Content Based Image Retrieval and Support Vector Machine Methods for Face Recognition

Anton Satria Prabuwono [1,2], Wendi Usino [2], Arif Bramantoro [1,2], Khalid Hamed S. Allehaibi [3], Hasniaty A. [4,5], Tomi Defisa [2]

[1]*Faculty of Computing and Information Technology Rabigh, King Abdulaziz University, Rabigh, Saudi Arabia* [2]*Faculty of Information Technology, Universitas Budi Luhur, Jakarta, Indonesia*
[3]*Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia* [4]*Institute of Visual Informatics, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia*
[5]*Faculty of Engineering, Universitas Hasanuddin, Makassar, Indonesia*

Abstract **– The development of biometrics is growing rapidly. The recognition as non-trivial element in biometrics is not only using fingerprints, but also human face. The purpose of this research is to implement both Content Based Image Retrieval (CBIR) and Support Vector Machine (SVM) methods in the face recognition system with a combination of features extraction. CBIR method interprets images by exploiting several features. The feature usually consists of texture, color, and shape. This research utilizes color, texture, shape and shape coordinate features of the image. The proposed algorithms are HSV Color Histogram, Color Level Co-Occurrence Matrix (CLCM), Eccentricity, Metric, and Hierarchical Centroid. SVM method is used to train and classify the extracted vectors. The result shows that the proposed system is 95% accurate in recognizing faces with different resolutions.**

Keywords **– Face recognition, Content-based image retrieval, Euclidean distance, Support Vector Machine.**

1. Introduction

Biometric system is a technology for self-recognition that uses human body parts or human behavior. Self-recognition is based on face, DNA, ears, eyes, nose, hands, fingerprints, and others. In computer science, there are several methods developed to recognize biometrics.

Biometric application is tailored to a specific purpose. One of the biometric applications is face recognition. In recent years, face recognition gains its popularity through the wide implementation in various fields of computing. However, computer is not human who is able to recognize human face. Therefore, computer requires artificial intelligence to be able to recognize face.

Artificial Intelligence is a general term that implies the use of computers for modeling intelligent behavior with minimum human intervention. To date, the methods of face recognition varies, such as Principal Component Analysis (PCA), Gabor Wavelet Transform, Content Based Image Retrieval (CBIR), and others. Face recognition using the Gabor Wavelet (GW) method, Wavelet Transformation (WT) and Principal Component Analysis (PCA) results in a system accuracy of 95.42% [1].

An extended method of PCA has been proposed in [2] to recognize the pictures in lower dimension that results in almost 82%. A combination between Wavelet Transform methods and Clustering Support Vector Machine in recognizing faces obtains an accuracy of 90% in [3].

Face recognition using color variants and the texture is considered as more accurate than using only one feature. This is because the more features are used in the recognition leads to more characterization from each feature precision. For example in [4], the recognition by using one feature results in 0.2 precision, while using two features gets a result of 0.6 precision.

Hence, this research proposes a face recognition system using a combination of extraction features in the Content Based Image Retrieval (CBIR) and Support Vector Machine methods. It proposes to include the features extracted from the image, such as color, texture, shape, and shape coordinates.

The algorithms in this research are HSV Color Histogram, Color Level Co-Occurrence Matrix (CLCM), Eccentricity, Metric, and Hierarchical

DOI: 10.18421/TEM82-10
https://dx.doi.org/10.18421/TEM82-10

Corresponding author: Anton Satria Prabuwono,
Faculty of Computing and Information Technology Rabigh, King Abdulaziz University, Rabigh, Saudi Arabia
Email: antonsatria@eu4m.eu

Received: 14 February 2019.
Accepted: 28 March 2019.
Published: 27 May 2019.

© 2019 Anton Satria Prabuwono et al; published by UIKTEN. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 License.

The article is published with Open Access at www.temjournal.com

Centroid. The HSV Color Histogram algorithm is used for image extraction based on its color. For extraction based on image texture, the algorithm is Color Level Co-Occurrence Matrix (CLCM). To get the form values from image, Eccentricity and Metric algorithms are used. Eccentricity is an algorithm to measure the magnitude of the length ratio between the major and the minor axis on the object shape, while Metric is an algorithm to calculate the amount of the roundness in particular object [5].

The results of extraction are in the form of vector values. This value is used to recognize different faces. The collection of images is divided into two categories: training and test image. The training image is an image that is used as a reference during the comparison with the test image. The test image is an image that matches with a similar image found in the database. Theoretically, the number of training image is more than the number of test image. The similarity between the training and the test images is calculated by utilizing Euclidean distance. Finally, the distance result is trained and classified by utilizing Support Vector Machine method.

The scope of this research is limited to the capture of images by using a smartphone camera. The object used for the testing in this research is 100 photos of 10 faces in different poses. The original image has the size of 512x512 pixels and image resolution of 256, 128, 64, 32, 16 pixels. In addition, the system is built by using the Matlab programming language in Matlab R2013a auxiliary system and the RBF kernel function.

2. Theoretical background

A. Content based image retrieval

Content Based Image Retrieval (CBIR) is an image capturing technique that uses visual features of an image, such as color, shape, texture features, and others. CBIR is also known as Query by Image Content (QBIC) and Content-based Visual Information Retrieval (CBVIR). There are two important issues in CBIR: the measurement of similarity and the representation of visual features [5].

CBIR is used to extract features. It then indexes the features by using the right and efficient structure to answer the user requests. There are several workflows in CBIR. The simple one is proposed in [6] as illustrated in Figure 1.

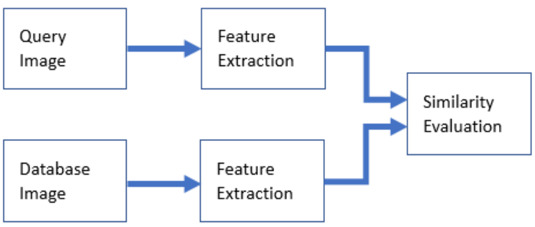

Figure 1. CBIR workflow [6]

The processing flow of the CBIR is explained as follows. At the beginning, image database is considered as a database that contains a collection of images in the database. Query image is obtained as an input image used to find the similarity of the image in the database. The query image is then extracted to obtain its features. In parallel, the database image is also extracted. At the end, the similarity evaluation is conducted based on the extracted features from these two images.

CBIR is a technique that allows users to extract the required images from databases which contains a large number of images. Hence, it searches the image based on similarities between database images and inputted images. It relies on a vector mechanism image feature as a guide to process the image.

B. Euclidean distance

In mathematics, Euclidean distance is considered as a regular distance measured from a given couple of coordinates. This distance can be manually calculated by using a mechanical measurement and the Pythagorean rule. Euclidean distance between two coordinates of *a* and *b* is basically the gap between the connecting line fragments.

Euclidean distance between two vectors is calculated in the square root of the total of differences. Euclidean Distance can be represented in the following formula.

$$D_E = \sqrt{\sum_{i=1}^{n} (|I_i - D_i|)^2} \qquad (1)$$

C. Support vector machine

Support Vector Machine is a one of the known methods to classify the datasets by discriminating two of them which have the highest distance. This method discriminates two datasets by optimizing the Hyperplane separation between them. The boundary between datasets and optimized separation is known as support vector [7].

Support Vector Machine is initially proposed by Vapnik [8], which has a better performance in image classification than other classification algorithms. Image classification is basically a process of collecting similar types of images in data set [9].

The mathematical description of SVM has several approaches, such as Optimal separating hyper planes, Linearly Separable, Linearly non-separable, and nonlinear SVM.

In high dimensional space, $(\phi(x_i).\phi(x))$ multiplication is hard to solve. For this reason, the kernel function is used with the following formula.

$$K(x_i, x) = (\phi(x_i).\ \phi(x)) \qquad (2)$$

There are two kernel functions that are widely used: Polynomial and Radial Basis Function (RBF) kernel. Polynomial kernel is represented in the following formula.

$$K(x_i, x) = (x_i . x + 1)^p \quad (3)$$

RBF kernel is represented in the following formula.

$$K(x, x_i) = \exp[-\gamma \parallel x - x_i \parallel^2] \quad (4)$$

A classification using the Support Vector Machine (SVM) method has advantages and disadvantages. The advantage of SVM includes generalization, curse of dimensionality, and feasibility.

The disadvantage of SVM is that the method is theoretically developed for classification problems that use two classes.

D. Features extraction

Feature extraction is the most important step in the CBIR system because it directly affects the level of retrieval. This process involves extracting image features to a level that can be distinguished for a better classification [10].

Visual features are color, shape and texture. These are considered as non-trivial features of an image for pattern recognition [6]. Amongst other features, color is the most significant feature of an image. Each pixel in a particular image holds a distinct color vector and a specific color space to model color information in three dimensions.

Texture represents tangible patterns and implies significant knowledge regarding the morphological organization of surfaces, such as clouds, houses, plants, hair and cloths. In addition, it also provides the association with the neighboring situation.

Shape provides a significant knowledge for human mind to spot and distinguish matters. Human has a capability in interpreting basic geometric shapes, such as a line of zero curvature facing to various ways. There are two methods in extracting shape: region and contour-based [5]. The representation and extraction techniques for shape feature are summarized in Figure 2.

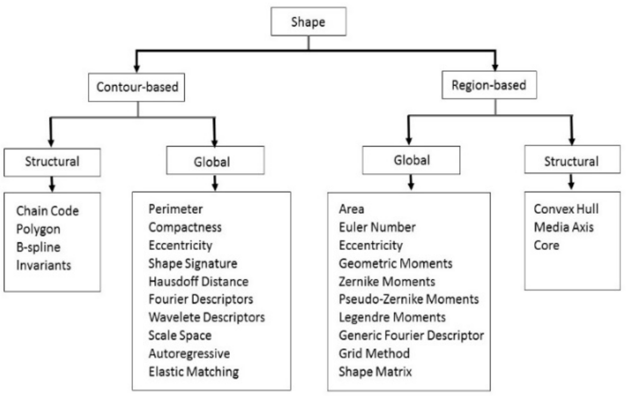

Figure 2. Representation and shape extraction techniques

E. HSV color histogram

HSV has a meaningful correlation amongst red, green, blue color space. *H* denotes Hue, *S* denotes Saturation, and *V* denotes Value. *H* value indicates the variation amongst colors. *S* value is used to characterize the color transparency. If the *S* value is 0, the image is illustrated as gray. Accordingly, white, black and other gray colors are also represented by the same value. Each color has the purest color if the saturation value reaches its maximum. Lastly, the color vividness is implied in *V* value. If V value is equal to 0, the image is illustrated as black. Figure 3. conceptualizes the model of HSV.

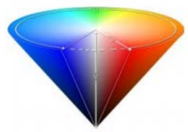

Figure 3. HSV Model

F. Color level co-occurrence matrix

Color Level Co-Occurrence Matrix is a two-dimensional matrix on the combined probability between two adjacent pixels, separated by distance d in particular direction Θ. To describe the texture state in a more intuitive way, there are several matrix parameters derived from the co-occurrence matrix. Energy is represented in the following formula.

$$F_1 = \sum_{i,j} p(i, j, d, \theta)^2 \quad (5)$$

Contrast is represented in the following formula.

$$F_2 = \sum_{i,j} |i - j|^2 \, p(i, j, d, \theta) \quad (6)$$

Homogeneity is represented in the following formula.

$$F_3 = \sum_{i,j} \frac{p(i, j, d, \theta)^2}{1 + |i - j|} \quad (7)$$

Entropy is represented in the following formula.

$$F_4 = \sum_i \sum_j p(i, j, d, \theta) \log p(i, j, d, \theta) \quad (8)$$

G. Eccentricity and metric

Eccentricity is the value of the ratio between the distance of a minor foci ellipse and the major foci ellipse of an object. Eccentricity has a range of values between 0 and 1. The model of eccentricity values is represented in Figure 4.; where *e* is eccentricity, *a* is major axis, and *b* is minor axis.

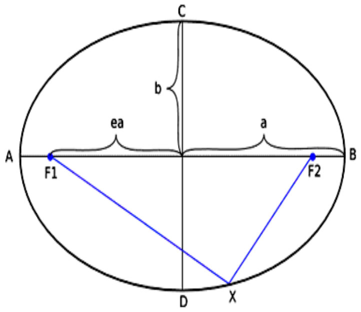

Figure 4. Eccentricity model

Metric is a comparison value between area and circumference of an object. Similar to Eccentricity, Metric has a range of values from 0 to 1. Objects that are elongated or close to a straight line have the metric value close to 0. On the other hand, objects that are round or circular have the metric value close to 1. The model of metric is illustrated in Figure 5.

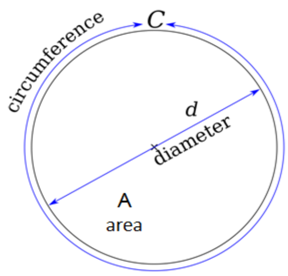

Figure 5. Metric model

H. Hierarchical Centroid

Hierarchical Centroid is used to represent the descriptors of an image. This method is based on a recursive subdivision of binary image input by measuring the centroid in each division and producing vector features in a fixed length. The vector of image is then normalized according to image size. Therefore, the size of the image does not necessarily affect the final result set. In this method, there are two processes that will be carried out: feature extraction with x-coordinate and feature extraction with y-coordinate. Binary image's x-coordinate are calculated and then the image is divided into two sub-images by x-coordinate. The vector feature length is provided by computing 2^d-1, where *d* is a parametric value that indicates recursive or division depth [11].

3. System design

In order to fulfill the objective of the research, it is required to have a comprehensive framework and a detail design. Figure 6. illustrates the conceptual framework in this research.

In this conceptual research framework, there are six steps in the conceptual framework: testing the data, training the data, calculating the similarity distance, training with SVM, classification with SVM and producing image result. The data testing and training are detailed into several sub steps. Figure 7. represents the design of training data in this research, while Figure 8. illustrates the design of test data utilized in this research.

4. Results and discussion

Once the face image is taken, then the image separation is conducted to separate training and test data. For each person, his six images are used as training data and four images are used for test data in various poses. Hence, there are 60 images as training data and 40 images as test data proposed in this research. It is important to note that the image is extracted based on color, texture, shape and shape coordinates features. Figure 9. shows several examples of face images used in this research.

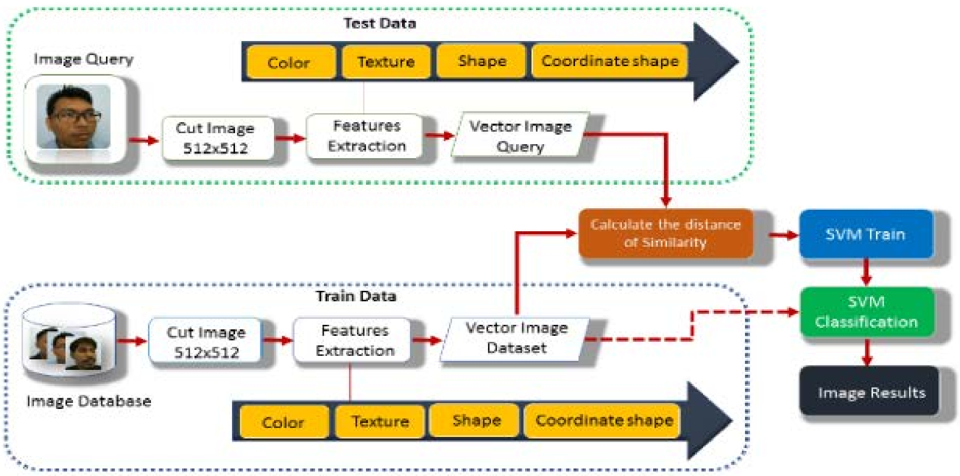

Figure 6. Conceptual research framework

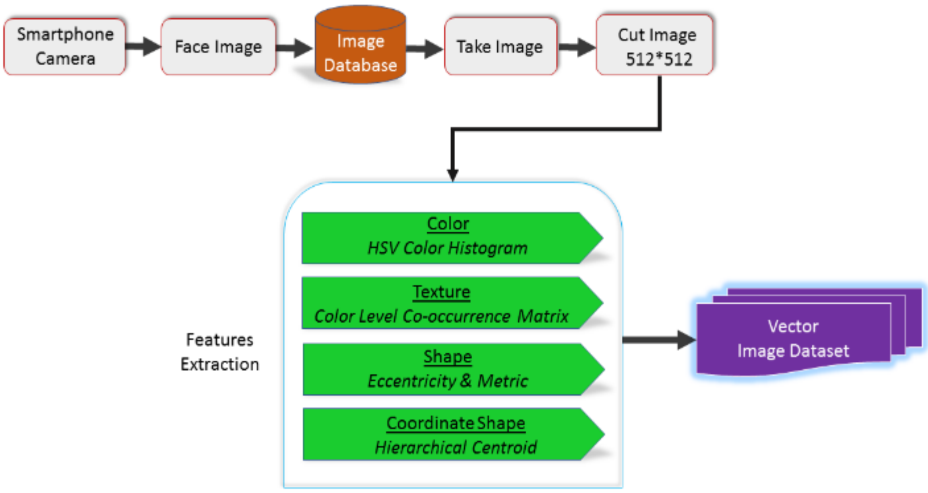

Figure 7. Data training design

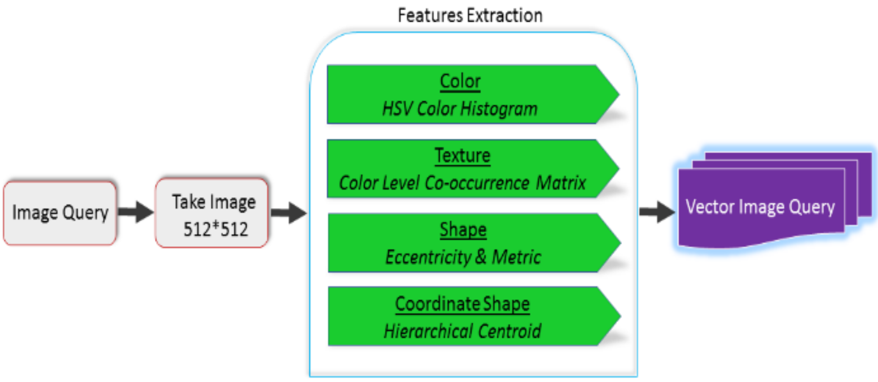

Figure 8. Data testing design

Figure 10. provides an image which results with the color extraction using the HSV Color Histogram algorithm of Hue 0.3195, Saturation 0.3259, Value 0.5563.

For texture feature extraction, the values taken from the image are Energy, Contrast, Homogeneity, and Entropy with a distance of 1 at an angle of 0, 45, 90, and 135 degrees. To obtain the Energy, Contrast, and Homogeneity values, the gray level co-occurrence matrix property is used by converting the image into grayscale and then calculating its co-occurrence matrix. The value of each parameter is based on the predetermined distance and angle. The average of these values is calculated and used for the testing phase. The results are 0.1715 for Contrast, 0.1507 for Energy, 0.9555 for Homogeneity, and 7.5172 for Entropy.

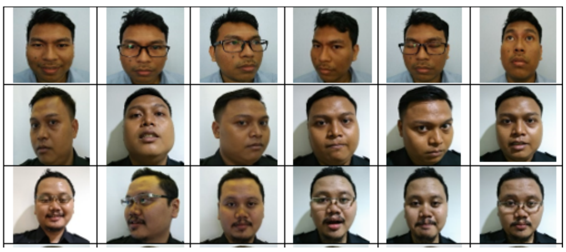

Figure 9. Face Images Samples

Figure 10. Image with Color Feature Extraction

There are two features used for shape extraction: eccentricity and metrics. The result is 0.5675 for Eccentricity, and 0.0611 for Metric. Figure 11. shows the example of an image with its shape extraction.

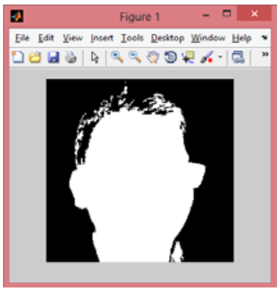

Figure 11. Image with shape extraction

For shape coordinate feature extraction, this research uses a depth value of binary digits in the grayscale image of 6 bits. The bigger the depth value of an image, the brighter the image. Image extraction results are based on coordinate shapes using centroid hierarchical as illustrated in Figure 12.

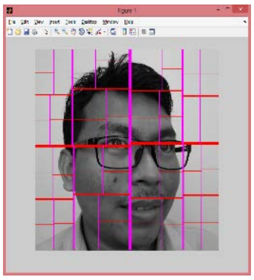

Figure 12. Image with shape coordinate extraction

The number of images used for testing in this research is ten images with different faces. Each image has four different poses which has the size of 512x512 pixels. Figure 13. shows the test results on the face recognition system using the original image.

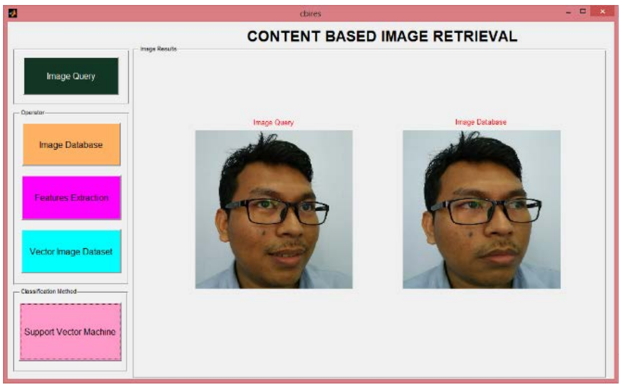

Figure 13. Original image testing

During the experiment, the query image was successfully matched with the image in the database. Table 1. shows the examples of test results.

Table 1. Original image testing results

No	Name	Image Test	Distance	Results
1	Tomi	1	0.4211	Success Identified
		2	0.2068	Success Identified
		3	0.5138	Success Identified
		4	0.5511	Success Identified
2	Viki	1	0.1072	Success Identified
		2	0.2277	Success Identified
		3	0.3345	Success Identified
		4	0.2011	Success Identified

Out of 40 total number of test images, 38 images have been successfully recognized by the system. In other words, the percentage of facial recognition system accuracy using the original image test data is 95% as illustrated in Figure 14.

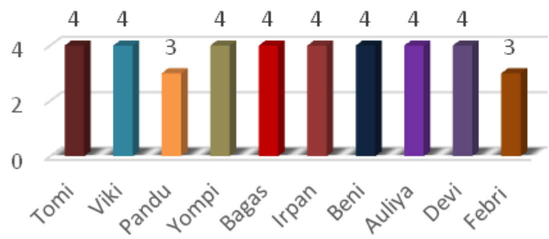

Figure 14. Original image testing results

The test was conducted 50 times because each face has five different resolutions from ten images. As previous experiment indicated that the reorganization can be conducted in different environment [12], the changing resolution is also important to be investigated for face recognition. The variations of image resolution used in this test are 256, 128, 64, 32 and 16 pixels. Figure 15. shows the test results on the face recognition system using different resolution.

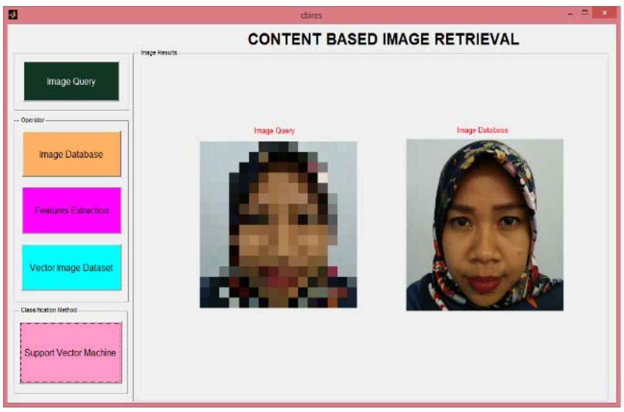

Figure 15. Different resolution image testing

Based on the test results, it can be inferred that the system is able to match query images in different resolutions with the images in the database. The system recognizes an image that has the smallest Euclidean distance value amongst other images available in the database. The sample results of testing for different resolution test images are provided in Table 2.

Table 2. Different resolution image testing results

No	Name	Test Image Resolution	Distance	Results
1	Tomi	256	0.1937	Success Identified
		128	0.2045	Success Identified
		64	0.2264	Success Identified
		32	0.2677	Success Identified
		16	0.3776	Success Identified
2	Viki	256	0.2359	Success Identified
		128	0.0822	Success Identified
		64	0.2359	Success Identified
		32	0.3131	Success Identified
		16	0.8465	Success Identified

Out of 50 images in different resolution used for testing, there are 50 images that can be recognized by the system. In other words, the accuracy of face recognition system with a different resolution is 100%. Figure 16. and Figure 17. show the result for each experiment and resolution.

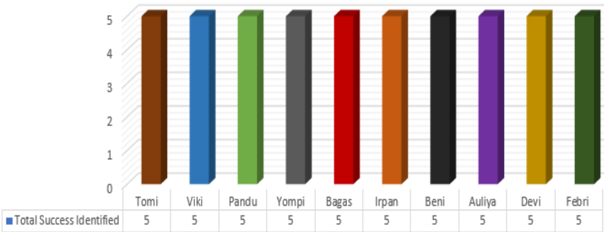

Figure 16. Recognition result of image with different resolution

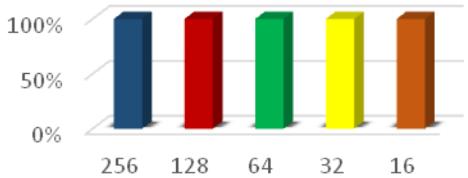

Figure 17. Percentage of difference resolution in pixel size

5. Conclusion

This research proposed a new face recognition system. Based on the results of the research and testing that has been done, it can be concluded that by combining feature extraction and Support Vector Machine (SVM) classification using RBF kernel, the accuracy of the system is 95%. In addition, this model is able to recognize face images in different resolutions. It can be inferred that using classification for processing large image data provides more accurate results than the one without classification.

For further research in the area of facial recognition, it is suggested to use the extraction of other features in the Content Based Image Retrieval method to improve accuracy. It is interesting to see the implementation of the Content Based Image Retrieval method by combining it with other face recognition methods in Android system. It is also suggested to test the image captured in different lighting in order to evaluate system performance.

References

[1] Devi, N. S., & Hemachandran, K. (2017). Content based feature combination method for face image retrieval using neural network and SVM classifier for face recognition. *Indian Journal of Science and Technology, 10(24).*

[2] Saif, A. S., Prabuwono, A. S., Mahayuddin, Z. R., & Mantoro, T. (2013). Vision-based human face recognition using extended principal component analysis. *International Journal of Mobile Computing and Multimedia Communications (IJMCMC)*, *5*(4), 82-94.

[3] Elahi, M., & Gharaee, M. (2015). Performance comparison for face recognition using PCA and DCT. *Journal of Electrical and Electronic Engineering*, *3*(2-1), 62-65.

[4] Afifi, A. J., & Ashour, W. M. (2012, December). Content-based image retrieval using invariant color and texture features. In *2012 International Conference on Digital Image Computing Techniques and Applications (DICTA)* (pp. 1-6). IEEE.

[5] Kandalkar, N., Mani, A., Pandey, G., & Soni, J. (2015). Content based image retrieval using color, shape and texture extraction techniques. *International Journal of Engineering and Technical Research (IJETR)*, *3*(5), 74-79.

[6] Malode, P., & Gumaste, S. V. (2015). A review paper on content based image retrieval. *International Research Journal of Engineering and Technology (IRJET)*, 883-885.

[7] Gao, X., Fan, L., & Xu, H. (2018). Multiple rank multi-linear kernel support vector machine for matrix data classification. *International Journal of Machine Learning and Cybernetics*, *9*(2), 251-261.

[8] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273-297.

[9] Bansal, S., & Kaur, E. R. (2014). A review on content based image retrieval using SVM. *International Journal of Advanced Research in Computer Science and Software Engineering, 4(7).*

[10] Mistry, Y., Ingole, D. T., & Ingole, M. D. (2017). Content based image retrieval using hybrid features and various distance metric. *Journal of Electrical Systems and Information Technology*, *5*(3), 874-888.

[11] Ilunga-Mbuyamba, E., Avina-Cervantes, J. G., Lindner, D., Guerrero-Turrubiates, J., & Chalopin, C. (2016, February). Automatic brain tumor tissue detection based on hierarchical centroid shape descriptor in Tl-weighted MR images. In *2016 International Conference on Electronics, Communications and Computers (CONIELECOMP)* (pp. 62-67). IEEE.

[12] Prabuwono, A. S., Burhanuddin, M. A., & Said, S. M. (2008). Autonomous contour tracking using staircase method for industrial robot. *10th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, 2272-2276.