Journal of Applied Engineering and Technological Science
Vol 4(1) 2022 : 180-189

HEALTH DETECTION OF BETAL LEAVES USING SELF-ORGANIZING MAP AND THRESHOLDING ALGORITHM

Dadang Iskandar Mulyana[1], Ahmad Saepudin[2*], Mesra Betty Yel[3]
Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika[1,2,3]
ahmadsaepudin0205@gmail.com

Received : 24 August 2022, Revised: 18 September 2022, Accepted : 18 September 2022
**Corresponding Author*

ABSTRACT
Betel leaf is one of the plants that is widely used as a natural or traditional medicine by the community, natural treatment with the use of plants is relatively safer. But there is a problem when we choose healthy betel leaves because of our mistakes in choosing which betel leaves are healthy and which are not. With this research the authors aim to detect healthy and sick betel leaves using data collection. Feature extraction used is the value of Red, Green, and Blue (RGB) and Hue, Saturation, and Value (HSV) to get the characteristics of the color image. Then the results of the feature extraction are used to classify the health of green betel leaves using the Self-Organizing Maps method. The green betel leaf data used is 1500 images for train data and 450 images for testing data are image test data, test data that produces an evaluation value with an accuracy value of 97.20% on the Self-Organizing Maps method.
Keywords : *Betel Leaf, Thresholding, Self-Organizing Map*

1. Introduction

Indonesian is a country with abundant plant diversity, each has its own advantages. Plants are used in agriculture, plantation, forestry, industrial materials and more. However, the plant can also be used as a natural/traditional medicine. Natural treatment with the use of plants is relatively safe from the side effects of long-term consumption and includes psychological motivation to increase belief in healing. Betel leaf has long been used as a traditional medicine by the people of Indonesia. Betel plant (Piper bettle L.) is a plant that has antibacterial properties, this is due to the substances contained in its (Iswandana & Sihombing, 2017).

There are various types of betel, namely green betel, red betel. Green betel is often used to treat body and mouth odor, canker sores, nosebleeds, itching and sores, and treat vaginal discharge in women. In addition to green betel, one other type of green betel that is currently widely used for health is red betel. Although the chemical content of this plant has not been studied in detail, the results of the crematoGram of red betel leaf contain flavonoids, polevenolad compounds, tannins and essential oils. These compounds are believed to be able to treat various diseases (Khotimah & Darsin, 2020).

The need for drugs in the community is relatively high, but more frequent use of chemical drugs can have a negative impact on health if consumed excessively. For some symptoms that appear, such as body weakness, weakened brain, indigestion, palpitations, joint or muscle pain. There are various types of betel, namely green betel, red betel. Green betel is often used to treat body and mouth odor, canker sores, nosebleeds, itching and sores, and treat vaginal discharge in women. In addition to green betel, one other type of green betel that is currently widely used for health is red betel. Although the chemical content of this plant has not been studied in detail, the results of the crematoGram of red betel leaf contain flavonoids, polevenolad compounds, tannins and essential oils. These compounds are believed to be able to treat various diseases (Ardiansah et al., 2021).

Along with current technological developments that are increasingly rapidly, especially in the field of digital imagery in determining the health of betel leaves. One of the medicinal plants known by the general public is green betel leaf which is included in the group of medicinal plants which reaches more than 1000 species. Self-organizing maps (SOMs) is one of the clustering algorithms based on the characteristics or characteristics of the data. SOM is a form of topology that does not require monitoring in the training process (target output) (Sinaga, Adikara, and Sari 2020). The advantage of using SOM when grouping objects is that it can map high-dimensional

input vectors to 2D space. In addition, SOM can group categorical data and incomplete input data (Juliansa hengki & Sumijan, 2017; Angulo-Saucedo et al., 2022).

Machine learning is a computational algorithm or computer process that works based on historical data to improve performance in creating predictors. In machine learning, there are three learning methods, namely unsupervised learning, supervised learning, and reinforcement learning. In unsupervised learning, the training data used does not yet have a class, so the data are grouped based on the same characteristics. Supervised learning is a learning method for training data that already has classes. Furthermore, in reinforcement learning, the right steps will be sought in order to obtain the right predictions and in accordance with the existing conditions (Adenugraha et al., 2022; Greener et al., 2022; Malekloo et al., 2022).

In the process of implementing the Self-Organizing Map (SOM) learning algorithm, it is necessary to do cluster validation to determine whether the cluster model is suitable for use as inference. There are several methods that can be used to carry out the cluster verification process, one of which is the Davies Bouldin Index (DBI) value method. This method makes it possible to correctly display index values up to the number of groups formed (Kania et al., 2019; Ashari et al., 2022).

Thresholding is the process of changing the image from gray level to binary form. The thresholding process is also called binaryization. Thresholding is used to adjust the number of degrees of gray in the image. By using thresholding, the degree of gray can be changed as desired Value (Nafi'iyah & Fatichah, 2017).

The Kohonen SOM network was introduced in 1982 by the Finnish researcher Tuevo Kohonen. Kohonen SOM is a network that does not require special supervision, hence the name self-organizing. The word maps means that this method uses maps in the weighting of input data. Each node in the SOM network tries to be like the input that has been given to the network. SOM is also commonly referred to as Self Organizing Feature Maps, which means that SOM uses the principle of "features" or special features in its basic principles that make it different from other methods (Halim & Widodo, 2017; Alazzam et al., 2022).

Image is a two-dimensional matrix that is produced from a continuous two-dimensional analog image into a discrete image through a sampling process. Digital image processing is image processing using computer equipment. This is done so that it can be easily interpreted by humans or machines (Utari et al., 2019; Wang et al., 2022).

Based on the problems above, the author tries to make software about the health detection of green betel leaves to determine whether green betel leaves are suitable for use as medicine or not based on color by using Self Organizing Maps (SOM).

2. Research Methods

This survey methodology is based on PICOC (Population, Intervention, Comparation, Outcomes, Context) as an identification of information needs from previous research sources in table 1 PICOC Review.

Table 1 - Review PICOC

HEALTH DETECTION OF BETAL LEAVES USING SELF-ORGANIZING MAP AND THRESHOLDING ALGORITHM	
Population	Health Detection of Betal Leaves Using Self-Organizing Maps and Thresholding Algorithm
Intervention	1. How is the application of preprocessing on the sample image which is then used as a dataset for the SOM training process? 2. How the right number of datasets can get the best detection accuracy? 3. What are the results of the health detection of betel nut from the model obtained?
Comparation	Self-organizing map
Outcomes	Detection accuracy in digital images
Context	Datasets used by the public.

In searching for research sources before the stages in obtaining research journals that have been published, a Study Selection Strategy is carried out, namely carrying out the following stages:

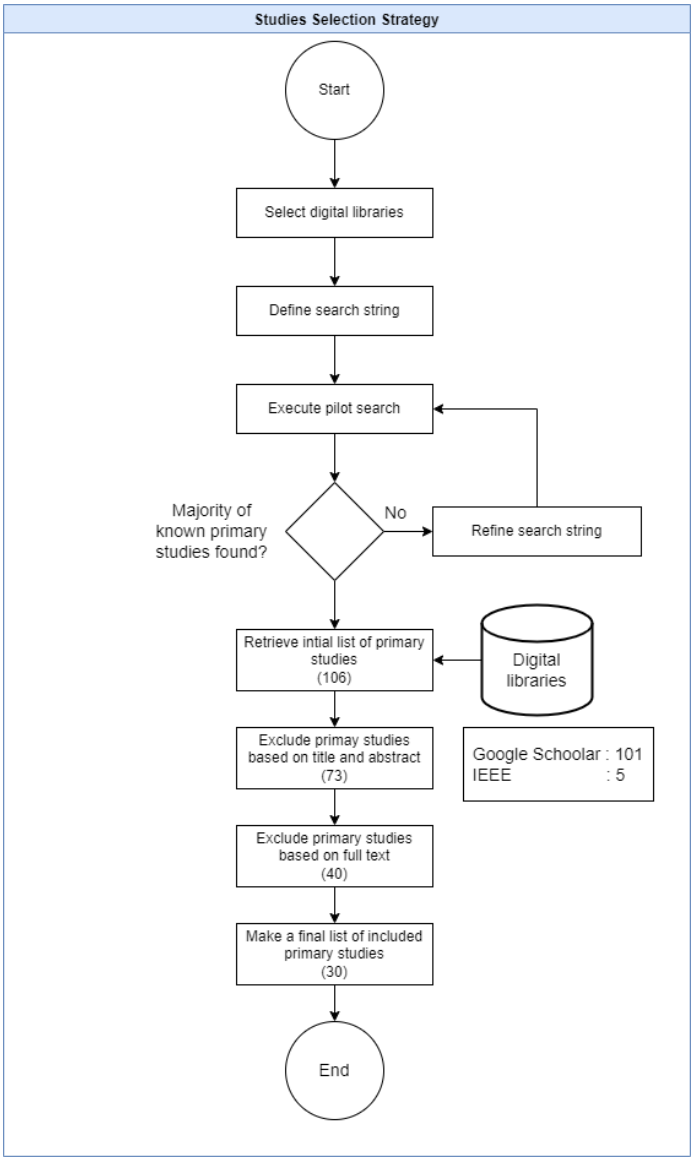

Fig 1. Studies Selection Strategy

The research methodology used in this study uses the Self-Organizing Maps algorithm. There are several stages in the process of identifying the health of betel leaf by applying the SOM (Self Organizing Maps) method, namely the preprocessing, feature extraction, and classification stages. The software used in the analysis of betel leaf image data is Matlab software(Nayak et al., 2022). The stages carried out in this research are as follows:

2.1 Data Sheet

The process of collecting data in this study used the scraping method. The scraping method used is to download images related to the object in this study, namely the betel leaf images in this research variable which were obtained on the google image site and the https://www.kaggle.com/datasets/lekhal/ site. bethel-leaves-diseased. As for the sample in this study, taking sick and healthy green betel leaf types with a total sample of 1500 images for train data and 450 images for testing data with 225 images each for test data for each healthy and sick green betel leaf type.

Table 2 - Research Variable

No	Variable	Train	Test	Definition
1	Sick	750	225	Image of a sick green betel leaf
2	Healtt	750	225	Image of a healthy green betel leaf

Source: Image Processed by Own Data from Data Processing

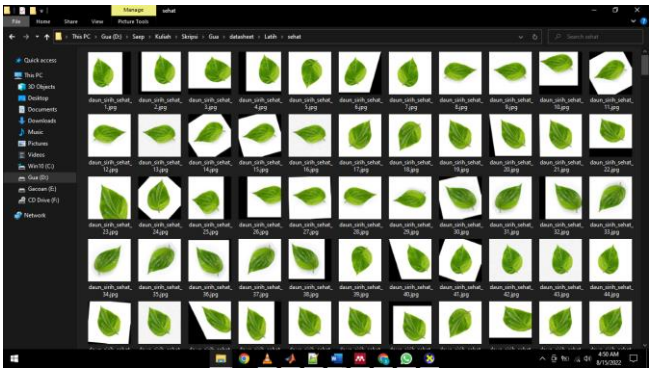

Fig 2. Image Download Results

Source: Public Data

2.2 Test Design

In this study, the Self-Organizing Maps model was used

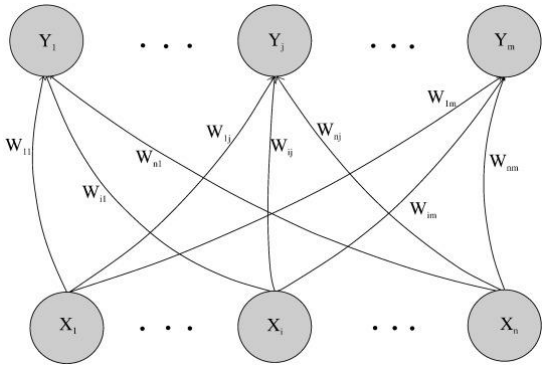

Fig 3. Classification of Self-Organizing Maps

Source: https://pemrogramanmatlab.com/2021/05/08/deteksi-kematangan-buah-sawit-menggunakan-self-organizing-maps-som/

The test design is carried out to evaluate each model generated by the Self-Organizing Map. The flow of this testing method is carried out in two stages, namely the training stage in Figure 4 and the testing stage in Figure 5. Both the training and testing processes go through the preprocessing and feature extraction stages before entering the classification stage. The relationship between the training and testing processes can be seen more clearly in the data flow diagram (DFD) level 1 which is shown in Figure 6 (Asri & Wulanningrum, 2021). The training stage is the stage where the Hold-out model. The distribution of training data and testing data is done randomly with a percentage of 80% for training data and 20% for testing data. For each scenario, 5 trials were carried out. The training stage is the stage where the SOM model is tested with the provided training data. The amount of training data provided is 1500 betel leaf images. The training data is then divided into two, namely training and validation. At this stage the model is tested with different images with the aim of testing whether the model has produced good performance and accuracy values in classifying an image.

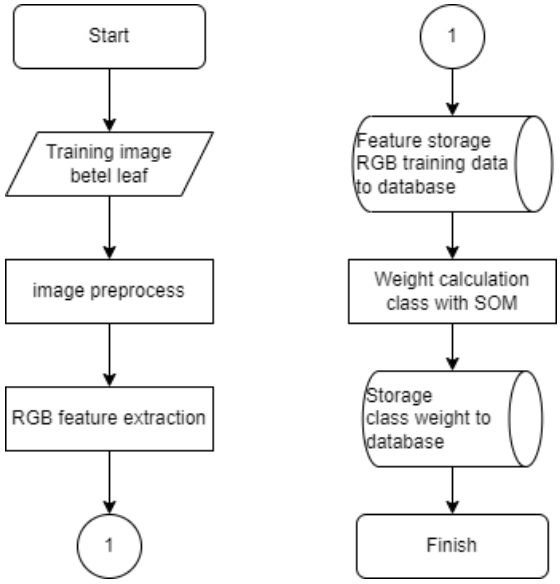

Fig 4. Training process diagram

Source: Image Processed by Own Data from Data Processing

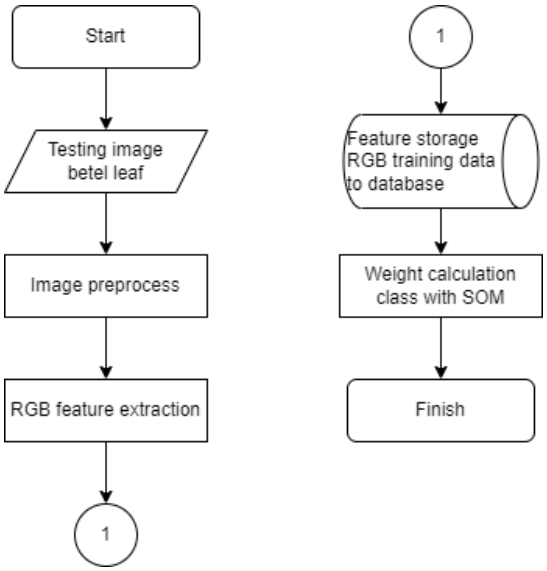

Fig 5. Testing process diagram

Source: Image Processed by Own Data from Data Processing

The user enters the betel leaf image training data folder to be used in the training process. Furthermore, the training data will go through a preprocessing and feature extraction process so that the RGB feature value of the training data is obtained. The value of the RGB feature is used in the SOM training process so as to produce weight data for healthy and sick classes.

Users also enter an image of a betel leaf as testing data. The testing data will go through a preprocessing and feature extraction process to produce RGB feature values from the testing data. Using the RGB feature value of the testing data and the weight data of healthy and sick classes obtained as the output of the training process, the SOM testing process will be carried out so that the healthy and sick classes from the testing data can be identified.

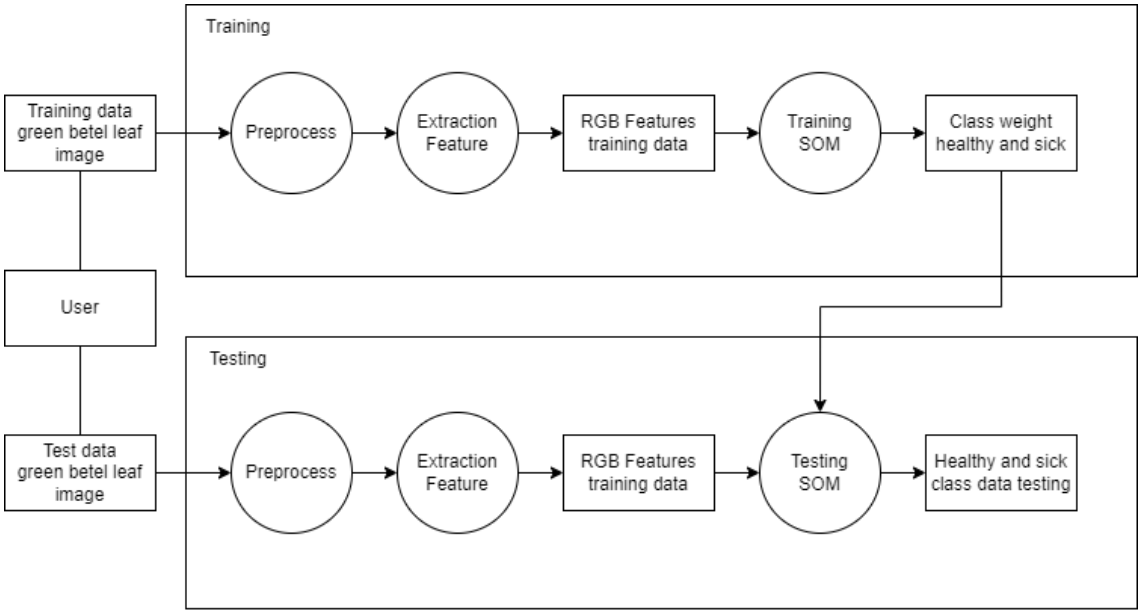

Fig 6. DFD level 1 detection system for green betel leaf health using the SOM method

Source: Image Processed by Own Data from Data Processing

4. Results and Discussions

Testing with a GUI using one image in determining the image processing, the user is asked to enter the address of the folder containing the tomato image for training data by pressing the "Open Image" button, then pressing the "Segmentation" button to view the segmentation results, pressing the "Feature Extraction" button to see feature extraction based on 6 features, namely RGB, and HSV, then press the "Classification" button to see the health results of green betel leaves.

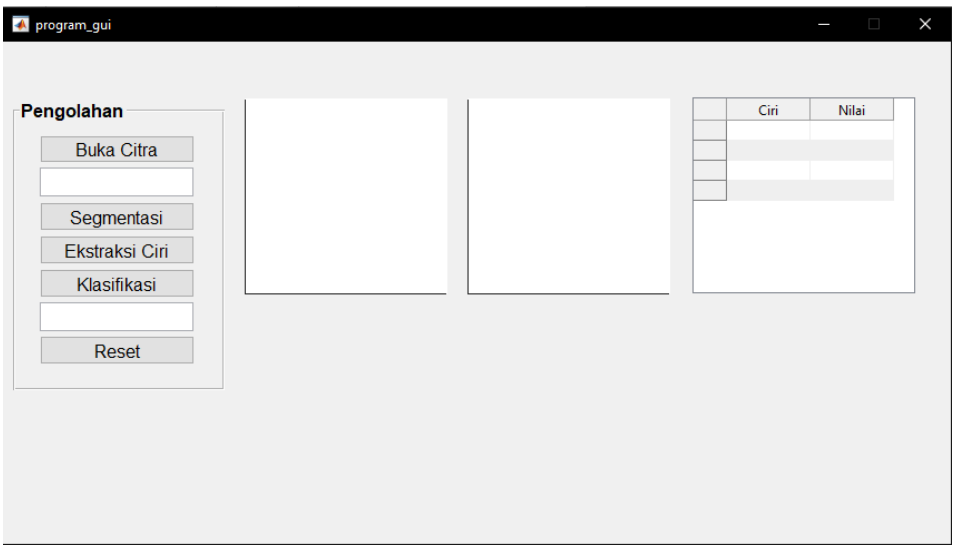

Fig 7. Interface Program

Source: Image Processed by Own Data from Data Processing

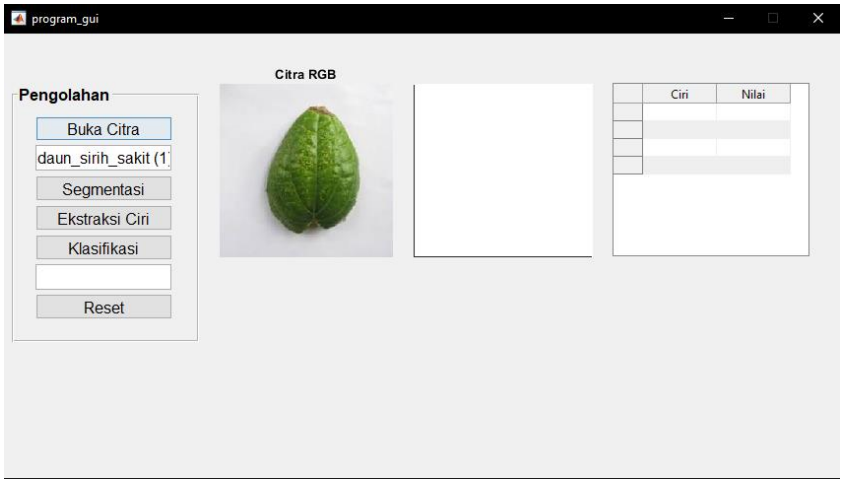

Fig 8. Open Image

Source: Image Processed by Own Data from Data Processing

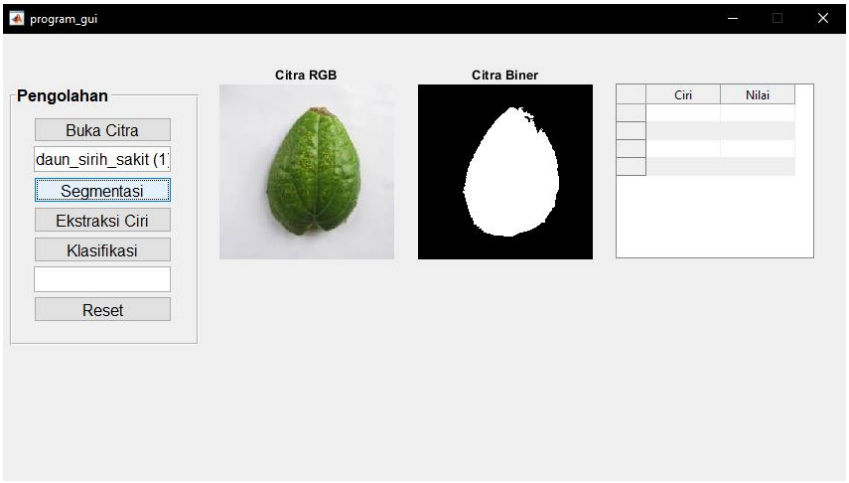

Fig 9. Results of Segmentation

Source: Image Processed by Own Data from Data Processing

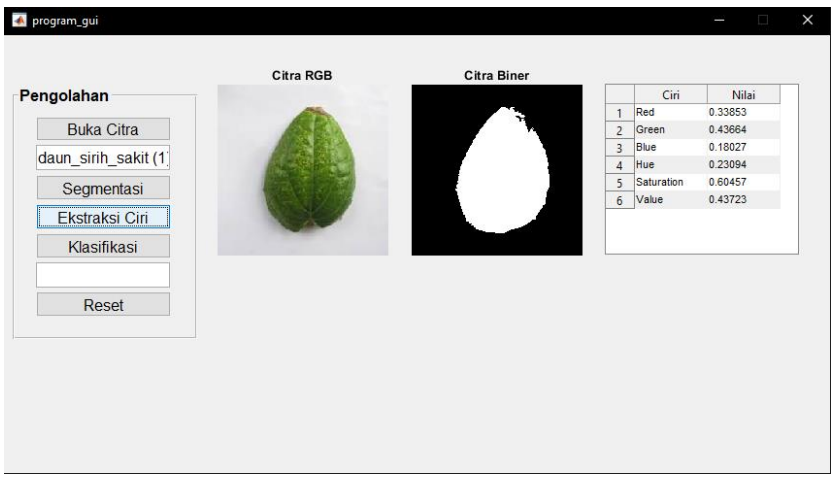

Fig 10. Determining Feature Extract

Source: Image Processed by Own Data from Data Processing

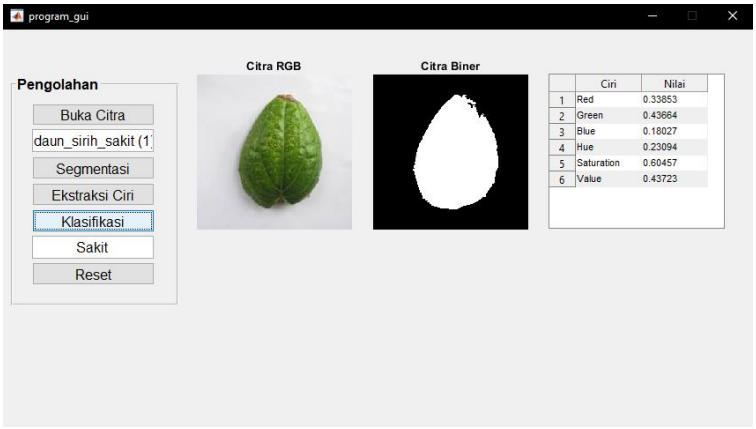

Fig 11. Result of Classification

Source: Image Processed by Own Data from Data Processing

To see the total images tested, the author uses a program with a script. The image below is the code used to see the evaluation of SOM performance, from the image it can be seen the evaluation of the accuracy generated from the test data with a training accuracy value of 96.86%.

```
170     % membaca nilai keluaran hasil pelatihan
171 -   [~,Groupmin] = min(Group,[],2);
172
173     % menghitung nilai akurasi pelatihan
174 -   akurasi = (sum(target_latih==Groupmin)/numel(target_latih))*100;
175
176     % menampilkan nilai akurasi pelatihan
177 -   disp(['Akurasi Pelatihan = ',num2str(akurasi),' %'])
178
```

```
Command Window
  Akurasi Pelatihan = 96.8667 %
fx >>
```

Fig 12. Training Accuracy Results

Source: Image Processed by Own Data from Data Processing

To see the total images tested, the author uses a program with a script, the image below is the code used to view the evaluation of SOM performance, from the image it can be seen that the evaluation of the accuracy generated from the test data with a test accuracy value of 97.20%.

```
156     % membaca nilai keluaran hasil pengujian
157 -   [~,Groupmin] = min(Group,[],2);
158
159     % menghitung nilai akurasi pengujian
160 -   akurasi = (sum(target_uji==Groupmin)/numel(target_uji))*100;
161
162     % menampilkan nilai akurasi pengujian
163 -   disp(['Akurasi Pengujian = ',num2str(akurasi),' %'])
164
```

```
Command Window
  Akurasi Pengujian = 97.2093 %
fx >>
```

Fig 13. Testing Accuracy Results

Source: Image Processed by Own Data from Data Processing

5. Conclusion

From the design results of this green betel leaf health detection, an application is produced that can detect the health of green betel leaf according to its class, namely healthy and sick classes. This application also makes it easier for users or users to sort betel leaves according to their color quickly, precisely, and efficiently so as to speed up the process of using them. From the results of the tests carried out, the accuracy level of 96.86% with training data is 1500, the accuracy level is 97.20% with testing data as much as 450 green betel leaves. The distribution of the amount of training data and testing data can affect the accuracy results, the shooting distance affects the results, and the effect of light is very vulnerable to the test results.

References

Adenugraha, S. P., Arinal, V., & Mulyana, D. I. (2022). Klasifikasi Kematangan Buah Pisang Ambon Menggunakan Metode KNN dan PCA Berdasarkan Citra RGB dan HSV. *Jurnal Media Informatika Budidarma*, *6*(1), 9. https://doi.org/10.30865/mib.v6i1.3287

Angulo-Saucedo, G. A., Leon-Medina, J. X., Pineda-Muñoz, W. A., Torres-Arredondo, M. A., & Tibaduiza, D. A. (2022). Damage Classification Using Supervised Self-Organizing Maps in Structural Health Monitoring. *Sensors*, *22*(4), 1484.

Ardiansah, F., Silawati, Yesi, Yuhani, & Fatayah Nur Isnani. (2021). Pelatihan Pembuatan Handsanitizer Alami Dengan Ekstrak Daun Sirih Dan Jeruk Nipis Untuk Mencegah Penyebaran Covid-19 Di Desa Cupat Parittiga. *Jurnal Abdimas Bina Bangsa*, *2*(1), 198–203. https://doi.org/10.46306/jabb.v2i1.69

Ashari, I. F., Banjarnahor, R., Farida, D. R., Aisyah, S. P., Dewi, A. P., & Humaya, N. (2022). Application of Data Mining with the K-Means Clustering Method and Davies Bouldin Index for Grouping IMDB Movies. *Journal of Applied Informatics and Computing*, *6*(1), 07-15.

Asri, P. P., & Wulanningrum, R. (2021). Implementasi SOM (Self Organizing Maps) untuk Identifikasi Kematangan Buah Tomat. *JTECS : Jurnal Sistem Telekomunikasi Elektronika Sistem Kontrol Power Sistem& Komputer*, *1*(2), 185–192.

Alazzam, M. B., Mohammad, W. T., Younis, M. B., Al Sayeh, A. M., Hajjej, F., AlGhamdi, A. S., & Rahman, M. A. (2022). Studying the effects of cold plasma phosphorus using physiological and digital image processing techniques. *Computational and Mathematical Methods in Medicine*, 2022.

Greener, J. G., Kandathil, S. M., Moffat, L., & Jones, D. T. (2022). A guide to machine learning for biologists. *Nature Reviews Molecular Cell Biology*, *23*(1), 40-55.

Halim, N. N., & Widodo, E. (2017). Clustering dampak gempa bumi di indonesia menggunakan kohonen self organizing maps. *Prosiding SI MaNIS (Seminar Nasional Integrasi Matematika Dan Nilai Islami)*, *1*(1), 188–194. http://conferences.uin-malang.ac.id/index.php/SIMANIS/article/view/62

Iswandana, R., & Sihombing, L. K. (2017). Formulasi, Uji Stabilitas Fisik, dan Uji Aktivitas Secara In Vitro Sediaan Spray Antibau Kaki yang Mengandung Ekstrak Etanol Daun Sirih (Piper betle L.). *Pharmaceutical Sciences and Research*, *4*(3), 121–131.

Juliansa hengki, D. sarjon, & Sumijan. (2017). Uji Validasi Algoritma Self Organizing Map (SOM) dan K-Mens untuk Pengelompokan Pegawai. *Resti*, *1*(1), 19–25.

Kania, S., Rachmatin, D., & Afgani Dahlan, J. (2019). Program Aplikasi Pengelompokan Objek Dengan Metode Self Organizing Map Menggunakan Bahasa R. *Jurnal EurekaMatika*, *7*(2), 17–29.

Khotimah, T., & Darsin, D. (2020). Clustering Perkembangan Kasus Covid-19 Di Indonesia Menggunakan Self Organizing Map. *Jurnal Dialektika Informatika (Detika)*, *1*(1), 23–26. https://doi.org/10.24176/detika.v1i1.5596

Malekloo, A., Ozer, E., AlHamaydeh, M., & Girolami, M. (2022). Machine learning and structural health monitoring overview with emerging technology and high-dimensional data source highlights. *Structural Health Monitoring*, *21*(4), 1906-1955.

Nafi'iyah, N., & Fatichah, C. (2017). *Fuzzy self organizing map untuk proses thresholding pada citra dental panaromic*. *September*, 511–524.

Nayak, P., Swetha, G. K., Kaushal, P., & Padhan, D. G. (2022). Cluster Formation Algorithm in

WSNs to Optimize the Energy Consumption Using Self-Organizing Map. In *IoT and Analytics for Sensor Networks* (pp. 11-22). Springer, Singapore.

Utari, E. L., Listyalina, L., & Puspaningtyas, D. E. (2019). Aplikasi self-organizing mapping sebagai alat deteksi anemia pada citra sel darah merah. *Jurnal Gizi Klinik Indonesia*, *16*(2), 64. https://doi.org/10.22146/ijcn.39560

Wang, R., Shi, T., Zhang, X., Wei, J., Lu, J., Zhu, J., ... & Liu, M. (2022). Implementing in-situ self-organizing maps with memristor crossbar arrays for data mining and optimization. *Nature communications*, *13*(1), 1-10.