Face Recognition using Self Organizing Map Based on Multi-Level Thresholding and Features Extraction

A Saleh¹*, NP Dharshini¹, F Azmi¹

Faculty of Technology and Computer Science, Prima Indonesia University, North Sumatra, Indonesia

*amirsalehnst1990@gmail.com

Abstract. Facial recognition has played a very important role in security systems in recent decades. Various studies have been conducted to find a good method for this recognition to be effective. The facial recognition process in this study uses the self-organizing maps (SOM) method which is based on Multi-Level Thresholding segmentation and extraction of image features. The obtained face image will be segmented using Multi-Level Thresholding and the face will be generated without involving the background. The image that has been segmented will be calculated the values of texture feature extraction using GLCM and color feature extraction using HSV. The resulting image will be carried out by a training and testing process for facial recognition using the self-organizing map (SOM) method. The test results in the proposed study showed good results with an introduction rate of 78.44%. This percentage is higher than face recognition without a segmentation process with a recognition rate of 65.78%. This shows that the proposed method has better performance for the face recognition process without involving the segmentation process.

1. Introduction

Facial recognition has many functions in various life, such as: bank security system, attendance, verification and so on [1]. The main function of the facial recognition system is to validate that the person doing the recognition has access rights to a system. The system is said to be reliable if face validation can be done appropriately during the recognition process. The face recognition process has a weak point where the image acquisition process by the camera involves the background around the face which causes it to be not optimal because the image is influenced by several factors, such as noise and environmental objects [2]. This will affect the results of facial recognition because the recognition process will affect the value of facial feature extraction. To overcome this, various methods have been offered such as face detection and segmentation. The process of face segmentation is necessary so that the characteristics of an image can be defined perfectly.

Face segmentation is a method used to detect facial parts so that the recognition process can be carried out [3]. There are several methods proposed for the image segmentation process, such as Fuzzy C-Means Clustering, which show good results. The segmentation research with Fuzzy C-Means Clustering shows good segmentation quality results in a short time [4]. This shows that the proposed method can effectively solve the image segmentation problem. Meanwhile, research with image
segmentation using adaptive k-means clustering shows the ease of segmentation process, accuracy, and effectiveness of the proposed method [5].

In this study, the segmentation method used is facial image segmentation using Multi-Level Thresholding. This process is carried out to obtain facial parts based on skin color. Then, the facial recognition process will be carried out using the Self Organizing Map (SOM) neural network method. Many studies have carried out the recognition process with SOM and obtained very good recognition accuracy. Research on facial recognition using the SOM network with DDCT features extraction obtained 100% pattern recognition results [6]. Meanwhile, research that examines facial expression recognition with a combination of feature extraction and self-organizing maps results in pattern recognition with an average recognition rate of 96.68% [7].

2. Materials

2.1. Face Recognition

Face recognition is a process to identify unknown face images using computational methods and compare it with previously stored facial data [8]. Face recognition has three processes, namely: face detection, feature extraction and facial image classification or face recognition.

2.2. SOM (Self Organizing Map)

The SOM network architecture has two layers (layers), namely: the input layer (x) and the output layer (y). Each neuron in the input layer is connected to every neuron in the output layer. Each neuron or node in the output layer represents a class of a given input. The SOM network architecture can be shown in Figure 2 below.

The steps carried out in the clustering and classification stages using the SOM network can be explained as follows [9].

1) Initialize a randomly selected initial weight.
2) Set the epoch value and learning rate (α).
3) Prepare training data (input vectors) from the dataset.
4) Calculate the proximity of the distance (j) between the input vector and the weight (wij) at each node. The weighting result that is close to the wij value is called the best matching unit (BMU). The equation used is as follows:

$$d_{ij} = \sqrt{\sum (x_{ij} - w_{ij})^2}$$  (1)
5) Update (change in weight) to the closest value based on the BMU value and make changes to the learning rate (α). The weight change equation can use the following equation:

\[
\begin{align*}
    w_{ij}(t + 1) &= w_{ij}(t) + \alpha(t)[x(t) - w_{ij}(t)] \\
    \text{Where,} \ t &= \text{Iteration,} \ x(t) &= \text{Vector input,} \ \alpha(t) &= \text{Learning rate,} \ (t) &= \text{Weight}
\end{align*}
\]  

(2)

6) Perform the next calculation as in step 4 onwards based on the number of epochs set. Network learning will be stopped when the maximum epoch value is reached.

2.3. Image Segmentation

In the image processing process, sometimes not all parts of the image are needed. The process for separating the two parts generally uses the thresholding method. Before doing the thresholding process, the image is converted into grayscale form. This process converts the RGB (Red, Green, Blue) image into a gray image. In general, RGB images have high pixel color content values, making them difficult to analyze. RGB values contain 24 bits for each pixel, while grayscale values contain 8 bits for each pixel. This is the reason for converting the RGB value to grayscale so that it is easier to analyze.

The threshold process is carried out after the grayscale stage by using a threshold value that is used to change the pixel value in the gray image to black or white. If the pixel value in the gray image is greater than the threshold, then the pixel value will be replaced by 1 (white), on the other hand, if the pixel value in the gray image is smaller than the threshold, the pixel value will be replaced with 0 (black). This process is also called the binaryization process so that only black and white are displayed with values of 0 and 1.

2.4. Image Feature Extraction

In the feature extraction process, parameters that characterize the object to be recognized will be generated. One method that can be used is a statistical method based on the Gray Level Cooccurrence Matrix (GLCM). These characteristics include [10].

\[
\begin{align*}
    \text{Contrast} &= \sum \sum (i - j)^2 p(i,j) \\
    \text{Correlation} &= \frac{\sum \sum i\cdot j\cdot p[i,j]-\mu_i\mu_j}{\sigma_i\sigma_j} \\
    \text{Energy} &= \sum \sum p[i,j]^2 \\
    \text{Homogeneity} &= \sum \sum \frac{p[i,j]}{1+|i-j|}
\end{align*}
\]  

(3) \hspace{1cm} (4) \hspace{1cm} (5) \hspace{1cm} (6)

The color feature in the image is done by looking for the extraction of color features using the HSV method. Where this method is used to define RGB colors which are converted into Hue, Saturation, and Value values. The formula used is [11].

\[
\begin{align*}
    H (\text{Hue}) &= 60^0 \times \left( \frac{\frac{G-B}{\Delta} \mod 6}{\Delta} \right), \ C_{max} = R' \\
    S (\text{Saturation}) &= \begin{cases} 
    0 & , \Delta = 0 \\
    \frac{\Delta}{C_{max}} & , \Delta <> 0
\end{cases} \\
    V &= C_{max}
\end{align*}
\]  

(7) \hspace{1cm} (8) \hspace{1cm} (9)
3. Research Method

This study will test the multilevel thresholding segmentation process in image processing techniques for face recognition using the SOM network method. The research steps can be seen in Figure 3 as follows.

![Figure 3](image.png)

**Figure 3.** Face recognition process with SOM and Multi-level Thresholding.

In the picture above tells the face recognition process of the SOM network based on the multi-level thresholding segmentation. The complete process can be seen from the description as follows:

1. **Image database**, is a face image obtained from the internet with 450 images with an image resolution size of 640 x 480 px.
2. **Image segmentation**, is the pre-processing stage which is carried out by image processing. This process will use multilevel thresholding techniques to find the face part of the input image. The results of this process can be divided into two, namely:
   a. **Face image**, is part of the image which is part of the face of an image. This image will be stored as input for learning and training the SOM network.
   b. **Background image**, is an unnecessary part of the facial recognition process.
3. **Features extraction**, is the stage of taking the values of facial image characteristics that have been done previously. This process is carried out by the following methods:
   a. **GLCM**, is the process of extracting the features of an image texture. With this process, 4 features will be generated, namely: contrast, correlation, energy and homogeneity.
   b. **HSV (Hue Saturation Value)**, is the process of extracting image color features. This process will produce 4 features, namely: Hue, Saturation, and Value.
4. **Image training**, is an image of learning or training obtained from the extraction of facial features.
5. **Image testing**, is a test image obtained from the extraction of facial image features. This process is also carried out to determine the performance of the method proposed in this study.
6. **Recognition using SOM (Self Organizing Map)**, is the facial recognition stage with the SOM network. This stage is the process of grouping data according to the closest proximity...
between the given image input and the resulting image output. The closer the data is, the stronger the data will be grouped into the same data group.

7. Result (Face recognition accuracy), is a process to calculate the accuracy of the calculation data from the proposed method with the actual results. This is done to determine how influential the method has been so that conclusions can be drawn.

4. Result and discussion

4.1. The results of testing the accuracy of the SOM network without segmentation

This test is conducted to determine the accuracy of facial recognition using the SOM network without involving the segmentation process on facial images. The image that has been obtained will be calculated the value of feature extraction using GLCM-based texture feature extraction and HSV-based color feature extraction. These two values are used as network inputs to be learned on the network to get the final weight of the network. Before entering the network, firstly, a data for each image class is selected to be used as the initial weight of the network that is randomly or randomly selected and the weight of the specified epoch is updated. The results of facial recognition using the SOM network without involving the segmentation process can be seen in table 1 below.

| Information       | Testing Image |
|-------------------|---------------|
| The amount of data| 450           |
| Correct Data      | 296           |
| Accuracy          | 65.78%        |

From table 1 above, it can be seen the accuracy of the SOM network testing without the segmentation process for facial recognition. The epoch determined in the network learning process is 100 and alpha (α) is 0.001 with a reduction of 0.5.

4.2. The results of testing the accuracy of the SOM network with Multilevel Thresholding Segmentation

This test is conducted to determine the accuracy of facial recognition using the SOM network by involving the Multi-Level Thresholding segmentation process on facial images. The results of facial recognition using the SOM network by involving the Multi-Level Thresholding segmentation process can be seen in table 2 below.

| Information       | Testing Image |
|-------------------|---------------|
| The amount of data| 450           |
| Correct Data      | 353           |
| Accuracy          | 78.44%        |

From table 4.2 above, it can be seen a comparison between training accuracy and SOM network testing accuracy with the Multi-Level Thresholding segmentation process for face recognition. The epoch determined in the network learning process is 100 and alpha (α) is 0.001 with a reduction of 0.5. In the training process, the final weight generated will be stored in the network and will be used again for the SOM network testing process.

4.3. Research Discussion

The testing process is carried out by comparing the same SOM network parameters, namely: network input (face image), epoch, alpha (α), and dec alpha (α) must be the same so that the comparison of methods without segmentation and with segmentation can be compared properly. Setting the value is
done by testing with the appropriate parameters. This parameter value is the reference in the network so that it can be seen to what extent the image segmentation process affects the accuracy of facial recognition using the SOM network. Comparison of the percentage of facial recognition accuracy can be seen in table 3 below.

Table 3. Face recognition accuracy comparison.

| Method                     | Accuracy |
|----------------------------|----------|
| SOM                        | 65.78%   |
| SOM - Multilevel Thresholding | 78.44%   |

From table 3 above, the facial recognition process with the SOM network requires a segmentation process so that images can be recognized properly. The process to produce features on an image must be done properly, such as involving segmentation for image processing. This is done so that the image that will be recognized only involves the object under study during the image processing. Image without segmentation will produce feature readings for all parts of the image. This process will involve reading the background (background) and characterizing a digital image. Environmental influences such as background will affect the extraction results of image features and have an impact on decreasing pattern recognition accuracy. So, the segmentation process will separate the object and background from an image.

5. Conclusion

The results of the SOM network training for facial recognition without segmentation obtained an accuracy of 65.78% on network testing. Face recognition using the SOM method with Multi-Level Thresholding segmentation obtained a better accuracy of 78.44% for the network testing process. Thus, the conclusion is that the segmentation method using Multi-Level Thresholding is better than facial recognition without using image segmentation with an increase in accuracy of 12.66% for network testing.

6. References

[1] Parmar, D. N & Mehta, BB 2013. Face Recognition Methods & Applications. International Journal Computer Technology & Applications, Vol 4 (1), 84-86
[2] Jun, K. & Xiaorui, W. 2018. Study on Image Recognition Based on PCNN and LVQ. IEEE International Conference on Automation, Electronics and Electrical Engineering
[3] Abbadi, NKE & Qazzaz, AAA 2015. Detection and Segmentation of Human Face. International Journal of Advanced Research in Computer and Communication Engineering
[4] Zhang, S. 2014. An Approach of Color Image Segmentation Based on Fuzzy Clustering. International Conference on Audio, Language and Image Processing
[5] Zheng, X., Lei, Q., Yao, R., Gong, Y. & Yin, Q. 2018. Image segmentation based on adaptive K-means algorithm. EURASIP Journal on Image and Video Processing
[6] Jadon, A., Kumar, M. & Rathi, Y. 2015. Face Recognition Using Som Neural Network With Ddct Facial Feature Extraction Techniques. International Conference on Communications and Signal Processing (ICCSSP)
[7] Farooq, F., Ahmed, J. & Zheng, L. 2017. Facial Expression Recognition using Hybrid Features and Self-Organizing Maps. Proceedings of the IEEE International Conference on Multimedia and Expo (ICME)
[8] Singh, A. & Singh, SK 2013. Effect Of Face Tampering On Face Recognition. Signal & Image Processing: An International Journal (SIPIJ) Vol.4, No.4
[9] Lumchanow, W. & Udomsiri, S. 2017. Chicken embryo development detection using Self-Organizing Maps and K-mean Clustering. International Electrical Engineering Congress
[10] Zulfaezal, M., Azemin, C., Izzuddin, M., Tamrin, M., Hilmi, MR, & Kamal, KM (2015). GLCM Texture Analysis on Different Color Space For Pterygium Grading, 10 (15), 6410–6413.
[11] Deswal, M., & Sharma, N. (2012). A Fast HSV Image Color and Texture Detection and Image Conversion Algorithm. International Journal of Science and Research (IJSR).