Facial Recognition-Based Automatic Door Access System Using Extreme Learning Machine

R F Rahmat¹, E N Zai¹, I Fawwaz² and I Aulia¹

¹ Department of Information Technology, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Medan, Indonesia
² Universitas Prima Indonesia, Medan, Indonesia

Email: romi.fadillah@usu.ac.id

Abstract. Facial recognition is one of the best forms of security since it is biometric-based security that uses biological features of the face. One of the artificial neural networks that can be implemented in the face recognition case is Extreme Learning Machine with the assistance of Local Binary Pattern in obtaining the facial features. The testing process in this study used 150 face images for training data and 60 face images for test data. The test was carried out using several three parameters of the hidden neuron numbers, namely 10, 30, and 50, and also five parameters of the image condition, namely feature, expression, face direction, lighting, and webcam distance. The system achieved the highest accuracy of 90% using 50 hidden neurons.

1. Introduction

Security is crucial for daily activities, especially when it comes to authority or privacy. Conventional security systems require keys, passwords, RFID cards, or ID cards to have access. The disadvantages of these methods are they are either hard to remember, can be duplicated, or stolen by others. Therefore, we need a way to improve the security of certain rooms involving authority or privacy. One of the security forms that is difficult to imitate, open, or modify by others is to use the automatic characterization of biological characteristics that are possessed by each individual. This characterization is known as biometrics. These characteristics can be seen from physical characteristics, such as fingerprints, facial features, eye retina, and voice. Facial recognition is one of the biometric methods that is quite popular. Faces tend to difficult to emulate, modify, or steal compared to keys, passwords, or non-biometric security. Therefore, a face recognition system to access a door for a particular room is necessary in effort to increasing the security. Authors proposed Extreme Learning Machine as the method to allow the system to perform face identification.

Several previous studies related to facial recognition for door security have been conducted. Januzaj et al. used a haar-like feature algorithm to detect faces, and PCA to recognize faces. The research succeeded in making a tool that was able to carry out the training process and face identification directly, without being connected to other equipment. The results indicated that the accuracy could be improved by considering some factors such as background, lighting, and the amount of training data [1].

Another study was conducted by Lwin et al. The face detection and recognition processes were performed on a PC with Matlab. Microcontroller PIC 16F887 was used to control the door access system, which depends on the incoming data sent from the PC. The door will open automatically
immediately once the face is successfully verified. After 2 seconds, the door will close automatically. The viola-jones method was implemented to detect faces in an image. This method has limits on head orientation. This method can only detect faces with foresight. Meanwhile, PCA method was applied to identify the face. The PCA method performed feature extraction on the face images and then used euclidian distances to recognize images [2].

Another research had also been carried out by Yang et al. to create an automatic door control system using a DSP-based multimedia module. In the study, the system was made to identify people through face detection, and then analyzed the trajectory path to determine whether the person has the intention to access the door or not. This was what controls the door. The system has a low error rate (up to 0%), a high correct activation rate of 99.6%, and a fast response time (within 2 seconds) in detecting a target, intention confirmation, and opening the door [3].

2. Methodology

In this study, the input data was acquired using a webcam connected to a PC; then, the data was processed for facial recognition. The connected microcontroller and the servo motor would receive a signal in the form of 1 or 0, depending on the output of the processing results on the PC, and the servo motor would move the door.

The webcam is a VGA camera with 0.3-megapixel resolution connected to the laptop with the capability of capturing images with a maximum dimension of 640 x 480. The microcontroller was a microcontroller type Arduino Uno (ATMega328), and the servo motor was the Tower Pro SG90 micro servo.

The face recognition process began with the collection of datasets for training and testing. The process continued to pre-processing to adjust the size of the entire image data to a size of 100x100, and convert those images to grayscale forms. After grayscale images obtained, the images would go through the feature extraction using LBP to get the characteristic value of each image. The feature extraction results (59 feature values) of the training image were stored as a comparison of the classification process using ELM with the feature extraction results from the test image. The output of the classification results in the recognition process would serve as the reference to control the door automatically.

This facial recognition program was created using Matlab r2016a. The training process (training) was conducted manually and simultaneously to all trained image data, while the testing process was performed directly through the program interface. The general architecture of this study can be seen in Figure 1.

2.1. Face Detection

At this stage, the input data was the captured image from the webcam. The aim is to detect whether there were faces in the input data.

Input : Image Data
Method : Viola-Jones
Output : Facial Image Data

2.2. Pre-processing

Pre-processing is a stage in image processing to produce a better image quality. In this research, there are two steps to this stage, which are resizing and grayscaling.

Resizing aims to obtain the same size images for the input data. The specified size for the image in this study is 100 x 100 pixels.

Input : Facial Image Data
Method : Bicubic interpolation
Output : Facial Image Data (100x100px)
Grayscaling is a process of converting RGB images to grayscale forms.
Input : Facial Image Data (100x100px)
Method : Grayscale
Output : Grayscaled Facial Image (100x100px)

Figure 1. General architecture

2.3. Feature Extraction
At this stage, the data will be processed by extracting the features in the data. The proposed method was LBP (Local Binary Pattern). LBP is defined as a measure of invariant grayscale texture, derived from the general definition of texture in the surrounding area. LBP operators can be seen as a unitary approach with traditional statistical and structural models different from texture analysis. Simply stated, LBP is a binary code that describes local texture patterns. It is built with a limited environment of a gray value from its center [4]. The result of feature extraction per face image data is 59 feature values. These values will serve as the presentation of the image data.

Input : Grayscaled Facial Image (100x100px)
Method : LBP
Output : Extracted values of the faces (59 values per image).

2.4. Facial Recognition
The face recognition test will be carried out after feature extraction. The extraction results of the test data feature will be compared with the training data results using Extreme Learning Machine. Extreme Learning Machine (ELM) is supervised learning of neural network types a. ELM is one of the Feed-Forward Neural Networks, which has one single hidden layer. The ELM method is
recommended to overcome the learning speed problem that has been happening in the implementation of other Feed-Forward Neural Network methods [5], a little bit similar approach with Deep Learning [8]. The activation function was sigmoid, while the numbers of hidden neurons were 10, 30, and 50.

Input: 59 extracted values
Method: LBP
Output: Facial recognition result (success or failed)

2.5. Dataset

The dataset of this study consisted of 210 face images with 150 images for training data and 60 for test data. The facial images were taken in varied conditions, namely normal face, face with certain expressions, face with sided, and tilted pose orientation. Training data and test data were acquired separately. The training data were obtained through a webcam. Then after being collected, the training data will be trained manually using ELM. The training results were the feature extraction values (in .mat format) and will be used as a comparison for the face classification. As for the test data, they were acquired in real-time via a webcam when the program runs. A total of 50 test data were the images from the trained facial images and 10 data for random facial images. Variations in the face conditions for the test data are normal, expressive faces, face-to-side orientation, lighting, and distance to the webcam.

| No | Dataset | Amount |
|----|---------|--------|
| 1  | Training| 150    |
| 2  | Test    | 60     |

Samples of the data training can be seen in Figure 2.

3. Result and Discussion

System testing was performed to determine the ability of the built system. The system capability depended on the data training. The testing parameters were different hidden neurons of 10, 30, and 50, while the applied activation function was sigmoid. As for the image condition parameters, there were images with normal lighting (≥200 lux) and images with low lighting (≤20 lux), expressive faces, face orientation, and facial distance to the webcam (> 40 cm). The test data amounted to 60 facial images consisting of 12 facial images for each image condition. The output of the process can be seen in Figure 3.
Figure 3. Result of data processing

The initial image captured by the webcam can be seen in Figure 3. Then, the images were examined to detect whether there is a face in the images, and cropping will be performed to separate the face from the background. The image was processed in the pre-processing stage to resize the images to a size of 100x100 pixels and to transform the RGB image into a grayscale image. Furthermore, the image feature extraction was conducted to get the value of the image characteristics. There were 59 characteristic values of each facial image.

| Label | Citra Awal | Detection & Cropping | Grayscale & Resizing | Elastroski Fitur (LBP) |
|-------|------------|----------------------|----------------------|----------------------|
| Des   | ![Image](image1.png) | ![Image](image2.png) | ![Image](image3.png) | ![Image](image4.png) |
| 0,2540035494490590  | 0,6245695634536565  | 0,2540035494490590  | 0,6245695634536565  |
| 0,7033091365095285  | 0,1834890016019104  | 0,7033091365095285  | 0,1834890016019104  |
| 0,2540035494490590  | 0,6245695634536565  | 0,2540035494490590  | 0,6245695634536565  |
| 0,1834890016019104  | 0,7033091365095285  | 0,1834890016019104  | 0,7033091365095285  |

Table 2. System testing using 10 hidden neurons

|                | Normal | Expressive | Face Orientation Direction | Bad lighting ≤ 20 lux | Distance from Webcam (> 40 cm) |
|----------------|--------|------------|----------------------------|-----------------------|-------------------------------|
| Succeed        | 58     | 57         | 58                         | 51                    | 57                            |
| Failed         | 2      | 3          | 2                          | 9                     | 3                             |
| Accuracy (%)   | 96,67  | 95         | 96,67                      | 80,75                 | 95                            |

Table 3. System testing using 30 hidden neurons

|                | Normal | Expressive | Face Orientation Direction | Bad lighting ≤ 20 lux | Distance from Webcam (> 40 cm) |
|----------------|--------|------------|----------------------------|-----------------------|-------------------------------|
| Succeed        | 60     | 56         | 60                         | 57                    | 58                            |
| Failed         | -      | 4          | -                          | 3                     | 2                             |
| Accuracy (%)   | 100    | 93,3       | 100                        | 95                    | 96,67                         |

Table 4. System testing using 50 hidden neurons

|                | Normal | Expressive | Face Orientation Direction | Bad lighting ≤ 20 lux | Distance from Webcam (> 40 cm) |
|----------------|--------|------------|----------------------------|-----------------------|-------------------------------|
The accuracy of the facial identification system was influenced by the condition of the test image, as shown in Table 2, Table 3, and Table 4, as well as Figure 4. Changes in expression, the direction of facial poses, face distance to the webcam, and lighting affect the test results. Some misidentifications were found in images that were unknown or not included in the training data. However, the more the number of hidden neurons, the higher the percentage of recognition accuracy.

The performance of the system was reviewed using an evaluation standard where the assessment is based on True Positive (TP), False Positive (FP), True Negative (TN), False Negative (FN). True Positive (TP) is the number of faces that are correctly classified. TP shows precisely identified facial images in accordance with the database. FP shows the image that should be correctly identified but misidentified by the system. TN is the number of facial images that are not contained in the database and are not recognized. FN is the facial image that is not in the database but is identified as someone else's image. The values to be calculated for the comparison were the True Positive Rate (TPR), the False Positive Rate (FPR), and the Positive Predicted Value (PPV). The higher the value of TPR and PPV, the better the system performance and vice versa. Whereas FPR, if it is higher, then the system performance decreases and vice versa. The calculation results of the values of TPR, FPR, and PPV are shown in Table 5.

**Table 5. The evaluation result of system performance**

| Number of Hidden Neuron | 10   | 30   | 50   |
|-------------------------|------|------|------|
| TPR                     | 80,77 % | 89,65 % | 94,91 % |
| FPR                     | 100 %    | 42,86 % | 25 %    |
| PPV                     | 84 %     | 94 %    | 96,43 % |
The results of the overall system accuracy can be shown in Figure 5. It can be seen that the amount of hidden neurons affects facial recognition accuracy. In addition, these results were also influenced by the condition of the test and trained data.

![Result of System Accuracy](image)

**Figure 5.** Result of system accuracy

After the face is recognized, the signal will be sent to the microcontroller to open the door. On the contrary, if the face fails to be recognized, then the door remains closed. The illustration of this process is shown in Figure 6.

![System Prototype](image)

**Figure 6.** System prototype

### 4. Conclusion

The conclusion that can be drawn based on the test results of the automatic door access system using Extreme Learning Machine is that the door access system successfully opens the door if the system recognizes the face image. In contrast, the door will remain closed when the face image is not recognized or failed to be identified. So the system can be used as a security feature on a private room door. The Local Binary Pattern method served as the method for texture-based feature extraction worked well in facial feature extraction. These methods cannot distinguish facial images with many similarities, which results in the downfall of the percentage. Therefore, the system greatly depends on the ability of the classification method. As for the Extreme Learning Machine (ELM), the method can perform facial recognition through facial images very well. However, the accuracy is low when processing facial images which were not included in the database. The results showed the system accuracy is at 70% with 10 hidden neurons, 85% with 30 hidden neurons, and 90% with 50 hidden neurons.
Future research can be developed by adding the amount of training data to increase the system's ability to recognize faces and implementing a combination of other methods at the feature extraction stage to get a more accurate feature value.

References

[1] Januzaj Y, Luma A, and Ramaj V 2015 Real-Time Access Control Based on Face Recognition. *International Conference on Network Security and Computer Science (ICNSCCS-15)*.

[2] Lwin H H, Khaing A S, Tun H M 2015 Automatic Door Access System Using Face Recognition. *International Journal of Scientific & Technology Research, 4*(6).

[3] Yang J C, Lai C L, Sheu H T, Chen J J 2013 An Intelligent Automated Door Control System Based on a Smart Camera. Sensors (Basel, Switzerland).

[4] Ahonen T, Hadid A, and Pietik A M 2006. Face Description With Local Binary Patterns: Application to Face Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 28*(12): 2037-2041.

[5] Huang G B, Zhu Q Y, Siew C K 2006 Extreme Learning Machine: Theory and Applications. *Neurocomputing, 70*: 489-501.

[6] Kasar M M, Bhattacharyya D, Kim T H 2016 Face Recognition Using Neural Network: A Review. *International Journal of Security and Its Applications, Vol 10*(3): 81-100.

[7] Nazeer S A, Omar N, and Khalid M 2017 Face Recognition System using Artificial Neural Network Approach. *International Conference on Signal Processing, Communication, and Networking (ICSCN)*, 420-425.

[8] Rahmat R F, Dennis, Sitompul O S, Purnamawati S, Budiarto R 2019 Advertisement billboard detection and geotagging system with inductive transfer learning in deep convolutional neural network *Telkomnika 17*(5) pp 2659-2666