





# Comparison of Backpropagation and Kohonen Self Organising Map (KSOM) Methods in Face Image Recognition

Lady Silk Moonlight<sup>1)\*</sup> , Fiqqih Faizah<sup>2)</sup> , Yuyun Suprpto<sup>3)</sup> , Nyaris Pambudiyatno<sup>4)</sup> 

*Politeknik Penerbangan Surabaya, Indonesia*

*Jemur Andayani 1 No 73, Surabaya*

<sup>1)</sup>lady@poltekbangsby.ac.id, <sup>2)</sup>fiqqihfaizah@poltekbangsby.ac.id, <sup>3)</sup>yuyunsuprpto@poltekbangsby.ac.id, <sup>4)</sup>nyaris@poltekbangsby.ac.id

---

## Abstract

**Background:** Human face is a biometric feature. Artificial Intelligence (AI) called Artificial Neural Network (ANN) can be used in recognising such a biometric feature. In ANN, the learning process is divided into two: supervised and unsupervised learning. In supervised learning, a common method used is Backpropagation, while in the unsupervised learning, a common one is Kohonen Self Organizing Map (KSOM). However, the application of Backpropagation and KSOM need to be adjusted to improve the performance.

**Objective:** In this study, Backpropagation and KSOM algorithms are rewritten to suit face image recognition, applied and compared to determine the effectiveness of each algorithm in solving face image recognition.

**Methods:** In this study, the methods used and compared in the case of face image recognition are Backpropagation dan Kohonen Self Organizing Map (KSOM) Artificial Neural Network (ANN).

**Results:** The smallest False Acceptance Rate (FAR) value of Backpropagation is 28%, and KSOM is 36%, out of 50 unregistered face images tested. While the smallest False Rejection Rate (FRR) value of Backpropagation is 22%, and KSOM is 30%, out of 50 registered face images. The fastest time for the training process using the backpropagation method is 7.14 seconds, and the fastest time for recognition is 0.71 seconds. While the fastest time for the training process using the KSOM method is 5.35 seconds, and the fastest time for recognition is 0.50 seconds.

**Conclusion:** Backpropagation method is better in recognising face images than KSOM method, but the training process and the recognition process by KSOM method are faster than Backpropagation method due to the hidden layers.

**Keywords:** Artificial Neural Network (ANN), Backpropagation, Kohonen Self Organizing Map (KSOM), Supervised learning, Unsupervised learning

**Article history:** Received 12 July 2021, first decision: 18 August 2021, accepted 29 September 2021, available online 28 October 2021

---

## I. INTRODUCTION

Face images, fingerprints and voices are biometric features that can be used as authentication methods. With a computer-based biometric recognition system, digital images containing biometric features with certain characteristics are processed using the Artificial Intelligence (AI) to identify a person. Artificial neuronal network (ANN) is a method that represents the working system of human brain by simulating the learning process. It can solve digital image processing problems in identification, classification, authentication, optimisation, diagnostics and approximation [1]. ANN is the best recognition method, compared to fuzzy and other hybrid techniques [2]. ANN can learn from experience, generalise the examples it gets and abstract the input characteristics, even for irrelevant data. It is important to note that ANN is not programmed to produce a specific output. All outputs or conclusions drawn by the network are based on their experiences during the learning process. In the learning process, the input and output patterns are inserted into the ANN, then the network will be taught to provide acceptable answers. [3].

In ANN, the learning process methods are divided into two: supervised and unsupervised learning. Supervised learning takes place in artificial neural networks where the expected output value is known in advance. In the learning or training process, the resulting output patterns will be compared with the target output patterns. If the difference is too large, learning will be carried out again, until it reaches the minimum tolerable error value [4]. Unsupervised

---

\* Corresponding author

learning takes place in artificial neural networks that does not require an output target. The purpose of this lesson is to group units that are almost the same or to classify patterns [5].

Face image recognition is a problem that is often encountered in previous studies, even using the Backpropagation or KSOM method. Therefore, in this research, the Backpropagation and KSOM algorithms were rewritten with sequential steps that had been adapted to the application of face image recognition. The Backpropagation method was chosen because it is one of the most popular supervised learning ANN methods, with excellent performance. In previous research, the success rate of face image recognition using the Backpropagation method was 98% [6]. Also, other research concluded that the Backpropagation Method with improvised Scaled Conjugate Gradient (SCG) generates average recognition rate of 93% and an average number of iterations of 815 [7]. The KSOM method was chosen because it is one of ANN the most widely used unsupervised learning methods, with a high accuracy value, and a short recognition time process. In previous studies, the KSOM method has been able to recognise face images with an accuracy value of 98% [8] within 0.008–0.03 sec [9].

The problem in this research is how to apply the steps of the Backpropagation and KSOM algorithms in terms of face image recognition, analysing the performance of each method measured based on FAR, FRR and the success rate in recognising face images. The novelty in this research compared to other studies is that the performance of the Backpropagation and KSOM methods is measured based on the success rate based on the level of face tilt (turning face), face expressions and accessories used, and the time needed to carry out the learning and recognition process. Secondly, the novelty in this research is that the face image database used is a combination of the database from The Olivetti Research Laboratory (ORL) Database, and face images taken from a web camera. Thirty face image identities were obtained from ORL dataset, and 20 face image identities were taken using a web camera. The total is 50 face image identities, with 10 face images in each sample, so the total is 500 face images.

## II. LITERATURE REVIEW

### A. Digital Image

Digital image is a matrix consisting of rows and columns, where the value of each row and column index represents a point on the image. True colour image is an image representation that has three main component values, namely red, green and blue (RGB). Each component in the true colour image has 256 possible values, so that the overall true colour image has a total of 16,777,216 possible colours. Grayscale images are also called 8-bit images because they have 28 (256) possible values for each pixel. The values start at zero for black and 255 for white. The image that will be used and processed by the system is a grayscale image. Therefore, true coloured image must be converted into grayscale image. To change a true coloured image to grayscale, it takes the values of R, G, B from the true coloured image. The values start at zero for black and 255 for white.

### B. Artificial Neural Network (ANN)

Artificial neuronal network (ANN) is a computer science method that represents the working system of human brain by simulating the learning process. ANN can solve digital image processing problems in identification, classification, authentication, optimization, diagnostics and approximation [1]. ANN is the best recognition method, compared to fuzzy and other hybrid techniques [2]. ANN can learn from experience, generalise the examples it gets and abstract the input characteristics, even for irrelevant data. ANN is not programmed to produce a specific output. All outputs or conclusions drawn by the network are based on their experiences during the learning process. In the learning process, input (and output) patterns are inserted into the ANN, then the network will be taught to provide acceptable answers.

### C. Supervised Learning

Supervised learning is a learning method in artificial neural networks, where the expected output value is known in advance. In the learning or training process, the resulting output patterns will be compared with the target output patterns. If the difference is too large, learning will be carried out again, until it reaches the minimum tolerable error value. [5]

### D. Unsupervised Learning

Unsupervised learning is a learning method in artificial neural networks that does not require an output target. The purpose of this lesson is to group units that are almost the same or to classify patterns. [6]

### III. METHODS

#### A. Convert True Coloured Image to Grayscale

The conversion of true coloured image to grayscale image changes the pixel value which originally had 3 values, namely R, G, B into one value, namely gray. The following (1) is used to get the gray value [10]:

$$K_i = W_R R_i + W_B B_i + W_G G_i \quad (1)$$

where  $K_i$  is the gray value in the  $i^{\text{th}}$  pixel,  $W_R$  is the weight for the red element,  $W_B$  is the weight for the blue element,  $W_G$  is the weights for the green colour elements,  $R_i$  is the intensity value of the red colour element,  $B_i$  is the intensity value of the blue colour element, and  $G_i$  is the green element intensity value.

NTSC (National Television System Committee) defines the weights for converting true coloured images to greyscale as follows:  $W_R=0.299$ ,  $W_B=0.587$ ,  $W_G=0.114$ . Input data is in the form of true coloured images and the output data is in the form of grayscale images.

#### B. Image Size Normalisation

Image normalisation is performed to uniform the size of the face image to be processed. The size used is 92x112 pixels, according to the standard 'The ORL Database of Faces', AT&T Laboratories Cambridge. The normalisation process is carried out using the bilinear interpolation method. [11]

A digital image with dimensions of  $m \times n$  pixels is defined as a two-dimensional matrix, with the value  $ya[1 \dots m][1 \dots n]$  which contains the grayscale level of the pixels in the image. Location information for these pixels is stored in the  $x_1a[1 \dots m]$  and  $x_2a[1 \dots n]$  arrays. The relation between these values can be written as (2):

$$ya_{[j][k]} = y(x_1a_{[j]}, x_2a_{[k]}) \quad (2)$$

By interpolation, we estimate the value of the function  $y(x_1, x_2)$  at a location  $(x_1, x_2)$ , where  $x_1$  and  $x_2$  are not integer values between  $1 \dots m$  and  $1 \dots n$ . The formula for performing bilinear interpolation are (3) (4) and (5):

$$t = \frac{x_1 - x_1a_{[j]}}{x_1a_{[j+1]} - x_1a_{[j]}} \quad (3)$$

$$u = \frac{x_2 - x_2a_{[k]}}{x_2a_{[k+1]} - x_2a_{[k]}} \quad (4)$$

and

$$y(x_1, x_2) = ((1-t)(1-u)y_1) + (t(1-u)y_2) + tuy_3 + ((1-t)uy_4) \quad (5)$$

where

$$y_1 = ya_{[j][k]}, y_2 = ya_{[j+1][k]}, y_3 = ya_{[j+1][k+1]}, \text{ and } y_4 = ya_{[j][k+1]} \quad (6)$$

are the points around the to-be-interpolated point.

In the normalisation process, the input image is a Bitmap image with various sizes and the output image is a Bitmap image with a size of 92x112 pixels.

#### C. Binary and Bipolar Sigmoid Activation Functions

The results of initialisation process are converted into a range  $[1, -1]$  with the bipolar sigmoid function formula to make it easier and faster in the training process using the Backpropagation algorithm as shown in Fig. 1 [12] [13].

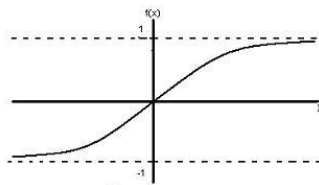


Fig. 1 Activation Function of Bipolar Sigmoid

Equation (7) for the conversion process is:

$$y = f(x) = \frac{2}{1+e^{-x}} - 19 \quad (7)$$

whereas in training using the KSOM algorithm, the binary sigmoid function is used which has a value range of 0 to 1. Fig. 2 shows the activation function of binary sigmoid [12] [14].

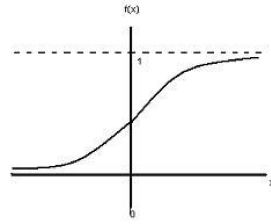


Fig. 2 Activation Function of Binary Sigmoid

Equation (8) for the conversion process is:

$$y = f(x) = \frac{1}{1+e^{-x}} \quad (8)$$

#### D. Discrete Cosine Transform (DCT)

Discrete Cosine Transform (DCT) is an image transformation based on cosine. The properties of DCT are the frequency in this transformation is real, orthogonal, separable, with an efficient computation process. In this system, two-dimensional (2D) DCT is used to extract features or obtain certain features in face images and also to reduce dimensions in face images to speed up the training process [15] [16].

The 2D DCT process is formulated by (9):

$$F(u, v) = \frac{2}{\sqrt{MN}} \cdot \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x, y) \cdot \cos \left[ \frac{(2x+1)u\pi}{2N} \right] \cos \left[ \frac{(2y+1)v\pi}{2M} \right] \quad (9)$$

with  $\alpha(n)$  value is described in (10)

$$\alpha(n) = \frac{1}{\sqrt{2}}; n = 0 \quad (10)$$

$$\alpha(n) = 1; n \neq 0 \quad (11)$$

where  $u$  and  $v$  are the image pixel points,  $m$  is the image height, and  $n$  is the image width.

This retrieval method is caused by the DCT coefficient, which is scattered in the upper left corner and the lower the value is getting smaller and can be ignored. In accordance with the Backpropagation theory, training will be faster if the input value is within the range of the activation function. Since the activation function used is a bipolar sigmoid function, the input layer values will be converted between ranges [1, -1]. The conversion process can be formulated:

$$z_{(i)(j)} = 2 \frac{x_{(i)(j)} - a}{b - a} - 1 \quad (12)$$

where  $i = 0, 1, \dots$  is the sum of the DCT-1 coefficients,  $j = 0, 1, \dots$  is the number of data sets -1,  $z$  is the input layer,  $x$  is the DCT coefficient,  $a$  is the lowest DCT coefficient value, and  $b$  is the highest DCT coefficient value.

#### E. Viola-Jones Face Detection Algorithm

The Viola-Jones algorithm is the most widely applied method to detect face image with fast, accurate and efficient performance. The main processes of Viola-Jones are:

1. Haar-like feature selection  
 The image to be processed is classified based on the feature value, to separate the images that are not needed. In this process, the background in the image is not counted. There are three types of features that are used based on the number of rectangles contained in it: two, three, four rectangles.
2. Creating integral image  
 Integral image is a data structure and method that adds up the subset values in the image matrix.
3. AdaBoost training  
 The AdaBoost algorithm works to find features that have a high level of differentiation. This is done by evaluating each feature between the face and non-face parts that are considered as the best feature.
4. Cascading Classifier

The characteristic of the Viola-Jones method is the cascading classifier. Classification in this method consists of several levels and each level produces sub-images that are believed not to be part of the face. This is done because it is easier to assess sub-images that are not part of the face compared to the face. [17] [18]

#### F. Self-Organising Map (KSOM) Method

The Kohonen Self-Organising Map (KSOM), also known as Kohonen map/network learning method is an unsupervised learning method, so this network structure does not require an output target. It modifies the weight of an ANN without needing to determine the output for a specific input pattern. The advantage is that it enables the network to find solutions, thus making it more effective and efficient with pattern connections. The main disadvantage is that interpreting the output must be done correctly [2] [19].

The activation function used in KSOM is Sigmoid Binary (Logsig), so before entering the learning process, the input data must be changed so that the value is in the range 0 to 1. All face images will be trained according to the algorithm of this method to update the weights which will be the final weights used for the recognition process.

The initial values needed for entering the KSOM algorithm such as maximum epoch, learning rate/alpha ( $\alpha$ ), data training matrix, random initial weight matrix and alpha reduction ( $\delta$ ). The algorithm for the KSOM is described as follows: [14] [20]

- Initialisation: determine the initial weight random value of  $W_{ij}$
- Do this if the stop condition is FALSE
  - 1) For every  $j$ , equation (13) for calculating:

$$D_j = \sum_i (W_{ij} - X_i)^2 \quad (13)$$

- 2) Specify  $j$ , up to  $D(j)$  the smallest/minimum value
- 3) For unit  $j$ , for each  $i$ , equation (14) for counting:

$$W_{ij}(\text{new}) = W_{ij}(\text{old}) + \alpha (X_i - W_{ij}(\text{old})) \quad (14)$$

- 4) Equation (15) for improving learning rate:

$$\alpha(\text{new}) = \alpha(\text{old}) \cdot \delta \quad (15)$$

- Do this until the epoch value is reached or the condition test stops.

Equation (16) is used for calculating the Euclidian distance, in the matching or recognition process:

$$d = \sqrt{\sum_i^n (W_i - X_i)^2} \quad (16)$$

where  $d$  is the Euclidean distance,  $W_i$  is the weight of neuron  $i$  (final weight), and  $X_i$  is the vector input to  $X_i$ . The minimum Euclidian distance is the result of recognition that best matches the stored face image. The threshold is used to limit the Euclidean Distance in the matching/recognition process. While the similarity distance serves to limit the iteration of changes in weight so that the best weight is obtained, even though the maximum epoch has not been fulfilled.

#### G. Backpropagation Method

One of the most widely used artificial neural network training algorithms in the field of pattern recognition is Backpropagation. This algorithm is generally used in multi-layer feed-forward neural networks, which are composed of several layers and the signal is flowed in a unidirectional direction from input to output [21] [22]. The Backpropagation training algorithm basically consists of three stages:

- a. Input the value of the training data so that the output value is obtained
- b. Backpropagation of the error value obtained
- c. Weight connection adjustment to minimize error value

These three stages are repeated continuously until the desired error value is achieved. After the training is complete, only the first step is needed to utilise the neural network. Error information is propagated sequentially starting from the output layer and ending at the input layer, so this algorithm is named Backpropagation [13] [23] [24].

In training artificial neural networks using the backpropagation algorithm, the steps are as follows:

- a. Initialise the weights with a random value between -0.5 to 0.5
- b. Determine the learning rate ( $\alpha$ ).
- c. Specify the error tolerance value or threshold value (when using the threshold value as a stop condition) or the maximum set of epochs (when using the number of epochs as a stop condition).
- d. Perform the following steps as long as the stop condition has not been met (value FALSE)

1) For each pair of training patterns, do the following.

a) Feedforward

- a. Each input unit  $X_i$  (from the 1<sup>st</sup> unit to the n<sup>th</sup> unit in the input layer,  $i = 1, \dots, n$ ) sends an input signal to all units in the upper layer (to the hidden layer);
- b. For each unit in the hidden layer  $Y_j$  (from the 1<sup>st</sup> unit to the p<sup>th</sup> unit;  $j = 1, \dots, p$ ) the hidden layer output signal is calculated by applying the activation function to the sum of the input signals weighs  $X_i$ . In this study, the bipolar sigmoid activation function is used (17);

$$Y_j = f(V_{0j} + \sum_{i=1}^n X_i V_{ij}) \quad (17)$$

then sent to all overlay units.

- c. Each unit in the output layer  $Z_k$  (from the 1<sup>st</sup> unit to the m<sup>th</sup> unit,  $i = 1, \dots, n$ ;  $k = 1, \dots, m$ ) is calculated the activation function (18) of the  $Z_j$ -weighted sum of the input signals for this layer:

$$Z_k = f(W_{0k} + \sum_{j=1}^p Y_j W_{jk}) \quad (18)$$

then sent to all overlay units.

b) Backpropagation

- (1) Each unit of output  $Z_k$  (from 1<sup>st</sup> unit to m<sup>th</sup> unit  $j = 1, \dots, p$ ;  $k = 1, \dots, m$ ) receives a target pattern  $t_k$  and then the output layer error information ( $\delta_k$ ) is computed by (19).  $\delta_k$  is sent to the layer below it and is used to calculate the weight and bias correction ( $\Delta W_{jk}$  and  $\Delta W_{0k}$ ) between the hidden layer and the output layer. It is shown on (20) (21):

$$\delta_k = (T_k - Z_k) f'(W_{0k} + \sum_{j=1}^p Y_j W_{jk}) \quad (19)$$

$$\Delta W_{jk} = \alpha \cdot \delta_k \cdot Y_j \quad (20)$$

$$\Delta W_{0k} = \alpha \cdot \delta_k \quad (21)$$

- (2) For each unit in the hidden layer (from 1<sup>st</sup> unit to p<sup>th</sup> unit  $i = 1, \dots, n$ ;  $j = 1, \dots, p$ ;  $k = 1, \dots, m$ ) the layer error information is calculated by (22). Hidden ( $\delta_j$ ).  $\delta_j$  is then used to calculate the amount of weight and bias correction ( $\Delta V_{ij}$  and  $\Delta V_{0j}$ ) between the input layer and the hidden layer. It is shown on (23) (24):

$$\delta_j = (\sum_{k=1}^m \delta_k W_{jk}) f'(V_{0j} + \sum_{i=1}^n X_i V_{ij}) \quad (22)$$

$$\Delta V_{ij} = \alpha \cdot \delta_j \cdot X_i \quad (23)$$

$$\Delta V_{0j} = \alpha \cdot \delta_j \quad (24)$$

c) Weights and Bias Updates

- (1) For each unit of  $y_k$  output (from the 1<sup>st</sup> unit to the m-unit), the bias and weight ( $j = 0, \dots, p$ ;  $k = 1, \dots, m$ ) are corrected so that the new bias and weights become (25):

$$W_{jk}(new) = W_{jk}(old) + \Delta W_{jk} \quad (25)$$

From the 1<sup>st</sup> unit to the p-unit in the hidden layer, updates are also made to the bias and the weights ( $i = 0, \dots, n$ ;  $j = 1, \dots, p$ ) by the following formula (26):

$$V_{ij}(new) = V_{ij}(old) + \Delta V_{ij} \quad (26)$$

- 2) Calculate the MSE (Mean Squared Error) with the formula (27):

$$MSE = (T_k - Z_k)^2 \quad (27)$$

Difference of squares of  $T_k$  Output Target and  $Z_k$  output value.

- 3) Stop condition test: stop condition is TRUE, if the MSE value is more than or equal to the error tolerance value, or the number of Epochs has not exceeded the maximum epoch.  
 In the matching/recognition process, do the feedforward process using the final weights.

#### IV. RESULTS

The process of testing the Kohonen Self Organizing Map (KSOM) Neural Network, used training variables as follows: training rate ( $\alpha$ ) = 0.6; alpha reduction ( $\delta$ ) = 0.5; threshold = 0.02; and similarity distance = 0.000000000000001. The threshold function is to limit the Euclidean distance in the matching/recognition process. While the similarity distance serves to limit the iteration of changes in weight so that the best weight is obtained, even though the maximum epoch has not been fulfilled.

While the Backpropagation Neural Network testing process, used the following training variables: training rate ( $\alpha$ ) = 0.008, momentum ( $\mu$ ) = 0.02; error tolerance = 0.01; number of hidden layers = 30.

##### A. Flowchart

Fig. 3 illustrates the flowchart of the training process and face image recognition system.

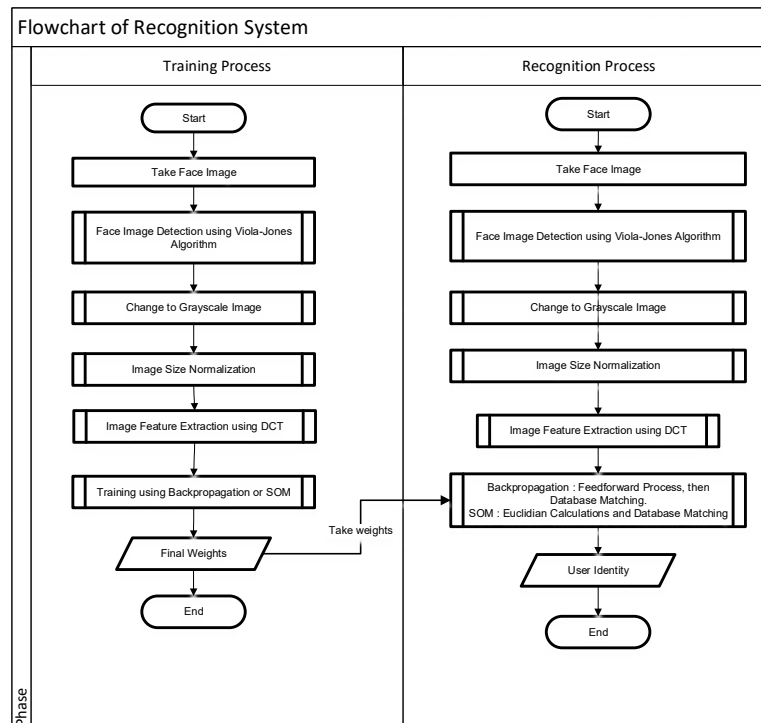


Fig. 3 Flowchart System

**B. Data for Training and Testing**

In the training process, 50 identities were used with 10 face images each, so that a total of 500 face images were used as training data. Meanwhile, during the testing process, the face image used is different from the stored face image in the training process. To calculate FAR, 50 face images—whose identities were not stored in the database and had different identities—were used as testing data. Meanwhile, to calculate FRR, 50 different face images—whose identities are stored in the database—were used.

The input image as the training data, is a face image, or an image that has a background. The image that had been taken was processed by face detection using the Viola Jones algorithm. Cropping was done to take a face image, by removing the background. Then, the true coloured image was converted to grayscale image. The grayscale image was normalised to a size of 92x112 pixels. Feature extraction was carried out to obtain image characteristics (59 values), which were used as input data in the input layer. Then the image was ready to be trained using the Backpropagation or KSOM algorithm. After the training process, the testing was carried out.

The face image database used is 500 face images, consisting of 50 identities with 10 face images for each identity. From the 50 identities, 30 identities came from ORL face dataset [11], and 20 identities were taken using a web camera. The 10 face images were taken with several tilt angles and face expressions. These were sample training/learning and testing images that had been cropped and converted to grayscale. Fig. 4 shows the images from ORL database; and Fig. 5 shows sample images taken from a web camera.



Fig. 4 Training and Testing Images taken from The ORL Database



Fig. 5 Training and Testing Images taken from Web Camera

Face images taken from a web camera were taken in a closed room with a room area of 12m<sup>2</sup> and a room height of 4m, with a lamp, namely LED Cool Daylight 6-Watt, brightness level of 470 Lumen. The distance between the camera and the face is 40-50cm. Because the distance and lighting levels have been limited, this system only allows to recognise skin colour according to the training image stored in the database. Table 1 is an example of the matrix used in the training process, adjusted for the number of input layers and the number of registered face image identities.

TABLE 1  
 MATRIX INPUT OF TRAINING PROCESS

Image	Layer Input (as many as the DCT coefficient)				
	1	2	3	.....	59
1	$W[i][j]$				
2					
.....					
500					

The database used in this system is simple. The tables used are the training table and the identity table. The identity table shown in Fig 6 is used to store the identity of the face image. While the training table shown in Fig 7 is used to store face images used in the training process.

Field Name	Data Type
id	Text
nama	Text
alamat	Text
pekerjaan	Text

Fig. 6 Identity Table

Field Name	Data Type
kodewajah	Text
id	Text
path	Text

Fig. 7 Training Table



### C. Application of Face Image Recognition System

Following are the results of the data input form shown in Fig 8, training form with Backpropagation shown in Fig 9, training form with KSOM shown in Fig 10 and face image recognition form shown in Fig 11.



Fig. 8 Identity and Face Input Form

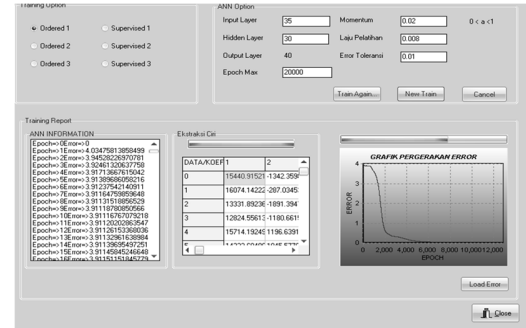


Fig. 9 Backpropagation Training Form

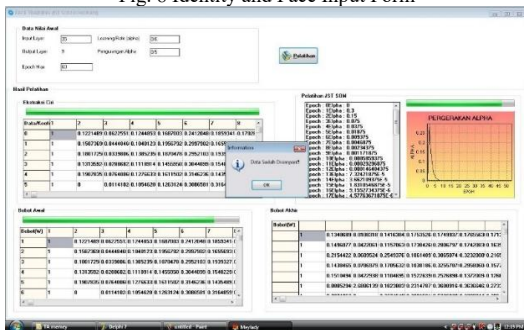


Fig. 10 KSOM Training Form



Fig. 11 Recognition Form

### D. Definition of FAR and FRR

In the case of matching or recognition, the accuracy of the system is measured by Accuracy Rate, False Acceptance Rate (FAR) and False Rejection Rate (FRR). False Acceptance (FA) occurs when the system accepts a face image whose identity is not recorded in the database. False Rejection (FR) occurs when the system rejects a face image whose identity is stored in the database. The success rate is the percentage of the system's success in recognising the right face, namely accepting face images stored in the database, or rejecting face images that are not stored in the database. The rate of Accuracy, FA and FR is calculated using (28), (29) and (30) [25] [26] [27]. Therefore, if the FAR or FRR is lower, then the success rate will be higher.

$$FAR = \frac{\text{number of FA}}{\text{number of test}} \times 100\% \quad (28)$$

$$FRR = \frac{\text{number of FR}}{\text{number of test}} \times 100\% \quad (29)$$

$$\text{Success Rate} = \frac{\text{the number of system success}}{\text{number of experiments}} \times 100\% \quad (30)$$

### E. FAR and FRR of Backpropagation based on the Number of Hidden Layers

In the testing process of determining the best hidden layer on the Backpropagation, the following training variables are used: input layer = 35, training rate ( $\alpha$ ) = 0.008, momentum ( $\mu$ ) = 0.02, error tolerance = 0.01. The face image used is in accordance with training and testing data. Table 2 shows the comparison of the calculation of the FAR and FRR values of the face image recognition system using the Backpropagation method based on the number of hidden layers. To calculate FAR, 50 face images whose identities were not stored in the database were used; and to calculate

different FRRs, 50 face images whose identities were stored in the database were used. From the testing results, it can be concluded that the best number of hidden layers in the face image recognition system using the Backpropagation method is 30 hidden layers.

TABLE 2  
 BACKPROPAGATION TESTING RESULTS BASED ON THE NUMBER OF HIDDEN LAYERS

Number of Hidden Layer	The Number of Image	FA	FAR	Success Rate	FR	FRR	Success Rate
20	50	17	34%	66%	15	30%	70%
<b>30</b>	<b>50</b>	<b>15</b>	<b>30%</b>	<b>70%</b>	<b>13</b>	<b>26%</b>	<b>74%</b>
40	50	16	32%	68%	14	28%	72%
50	50	18	36%	64%	15	30%	70%
60	50	17	34%	66%	18	36%	64%
70	50	19	38%	62%	17	34%	66%
80	50	20	40%	60%	18	36%	64%
90	50	20	40%	60%	18	36%	64%
100	50	21	42%	58%	20	40%	60%

F. Comparison of the Backpropagation and KSOM

The testing used the KSOM method, the training variables were used as follows: training rate ( $\alpha$ ) = 0.6; alpha reduction ( $\delta$ ) = 0.5; threshold = 0.02 and similarity distance = 0.000000000000001. Table 3 shows a comparison of the FAR and FRR of the face image recognition process using the Backpropagation or KSOM method. In the first experiment, to calculate FAR, 50 face images whose identities were not stored in the database were used; and to calculate different FRRs, 50 face images whose identities were stored in the database were used.

TABLE 3  
 FAR AND FRR OF BACKPROPAGATION AND KSOM METHOD

Method	FA	FAR	Success Rate	FR	FRR	Success Rate
Backpropagation	14	28%	72%	11	22%	78%
KSOM	18	36%	64%	15	30%	70%

From the test results, it can be seen that the FAR of Backpropagation is 28%, and the FAR of KSOM is 36%, while the FRR of Backpropagation is 22%, and the FRR of KSOM is 30%. The average success rate of the Backpropagation is 75%, and the average success rate of the KSOM is 67%. Because the FAR and FRR of Backpropagation are lower than KSOM, and the success rate of Backpropagation is higher than KSOM, so it can be concluded that the Backpropagation method is better in recognising face images than the KSOM method. In the second experiment, a recognition test was conducted based on the level of face tilt, such as tilting right, left, up and down. In this experiment, 20 face images were used whose identities were stored in the database. The result of the second experiment is shown in Table 4.

TABLE 4  
 SUCCESS RATE BASED ON FACE TILT LEVEL

Face Tilt Level	Backpropagation		KSOM	
	Number of Faces Recognized	Success Rate	Number of Faces Recognized	Success Rate
Right 15°	16	80%	15	75%
Right 30°	15	75%	14	70%
Right 45°	14	70%	13	65%
Left 15°	16	80%	15	75%
Left 30°	15	75%	14	70%
Left 45°	14	70%	13	65%
Below 15°	14	70%	14	70%
Below 30°	12	60%	11	55%
Below 45°	10	50%	10	50%
Above 15°	15	75%	14	70%
Above 30°	14	70%	13	65%
Above 45°	12	60%	12	60%

From the experiment results that is written in Table 5, it is known that the Backpropagation method has a higher success rate based on the level of face tilt than KSOM method. Face image with right or left tilt levels are easier to identify than up and down tilts. The degree of tilt to the left or right of 15° is the most easily recognized face tilt. In the

third experiment, a recognition test was conducted based on face expressions and the accessories used. In this experiment, 20 face images were used whose identities were stored in the database.

TABLE 5  
 SUCCESS RATE BASED ON EXPRESSION OR ACCESSORIES USED

Expression or Wearing Accessories	Backpropagation		KSOM	
	Number of Faces Recognized	Success Rate	Number of Faces Recognized	Success Rate
Smile	16	80%	15	75%
Laugh	16	80%	15	75%
Angry	16	80%	14	70%
Sad	16	80%	14	70%
Closed Eyes	14	70%	12	60%
Wearing Glasses	3	15%	2	10%
Wearing a Mask	0	0%	0	0%
Wearing a Hat	4	20%	2	10%

Based on face expressions, the average success rate of the Backpropagation method is 78%, while the average success rate of the KSOM method is 70%. The Backpropagation method can only recognise faces by wearing a hat with a success rate 20%, while the average success rate of the KSOM method is 10%. The Backpropagation method can only recognise faces by wearing a pair of glasses with a success rate 15%, while the average success rate of the KSOM method is 10%. Backpropagation and KSOM methods cannot recognise faces by using mask accessories. Therefore, it can be concluded that the success rate of the Backpropagation method is higher than KSOM method based on face expressions and the wearing a hat or glasses. In the fourth experiment, a recognition test was conducted based on the time of the training process and the recognition process. The training process is carried out by processing 500 face images from 50 identities that have been stored in the database; and the recognition process is carried out by processing the recognition of one face image.

TABLE 6  
 COMPARISON OF TRAINING AND RECOGNITION PROCESS TIME

Method	Time of Training Process (Second)	Time of Recognition Process (Second)
Backpropagation	00:07:14	00:00:57
KSOM	00:05:49	00:00:50

From the Table 6, it can be seen that the time for the training process using the backpropagation method is 7.14 seconds, and the time for recognition is 0.57 seconds. The time for the training process using the KSOM method is 5.49 seconds, and the time for recognition is 0.50 seconds. Therefore, it can be concluded that the training process and the recognition process of the KSOM method are faster than the Backpropagation Method. This is because in the Backpropagation method, there are hidden layers that demands the computation of the feed forward.

## V. DISCUSSION

Firstly, the face image is detected using the Viola-Jones Algorithm, and then cropping process on the face image [17]. After that, the true coloured image is converted to the grayscale image [10]. Then the grayscale face image is normalised, so that all face images have the same size. Feature extraction on face image is carried out using the DCT (Discrete Cosine Transform) method [16]. The Matrix result from the feature extraction process, becomes input data in the training process. Backpropagation or KSOM method is used as a training process, using predetermined variables. The variables used in the KSOM method are: training rate ( $\alpha$ ), alpha reduction ( $\delta$ ), threshold, similarity distance, and maximum epoch [20]. While the variables used in the Backpropagation Method are: training rate ( $\alpha$ ), momentum ( $\mu$ ), error tolerance, number of hidden layers, and maximum epoch [22]. After the training process, the recognition process can be carried out. If the training process uses the KSOM method, then the matching in the recognition process uses the Euclidean distance. If the training process uses the Backpropagation method, then the matching in the recognition process uses a feedforward process.

It is important to note that, instead of binary, this research uses Backpropagation with a bipolar sigmoid function, because the results of using a bipolar sigmoid function are more precise and suitable to be applied. The input layer values of the bipolar sigmoid function are converted between ranges [1, -1] [13]. Secondly, in the KSOM method, the activation function recommended is a binary sigmoid function, so that the input layer values are converted between ranges [0,1] [14]. Third, in the application of the KSOM method, determining the threshold value also greatly influences the matching/recognition process, because the threshold value is used as the limit of the Euclidean distance

in determining the identity of the image [9]. Determining the minimum value of the Similarity distance affects the number of iterations or the maximum epoch in determining the weight, so the iteration will stop when the similarity distance is reached even though the maximum input value for epoch has not been fulfilled [19]. Fourth, the limitations in taking face images using a camera are: the face image is taken from a web camera, taken in a closed room with an area of 12m<sup>2</sup> and a room height of 4m, with LED Cool Daylight 6-Watt, brightness level of 470 Lumen. The distance between the camera and the face is 40-50cm. Because the distance and lighting levels have been limited, this system only allows to recognise skin colour according to the training image stored in the database. If in the testing process, the image to be tested is an image using glasses, then one of the images using glasses must be entered into the training image database. As in the case with face expressions, and the level of face tilt, there must be an image of a face with several face expressions and the level of face tilt entered in the training image database [7].

## VI. CONCLUSIONS

Based on the FAR and FRR values of the hidden layer comparison by using the Backpropagation method, the best number of hidden layers is 30. Based on the FAR and FRR values of the face image recognition process, the Backpropagation method is better than the KSOM method. Based on the level of the face tilt to the right, left, up, and down, the Backpropagation method shows a higher success rate than the KSOM method. Face images with a tilt turn to the right or left are easier to identify than those with an upward and downward tilt. The degree of tilt to the left or right of 15° is the most easily recognised face tilt. Based on the face expression and use of accessories, the success rate of the Backpropagation method is higher than the KSOM method. However, the Backpropagation and KSOM methods cannot recognise faces with a mask. Based on the time of the training process and the recognition process, the KSOM method is faster than the Backpropagation Method.

**Author Contributions:** *Lady Silk Moonlight:* Conceptualization, Methodology, Writing - Original Draft, Writing - Review & Editing, Supervision, Software, Investigation, and Data Curation, *Fiqqih Faizah:* Writing - Original Draft and Writing - Review & Editing, *Yuyun Suprpto:* Investigation and Data Curation, *Nyaris Pambudiyatno:* Investigation and Data Curation.

**Funding:** This research received no specific grant from any funding agency.

**Acknowledgments:** We are grateful to Mr. Alfian Naufal, S.Kom and Mr. Mula'ab, S.Si, M.Kom for all the advice in completing the research.

**Conflicts of Interest:** The authors declare no conflict of interest.

## REFERENCES

- [1] A.-E. C. Pandelea, G. Covatariu and M. Budescu, "Image Processing Using Artificial Neural Networks," *Bulletin of the Polytechnic Institute of Jassy, Section*, no. Constructions Architecture, 2015.
- [2] O. I. Abiodun, K. V. Dada, A. Jantan, A. E. Omolara, H. Arshad, A. M. Umar, O. U. Linus, M. U. Kiru, A. A. Kazaure and U. Gana, "Comprehensive Review of Artificial Neural Network Applications to Pattern Recognition," *IEEE Access*, vol. 7, pp. 158820-158846, 2019.
- [3] A. Wosiak, A. Zamecznik and K. N. Jarosik, "Supervised and unsupervised machine learning for improved identification of intrauterine growth restriction types," in *Federated Conference on Computer Science and Information Systems (FedCSIS)*, Gdansk, Poland, 2016.
- [4] M. Seetha, I. V. Muralikrishna, B. L. Deekshatulu and B. L. Malleswari, "Artificial Neural Networks and Other Methods of Image Classification," *Journal of Theoretical and Applied Information Technology (JATIT)*, pp. 1039-1053, 2013.
- [5] A. Olaode, G. Naghdy and C. Todd, "Unsupervised Classification of Images: A Review," *International Journal of Image Processing (IJIP)*, vol. VIII, no. 5, pp. 325-342, 2014.
- [6] R. A. Priyanka, C. Ashwitha, R. A. Chakravarthi and R. Prakash, "Face Recognition Model Using Back Propagation," *International Journal of Engineering & Technology*, 2018.
- [7] Q. Huang and L. Cui, "Design and Application of Face Recognition Algorithm Based on Improved Backpropagation Neural Network," *Revue d'Intelligence Artificielle*, vol. 33, no. 1, pp. 25-32, 2019.
- [8] P. Bose and S. K. Bandyopadhyay, "Human Face Detection: Manual vs. Kohonen Self Organizing Map," 11 September 2020. [Online]. Available: <https://www.preprints.org/>.
- [9] V. Skuratov, K. Kuzmin, I. Nelin and M. Sedankin, "Application Of Kohonen Self Organizing Map to Search for Region of Interest in the Detection Interest in The Detection Of Objects," *EUREKA: Physics and Engineering*, 2020.

- [10] ihritik, "RGB Image to Grayscale Image Conversion," GeeksforGeeks, 25 June 2018. [Online]. Available: <https://www.geeksforgeeks.org/matlab-rgb-image-to-grayscale-image-conversion/>. [Accessed 18 September 2021].
- [11] KAGGLE, "The ORL database for training and testing: The Olivetti Research Laboratory (ORL) face dataset," KAGGLE, 2020. [Online]. Available: <https://www.kaggle.com/tavarez/the-orl-database-for-training-and-testing>. [Accessed 16 September 2021].
- [12] D. Gupta, "Fundamentals of Deep Learning – Activation Functions and When to Use Them?," Analytics Vidhya, 30 January 2020. [Online]. Available: <https://www.analyticsvidhya.com/blog/2020/01/fundamentals-deep-learning-activation-functions-when-to-use-them/>. [Accessed 2021 September 18].
- [13] L. S. Moonlight and A. S. Prabowo, "Forecasting System for Passenger, Airplane, Luggage and Cargo, Using Artificial Intelligence Method-Backpropagation Neural Network at Juanda International Airport," *Warta Ardhia Jurnal Perhubungan Udara*, vol. 45, no. 2, pp. 99-110, 2019.
- [14] L. S. Moonlight, "Sistem Pengenalan Wajah Berbasis Jaringan Syaraf Tiruan Self Organizing Map (SOM) Dengan Pemrosesan Awal Discrete Cosine Transform (DCT)," *Jurnal Penelitian Politeknik Penerbangan Surabaya*, vol. IV, no. 3, pp. 29-39, 2019.
- [15] H. Mohamed, "DCT-Based Image Feature Extraction and Its Application in Image Self-Recovery and Image Watermarking," The Department of Electrical and Computer Engineering, Concordia University, Canada, 2016.
- [16] Ö. Aydogdu and M. Ekinici, "An Approach for Streaming Data Feature Extraction Based on Discrete Cosine Transform and Particle Swarm Optimization," *Symmetry*, vol. 12, no. 299, 2020.
- [17] I. G. N. M. K. Raya, A. N. Jati and R. E. Saputra, "Analysis realization of Viola-Jones method for face detection on CCTV camera based on embedded system," in *International Conference on Robotics, Biomimetics, and Intelligent Computational Systems (Robionetics)*, Bali, Indonesia, 2017.
- [18] Ibrahim, "Analisis Akurasi Pengenalan Wajah menggunakan Algoritma Viola Jones dan Modified Self Organizing Map," Prodi S2 Teknik Informatika Universitas Sumatera Utara, Medan, 2020.
- [19] L. Tan, C. Li, J. Xia and J. Cao, "Application of Self-Organizing Feature Map Neural Network Based on K-means Clustering in Network Intrusion Detection," *Computers, Materials & Continua (CMC)*, vol. 61, no. 1, pp. 275-288, 2019.
- [20] S. Aly and S. Almotairi, "Deep Convolutional Self-Organizing Map Network for Robust Handwritten Digit Recognition," *IEEE Access*, vol. 8, pp. 10735-10745, 2020.
- [21] S. Rawat, "Face Recognition Using Back Propagation Neural Network," *International Journal of Innovations in Engineering and Technology (IJJET)*, vol. X, no. 4, pp. 125-132, 2018.
- [22] Yuhandri, E. Prasetyo, S. Madenda and Karmilasari, "Pattern Recognition and Classification using Backpropagation Neural Network Algorithm for Songket Motifs Image Retrieval," *International Journal on Advanced Science Engineering Information Technology*, vol. 7, no. 6, pp. 2343-2349, 2017.
- [23] A. N. Gomez, M. Ren, R. Urtasun and R. B. Grosse, "The Reversible Residual Network: Backpropagation Without Storing Activations," in *31st Conference on Neural Information Processing Systems (NIPS)*, CA, USA, 2017.
- [24] J. Brownlee, "How to Code a Neural Network with Backpropagation In Python (from scratch)," 7 November 2017. [Online]. Available: <https://machinelearningmastery.com/>. [Accessed 26 August 2021].
- [25] Recogtech, "FAR and FRR: Security Level Versus User Convenience," Recogtech, 6 December 2017. [Online]. Available: <https://www.recogtech.com/en/knowledge-base/security-level-versus-user-convenience>. [Accessed 20 September 2021].
- [26] A. A. Syukur, B. Pramadi and Y. Abdurrozaq, "Implementasi Webcam sebagai Pendeteksi Wajah pada Sistem Keamanan Perumahan Menggunakan Image Processing," *ELECTRICES*, vol. 2, no. 1, pp. 1-5, 2020.
- [27] S. Shaik and C. Konda, "Performance of Biometrics : False Acceptance Rate ( FAR ) False Rejection Rate ( FRR ) Image Classification Techniques : SVM Classifier K-Nearest Neighbour Classifier Naive Bayes Decision Trees," *International Journal of Scientific Development and Research (IJS DR)*, vol. 2, no. 7, pp. 287-292, 2017.