Accredited Ranking SINTA 2
Decree of the Director General of Higher Education, Research, and Technology, No. 158/E/KPT/2021
Validity period from Volume 5 Number 2 of 2021 to Volume 10 Number 1 of 2026

Published online on: **http://jurnal.iaii.or.id**

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 6 No. 2 (2022) 339 - 345 ISSN Media Electronic: 2580-0760

Identification of Vehicle Types Using Learning Vector Quantization Algorithm with Morphological Features

Rohmat Indra Borman[1], Yusra Fernando[2], Yohanes Egi Pratama Yudoutomo[3]
[1,2,3]Faculty of Engineering and Computer Science, Universitas Teknokrat Indonesia
[1]rohmat_indra@teknokrat.ac.id, [2]yusrafernando@teknokrat.ac.id, [3]yohanes_egi@teknokrat.ac.id

Abstract

The increase in the number of vehicles every year results in traffic jams. So it is necessary to identify the type of vehicle, so that the vehicle can be arranged according to the path. This study aims to develop a system that can identify the type of vehicle using the Learning Vector Quantization (LVQ) algorithm. In order for LVQ to work well in identifying, information in the form of characteristics of the object is needed. For this reason, the LVQ algorithm is combined with morphological feature extraction using the parameters of area, circumference, eccentricity, major axis length, and minor axis length to obtain shape features. Based on the test results using a confusion matrix by calculating precision, recall and accuracy, it is obtained that the precision value is 85%, recall is 82% and accuracy is 83%. This paper shows that for vehicle identification, the combination of morphological feature extraction and LVQ algorithm produces a model that can identify vehicles based on their shape and classify classes through competitive layers that are supervised by a single layer network architecture, this makes the computational process faster and does not burden the computational process.

Keywords: *image identification, learning vector quantization, morphological features*

1. Introduction

Vehicles are a means of transportation that are usually used in people's daily activities. Land vehicles that are often used include cars, motorcycles, and buses. Even vehicles are now a primary need for the community because they play an important role in facilitating community mobility. This has resulted in the last few years, the volume of vehicles in Indonesia has increased rapidly. The number of all types of two-wheeled and four-wheeled vehicles based on BPS data in 2020 reached 136,137,451 [1]. The increase in the number of vehicles every year results in traffic jams on the highway. There are several efforts made by the Government to prevent traffic jams by establishing special lanes that can be traversed by two-wheeled or four-wheeled vehicles. So it is necessary to identify the type of vehicle so that it is known whether the incoming vehicle is in accordance with the path. From these problems, a system is needed that can identify vehicles based on their image which makes it easier to get information about the type of vehicle. To overcome this, it will involve digital image processing, digital image processing. Digital image processing technology is the use of computers to manage images such as cleaning noise, improving quality, segmenting, identifying and others [2]. Image identification is the process of describing an image based on the main characteristics of the image [3].

The identification and classification of vehicle types has been carried out by previous studies. There is research on the identification of vehicle types with the Support Vector Machine (SVM) algorithm [4]. In this study, SVM can perform separable non-linear classification. The results of the accuracy test showed a value of 84.60%. The SVM method has a weakness that it is only maximal if it is used to classify two classes, because SVM is able to determine the best hyperplane to determine the separator of two classes.

Further research, regarding the classification of vehicle types using the Extreme Learning Machine (ELM) method [5]. In this study using the ELM method where this method is a feed forward artificial neural network approach, which can solve regression and classification problems. The results of the test show that the level of accuracy reaches 86.6%. In this study only classify vehicles which consist of two classes, namely cars and motorcycles. However, the ELM algorithm has a weakness, namely determining the number of hidden nodes used through a try and error process, this results

Accepted: 25-03-2022 | Received in revised: 28-04-2022 | Published: 30-04-2022

in the unknown number of hidden nodes to produce an optimal model.

Another study, regarding the classification of vehicle types with backpropagation neural networks [6]. The test results show an accuracy value of 87.5%. Backpopragation neural network algorithm can perform classification by using error output to change the weight value backwards. In this study, feature extraction only uses metric and eccentricity parameters, so that the information obtained from the image is only based on shape so that it is susceptible to occlusion if the image is not clearly visible.

This study aims to develop a system that can identify vehicle types using an artificial neural network with the Learning Vector Quantization (LVQ) algorithm. This algorithm was chosen because it has a smaller error value, can summarize large data sets into smaller ones so that the computational process is faster, and the model is more flexible [7]. This LVQ is a learning training approach on a supervised competitive layer [8]. A competitive layer will automatically learn to classify the input vectors. In order for LVQ to work well in identification, it requires information or object characteristics to make grouping easy. For this reason, the LVQ algorithm is combined with morphological feature extraction using the parameters area, perimeter, eccentricity, major axis length, and minor axis length. These parameters are used to obtain shape features so that an object can be distinguished between the object and other objects. The identified vehicles consist of four vehicles, namely: motorcycles, cars, buses and trucks.

2. Research Methods

This study will develop a vehicle type identification system and model using the Learning Vector Quantization (LVQ) algorithm with a combination of morphological feature extraction. To produce good research, research must be planned and structured through clear stages. The stages of the research carried out are presented in Figure 1.

2.1. Collecting Dataset

The first stage begins with collecting a dataset, which will collect images of the type of vehicle used for training data and test data. The types of vehicles used in this study include: motorcycles, cars, buses and trucks. The images are collected through internet searches by collecting images of motorbikes, cars, buses and trucks. The dataset distribution process uses a trial-and-error approach [9], where this approach determines the division of the number of datasets into 50% training and 50% testing. For datasets with a good level of data distribution, the composition of the amount of training and testing data will not provide fluctuating accuracy values [10]. For test cases and prototyping a small number of datasets can be used [6]. The vehicle sample used as a dataset is 160 images. Based on the trial-and-error dataset approach, the number of training data used is 80 images and the test data is 80 images. The number of classes used is 4 classes, so that each type of vehicle is 20 images. Figure 2 below is a sample data set collected on the internet to be used as test data and training data.

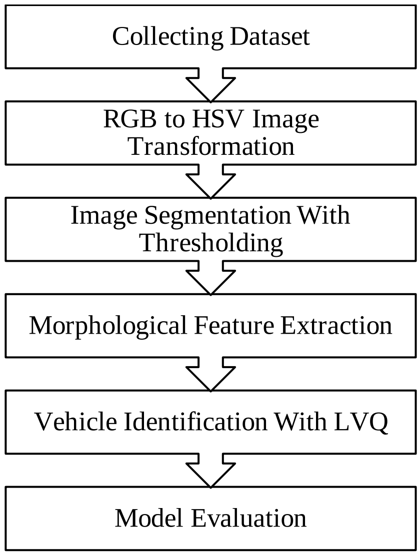

Figure 1. Research Stages

Tabel 1. Test Results For Each Vehicle Sample

Type of Vehicle	Sample Image Dataset
Motorcycle	
Car	
Bus	
Truck	

DOI: https://doi.org/10.29207/resti.v6i2.3954

2.2. RGB to HSV Image Transformation

To improve information before segmenting, the HSV color feature based on hue and saturation values is used. HSV is also called a type of perceptual color space, this is because in HSV there are cylindrical coordinates consisting of three color channels (hue, saturation, and value). [11]. RGB to HSV image transformation aims to obtain information from the colors in the image to facilitate the process of segmentation and feature extraction.

2.3. Image Segmentation With Thresholding

The image segmentation process is the stage where the object will be separated from other objects. The objects are separated based on the boundaries of the same region. The output of this process is a binary image, where the desired object will be separated from its background [12]. The image segmentation technique used is the thresholding method. Thresholding aims to find the right threshold value, so that the desired object can be distinguished from its background. The process of transforming the image into binary form makes it easy to perform feature extraction [13]. Thresholding process is actually an image quantization process using equations 1 and 2 below.

$$x = b * int(\frac{w}{b}) \tag{1}$$

$$b = int(\frac{256}{a}) \tag{2}$$

Based on equations 1 and 2 From the formula, w shows the value of the gray degree before Thresholding is carried out. Then x shows the gray degree value after thresholding.

2.4. Morphological Feature Extraction

The next stage is feature extraction, which at this stage will explore the existing characteristics of the object so that it can be distinguished from other objects. Feature extraction is an important stage, because the results of this process will be a source of information for machine learning to study certain characters or plots, making it easier to identify or classify [14], [15]. One of the characteristics that can be extracted is the shape feature. Feature extraction used is morphological feature extraction with parameters area, perimeter, eccentricity, major axis length, and minor axis length. The area is calculated based on the number of pixels that occupy the object in the image. While the perimeter is calculated based on the number of pixels around the object. Then the major axis length is the diameter of an area and the minor axis length is the shortest diameter of an area. Based on the major and minor axes, other morphological features can be calculated, namely eccentricity. Eccentricity is the length ratio between the major and minor axis. Ecentricity parameter can be calculated by equation 3.

$$e = \sqrt{1 - \frac{b^2}{a^2}} \tag{3}$$

Where, a represents the length of the major axis, while b represents the length of the minor axis.

2.5. Vehicle Identification With LVQ

Learning Vector Quantization (LVQ) is a training method for conducting learning at a supervised competitive layer whose network architecture is a single layer. The class obtained is used as a competitive layer obtained based on the distance between the input vectors. If the two input vectors obtain adjacent results, then the competitive layer will position the two input vectors in the same class. The advantage of using the LVQ artificial neural network is its ability to provide training to the competitive layer. LVQ algorithm starts from initialization of initial weight (W) and parameter LVQ. Then determine J such that $\|X_i - W_j\|$ at least using the Euclidian distance formula calculation with the equation 4.

$$D(j) = \sum\left(W_{ij} - x_i\right)^2 \tag{4}$$

Next, update the weights of W_{ij} with the condition that if $T = C_j$ then use equation 5, and vice versa if T C_j then use equation 6.

$$W_j(t + 1) = W_j(t) + \alpha(t)\left[x(t) - W_j\right] \tag{5}$$

$$W_j(t + 1) = W_j(t) - \alpha(t)\left[x(t) - W_j\right] \tag{6}$$

2.6. Model Evaluation

For evaluation in this study using a confusion matrix, where tests will be carried out based on the values of precision, recall, and accuracy. The test results will be entered into the confusion matrix. The confusion matrix consists of true positive, false positive, true negative and false negative to calculate precision, recall and accuracy [16]. Precision is a measure of the quality of information provided by the system. Then, recall is a measure of the success of the system in retrieving information. Furthermore, accuracy is the level of closeness between the identification results and the actual results. Calculating precision, recall, and accuracy can be formulated in the following 7, 8 and 9 equations.

$$Precision = \frac{TP}{TP+FP} \tag{7}$$

$$Recall = \frac{TP}{TP+FP} \tag{8}$$

$$Accuracy = \frac{TP}{TP+FP} \tag{9}$$

Where, TP is positive data that is correctly predicted. Then, TN is negative data that is predicted to be true. Furthermore, FP is negative data but the results are positive predictions. Meanwhile, FN is the opposite of FP, which is positive data but the prediction results are negative.

DOI: https://doi.org/10.29207/resti.v6i2.3954

3. Results and Discussions

To develop a system that can identify vehicle types, the first step is to build a model for training. The dataset used is 160 images of vehicles that are used as samples. The images used in the training are 80 image data and 80 image data as test data. After the dataset is collected, the next step is to prepare for the training. Training and testing is carried out using Matlab software. The initial stage is the RGB to HSV image transformation process. This process serves to obtain information from the colors in the image to facilitate the process of segmentation and feature extraction. The results of the RGB to HSV image transformation can be seen in Figure 2.

(a) (b)

Figure 2. (a) Original Image, (b) Transformation Result to HSV

The next process is image segmentation, which at this stage serves to separate the foreground from the background. The image segmentation technique used is the thresholding method. The output of this stage is in the form of a binary image, so that it can be distinguished between objects and backgrounds. The image segmentation process can be seen in Figure 4.

(a) (b)

Figure 3. (a) Original Image, (b) Image Segmentation Results

After segmenting the image, the next step is to perform morphological operations to improve the segmentation results. Then proceed with feature extraction, in order to obtain the characteristics of the object that distinguishes it from other objects. Feature extraction uses morphological features with parameters area, perimeter, eccentricity, major axis length, and minor axis length. The feature extraction process can be seen in Figure 5.

(a) (b)

Figure 4. (a) Image Segmentation, (b) Image with Morphological Operation

The result of feature extraction becomes input for vehicle type identification. Identification of vehicle types using an artificial neural network using the Learning Vector Quantization (LVQ) algorithm. The LVQ algorithm performs learning through a supervised competitive layer with its single layer network architecture. In LVQ the classes obtained are the result of a competitive layer based on the distance between the input vectors. The final aim of the LVQ algorithm is to find an appropriate weight value for grouping vectors into classes that have been initialized during the formation of the LVQ network. The process of identifying vehicle types with the LVQ algorithm is presented in the program flowchart in Figure 6.

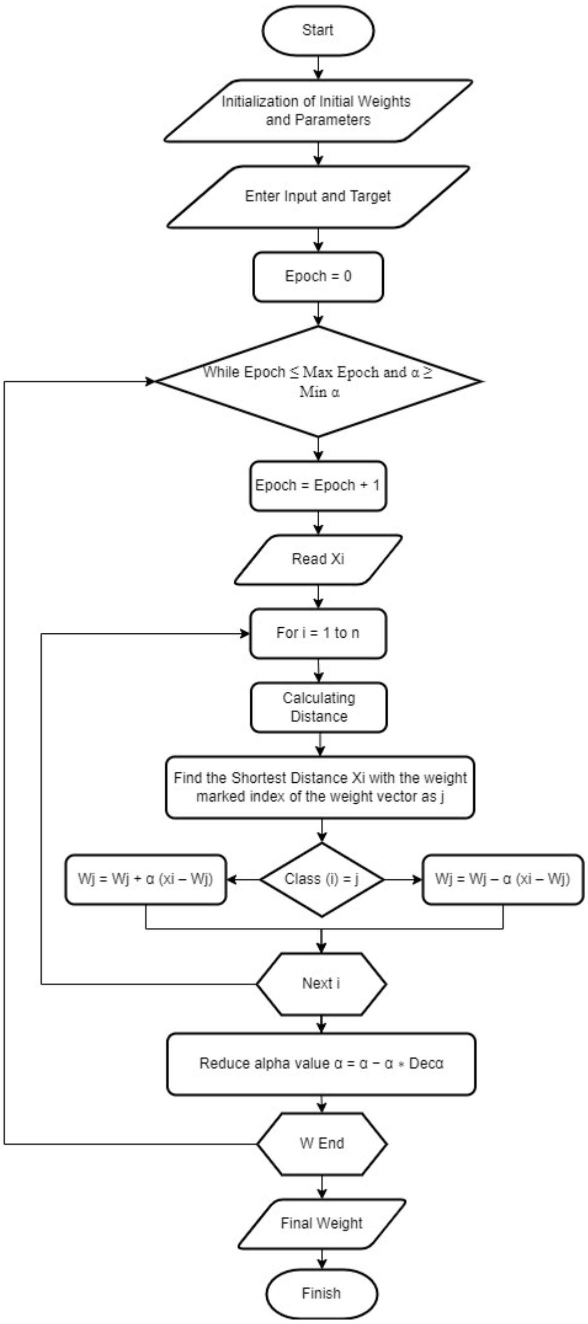

Figure 5. Identification Process Flowchart With LVQ Algorithm

Based on Figure 5, the LVQ learning algorithm begins with the initial weight (W) and the LVQ parameters,

DOI: https://doi.org/10.29207/resti.v6i2.3954

namely maxEpoch, , decα and minα. Then proceed by entering the input data (X) and the target class (T). If epoch max epoch and min , then proceed with the epoch+1 process, determining j to ‖x – wj‖ minimum and renew Wj. After Wj is updated further reduce the value of . The condition process will stop with the output of the most optimal weight. After training the LVQ network, this algorithm will classify the input vector by assigning it to the same class as the output unit, while those with reference vectors are classified as input vectors. In this case, a set of patterns whose classification is known and is given along with the initial distribution of the input vector in the same class as the output unit which has the weight closest to the input vector.

Based on the LVQ algorithm and morphological feature extraction, a training model architecture was built using Matlab software. The architecture of the built model will look for the maximum training accuracy value. The built model architecture will receive 80 training images as input with 4 classes as output (motorcycles, cars, buses and trucks). From the results of the LVQ network training the most optimal architecture is to use 100 epochs with 10 hidden layers. Figure 7 below is the most optimal model architecture in the training process.

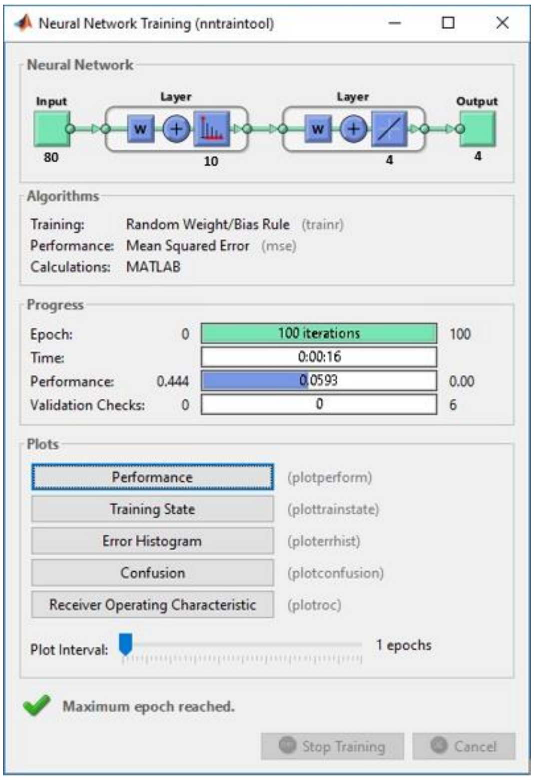

Figure 6. Learning Vector Quantization (LVQ) Architecture Used

From the architecture and model developed, further testing is carried out using test data. The test samples carried out can be seen in Table 2.

Table 2. Test Results For Each Vehicle Sample

Sample Image	Segmentation	Morphological Operation	Identification Result
			Motorcycle
			Car
			Bus
			Truck

To facilitate the use of the application, a GUI interface was created with Matlab software. The vehicle identification system interface with the implementation of the developed model can be seen in Figure 8.

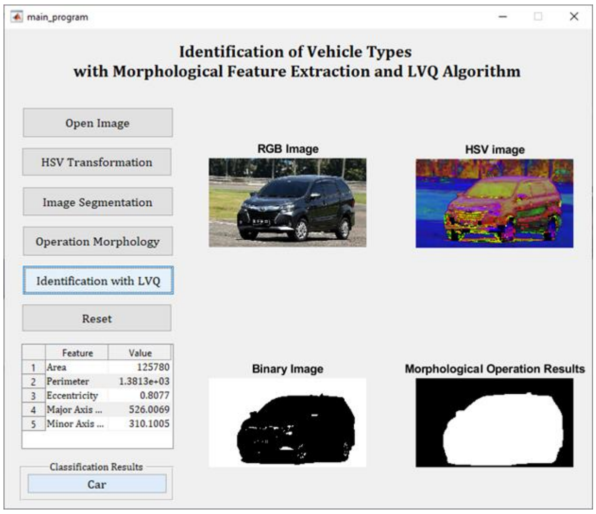

Figure 7. Vehicle Type Identification Application GUI Interface

Figure 8 shows that the application can input vehicle image data to be identified, convert the original image into an HSV image, perform image segmentation, improve segmentation results with morphological operations, perform morphological feature extraction and identify vehicle types with the LVQ algorithm.

After the model is implemented in Matlab, the next step is to evaluate the model. In the evaluation of the model using a confusion matrix, where the calculations performed include precision, recall, and accuracy. The test data used are 80 test data, then based on the results of the confusion matrix, precision, recall and accuracy can be calculated using equations 7, 8 and 9 which have been discussed previously. The results of the calculation of precision, recall and accuracy can be seen in Table 3.

Tabel 3. Precision, Recall and Accuracy Test Results

Precision	**Recall**	**Accuracy**
0.85	0.82	0.83

Based on Table 3, it can be seen that the precision value is 0.85 or 85%. This means that the level of accuracy between the information provided is 85%. While recall got a value of 0.82 or 82%, meaning that the success rate of the system in retrieving an information is 82%. Furthermore, for the accuracy of getting a value of 0.83 or 83%, it means that the level of closeness between the predicted value and the actual value is 83%. These results are converted into classification accuracy criteria with the following guidelines: Good, the value ranges between 76% to 100%; Enough, the value ranges between 56% to 75%; Less Good, the value ranges between 40% to 55%, and Less Good, if the result is below 40% [17]. When viewed from the average accuracy obtained is in the good category.

These results indicate that the developed model is able to classify vehicle types well. However, based on the tests that have been carried out, the average error value reaches 17%. From the results of the trials that have been carried out, this error rate is caused by the following factors: 1) The shape of the vehicle, especially buses and trucks when using morphological feature extraction, looks almost the same, making identification difficult; 2) the number of datasets used as training data and test data is still relatively small, so it is not optimal in the learning system; 3) Images of vehicles in various positions, for example the front or rear side, can affect the identification results; 4) The system has difficulty identifying image data with various backgrounds and is only able to identify images in which there is only one vehicle object.

This research produces an image processing model that can identify vehicles based on their type. In this study, the combination of morphological feature extraction and the LVQ algorithm produces a model that can identify vehicles based on their shape and classify classes through a competitive layer supervised by a single layer network architecture, this makes the computing process faster and does not burden the computational process. The model can be used for vehicle object detection to break down congestion by grouping vehicles by type. However, this is an initial research, it can be developed for further research in the form of a video surveillance system that utilizes CCTV.

4. Conclusion

This study develops a vehicle type identification system using a Learning Vector Quantization (LVQ) artificial neural network by combining morphological feature extraction. Morphological feature extraction can provide information on object characteristics based on area, perimeter, eccentricity, major axis length, and minor axis length parameters. The LVQ algorithm performs learning through a supervised competitive layer with its single layer network architecture. In LVQ the classes obtained are the result of a competitive layer based on the distance between the input vectors. Based on the test results using a confusion matrix by calculating precision, recall and accuracy, the precision value is 85%, recall is 82% and accuracy is 83%. These results indicate that the LVQ algorithm with morphological feature extraction can identify the type of vehicle well.

To improve further research, there are several suggestions that can be made, namely the shape of the vehicle, especially buses and trucks have almost the same shape, for that feature extraction is not enough just based on the shape. In addition, vehicle images with various positions such as the front or rear and images with various backgrounds can affect the identification results, it is necessary to develop learning using deep learning to be resistant to rotation and occlusion. The dataset used in this study is sourced from the internet, for further research it can use real data on the highway. So that learning outcomes can be optimal, it is necessary to try using a large number of datasets. In addition, it can be developed using data sources from CCTV and develop identification with low resolution.

Acknowledgment

The researcher would like to thank the LPPM of the Universitas Teknokrat Indonesia for the Internal Grant and the Universitas Teknokrat Indonesia which has provided support and funding for this research.

Reference

[1] BPS, "Perkembangan Jumlah Kendaraan Bermotor Menurut Jenis (Unit), 2018-2020," *Badan Pusat Statistik*, 2022. [Online]. Available: https://www.bps.go.id/indicator/17/57/1/jumlah-kendaraan-bermotor.html.

[2] C. Luo, Y. Hao, and Z. Tong, "Research on Digital Image Processing Technology and Its Application," in *International Conference on Management, Education and Information (MEICI)*, 2018, vol. 163, pp. 587–592.

[3] 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 *1st International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS) Identification*, 2021, pp. 12–17.

[4] M. Athoillah, "Klasifikasi Kendaraan Bermotor Dengan Multi Kernel Support Vector Mechine," *Buana Mat. J. Ilm. Mat. dan Pendidik. Mat.*, vol. 8, no. 1, pp. 1–8, 2018.

[5] R. Himilda and R. A. Johan, "Klasifikasi Jenis Kendaraan Menggunakan Metode Extreme Learning Machine," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 2, no. 4, pp. 237–243, 2021.

[6] H. Mayatopani, R. I. Borman, W. T. Atmojo, and A. Arisantoso, "Classification of Vehicle Types Using Backpropagation Neural Networks With Metric and

DOI: https://doi.org/10.29207/resti.v6i2.3954

Ecentricity Parameters," *J. Ris. Inform.*, vol. 4, no. 1, pp. 65–70, 2021.

[7] C. A. Wicaksana, D. E. Tyas, and L. Lufianawati, "Pengenalan Pola Vektor Tanda Tangan Citra Digital Menggunakan Metode Pembagian Wilayah dan Learning Vector Quantization (LVQ)," *J. Ilm. Setrum*, vol. 10, no. 2, pp. 24–29, 2021.

[8] U. Sudibyo, D. P. Kusumaningrum, E. H. Rachmawanto, and C. A. Sari, "Optimasi Algoritma Learning Vector Quantization (LVQ) Dalam Pengklasifikasian Citra Daging Sapi dan Daging Babi Berbasis GLSCM dan HSV," *J. SIMETRIS*, vol. 9, no. 1, pp. 1–10, 2018.

[9] Y. Chen, L. Song, Y. Liu, L. Yang, and D. Li, "A Review of the Artificial Neural Network Models for Water Quality Prediction," *Appl. Sci.*, vol. 10, no. 5776, pp. 1–49, 2020.

[10] W. Musu, A. Ibrahim, and H. Heriadi, "Pengaruh Komposisi Data Training dan Testing terhadap Akurasi Algoritma C4.5," in *Prosiding Seminar Ilmiah Sistem Informasi dan Teknologi Informasi*, 2021, vol. X, no. 1, pp. 186–195.

[11] K. Ayuningsih, Y. A. Sari, and P. P. Adikara, "Klasifikasi Citra Makanan Menggunakan HSV Color Moment dan Local Binary Pattern dengan Naïve Bayes Classifier," *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 3, no. 4, pp. 3166–3173, 2019.

[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 *ICOMITEE 2021*, 2021, pp. 46–50.

[13] M. Wati, Haviluddin, N. Puspitasari, E. Budiman, and R. Rahim, "First-order Feature Extraction Methods for Image Texture and Melanoma Skin Cancer Detection," *J. Phys. Conf. Ser.*, vol. 1230, no. 1, 2019.

[14] D. Riana *et al.*, "Identifikasi Citra Pap Smear RepoMedUNM dengan Menggunakan K-Means Clustering dan GLCM," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 1, pp. 1–8, 2022.

[15] R. I. Borman, B. Priopradono, and A. R. Syah, "Klasifikasi Objek Kode Tangan pada Pengenalan Isyarat Alphabet Bahasa Isyarat Indonesia (Bisindo)," in *Seminar Nasional Informatika dan Aplikasinya (SNIA)*, 2017, no. September, pp. 1–4.

[16] M. F. Arifin and D. Fitrianah, "Penerapan Algoritma Klasifikasi C4.5 Dalam Rekomendasi Penerimaan Mitra Penjualan Studi Kasus: PT Atria Artha Persada," *InComTech*, vol. 8, no. 2, pp. 87–102, 2018.

[17] D. Nurnaningsih, D. Alamsyah, A. Herdiansah, and A. A. J. Sinlae, "Identifikasi Citra Tanaman Obat Jenis Rimpang dengan Euclidean Distance Berdasarkan Ciri Bentuk dan Tekstur," *Build. Informatics, Technol. Sci.*, vol. 3, no. 3, pp. 171–178, 2021.

DOI: https://doi.org/10.29207/resti.v6i2.3954