Analyzing Of Different Features Using Haar Cascade Classifier

[1]Ratna Yustiawati, [2]Nyayu Latifah Husni, [3]Evelina, [4]Sabilal Rasyad, [5]Iskandar Lutfi, [6]Ade Silvia, [7]Niksen Alfarizal, [8]Adella Rialita
Electrical Engineering
Politeknik Negeri Sriwijaya
Palembang, Indonesia
[1]ratnayustiawati2@gmail.com, [2]nyayu_latifah@polsri.ac.id, [3]evelinaginting15@gmail.com, [4]Sabyad_ok@yahoo.com, [5]lutfiskandar@gmail.com, [6]ade_silvia@polsri.ac.id, [7]niksen_alfarizal@polsri.ac.id , [8]itsarialita@yahoo.com

Abstract*— Technology always developed all the time. It makes many humans utilize the technology for doing everything. Not only used for doing something good, but there are many humans use it for doing criminal, such as robbing. On the other hand, there is something found for preventing that problem. It is a tool for detecting face. Face recognition is made for identifying a face by digital image. This paper presents a research on security of a room. The room can only be opened by identifying ad matching the face of the one that has been inputted to the data base of the system. Haar cascade classifier is used in this research. The processes of identifying of the objects were discussed in this paper. From the experimental test, it can be concluded that the proposed method is effective enough in processing the features of the objects.

Keywords—face detection; face recognition; haar cascade classifier; security.

I. INTRODUCTION

Indonesia is one of state with the most criminal acts. Based on the criminal statistic 2015 that issued by *Badan Pusat statistik* (BPS) there is 10 province with the most criminal acts in Indonesia. The highest position is the capital city of Indonesia DKI Jakarta [1]. In 2016 *BPS* released the statistical data of criminal acts in Indonesia that has increasing 1,2% from the last year [2]. Theft and robbery are the criminal acts that ussualy happen in Indonesia. This criminal acts causing many casualties. Less of security on the door is basically the reason that crimes happened. Although, the door already equpped with password, it does not guarantee that it is already safe. According to Irvan Budiawan and Andriana [3] good security system is the one that use data identity as a security, and it is called Biometric identity technique.

Biometric identity technique is an identification tecnique based on physiological characteristics and behavioral that exist in human body itself like face characteristic, fingerprint, palm, iris and retina of eyes, and then DNA [4]–[6]. Biometric identity technique has more advantages to the other method. The people who does not have the ability to access cannot manipulate the data carelessly [6]. Fingerprint biometric technique is one of the most technique which has been widely used, for example for automatic gate [7] and ATM machines security [6]. However, one type of safety is considered to be not enough to face the increasing of theft and the advance of technology they used. Due to that reason, giving double safeties for the door is highly recommended. One of the techniques that can be used is face detection biometric technique. This technique also called as face detection. The recognition of face characteristic became the main thing that can be used in accessing the security.

As mention above, there are several types of algorithm that can be used to detect the face. Each of them has the advantages and disadvantages [8]. A few method involve the templates, filter and neural network, such as research conducted by Thai Hoang Le [9]–[11]. The drawback of this algorithm is it is too expensive. Beside that, this algorithm analyzes the pixels of images that depends on colour and light intensity value. In face detection process, to get those value, it needs a long time and sometimes it is heard to reach. Considering that human face detection having wide and different shape of pigmentation value, causes the pixels need to be reanalysised for the scale and accuracy of picture [8]. Therefore, on this paper, haar cascade classifier method is proposed. According to Snehal B. Jagtap [8], the algorithm from Viola and Jones, or called haar cascade classifier often used as face detection because it has the advantages, such as, be able to detect any objects, including human face, quickly. This method does not processs the image by looking from the RGB value in every pixel but process the image based on Haar-like features.

There are few research which have been made before using the same method namely [4], [5], [12]–[18]. One of them is the research made by Wahyu Sulistiyo [4] that made a design for attendance system. The hardware produced is just a prototype and used as attendance not as the security of the door. Beside that, the development of human face features who using hijab and spectacles has not been applied in that research.

Based on that reference mentioned above, the development of the security system in the door is really needed. Human face features who wearing hijab becomes one of the challenge of this reseach. In this research, an Haar-like features image processing using dimensional image 10x10 was used in order to get the value of integral image. This reseach was implemented and integrated at measurement laboratory of Politeknik Negeri Sriwijaya. From this study, it is hoped that it gives a contribution in giving a solution to the security system problem which is very crucial.

II. FACE DETECTION AND FACE RECOGNITION

A. Definiton

Face detection is the first step of face recognition and part of the face identification. Face detection is also a section of computer technology that determines the location and size of human faces in any digital images by ignoring the other objects. Face recognition is a technology that has been used since 1960. It was applied in security system, robotic and interface . Face recognition works by identifying human faces based on digital images or frame that have been trained before including on face database. Face database previously works

978-1-5386-5721-8/18/$31.00 ©2018 IEEE

to fit someone faces that will be identified with face that already be stored in the database. These techniques are included in biometric technique. Those techniques were needed on face recognition system. Face detection caught the object of human faces that have already been captured by camera and followed by recognizing the face base on detected face images. The next step, is face identification, where the matching of the face was done. It decided whether it has been appropriate with the stored database or not. If the object was already stored in database, then the object was identified, however if it was not appropriate, the object would be not identified.

B. Algorithm

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

In this research, the Haar cascade classifier algorithm was used as a face recogniton system. This method was found by Paul Viola and Michael Jhon therefore it was also called Viola and Jhon method. This method was usually used because it gave advantages, such as high accuracy and fast compution. It has simple value of features and just depends on the amount of pixels inside the square, not the value of pixels of an image. This algorithm consist of haar features, integral image, AdaBoost machine-learning and Cascade classifier.

Haar-like features consistuted by Haar wavelet. Haar wavelet is a single rectangle-shaped (one high interval and one low interval). In two dimensions, the square is illustrated with white and dark. The value of Haar-like features exists if the result is above treshold value, by reducing the pixel's average value in value in dark and white area.

The function of integral image in this study is to determine wheter a hundred of haar features is existing or not on the images. When it exists, it will accelerate the detection process of object. The value of features that has been occured will be processed by AdaBoost machine learning. It has purpose to present wheter there was a face in input image or not. By comparing the value of features with specified threshold value, the features could be determined. If the value of the features is the same or above the threshold value, the face could be detected. The last process was the process of features selection (also known cascade classifier).

III. EXPERIMENTAL SETUP

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections A-D below for more information on proofreading, spelling and grammar.

Fig 1 represents the flow system of how the process of taking a picture by camera could be obtained a match qualification using haar cascade classifier method happened.

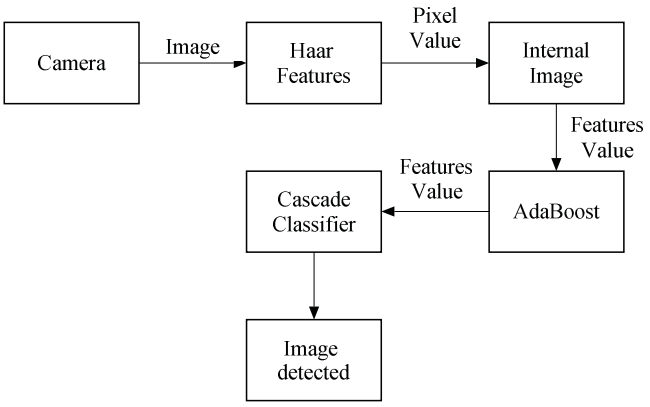

Fig 1. Data Flow Diagram

Before determining of the value of integral image on features, first it has to specify where the haar features with determined value of the dark and white be located.

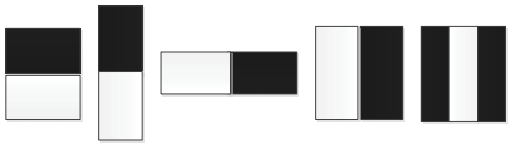

Fig 2. Haar Features

After that, determine wheter there is hundred of haar features or not an images. The dark area with the white one of haar features should be reduced using the formula (1).

$$\Delta = \text{dark} - \text{white} = \frac{1}{n}\sum_{\text{dark}}^{n} I(x) - \frac{1}{n}\sum_{\text{white}}^{n} I(x) \qquad (1)$$

Fig 3 is the example of input image with 10x10 dimension that was used in this research.

0,1	0,2	0,3	0,4	0,5	0,3	0,3	0,5	0,4	0,3
0,3	0,1	1,2	0,4	0,7	0,2	0,2	0,7	0,4	0,3
0.3	0,2	0,9	1,2	0,7	0,1	0,1	0,9	0,3	0,9
0.5	0,5	1,2	1,1	0,1	1,2	0,1	1,2	0,8	0,9
0,6	0,6	1,0	1,0	0,1	1,3	1,2	0,5	0,4	0,7
0,8	0,4	0,6	0,3	0,3	0,5	0,2	0,6	0,4	0,5
0,9	0,3	0,6	0,5	0,4	0,5	1,3	0,6	0,3	0,3
0,1	0,9	0,3	0,7	0,4	0,3	0,3	0,3	0,1	0,5
0,2	0,1	0,1	0,9	0,8	0,2	0,5	0,1	0,1	1,1
0,4	0,2	0,3	0,9	0,3	0,1	0,5	0,3	0,7	1,2

Fig 3. Dimensional of 10x10

From the Fig 3 above, the value of integral images using the value of pixel is shown in Fig 4.

0,1	0,2
0,3	0,1

Fig 4. Value of integral images using the value of pixel

For instance, the value of pixel on the second row and second column that obtaining with summing up the value of row pixel on the first column, the first row pixel (0.1) with second column, the first row pixel (0,2) with the first column, the second row pixel (0,3) with the second column, the second row pixel (0,1). The pixel value was then obtained with the value (0,7), where 0,1+0,2+0,3+0,1. The integral image from the image in Fig 4 is given in Fig 5.

0,1	0,3	0,3	1	1,5	1,8	2,1	2,6	3	3,3
0,4	0,7	1,2	3	4,2	4,7	5,2	6,8	7,2	8,7
0.7	1,2	3,6	5,6	7,5	8,1	8,7	11,2	11,3	14,3
1,2	2,2	5,8	8,9	10,9	12,7	13,4	17,1	18	21,9
1,8	3,4	8	12,1	14,2	17,3	19,2	23,4	24,7	29,2
2,6	4,6	9,8	14,2	16,6	20,2	22,3	27,1	28,8	33,8
3,5	5,8	11,6	16,5	19,3	23,4	26,8	32,2	34,2	39,5
4,4	6,8	12,9	18,5	21,7	26,1	29,8	35,5	37,6	43,4

4,6	7,1	13,3	19,8	23,8	28,4	32,6	38,4	40,6	47,5
5	7,7	14,2	21,6	25,9	30,6	35,3	41,4	44,3	52,4

Fig 5. Integral Image Citra 10x10

3.1 Without Integral Image

After the value of integral was obtained, the next step was to determine the value of haar features. The value of haar features could be obtained faster by summing up the value of integral image. The following is an example of searching the features value using integral image and without integal image. First, calculate the value of haar features without integral image. The value can be seen in Fig. 6.

0,1	0,2	0,3	0,4	0,5	0,3	0,3	0,5	0,4	0,3
0,3	0,1	1,2	0,4	0,7	0,2	0,2	0,7	0,4	0,3
0.3	0,2	0,9	1,2	0,7	0,1	0,1	0,9	0,3	0,9
0.5	0,5	1,2	1,1	0,1	1,2	0,1	1,2	0,8	0,9
0,6	0,6	1,0	1,0	0,1	1,3	1,2	0,5	0,4	0,7
0,8	0,4	0,6	0,3	0,3	0,5	0,2	0,6	0,4	0,5
0,9	0,3	0,6	0,5	0,4	0,5	1,3	0,6	0,3	0,3
0,1	0,9	0,3	0,7	0,4	0,3	0,3	0,3	0,1	0,5
0,2	0,1	0,1	0,9	0,8	0,2	0,5	0,1	0,1	1,1
0,4	0,2	0,3	0,9	0,3	0,1	0,5	0,3	0,7	1,2

Fig 6. Without Integral Image

Value of haar features

$$\Delta = \frac{1}{n}\sum_{\text{dark}}^{n} I(x) - \frac{1}{n}\sum_{\text{white}}^{n} I(x) \qquad (2)$$

Δ=|(total of black pixels)-(total of white pixels)|

= |(18,9)-(17,9)|

= 1

The final result was 1, and then calculate the value of haar features using integral image. If the final result is the same or closed to the determined value, then it can cloncluded that it has been closed to the ideal features from the value of haar features.

3.2 Using The Integral Image

0,1	0,3	0,3	1	1,5	1,8	2,1	2,6	3	3,3
0,4	0,7	1,2	3	4,2	4,7	5,2	6,8	7,2	8,7
0.7	1,2	3,6	5,6	7,5	8,1	8,7	11,2	11,3	14,3
1,2	2,2	5,8	8,9	10,9	12,7	13,4	17,1	18	21,9
1,8	3,4	8	12,1	14,2	17,3	19,2	23,4	24,7	29,2
2,6	4,6	9,8	14,2	16,6	20,2	22,3	27,1	28,8	33,8
3,5	5,8	11,6	16,5	19,3	23,4	26,8	32,2	34,2	39,5
4,4	6,8	12,9	18,5	21,7	26,1	29,8	35,5	37,6	43,4
4,6	7,1	13,3	19,8	23,8	28,4	32,6	38,4	40,6	47,5
5	7,7	14,2	21,6	25,9	30,6	35,3	41,4	44,3	52,4

Fig 6. Using The Integral Image

To calculate the value of pixel using integral image, the equation (3) was used.

$$D = D + A - (A + C) \qquad (3)$$

Where D is the value of pixel in right bottom, A is the value in top left of pixel D and C is the value of left pixel of pixel D. Fig 8 illustrates that pixels.

A	B	
C	D	

Fig 7. Integral Image Search Value of Pixel D On Integral Image

Value of haar features

$$\Delta = \frac{1}{n}\sum_{\text{dark}}^{n} I(x) - \frac{1}{n}\sum_{\text{white}}^{n} I(x) \qquad (4)$$

Δ=|(total of black pixels)-(total of white pixels)|

= |(30,6 + 0,7-(7,7+4,7))-(52,4+4,7-(30,6+8,7))|

= |18,9-17,8|

= 1,1

The ideal features is 1. The closer value, it means that a haar features has been found. After obtaining all the value of features from haar features and the value of integral image features, the next step was to determine the features of AdaBoost.

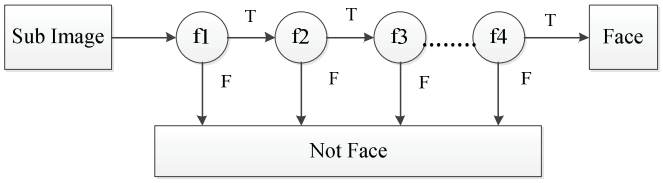

Fig 8. AdaBoost Machine

In fig 9 above, sub image was processed to determine whether the received features is true or false. If true feature was the feature that has already been stored in database, then the feature was identified and it is the face that matches with the feature. If it was not the feature, the feature would be discarded. It meant that it was not the face that matches with the database or sub image feature.

IV. Result And Discussion

In this research, some software were used, namely putty as monitoring of the program; netbeans as a place where the openvc program was run; and winscp as a database storage. The database has 50 pictures that consists of 20 pictures of woman who does not wearing hijab and 30 pictures of woman who wears hijab. All of the pictures was taken with different background and angle. The first test, the face of woman who does not wearing hijab was named with sub image user 1 in the database as shown shown in Fig 10.

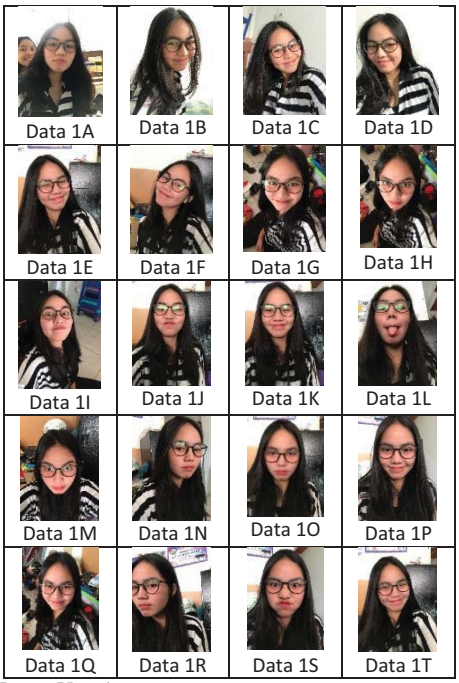

Fig 9. Sub Image User 1

The pictures of Fig 10 were taken wih different directions and styles. At first they did not quality using Haar cascade classifier method. After the pictures were matched with the

database system, the Haar cascade classifier method was used for determining the object as a face ot not. The result of qualification against sub image user 1 from Haar cascade classifier was shown Fig 11.

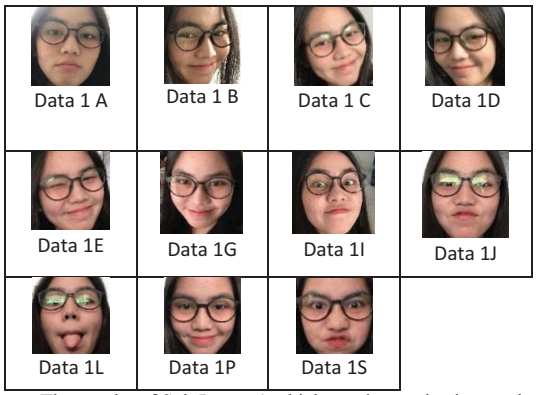

Fig 10. The results of Sub Image 1 which are detected using method

There were 11 features that werw recognized as a face of the qualifications. These features will be stored in the system database. In this test, 9 features were not recognized as a face (as shown in Fig 12).

Fig 11. The results of sub image 1 which are not detected using method

The features can not be recognized because the directions or styles make them only can be seen in one side while taking the object, for instances the pictures of 1F and 1R data. Beside that, it was also caused the object pictures were taken from top position like the pictures of 1I, 1N and 1R data. From the first test, the qualification with the proscess of Adaboost Machine inside the Haar Cascade Classifier method was appropriate.

After the qualification, trials face detection to the object sub image user 1 that already stored in the system database using putty was also conducted. Sub iamge user 1 was also known as no. Id 1526635892 on the database. Fig 13 shwos the data result of sub image test of user 1 ussing putty.

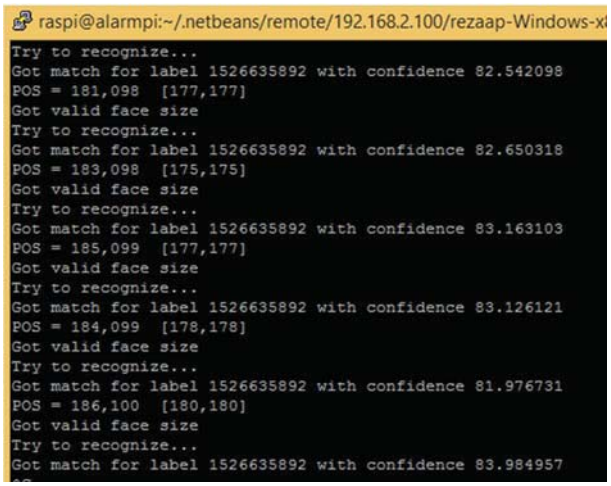

Fig 12. The data results of sub image user 1 in Putty

The data of Fig 13 shows the value of canfidence from the object of sub image of user 1 was ± 83. It indicated that the detected object was accurate and appropriate with the sub image of user 1 that has been stored in the database.

On the test, the object was a woman who wears hijab and named with sub image user 2. 30 pictures was taken from the second object. It was then, placed in the database that is shown in Fig 14.

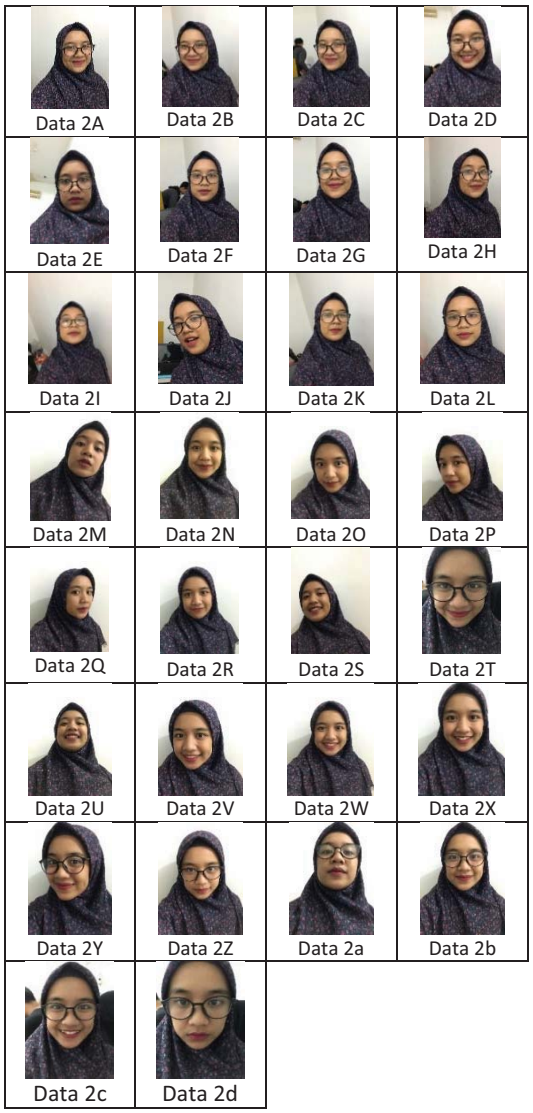

Fig 13. Sub Image User 2

The same like the first test, the pictures will be qualified using Haar Cascade Classifier method. Fig 15 shows the result of qualification of sub image of user 2.

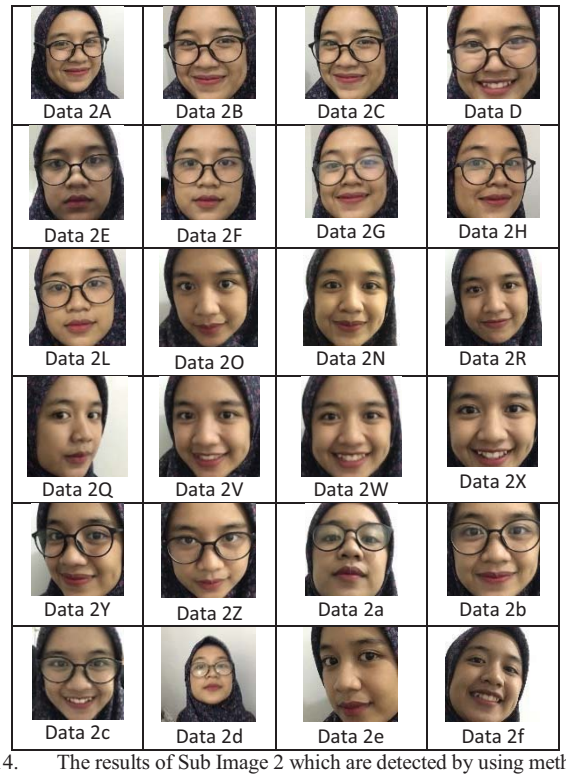

Fig 14. The results of Sub Image 2 which are detected by using method

There were 24 features that were recognized as a face from the qualifications, and it will be stored in the system database. From the test, 6 features were not recognize as a face (as shown in Fig 16).

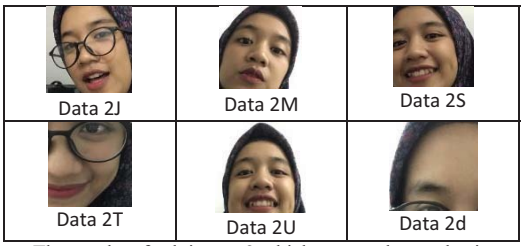

Fig 15. The results of sub image 2 which are not detected using method

The features cannot be recognized because the position while taking the pictures was only seen in one side, for instance, the data of 2J and 2S. Beside that, there were some features that could not be recognized because the background of the object also influenced the qualification of pictures, even though it has been seen clearly, such as the the data of 2M, 2U, 2T and 2 D. The process of Adaboost Machine also influenced the process of qualification of the pictures. The features that cannot be recognized from both test because it had a background that was more colorful if compared to the object, therefore the process of qualification was also processed the background of object.

In addition, the processed of qualification using Haar cascade classifier method detected the face directly from camera not from the object of sub image of user 2 that has been stored in database system. On the database features of sub image of user 2, it was known with no ID 1527664719, the result data is shown in fig 17.

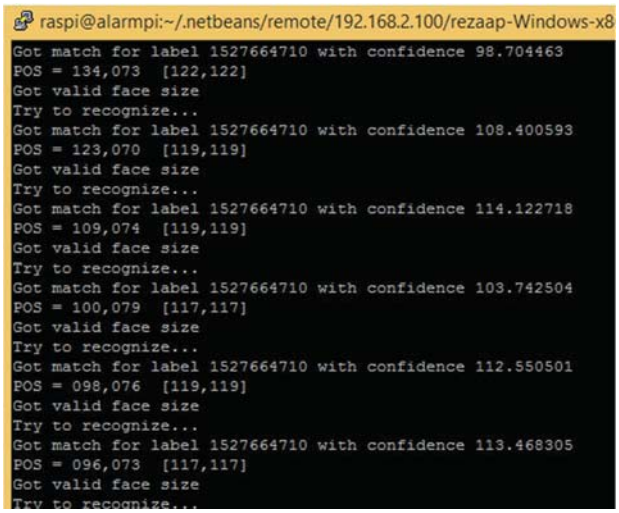

Fig 16. The result data of sub image of user 2 in Putty

The result data shows that the value of confidence from the object of sub image of user 2 is ± 100. this indicates that the detected object is less accurate and appropriate with the sub image user 2 that has been stored in the database.

V. Conclusion

The difference features in face detection can affect the number of accuracy. Taking the features of the woman who does not wear hijab as a database was easier than the woman who wears it. It was due to the woman with hijab made the number of dark and white more variable if it was compared to the other. Background of the features can also affected the accuracy of face detection. It was due to there were more varying backgrounds and colors. Therefore, it will be more difficult to distinguish dark and white between faces and background on the features. If taking the pictures was too closed and there were many different positions, it will be difficult to determine the object whether it was face or not. Besides that, it also can be caused by light and distance. It can be concluded that object, position, background and lighting on taking a picture could influence the accuracy of a feature that will be used as a system database. The resulting average confidence value between the object with hijab and without hijab was 80-100. The smaller value of confidence the better and more accurate the feature would be.

Acknowledgment

In this paper is also one of our bachelorprojects. Authors thank to the State of Polytechnic Sriwijaya, Electrical Engineering, Measurement Laboratory, Technician and all of author's friends who assist and support this research. The author also goes to the Cyborg IT Center who has provided a place for sharing and learning.

References

[1] K. D. Indonesia, "DKI Jakarta , Provinsi dengan Tindak Kriminal Terbanyak," p. 2015, 2015.

[2] K. D. Indonesia, "2016 , Jumlah Kejahatan di Indonesia Meningkat," p. 2016, 2016.

[3] I. Budiawan and Andriana, "Pengujian Pengenalan Wajah Menggunakan Raspberry Pi," vol. 6, no. 2, pp. 131–140, 2014.

[4] W. Sulistiyo, B. Suyanto, and I. Hestiningsih, "Rancang Bangun Prototipe Aplikasi Pengenalan Wajah untuk Sistem Absensi Alternatif dengan Metode Haar Like Feature dan Eigenface," *Jtet*, vol. 3, no. 2, pp. 93–98, 2014.

[5] H. H. Lwin, A. S. Khaing, and H. M. Tun, "Automatic Door Access System Using Face Recognition," *Int. J. Sci. Technol. Res.*, vol. 4, no. 6, pp. 294–299, 2015.

[6] T. ARIFIANTO Jurusan Teknik Informatika and S. Yadika Bangil, "Penerapan Fingerprint Recognition Dengan Metode Learning Vector Quantization (Lvq) Dalam Automatic Teller Machine (Atm)," vol. 9, no. 2, pp. 8–13, 2017.

[7] Y. El Anwar, N. Soedjarwanto, and A. Sadnowo Repelianto, "Prototype Penggerak Pintu Pagar Otomatis Berbasis Arduino Uno ATMEGA 328P dengan Sensor Sidik Jari," *Electr. J. Rekayasa Dan Teknol. Elektro*, vol. 9, no. 1, pp. 31–41, 2015.

[8] S. B. Jagtap, "Face Recognition and Biometric-(Fingerprint Detection)," vol. 5, no. 12, pp. 1413–1419, 2014.

[9] T. H. Le, "Applying Artificial Neural Networks for Face Recognition," *Adv. Artif. Neural Syst.*, vol. 2011, pp. 1–16, 2011.

[10] R. Dewi and Puspitasari, "Pengenalan Wajah Menggunakan Convolutional Neural Networks (Cnn) Skripsi," vol. 4, no. 3, pp. 4907–4916, 2016.

[11] A. Suryadi, "Sistem Pengenalan Wajah Menggunakan Metode Principal Component Analysis (Pca) Dengan Algoritma Fuzzy C-Means (Fcm)," *Mosharafa*, vol. 4, no. 2, pp. 58–65, 2015.

[12] H. Yang and X. A. Wang, "Cascade classifier for face detection," *J. Algorithms Comput. Technol.*, vol. 10, no. 3, pp. 187–197, 2016.

[13] V. Garg, "Face Recognition Using Haar Cascade Classifier," vol. 3, no. 12, pp. 140–142, 2016.

[14] R. Mustafa, Y. Min, and D. Zhu, "Obscenity detection using haar-like features and gentle Adaboost classifier," *Sci. World J.*, vol. 2014, 2014.

[15] R. Padilla, C. C. Filho, and M. Costa, "Evaluation of haar cascade classifiers designed for face detection," *J. WASET*, vol. 6, no. 4, pp. 323–326, 2012.

[16] M. Syarif, P. Studi, T. Informatika, F. I. Komputer, U. Dian, and N. Semarang, "Deteksi Kedipan Mata Dengan Haar Cascade Classifier Dan Contour Untuk Password Login," *Techno.com*, vol. 14, no. 4, pp. 242–249, 2015.

[17] V. Singh, V. Shokeen, and B. Singh, "Face Detection By Haar Cascade Classifier With Simple And Complex Backgrounds Images Using Opencv Implementation," *Int. J. Adv. Technol. Eng. Sci.*, vol. 1, no. 12, pp. 33–38, 2013.

[18] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Proc. 2001 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognition. CVPR 2001*, vol. 1, p. I-511-I-518, 2001.