RESULTS AND DISCUSSION

To develop a tea leaf disease classification model integrating Learning Vector Quantization (LVQ) and Linear Discriminant Analysis (LDA), the dataset must first be prepared for training and evaluation. This research uses the public "Tea Leaf Disease Classification" dataset from Kaggle , consisting of 5000 images, divided in an 80:20 ratio into 4000 training images (800 per class) and 1000 testing images (200 per class). This ensures balanced representation for each class during model development and validation.

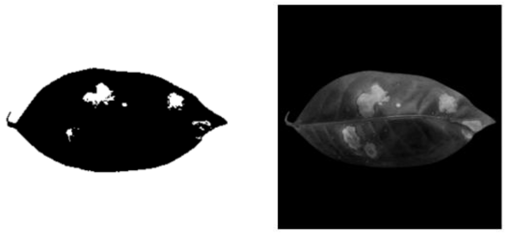
Training and testing are crucial for helping the model recognize unique patterns in different tea leaf types and assess its classification accuracy on unseen data.

(a) (b)

Source : (Research Results, 2024)
Figure 2. (a) Original Resulting Image and (b) CIELAB Image Transformation

The first step is pre-processing, where RGB images are converted to CIELAB due to its superior representation of human visual perception. This color space separates color from intensity, improving the algorithm's ability to extract key information and perform image segmentation, particularly for leaf-based image classification. Figure 2 shows an example of an RGB to CIELAB conversion.

Figure 2 shows the conversion of the original image to CIELAB, effectively separating color from intensity. The next step, segmentation, removes irrelevant elements like background noise and shadows, preparing the image for classification. Otsu's Thresholding was used to automatically split the image into foreground (leaf) and background by finding the optimal threshold that maximizes between-class variance. Figure 3 presents two outputs: the segmentation results and the segmentation results while retaining CIELAB characteristics.



(a) (b)

Source : (Research Results, 2024)
Figure 2. (a) Segmentation Results, and (b) Segmentation Results Preserving CIELAB Images

Figure 3 demonstrates the effectiveness of the segmentation process in clearly separating the main object from its background. After segmentation, the next step is feature extraction from the processed image, focusing on color and texture features. To analyze the color features of tea leaves, the Mean Color method was implemented. This technique essentially calculates the average color intensity for each channel in the color space used. In the CIELAB color space, Mean Color is calculated by averaging the intensity values of the L*, a*, and b* components. The specific results from the color feature extraction, which include values for each parameter, are comprehensively presented in Table 2.

Table 2. Example of Feature Extraction Results with Mean Color

Segmented Image	Color Feature Extraction	
	Color Channel	Mean Value
	L*	176.3553
	a*	135.3462
	b*	115.4992

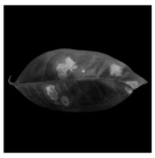
Source : (Research Results, 2024)

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 10. NO. 3 FEBRUARY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i3.6013

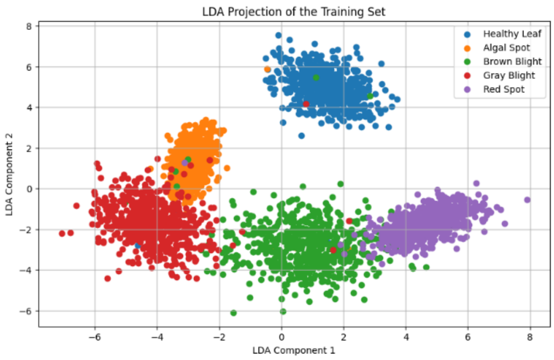
Furthermore, to obtain texture characteristics, the Gray Level Co-occurrence Matrix (GLCM) technique is used. GLCM analyzes the frequency of pixel pair occurrences with specific intensity values at a defined distance and orientation. Texture features are obtained based on parameters such as contrast, dissimilarity, homogeneity, energy, and correlation. The texture feature extraction results, summarized in Table 3, offer detailed insights into the texture characteristics of the analyzed tea leaf images.

Table 3. Example of Feature Extraction Results with GLCM

Segmented Image	Texture Feature Extraction Parameters	Value
	Contrast	86.0714
	Dissimilarity	0.9914
	Homogeneity	0.6521
	Energy	8.9166
	Correlation	0.0013

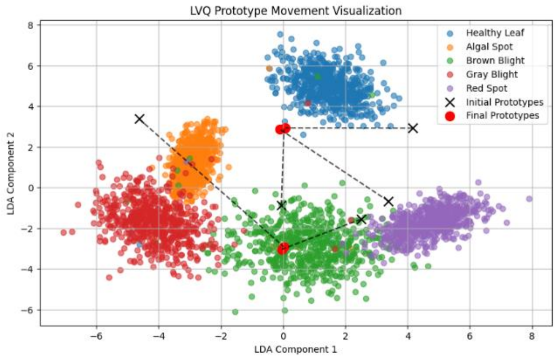
Source : (Research Results, 2024)

The features in Tables 2 and 3 are critical inputs for classification. Linear Discriminant Analysis (LDA) reduces dimensionality by retaining significant features, enhancing class separation, and minimizing within-class variability. Effective class separation is achieved by centering data around distinct centroids, as visualized in Figure 4 using 800 samples per class.



Source : (Research Results, 2024)
Figure 4. Plotting Data Reduction Using LDA

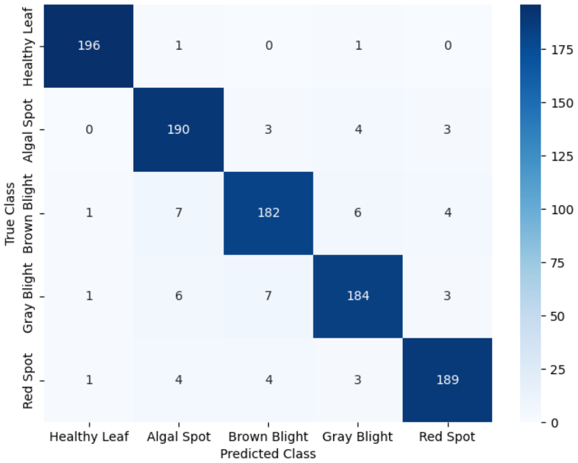
In the plot, samples from different classes are grouped, demonstrating LDA's success in distinguishing them. After LDA reduces the data, Learning Vector Quantization (LVQ) uses the most relevant features for learning and classification. LVQ maps each reduced feature vector to a prototype representing a class, initially assigned randomly or from the training data. It then compares each input sample with the prototype using Euclidean distance to find the closest match. The prototype is updated based on the classification's correctness, either moving closer to correct samples or further away from incorrect ones. This iterative process continues until the prototype effectively represents each class. Figure 5 visualizes the data distribution and prototype movement.



Source : (Research Results, 2024)
Figure 5. LVQ Prototype Movement

The dashed lines between the black and red dots indicate the trajectory of the prototype movement from the initial position to the final position during the training process. The prototype moves closer to the correct class data clusters during the iterations, as the updates occur each time new data is learned by the algorithm. This movement illustrates how LVQ optimizes the prototype position to better represent the data patterns of each tea leaf disease class.

After the training of the Learning Vector Quantization (LVQ) model is completed, the next crucial stage is the model performance evaluation. This evaluation process uses a series of metrics, including accuracy, precision, recall, and F1-score, to measure how effectively the model can classify various tea leaf diseases. To ensure objectivity and the model's generalization ability, the evaluation is conducted using a test dataset separate from the training data. This test dataset consists of 1000 images, with an even distribution of 200 images for each of the five tea leaf disease classes. The first step in the evaluation process is the formation of a confusion matrix, visualized in Figure 6.



Source : (Research Results, 2024)
Figure 6. Confusion Matrix Results

Figure 6 presents the confusion matrix, offering a detailed overview of the model's classification ability by showing the distribution of appropriate and inappropriate predictions for each class. Based on this information, various evaluation metrics, such as accuracy, precision, recall, and F1-score, can be calculated, offering in-depth insights into the model's strengths and weaknesses in classifying each type of tea leaf disease. The evaluation outcomes of the performance of the developed model are shown in Table 4.

Table 4. Model Performance Evaluation Results

Class Name	Precision	Recall	F1-Score	Accuracy
Healthy Leaves	98.49	98.99	98.74	
Algal Spot	91.35	95.00	93.14	
Brown Blight	92.86	91.00	91.92	94.10
Gray Blight	92.93	91.54	92.23	
Red Spot	94.97	94.03	94.50	

Source : (Research Results, 2024)

The model achieved an overall accuracy of 94.1%, outperforming several existing methods. For example, previous studies using Support Vector Machine (SVM) achieved 80% accuracy [4], Extreme Learning Machine (ELM) reported 84.67% accuracy [5], and Backpropagation neural network obtained an accuracy of 85.8% [6]. These methods faced challenges such as sensitivity to outliers (KNN), reliance on parameter tuning (SVM), and unstable performance due to random initialization (ELM). In contrast, the integration of LDA and LVQ in this study not only improved accuracy but also offered interpretability and computational efficiency.

While the proposed model achieves excellent results, its reliance on a dataset limited to 5000 images may not fully reflect the variability encountered in real-world conditions. Furthermore, disease classes with visual similarities may pose challenges in achieving even higher classification accuracy. Future research could address these limitations by incorporating data augmentation techniques to expand dataset variability and improve robustness. Exploring additional features, such as more intricate texture patterns or leaf morphology traits, could further enhance class discrimination. Moreover, experimenting with advanced neural network architectures, such as convolutional neural networks (CNNs), could enable better handling of high-resolution image data and improve overall classification performance. These advancements would strengthen the model's applicability in diverse and practical agricultural contexts.

CONCLUSION

This study demonstrates the effectiveness of combining Learning Vector Quantization (LVQ) and Linear Discriminant Analysis (LDA) for tea leaf disease classification, achieving an impressive overall accuracy of 94.1%. LDA reduced feature dimensionality by focusing on relevant color and texture attributes, while LVQ excelled in learning distinct patterns for each disease class. The model's high precision, recall, and F1-score across all classes highlight its robustness and sensitivity in accurately classifying diseases. In practical terms, the proposed model can significantly benefit tea farming by automating disease detection, reducing reliance on manual inspections, lowering labor costs, and enabling timely interventions to improve crop health and yield. However, the study's reliance on a limited dataset indicates the need for further validation on larger, more diverse datasets to enhance its generalizability. Future research could expand the dataset through data augmentation and field-sourced images, explore additional features like leaf morphology for better differentiation of visually similar diseases, and experiment with advanced machine learning models such as convolutional neural networks (CNNs) for handling high-resolution images. These advancements would enhance the model's scalability and real-world applicability, offering a more comprehensive solution for agricultural disease monitoring.

REFERENCE

[1] Directorate of Food Crops Horticulture and Estate Crops Statistics, *Indonesia Tea Statistics 2022*, vol. 16. BPS-Statistics Indonesia, 2023.

[2] A. K. Pandey, G. D. Sinniah, A. Babu, and A. Tanti, "How the Global Tea Industry Copes

VOL. 10. NO. 3 FEBRUARY 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v10i3.6013

With Fungal Diseases – Challenges and Opportunities," *Plant Dis.*, vol. 105, no. 7, pp. 1868–1879, Mar. 2021, doi: 10.1094/PDIS-09-20-1945-FE.

[3] A. M. Yatoo, M. N. Ali, Z. A. Baba, and B. Hassan, "Sustainable Management of Diseases and Pests in Crops by Vermicompost and Vermicompost Tea. A Review," *Agron. Sustain. Dev.*, vol. 41, no. 1, p. 7, 2021, doi: 10.1007/s13593-020-00657-w.

[4] Y. Sari, A. R. Baskara, and R. Wahyuni, "Classification of Chili Leaf Disease Using the Gray Level Co-occurrence Matrix (GLCM) and the Support Vector Machine (SVM) Methods," in *International Conference on Informatics and Computing (ICIC)*, IEEE, 2021, pp. 1–4.

[5] R. S. Septarini, M. Tonggiroh, and S. Nurhayati, "Identification of Potato Leaf Diseases Using Artificial Neural Networks With Extreme Learning Machine," *Pilar Nusa Mandiri J. Comput. Inf. Syst.*, vol. 20, no. 1, pp. 60–68, 2024, doi: 10.33480/pilar.v20i1.5307.

[6] Y. Lu *et al.*, "Recognition of rice sheath blight based on a backpropagation neural network," *Electron.*, vol. 10, no. 23, 2021, doi: 10.3390/electronics10232907.

[7] H. Gunawan, S. Soeheri, D. Adhar, H. Hardianto, L. Wahyuni, and C. B. Harahap, "Learning Vector Quantization (LVQ) For Colorectal Cancer Identification Based on Microscopic Network Image," in *International Conference on Cybernetics and Intelligent System (ICORIS)*, 2022, pp. 1–5. doi: 10.1109/ICORIS56080.2022.10031513.

[8] C. Diao, D. Kleyko, J. M. Rabaey, and B. A. Olshausen, "Generalized Learning Vector Quantization for Classification in Randomized Neural Networks and Hyperdimensional Computing," in *International Joint Conference on Neural Networks (IJCNN)*, 2021, pp. 1–9. doi: 10.1109/IJCNN52387.2021.9533316.

[9] F. Tang, P. Tiňo, and H. Yu, "Generalized Learning Vector Quantization With Log-Euclidean Metric Learning on Symmetric Positive-Definite Manifold," *IEEE Trans. Cybern.*, vol. 53, no. 8, pp. 5178–5190, 2023, doi: 10.1109/TCYB.2022.3178412.

[10] R. I. Borman, D. E. Kurniawan, Styawati, I. Ahmad, and D. Alita, "Classification of Maturity Levels of Palm Fresh Fruit Bunches Using the Linear Discriminant Analysis Algorithm," *AIP Conf. Proc.*, vol. 2665, no. 1, pp. 30023.1-30023.8, 2023, doi: 10.1063/5.0126513.

[11] S. Zhao, B. Zhang, J. Yang, J. Zhou, and Y. Xu, "Linear Discriminant Analysis," *Nat. Rev. Methods Prim.*, vol. 4, no. 1, p. 70, 2024, doi: 10.1038/s43586-024-00346-y.

[12] R. I. Borman, R. Napianto, N. Nugroho, D. Pasha, Y. Rahmanto, and Y. E. P. Yudoutomo, "Implementation of PCA and KNN Algorithms in the Classification of Indonesian Medicinal Plants," in *International Conference on Computer Science, Information Technology and Electrical Engineering (ICOMITEE)*, IEEE, 2021, pp. 46–50.

[13] A. Abdullah and R. H. Prince, "Tea Leaf Disease Classification," Kaggle. [Online]. Available: https://www.kaggle.com/datasets/mamun009/tea-leaf-disease-clasification

[14] R. I. Borman, F. Rossi, D. Alamsyah, R. Nuraini, and Y. Jusman, "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.

[15] S.-H. Baek, K.-H. Park, J.-S. Jeon, and T.-Y. Kwak, "Using the CIELAB Color System for Soil Color Identification Based on Digital Image Processing," *J. Korean Geotech. Soc.*, vol. 38, no. 5, pp. 61–71, 2022.

[16] 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.

[17] Z. Y. Tan, S. N. Basah, H. Yazid, and M. J. A. Safar, "Performance analysis of Otsu thresholding for sign language segmentation," *Multimed. Tools Appl.*, vol. 80, no. 14, pp. 21499–21520, 2021, doi: 10.1007/s11042-021-10688-4.

[18] R. I. Borman, F. Rossi, Y. Jusman, A. A. A. Rahni, S. D. Putra, and A. Herdiansah, "Identification of Herbal Leaf Types Based on Their Image Using First Order Feature Extraction and Multiclass SVM Algorithm," in *International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS)*, IEEE, 2021, pp. 12–17.

[19] M. Yogeshwari and G. Thailambal, "Automatic Feature Extraction and Detection of Plant Leaf Disease Using GLCM Features and Convolutional Neural Networks," *Mater. Today Proc.*, vol. 81, pp. 530–536, 2023, doi:

https://doi.org/10.1016/j.matpr.2021.03.700.

[20] S. Li, H. Zhang, R. Ma, J. Zhou, J. Wen, and B. Zhang, "Linear Discriminant Analysis with Generalized Kernel Constraint for Robust Image Classification," *Pattern Recognit.*, vol. 136, no. 109196, pp. 1–10, 2023, doi: 10.1016/j.patcog.2022.109196.

[21] A. Sucipto and A. K. Zyen, "Linear Discriminant Analysis for Apples Fruit Variety Based on Color Feature Extraction," in *International Seminar on Application for Technology of Information and Communication (iSemantic)*, IEEE, 2021, pp. 184–189.

[22] R. I. Borman, Y. Fernando, and Y. E. P. Yudoutomo, "Identification of Vehicle Types Using Learning Vector Quantization Algorithm with Morphological Features," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 2, pp. 339–345, 2022.

[23] J. E. T. Akinsola, F. O. Onipede, E. A. Olajubu, and G. A. Aderounmu, "Machine Learning Enabled Image Classification Using K-Nearest Neighbour and Learning Vector Quantization," in *Soft Computing and Its Engineering Applications*, K. K. Patel, K. C. Santosh, A. Patel, and A. Ghosh, Eds., Cham: Springer Nature Switzerland, 2024, pp. 148–163.

[24] A. F. Cobantoro, F. Masykur, and K. Sussolaikah, "Performance Analysis of Alexnet Convolutional Neural Network (CNN) Architecture With Image Objects of Rice Plant Leaves," *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 8, no. 2, pp. 111–116, 2023, doi: 10.33480/jitk.v8i2.4060.

[25] M. A. Sembiring, H. Saputra, R. A. Yusda, S. Sutarman, and E. B. Nababan, "Performance of Robust Support Vector Machine Classification Model on Balanced, Imbalanced and Outliers Datasets," *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 10, no. 1, pp. 208–215, 2024, doi: 10.33480/jitk.v10i1.5272.

[26] H. N. Syifa' and A. Nugroho, "Performance of the Yolov5 Algorithm To Detect Humans in the Rear Excavator Area," *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 10, no. 1, pp. 197–207, 2024, doi: 10.33480/jitk.v10i1.5576.