Router: Jurnal Teknik Informatika dan Terapan
Volume. 3 Nomor. 1 Tahun 2025
e-ISSN: 3032-3312; dan p-ISSN: 3026-3611, Hal 01-14
DOI: https://doi.org/10.62951/router.v2i4.385
Available Online at: https://journal.aptii.or.id/index.php/Router

Identification of Flower Type Images Using KNN Algorithm With HSV Color Extraction and GLCM Texture

Edhy Poerwandono[1] , M. Endang Taufik[2]
[1&2] Informatics Engineering Study Program, STIKOM Cipta Karya Informatics, Indoneisia
Jl. Radin Inten II, Duren Sawit, East Jakarta 13440
E-mail: edhypoerwandono@gmail.com[1] , endangtaufik@gmail.com[2]

Abstract. *Due to the variety of types of flowers that exist and having and tracking each variety, making plant lovers and cultivators difficult to distinguish in determining the type of flower, it takes a very long time to find out the type of flower if you only rely on the five senses. With the application of the K-Nearest Neighbor algorithm and feature extraction of color and texture, it is very helpful in image processing to identify flowers more easily and shorten the time, with the greatest accuracy of 71% using the K-7 value, the flower was successfully carried out.*

Keywords*: Identification, HSV, GLCM, K-Nearest Neighbor*

1. INTRODUCTION

Flowers consist of leaves and twigs that are around them and have undergone many changes. These changes are due to the results of a number of enzymes that are stimulated by a number of certain phytohormones [1]. The formation of a flower is very tightly controlled genetically and many types are induced by certain environmental changes, such as low temperatures, lack of sunlight, and available water. Flowers are almost always equilateral, which is often used as a characteristic of a taxon. There are two forms of flowers based on the symmetry of their shape: actinomorph and zygomorph. The actinomorphic form is very easy to find [2]. With the beauty of the flower, it causes plant lovers and cultivators are increasing in number and flower production continues to grow and develop [3].

Therefore, computer technology is developing rapidly as a tool to obtain information that can be used to identify types of flowers [4]. Computers can process data quickly, precisely and accurately with a high level of accuracy, so that they can be a tool to obtain the information needed according to human needs in various fields [5]. The development of image processing technology today provides the possibility for humans to create a system that can recognize something in a digital image [6]. Image processing is also a type of technology to solve problems regarding image processing [7]. In image processing, images are processed in such a way that the image can be used for further applications [8]. Therefore, the demand for information regarding types of flowers encourages the creation of a system to identify based on digital image processing. As in previous research that combines digital image processing with artificial intelligence to produce introduction absence face in school The Shafiyyatul

Received: Janauri 05, 2024; Revised: Januari 20, 2024; Accepted: Februari 09, 2024; Online Available : Februari 14, 2024

Amaliyahusing the K-NN method classification, the next study is an experiment and analysis of the results of testing the fish classification method using ORB and K-NN, in the next study is the detection of bacteria in the covid pandemic based on HSV color, there is also a study that discusses the detection of batik with texture extraction, the last study is using the KNN algorithm as a classification of chronic kidney disease, based on previous studies, researchers got the idea to identify flowers because there are so many plant lovers and plant cultivation problems

The existing problem based on the soaring production of ornamental plants in 2019 has resulted in the number of plant lovers continuing to grow and to minimize the problem of ignorance of the types of flowers that exist, it is necessary to create a system that helps to identify flower types.

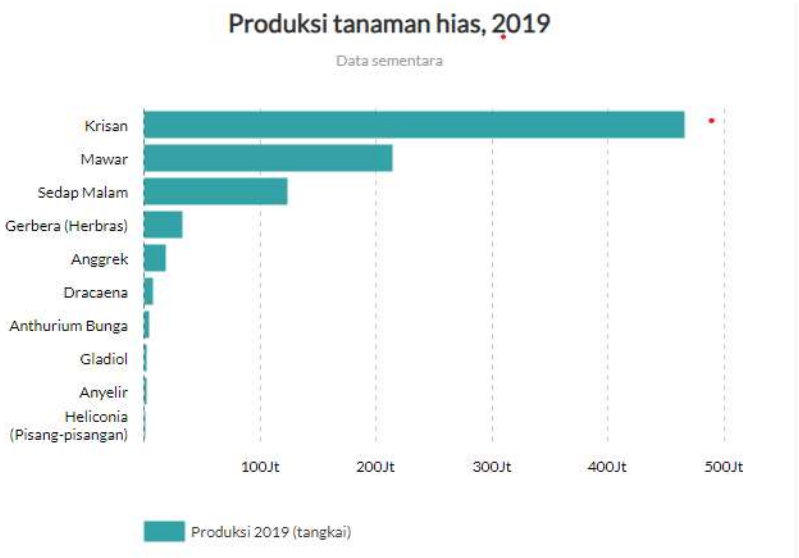


Figure 1. Ornamental Plant Production 2019 (Source: Ministry of Agriculture)

2. RESEARCH METHODOLOGY

This research methodology is based on the stages of digital image processing.

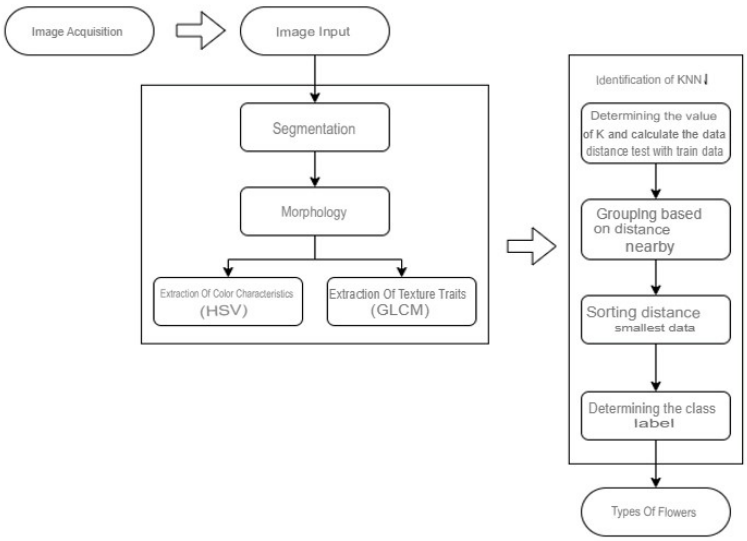


Figure 2.Research Flow

1. Data collection: The goal at this stage is to determine the data needed and select a digital image processing method [9]. The steps taken in this generally start from preparing the object to be imaged, to the imaging process [10]. In this process, image retrieval is carried out at https://www.kaggle.com/with the title Flower Color Image image data will be used as a sample for identification based on color and texture of flower types. The image data only uses 10 types of flowers [11].
2. *Pre Processing*: The pre-processing process is used to obtain images that will be continued and utilized for the designed system. This process consists of cropping, resizing, segmentation and removing noise from the image. Each flower type image has 100 data whose size is averaged to 500 x 500 with .png format. Then segmentation is carried out to change the RGB (Red, Green, Blue) flower type image into a grayscale image first with otsu thresholding [10]. Then continued with morphological operations as a noise remover.
   a. *Cropping : Cropping*is the process of cutting an image at a certain point in the image area. The engineering process used to determine exactly which part of the image contains the object area to be processed. In addition, this process can also change the size (resize) of the image to be more small, to speed up the calculation process.
   b. Segmentation: Image segmentation is a separator between objects and other objects,

in an image or between objects and the background in an image [12]. In the segmentation process, each object in the image can be taken individually so that it can be used as input for other processes [13]. There are two types of segmentation, namely full segmentation and partial segmentation.

c. Otsu Thresholding: Otsu thresholding is one of the simple segmentation methods in segmentation techniques, so that the process obtained is easier in the process of dividing homogeneous areas based on similarities in recognizing objects. Before observing the results of image segmentation, it must first go through the image input process, so that it is easy to enter the next process. The next step is to add brightness to the image to improve image quality [14]

d. Thresholding: Thresholding is a very simple method of other image segmentation. With grayscale image input, thresholding can be used to form a binary image. A binary image is a digital image in which each pixel has only two values, namely black and white based on the numbers 0 and 1.

e. Morphology: Morphology is used as the first and last step of the image analysis process. The method used in this research morphology is erosion operation. Erosion is the process of comparing each pixel in the input image with the value of the structural element so that the center of the structural element can match the pixel position in the processed image.

[15] Morphological operations are also useful for removing noise[16].

3. Extraction Features:

a. HSV Color: HSV is used as a feature extraction with color selection based on Hue, Saturation, and Value values. Hue is a pure color attribute. Saturation is the intensity of white light that affects the level of color dominance. Value is the difference in object brightness. The process of obtaining the value of each color to be displayed through a calculation process by converting RGB color space to HSV color space [17].

b. GLCM Texture: GLCM is a technique to obtain 2nd order statistical values by calculating the probability of proximity between two pixels at a certain distance (d) and angle (θ). The working process in the GLCM method is to form co-occurrence in image data, then determine the functional characteristics of the matrix between pixels [18].

4. *K-Nearest Neighbor*: KNN (K-Nearest Neighbor) is an object classification method based on the closest distance to an object or feature, the data most commonly used in learning data [19]. The training data is projected into a multi-dimensional space, where each

dimension has features in the data. The space is divided into several parts according to the classification of the training data. A point in the space

This is marked as category c, category c is the most common category among the k nearest neighbors of the points. The proximity or distance of neighbors is usually calculated based on the Euclidean distance with the following formula:

$$Dij = \sqrt{\sum_{K=1}^{n} (i - j)^2}$$

Information:

Dij : Euclidean distance between i and j.

i : data on the ith x for the calculation stage.

j : data on the jth y for the calculation stage.

3. RESULTS AND DISCUSSION

The results of this study are how to identify flower type images by applying HSV color feature extraction and GLCM texture K-Nearest Neighbor (K-NN) algorithm to identify flower type images into ten classes, and obtain the accuracy of the evaluation results in identifying flower type images for each type of flower as many as 10 images, with a dataset of 100.

1. Data Collection: At this stage, the type of flower will be imaged by visiting the website https://www.kaggle.com/with the title Flower Color Image then download the entire public data.

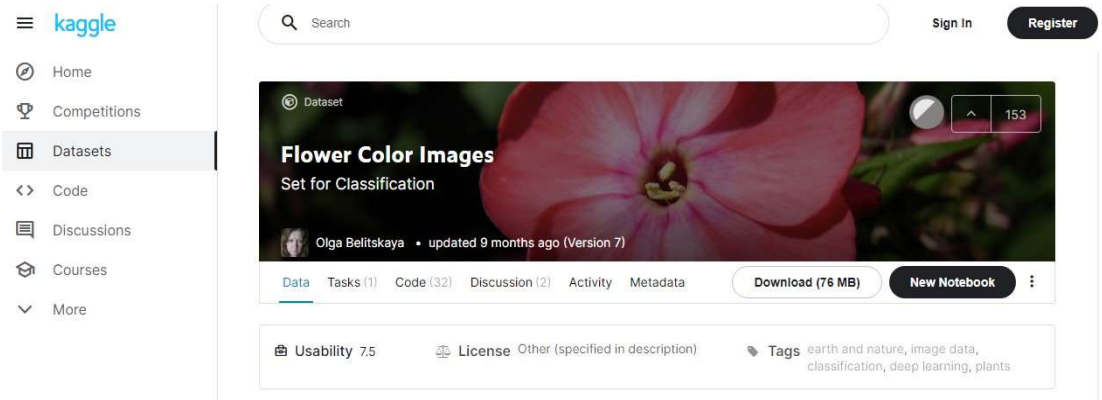
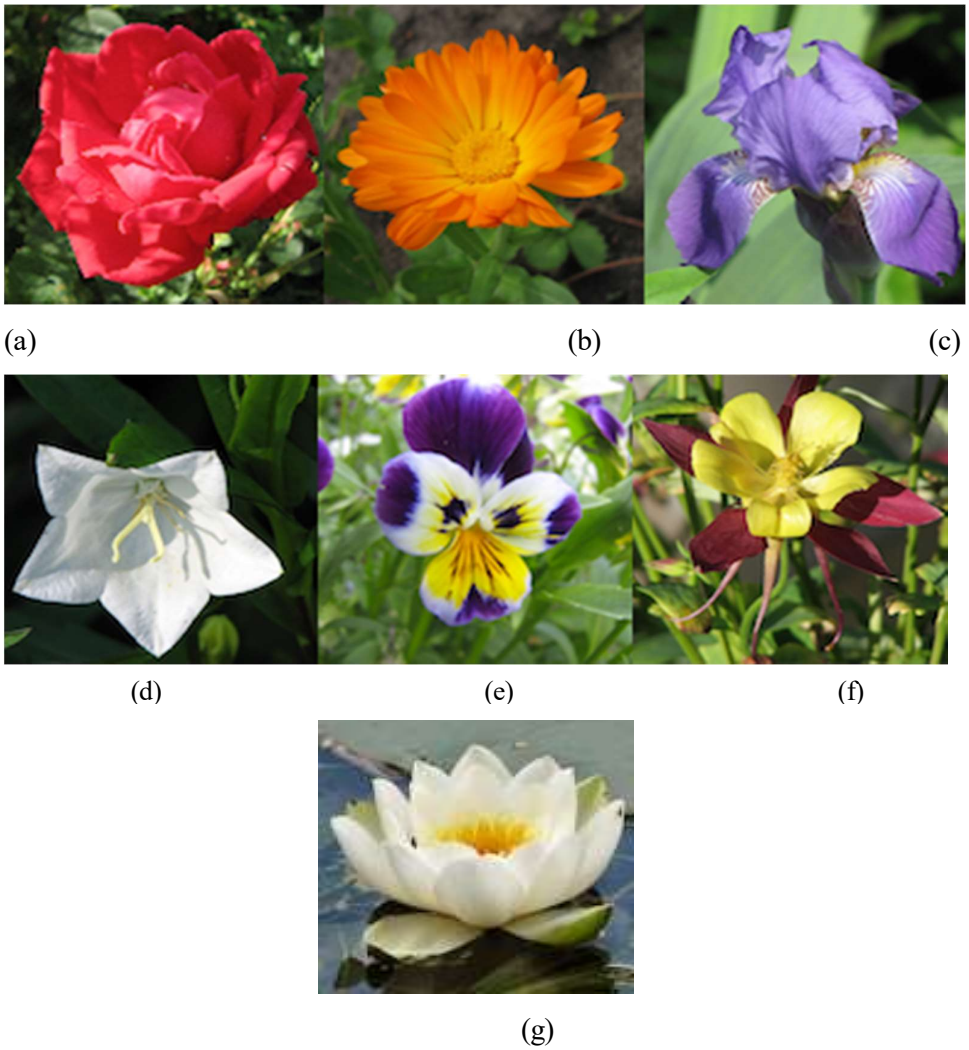


Figure 3. Kaggle Web Page Here are the results of image data that has been taken from the Kaggle web



(a) (b) (c)

(d) (e) (f)

(g)

Figure 4.(a) Rose flower (b) Calendula flower (c) Iris flower (d) Bellflower flower (e) Viola flower (f) Lilium flower (g) Water Lily flower

2. *Pre Processing: Pre Processing*used to simplify the process of identifying images. This stage consists of cropping, compressing image size, image segmentation and removing noise from the image. Cropping and compressing image size are done using third-party applications with an image size of 500 x 500 pixels and a ratio of 1:1.

Figure 5.Original Image

The obtained image is then converted into a grayscale image to carry out the segmentation process.

Figure 6. Grayscale Image

Then image segmentation using MATLAB software with Otsu thresholding to facilitate the image processing process to the next stage with images that have been converted from RGB images to grayscale images, then morphological segmentation is carried out to remove noise.

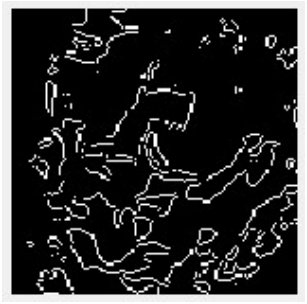


Figure 7.Binary Image

3. Feature Extraction: In this study, the feature extraction used is color feature extraction because the factor that affects the difference in flower types is color. The colors included for feature extraction are hue, saturation, and value as well as GLCM texture feature extraction which includes contrast, correlation, homogeneity and energy. Feature extraction has various variations, the more feature extractions that are relevant to the image are used, the better the accuracy results. The following is an HSV feature extraction image while for GLCM features using a grayscale image.

Figure 8.HSV Image

4. *K-Nearest Neighbor*: The stage of identifying flower type image data using the K-Nearest Neighbor (KNN) algorithm. The data from the previously standardized feature extraction stage has been obtained, then it will be divided into two in the identification process, namely training data and testing data. In dividing the data, the percentage split technique is applied where there are three data division scenarios. The following are details of the three scenarios of the percentage split technique and the results of class division on flower type images:

Table 1.*Percentage Split*

No	*Split Data*	*Training Data*	*Testing Data*
1	70% and 30%	70	21
2	60% and 40%	60	28

Table 2.Types of Flowers

No	ClassFlower	Type Flower
1	FlowerA	Rose Flower
2	FlowerB	Calendula Flower
3	FlowerC	Iris
4	BungaD	Bellflower
5	FlowerE	Viola Flower
6	FlowerF	Lilium Flower
7	FlowerG	Water Lily Flower

Steps in implementing the data sharing scenario that has been set, the identification process is carried out using the K-Nearest Neighbor algorithm with the k values used being 3, 5 and 7. The following are the results of the identification test.

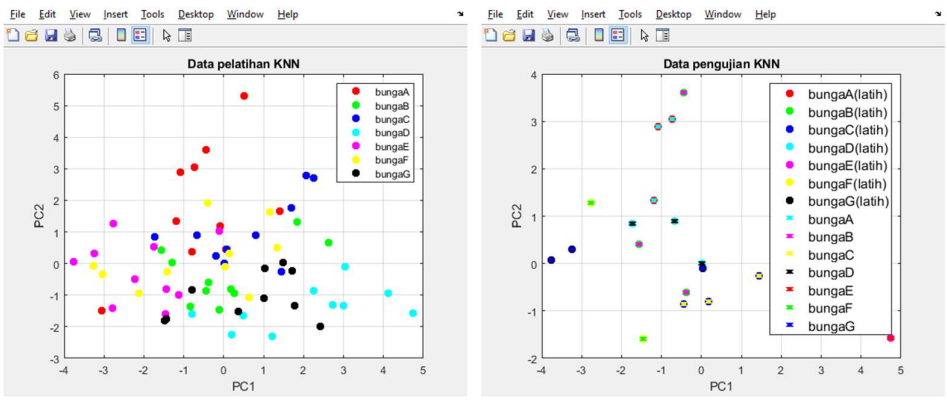


Figure 9. Visualization of KNN training and testing data results

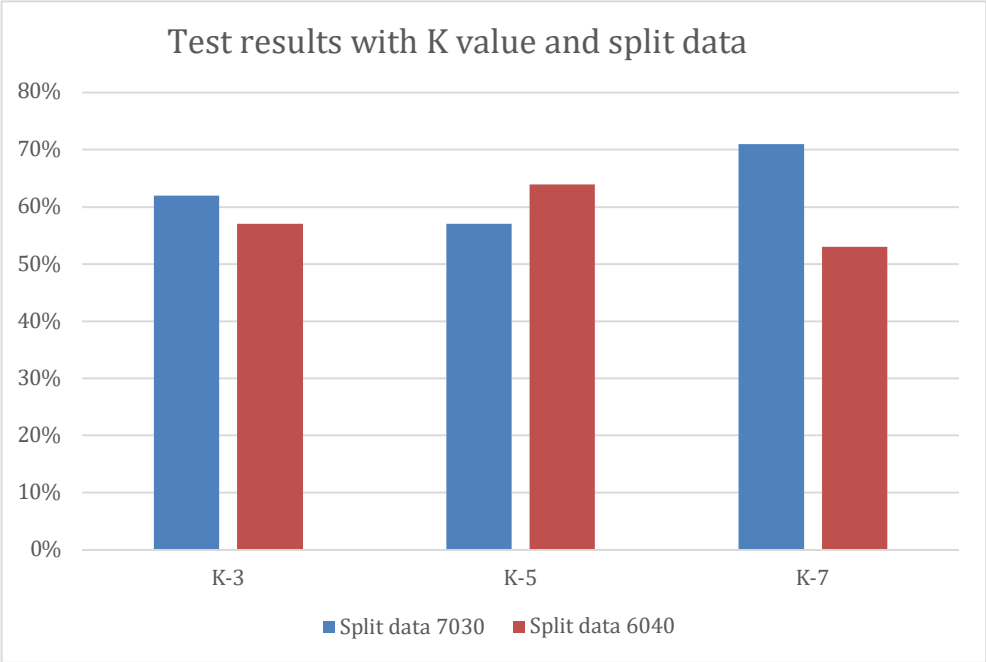


Figure 10. Test Results With K Value and Split Data

The first test is data division using 70% training data and 30% test data. From one hundred images with 10 images per type of flower divided into two data, namely 70 for training data and 21 test data, the results of identification accuracy were obtained with K-3 = 62%, K-5 = 57% and K-7 = 71%. While in the second test is data division using 60% training data and 40% test data. From one hundred images with 10 images per type of flower divided into two data, namely 70 for training data and 28 test data, the results of identification accuracy were obtained with K-3 = 57%, K-5 = 64% and K-7 = 53%.

Testing with K value and data division is done to determine the effect of a certain ratio or percentage of the amount of training data and testing data on the accuracy level of the KNN method [20]. After the test is successfully carried out, the best accuracy and the best K value are used for identification on a GUI in MATLAB. The following tests are carried out with the GUI.

In the GUI there are buttons with their respective functions, the first button is Image Input which functions to enter an image to be tested, can be seen in the following image.

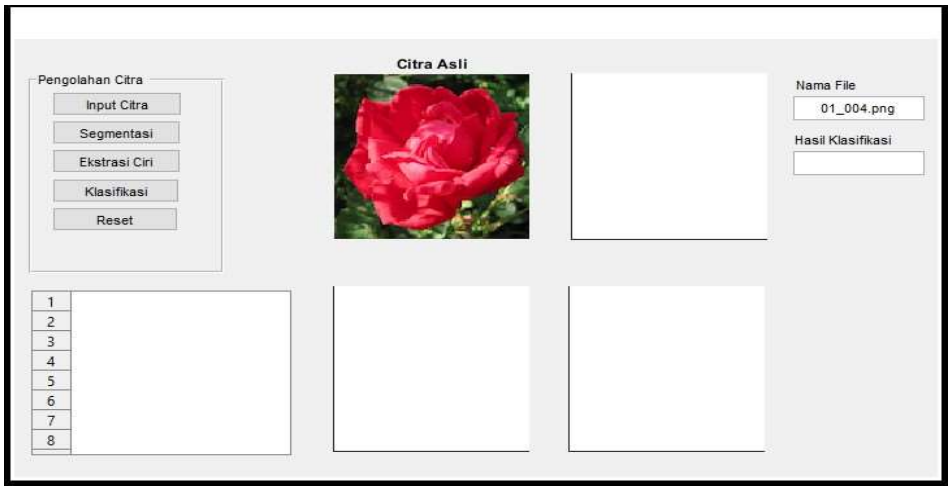


Figure 11. Image Input

After the image is successfully input, the image will appear on axes 1 or as the original image and on the right side there is the name of the file that has been input. Next there is a segmentation button that functions to segment the image with morphology to help eliminate noise, as seen in the following figure.

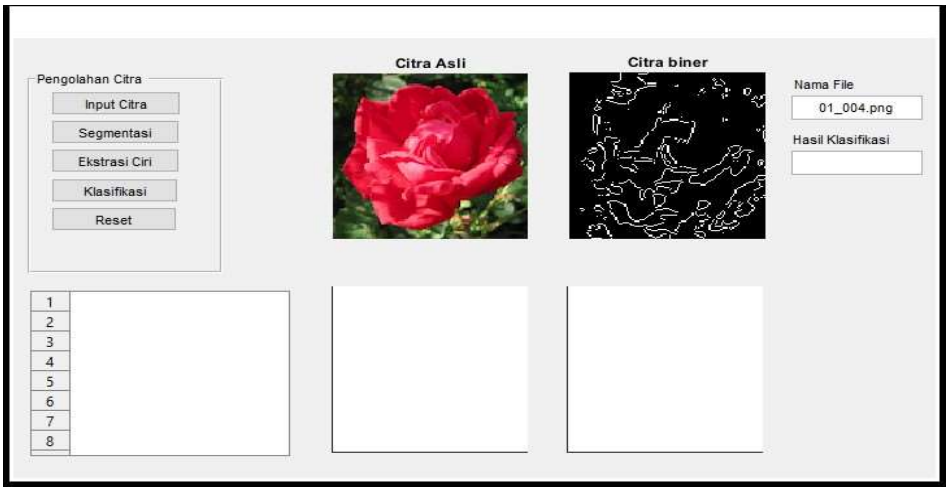


Figure 12. Image Segmentation

After carrying out the preprocessing stage, a feature extraction search will be carried out on the image, as can be seen in the following image.

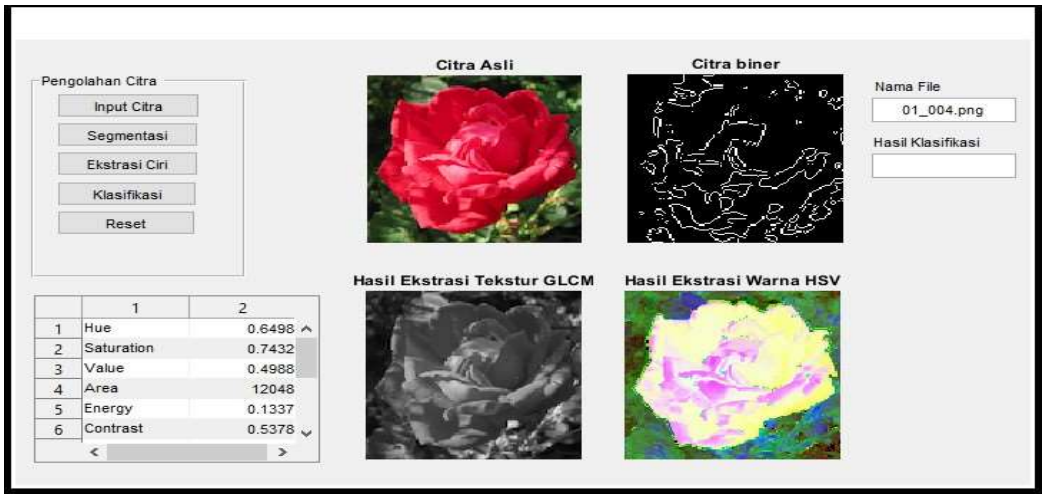


Figure 13. Feature Extraction

Because it uses GLCM texture features and HSV color features, there are grayscale image results for HSV texture and color features, in the table on the left side there are feature extraction values from an image. After getting the extraction value results on the image, an identification will be carried out on the image. Can be seen in the following figure.

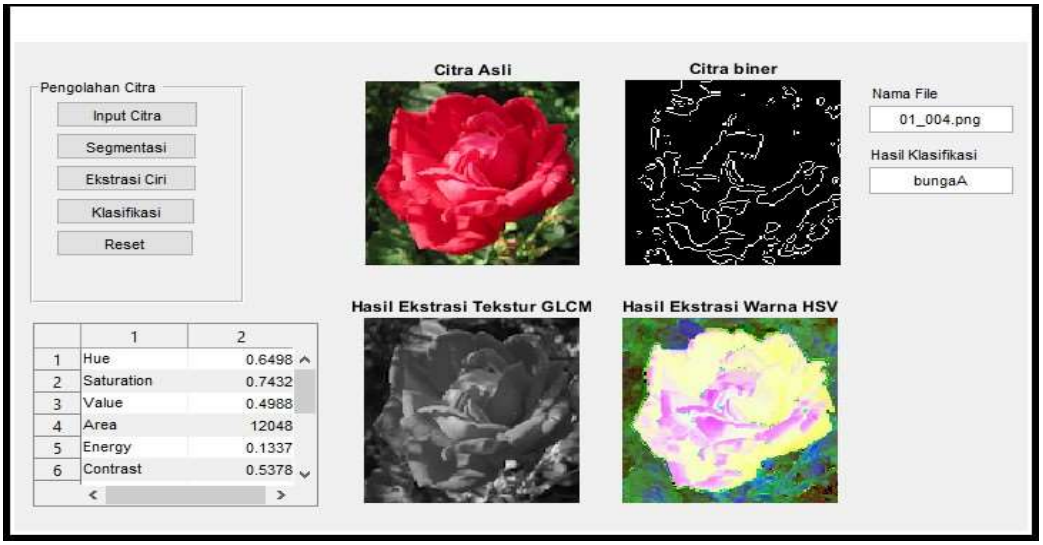


Figure 14. Identification Results

Identification results have been obtained and testing with GUI was successful was done for the image of the type of Rose flower identified as flower A can be seen

Table 2. Based on the test graph above, there are two scenarios of division between training data and testing data where the division details have been explained in Table 1. In each

of these scenarios, each experiment was carried out with various K, namely K-3, K-5 and K-7. From the results shown in Figure 10, it can be seen that the ratio of 70:30 and K-7 is a fairly good result compared to other tests. And testing with GUI succeeded in identifying the image of the type of flower according to its class

Table 3. Test Results with K ValueNo

	Split Data	K Value	Accuracy
1		K-3	62%
2	7030	K-5	57%
3		K-7	71%
4		K-3	57%
5	60 40	K-5	64%
6		K-7	53%

4. CONCLUSION

The conclusion obtained in the study of Identification of Flower Type Images using HSV Color and GLCM Texture is that digital image identification of flower types using the feature extraction method for color and texture was successfully carried out with data divisions of 70:30 and 60:40 and testing on K-3, K-5, K-7 values. Getting the best accuracy in data division 70:30 using a K-7 value of 71%, the results obtained were considered lacking because the data obtained was very limited to use.

5. SUGGESTION

It is hoped that further research can increase the number of input images of flower types to obtain better accuracy as well as the application of algorithms and other feature extractions that support the identification of flower types.

6. BIBLIOGRAPHY

[1] A. D. Williams, S. N. Abbott, N. Shoman, and W. S. Charlton, "Results from Invoking Artificial Neural Networks to Measure Insider Threat Detection & Mitigation," *Digit. Threat. Res. Pract.*, vol. 3, no. 1, 2022, doi: 10.1145/3457909.

[2] B. N. G. Koneru and V. Vasudevan, "Sparse artificial neural networks using a novel smoothed LASSO penalization," *IEEE Trans. Circuits Syst. II Express Briefs*, vol. 66, no. 5, pp. 848–852, 2019, doi: 10.1109/TCSII.2019.2908729.

[3] B. Isik *et al.*, "Neural Network Compression for Noisy Storage Devices," 2021, doi: 10.1145/3588436.

[4] E. O. Nascimento and L. N. Oliveira, "Numerical Optimization of Flight Trajectory for Rockets Via Artificial Neural Networks," *IEEE Lat. Am. Trans.*, vol. 15, no. 8, pp. 1556–

1565, 2017.

[5] F. Anders, M. Hlawitschka, and M. Fuchs, "Comparison of artificial neural network types for infant vocalization classification," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 29, pp. 54–67, 2021, doi: 10.1109/TASLP.2020.3037414.

[6] G. Yang, H. He, and Q. Chen, "Emotion-semantic-enhanced neural network," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 27, no. 3, pp. 531–543, 2019, doi: 10.1109/TASLP.2018.2885775.

[7] J. Bruinsma and R. Carloni, "IMU-Based Deep Neural Networks: Prediction of Locomotor and Transition Intentions of an Osseointegrated Transfemoral Amputee," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1079–1088, 2021, doi: 10.1109/TNSRE.2021.3086843.

[8] M. Hosseini and T. Mohsenin, "Binary Precision Neural Network Manycore Accelerator," *ACM J. Emerg. Technol. Comput. Syst.*, vol. 17, no. 2, 2021, doi: 10.1145/3423136.

[9] A. J. Paul, P. J. Collins, and M. A. Temple, "Enhancing Microwave System Health Assessment Using Artificial Neural Networks," *IEEE Antennas Wirel. Propag. Lett.*, vol. 18, no. 11, pp. 2230–2234, 2019, doi: 10.1109/LAWP.2019.2926932.

[10] S. V. Medina and U. P. Ajenjo, "Performance Improvement of Artificial Neural Network Model in Short-term Forecasting of Wind Farm Power Output," *J. Mod. Power Syst. Clean Energy*, vol. 8, no. 3, pp. 484–490, 2020, doi: 10.35833/MPCE.2018.000792.

[11] P. Staszewski, M. Jaworski, J. Cao, and L. Rutkowski, "A New Approach to Descriptors Generation for Image Retrieval by Analyzing Activations of Deep Neural Network Layers," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 33, no. 12, pp. 7913–7920, 2022, doi: 10.1109/TNNLS.2021.3084633.

[12] T. Liu, H. Mei, Q. Sun, and H. Zhou, "10.23919@Jcc.2019.10.014," pp. 1–6, 2019.

[13] W. C. Yeh, "A squeezed artificial neural network for the symbolic network reliability functions of binary-state networks," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 28, no. 11, pp. 2822–2825, 2017, doi: 10.1109/TNNLS.2016.2598562.

[14] Y. Gao, L. T. Yang, D. Zheng, J. Yang, and Y. Zhao, "Quantized Tensor Neural Network," *ACM/IMS Trans. Data Sci.*, vol. 2, no. 4, pp. 1–18, 2021, doi: 10.1145/3491255.

[15] Z. Chen, T. B. Xu, C. Du, C. L. Liu, and H. He, "Dynamical Channel Pruning by Conditional Accuracy Change for Deep Neural Networks," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 32, no. 2, pp. 799–813, 2021, doi: 10.1109/TNNLS.2020.2979517.

[16] P. Seminar and N. Riset, "Jaringan Syaraf Tiruan Backpropogation dalam Memprediksi Jumlah Tilang di Kejaksaan Negeri Simalungun," no. September, pp. 697–708, 2019.

[17] J. Antares, T. Infomasi, and U. Dharmawangsa, "Artificial Neural Network Dalam Mengidentifikasi Penyakit Stroke Menggunakan Metode Backpropagation (Studi Kasus di Klinik Apotik Madya Padang)," vol. 1, no. 1, pp. 6–14, 2020.

[18] H. Putra and N. Ulfa, "Jurnal Nasional Teknologi dan Sistem Informasi Penerapan Prediksi Produksi Padi Menggunakan Artificial Neural Network Algoritma Backpropagation," vol. 02, pp. 100–107, 2020.

[19] I. Engineering, S. Tunas, and B. Pematangsiantar, "Implementation of Artificial Neural

Networks In Predicting Students ' English Understanding Level Using The Backpropagation Method," vol. 1, no. 2, pp. 77–86, 2020, doi: 10.30596/jcositte.v1i2.5072.