International Journal of Information System & Innovative Technology 1(2) (2022) 1-10

GEVIVA-ETT
THE BEST SOLUTION

IJISIT

http://ejournal.enlightenlearner.com

IJISIT
INTERNATIONAL JOURNAL

Combination Of Haar Cascade Classifier And Convolutional Neural Networks For Classification Of Face Image

[1]Irwan Jani Tarigan, STMIK Methodist Binjai, Binjai - Indonesia
ijtarigan0105128105@gmail.com

Article Info

Article history:

Received August 31, 2022
Revised September 01, 2022
Accepted September 09, 2022

Keywords:

classification
face recognition
haar cascade classifier
convolution neural network

ABSTRACT

Humans are able to carry out face recognition in everyday of life, where faces are a unique identity, are a complex problem in machine learning. Face recognition technology is also very useful in various fields today, such as access control, attendance and security systems. This study aims to obtain a level of accuracy in classifying image data with a combination of algorithm haar-like feature with convolutional neural network, where haar-like feature is used to detect faces in image data, while convolutional neural network is used for face recognition. Image data using in this study, Japanes Female Facial Expression (JAFFE), the experimental results give the best accuracy results at the 500 epoch of 95.14% and a running time of 1.52 seconds.

1. INTRODUCTION

Facial recognition performed by humans in the computer world is the scope of biometric systems[6]. The biometric system is an important process, where the unique features of a person are stored in a database that will be used during identification and verification, one of the most widely applied biometric technologies, especially in security systems. Attendance systems with faces, recognizing criminals with CCTV are some of the applications of facial recognition. Efficiency and accuracy are the main factors that face recognition is widely applied [4]. Face recognition is a common problem in machine learning, this technology continues to be used in human life. For example, Facebook can automatically tag people's faces in pictures, and some mobile devices are equipped with facial recognition systems to protect personal security [2].

In the identification system [5], the biometric signature of an unknown person is presented into the system and then the new biometric signature is compared with a database of known individual biometric signatures.

The facial recognition method has been applied to become the main technology in various fields, such as access control in buildings, access control on personal and office computers, the use of Automated Teller Machines (ATMs) in daily affairs in terms of withdrawing money from bank accounts, dealings with post offices. , the field of criminal investigation and so on [3].

A previous study by [7], entitled Rapid Face Emotion Recognition Using Convolutional Neural Networks and Gabor Filters, carried out emotion recognition by detecting faces using the Convolutional Neural Networks algorithm and Gabor Filter. The experimental results of the CNN-Gabor method give better results in speed and accuracy when compared to the conventional CNN method. In the experiment the detection process with the best accuracy results is 91-92% obtained with a time of 9 minutes (541 seconds).

Looking at Zadeh et al's research above, author is interested in raising the title about Combination Of Haar Cascade Classifier And Convolutional Neural Networks (CNN) For Classification Of Face Image , where the haar cascade classifier method will be used to detect objects in digital images in this case the face because it has advantages, namely computational efficiency. very fast, while the Convolutional Neural Network method is used for face recognition in digital images through the image data training process based on the detection results of the haar cascade classifier method.

2. METHODS

2.1. Haar Cascade Classifier

Haar Cascade Classifier method is a proposed by Viola and Jones which trains machine learning in object detection in an image which can also be useful for detecting faces (Ivancic, 2019). Feature extraction and classification in the form of rectangles are haar features that present specific characteristics in an image where these features are used in simple object recognition and do not take into account the object's pixel value.

2.1.1 Haar like feature

Haar-Like Features[1] is a feature extraction method in recognizing facial images. Each Haar-Like Features is a combination of several features for training positive and negative image data. The types of features can be seen in Figure 1.

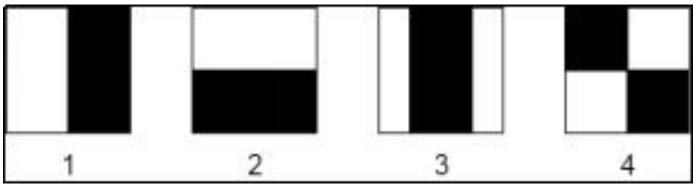

Figure 1. Haar like feature [1]

Haar rectangular feature formula is as follows:

$$F_{haar} = \sum R_{black} - \sum R_{white} \quad \text{...............} \quad (1)$$

In equation (1) above, the process of calculating the feature value of the haar rectangle is carried out by arithmetically reducing the value of the pixels in the darker area compared to the pixel value in the bright area, where if the difference value is above a predetermined threshold value, then means that the facial features are present.

2.1.2 Integral Image

Known as table area summation, this algorithm is considered faster and more efficient to calculate the number of pixels in an image, the grayscale value of an image pixel with a height of i and a width of j where the intensity of each pixel is the combined pixel value of the entire image. In Figure 2

Figure 2. Three Rectangle feature

For example, the grayscale value in the haar box can be seen in the table below.

Table 1. Grayscale Value from input image

Grayscale Value		
222	220	218
200	214	202
200	120	201

Integral image of the gray value can be calculated as in the table below.

Table 2. Calculated of integral image

Calculated of Integral Image		
222	222+220	222+220+218
222+200	222+200+220+214	222+200+2220+214+218+202
222+200+200	222+200+200+220+214+120	222+200+200+220+214+120+218+202+201

2.1.3 Adaboost

Adaptive-boosting functions is to improving the performance of feature classification where the algorithm combines a lot of performance from a weak classifier to get a strong classifier.

2.1.4 Cascade Classifier

Cascade Classifier is designed in such a way with the aim of improving performance in the form of detection levels, for each classifier level is a description of the acquisition of a boosting algorithm where each classifier level has several types of weak classifiers. Can be seen in a figure 3.

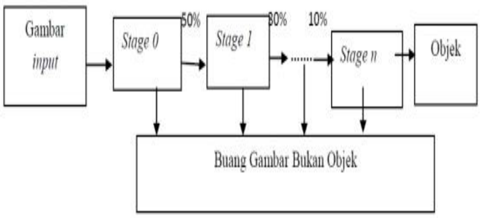

Figure 3. Haar Cascade Classifier Process

2.2. Convolutional Neural Network

Convolutional Neural Network (CNN) is developed from the Multi Layer Perceptron (MLP) algorithm platform which is designed to process two-dimensional information. This algorithm is included in the Deep Neural Network type because it has a higher network depth than MPL so it is widely used for digital image processing [8].

2.2.1 Architectured of CNN

The algorithm consists of two parts is a convolution layer and pooling layer can be seen in Figure 4.

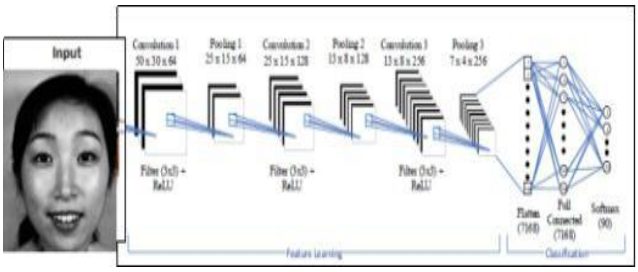

Figure 4. Architectured of CNN

a). Convolutional Layer

Process of calculating neurons and weights after the input layer on the network architecture using the CNN algorithm where the results of the convolution process at the output of the previous layer, the output of this convolution process can be calculated using equation (2) and convolutional process can be seen in figure 5.

$$(W - F + 2P)/S + I \ldots\ldots\ldots\ldots (2)$$

Where, in equation :

W = image dimension
F = filter dimension
P = value of padding
S = sliding or stride value process

2	6	3	2	2								
4	4	3	4	2		1	-1	1		-7	-7	-6
3	2	4	5	7	*	-1	1	-1	=	-2	-4	-11
2	1	5	2	9		1	-1	-1		-4	6	-9
1	3	1	1	3								
		5x5					3x3				3x3	

Figure 5. Convolutional Process

b). Pooling Layer

a technicque for reducing the size of the matrix, where in the process there is a certain filter and stride, which alternately shifts from top left to bottom right until the entire feature map area is a process in the pooling layer. Applying the pooling process in the max-pooling function which applies filtering to all the highest values in the kernel filter position with a dimension of 3 x 3 pixels and to use zero padding on the feature map with the amount of one pixel and performs 2 times the kernel stride filter.

2.2.2 Fully Connected

The stage in this process is to change the dimensions of the image data so that it can be classified in a linear way, where the input image is a fully connected layer where the previous output layer is still in the matrix. The initial stage for the fully connected layer on the input image data is a flattening process where the image matrix is converted into a vector and then the fully connected layer process is carried out with the following equation :

$$f_{ck} = b_k + \sum_{i=0}^{n} x_i w_{i,k} , k = 0,1,2,\ldots,t \quad \ldots\ldots\ldots. (3)$$

where,
f is a output value.
x is a flatten value.
b is a bias.
w is a weight value.
k is a amount of data on *flatten.*
t is a amount of target on fully connected layer.

2.2.3 Softmax Classifier

is a layer to identify with the Logistic Regression algorithm to group the results from the previous layer, namely the fully connected layer using the softmax activation function with the following softmax activation equation.

$$f(z)_j = \frac{e^{zj}}{\sum_{k=1}^{k} e^{zk}} = \frac{e^{(w^T x)j}}{\sum_{k=1}^{k} e^{(w^T x)k}} \quad \text{...........} \tag{4}$$

where,
f(z) is a output of function result
j is a many of looping as class
w^T is a value of *weight* already transposed
x is a value of data in hidden layer
K is a amount of a class.

The processes in this study are represented on the research block diagram, which shows an image-based face detection diagram. The series of steps in question are as follows: input image file; conversion of the color image into a gray image (grayscaling); feature calculation with integral image; object detection with Haar Like Feature or Haar Cascade Classifier and is called Cascade Classifier. Furthermore, it will display the part of the object that is detected as a face or not and proceed with the process of detecting the identity of the face. Block diagram of image-based face detection research as shown in Figure.6.

2.3 A Reasearch Block Diagram

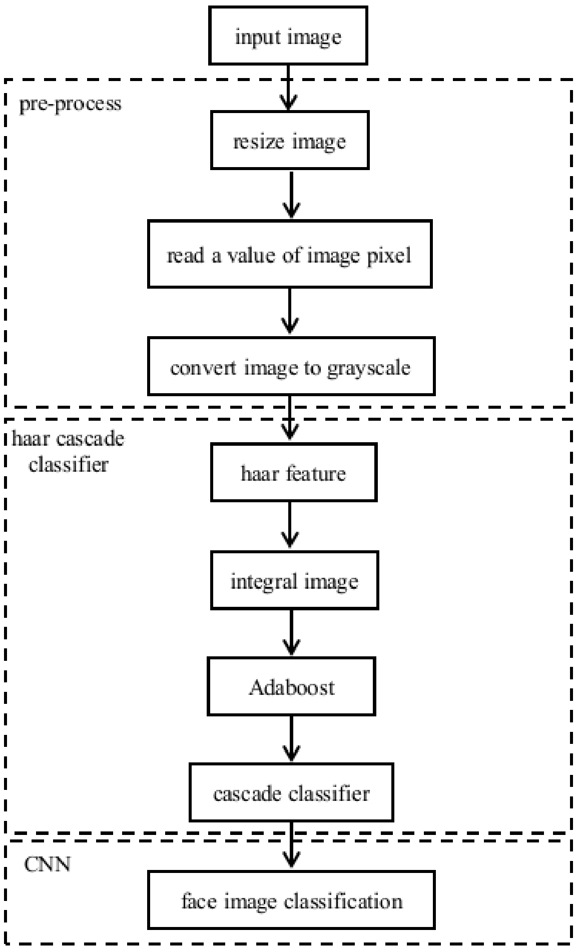

Figure 6. Research Block Diagram

2.4 Training Process

In this process, face image dataset for the training data that has been determined will be trained using the CNN algorithm so that during recognition it can produce an optimum accuracy value. This process consists of feedforward and backpropagation processes. In the feedforward process a number of layers are used, determining the size of the subsampling where the vector image is obtained in the pre-processing process. The feedforward process will work by performing a convolution process on vector images and max pooling to reduce the dimensions of the image and increase the number of neurons, so that many networks are formed that will vary the training data. The result that will be obtained from the feedforward process is a weight value that is useful in the evaluation process of the neural network. The flowchart of the network training process is shown in Figure 7.

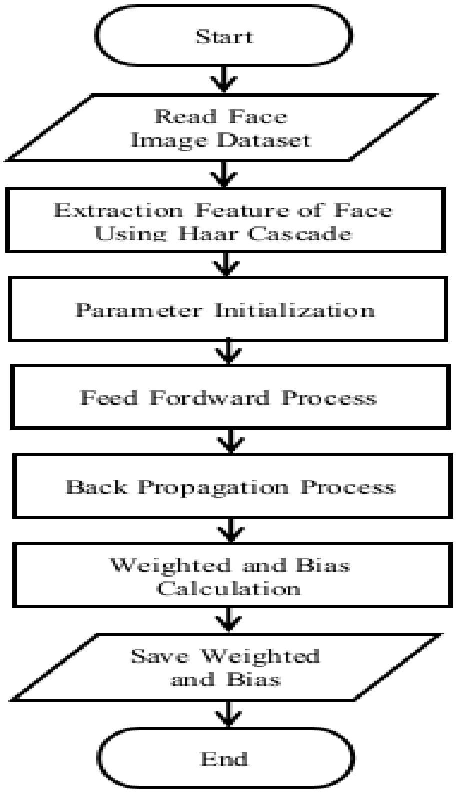

Figure 7. Flowchart of Training Process

2.5 Testing Process

Testing process is the process of classify the input face image by using the weight and bias values obtained from the training process. The process steps are the same as the training process, only difference being that the backpropagation process is abolished after feedforward. The output of this testing process gets the classification results in the form of images that are not successfully classified where the number of images that fail and the data formed from the feedforward process. The weight and bias values are updated to the latest ones, which will be used in the feedforward process and produce an output layer that is fully connected with the label provided. The testing process that is formed is as shown in Figure 8.

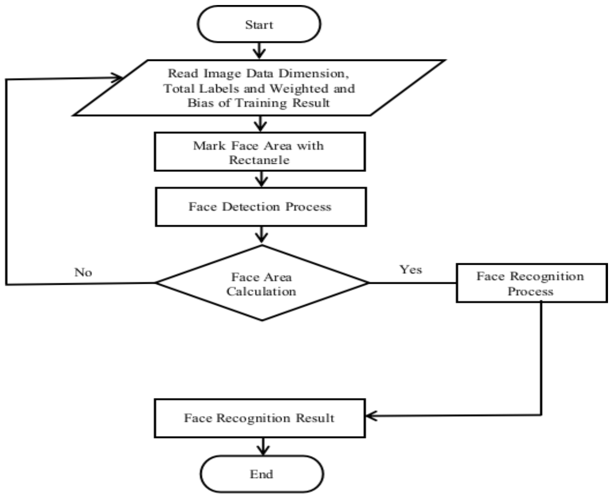

Figure 8. Flowchart of Testing Process

3. RESULT AND IMPLEMENTATION

3.1 Dataset

Dataset used in this paper is the Japanese Female Facial Expression (JAFFE) database which contains 213 facial images of Japanese women, which can be seen as an example of Figure 9.

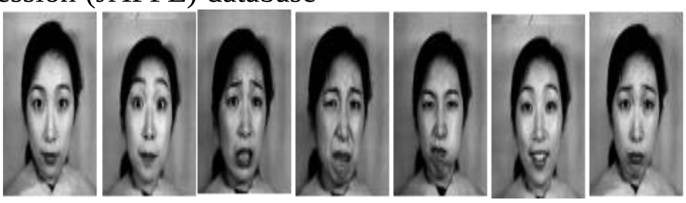

Figure 9. Example of JAFFE Dataset

3.2 Result

3.2.1 Face Detection

In this process serves to determine the face image in the form of a human face or not. Face detection used haar cascade algorithm using the Japanese Female Facial Expression (JAFFE) database as shown in Figure 10.

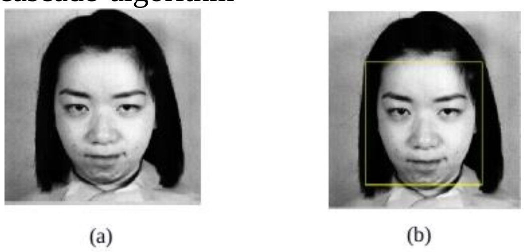

Figure 10. Face Detection Result

3.2.2 Face Recognition

In this paper, 70 facial data were tested, the experiment was carried out in stages of 100 epochs, 200 epochs, 300 epochs, 400 epochs and 500 epochs. Example of test results can be seen in Figure 11 dan 12.

Figure 11. Face Recognition in stages 100 epochs

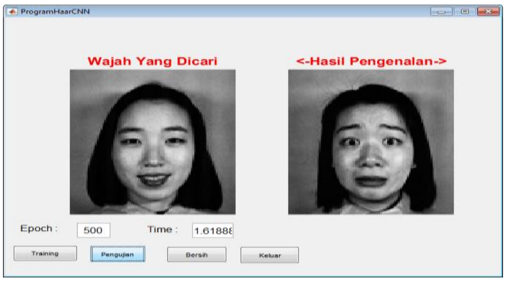

Figure 12. Face Recognition in stages 500 epochs

results of face image recognition with the same image search process so that the experimental value of image-1 is 1 and so on until the 70 facial data and the result can be seen as in Table 3 and Table 4.

Tabel 3. Experimental Face Recognition Result 100 epochs

		Objek Wajah										Annotation		
		1	2	3	4	5	6	7	8	9	10	Amount of Test Data	Time	Accuracy
Objek Wajah	1	45	2	1	5	2	3	3	3	3	3	70	1.54	64.29
	2	3	42	2	3	3	3	4	3	3	4	70	1.53	60
	3	2	3	47	4	4	3	2	1	2	2	70	1.56	67.14
	4	1	3	5	40	3	6	3	3	2	4	70	1.47	57.14
	5	2	3	4	3	44	3	2	5	2	2	70	1.52	62.86
	6	5	4	2	4	4	40	0	4	3	4	70	1.5	57.14
	7	3	5	4	2	3	3	47	0	2	1	70	1.47	67.14
	8	3	3	5	2	3	0	4	50	0	0	70	1.58	71.43
	9	3	2	3	1	3	4	2	3	49	0	70	1.57	70
	10	3	3	4	5	5	3	4	0	1	42	70	1.65	60

From the table above, it can be seen from 70 tests for each face object, then face object 1 is recognized as many as 45, face object 2 is recognized as many as 42, face object 3 is recognized as many as 47, and so on until face object 10, with an average processing time of 1.54 seconds and the average value of accuracy is 63, 71 %.

Tabel 4. Experimental Face Recognition Result 500 epochs

		Objek Wajah										Annotation		
		1	2	3	4	5	6	7	8	9	10	Amount of Test Data	Time	Accuracy
Objek Wajah	1	68	0	0	0	0	1	1	0	0	0	70	1.53	97.14
	2	0	64	0	1	0	1	1	1	1	1	70	1.52	91.43
	3	0	0	68	1	0	0	1	0	0	0	70	1.55	97.14
	4	0	1	1	66	0	0	1	0	0	1	70	1.49	94.29
	5	1	0	1	1	64	1	1	0	1	0	70	1.48	91.43
	6	0	0	1	0	1	67	0	0	1	0	70	1.5	95.71
	7	0	1	0	0	0	1	68	0	0	0	70	1.5	97.14
	8	0	1	0	0	0	0	1	68	0	0	70	1.54	97.14
	9	0	1	1	0	0	1	0	1	66	0	70	1.55	94.29
	10	0	0	0	1	0	1	0	0	1	67	70	1.56	95.71

From Table. 4 above, it can be seen that for object 1 is 68 means that from a total of 70 times the test was obtained 68 face object success for recognise, with processing time of 1.53 seconds and the accuracy value is 97.14 %.

From the entire face object data test, starting with the stages of 100 epochs, 200 epochs, 300 epochs, 400 epochs and 500 epochs that have been carried out, the test results are obtained in Table 5.

Table 5. Results of Test

No	*Epoch*	Time	Accuracy (%)
1	100	1.48	63.71
2	200	1.54	76.14
3	300	1.55	90
4	400	1.53	93.43
5	**500**	**1.52**	**95.14**

and from the table above, it can be obtained the results of test data visualization in Figure 13.

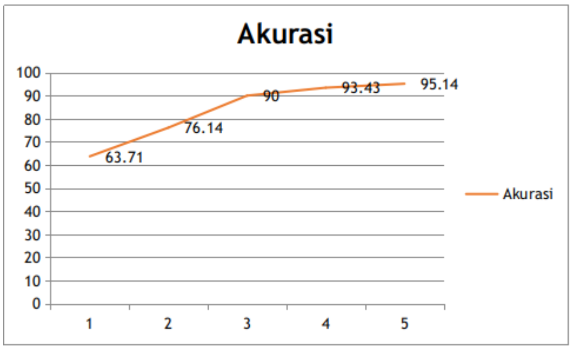

Figure 13. Results of Test Data Visualization

4. CONCLUSION

Combination of the Haar-Cascade Classifier and Convolutional Neural Network algorithm to perform facial recognition using image files from the Japanese Female Facial Expression (JAFFE) dataset, facial recognition experiments were carried out 70 times in each epoch, the results obtained in the 100 epoch test with an average - an average accuracy of 63.71% and a running time of 1.54 seconds, in the 200 epoch test with an average accuracy of 76.14% and a running time of 1.48 seconds, in the 300 epoch test with an average accuracy of 90% and a running time of 1.55 seconds, in the 400 epoch test with an average accuracy of 93.43% and a running time of 1.53 seconds and in the 500 epoch test with an average accuracy of 95.14% 1.52 second. The processing time (running time) is due to the face recognition process that begins with the face detection process or not with the Haar-Like Feature algorithm, so that the processed image only has faces which results in shorter processing time. From the results of facial recognition experiments in this study, the best accuracy at the 500 epoch is 95.14% and the running time is 1.52 seconds.

6. REFERENCES

[1] Cahu S., Banjarnahor J., Irfansyah D., Kumala S., & Banjarnahor J., 2019. Analisis Pendeteksian Pola Wajah Menggunakan Metode Haar-Like Feature. Journal of Informatics and Telecommunication Engineering.

[2] Ivancic K., 2019. *Traditional Face Detection With Python. (Online)* https://realpython.com/traditional-face-detection-python/. Diakses (26 Juli 2019).

[3] Jalled, F. 2017. *Face Recognition Machine Vision System Using Eigenfaces*. *(Online)* https://arxiv.org/abs/1705.02782.

[4] Kumar, G., Saurabh, S. 2016. A *Comparative Analysis of Face Recognition Algorithms*. *Journal of Recent Research Aspects*. *(Online)* https://www.ijrra.net/Vol3issue2/IJRRA-03-02-42.pdf.

[5] Novamizanti, L., De Lima, N., & Susatio, E. (2019). Sistem Pengenalan Wajah 3D Menggunakan ICP dan SVM. *Jurnal Teknologi Informasi dan Ilmu Komputer, 6*(6), 601-610. doi:http://dx.doi.org/10.25126/jtiik.2019661609

[6] Philip P.J., Martin A., Wilson C. L. & Przybocki M., 2000. An Introduction and Evaluating Biometrics Systems. Journal of IEEE *(online)* https://ws680.nist.gov/publication/get_pdf.cfm?pub_id=151436.

[7] Reddy, T.M., Prasad, F., Ramanaiah, N.V. 2015. *Face Recognition Based Attendance Management System By Using Embedded Linux*. *International Journal of Engineering Research and Science & Tekcnology*.

[8] Lyons Michael, Kamachi Miyuki & Gyoba Jiro, 1998. The Japanese Female Facial Expression (JAFFE) Database. (Access : 01 September 2022)

[9] Viola V. & Jones M., 2001. Robust Realtime Object Detection. International Workshop On Statistical And Computational Theories of Vision - Modelling, Learning Computing And Sampling.

[10] Zadeh, M. M. T, Imani, M. & Majidi, B.2019. *Fast Facial emotion recognition Using Convolutional Neural Networks and Gabor Filters*. 5th Conference on Knowledge-Based Engineering and Innovation, Iran University of Science and Technology, Tehran, Iran.