Recognition of Criminal Perpetrators using Multi Otsu Thresholding and Content-based Image Retrieval Approach

Suhenrro Y. Irianto
Institute Informatics and Business Darmajaya
Jl. Z.A. Pagar Alam No.93
Bandar Lampung. Indonesia

Sri Karnila
Informatics and Business Darmajaya
Jl. Z.A. Pagar Alam No.93
Bandar Lampung. Indonesia

Adimas Aglasia
Institute Informatics and Business Darmajaya
Jl. Z.A. Pagar Alam No.93
Bandar Lampung. Indonesia

ABSTRACT
The development of human face recognition techniques is highly complex. multidimensional. and often subject to changes based on environmental and psychological conditions. The creation of a system is urgently needed and crucial to assist law enforcement. such as determining photos of faces suspected of being involved in criminal activities. With automated tools. it becomes possible to provide or display suspected faces in accordance with desired queries. In the legal domain. searching for the faces of criminals or fugitives is essential because not all criminal activities are captured by CCTV and other means. Therefore. sketch images based on eyewitness accounts are employed. Law enforcement typically seeks the assistance of skilled artists. especially facial sketch artists. to create facial sketches of criminal suspects based on information provided by eyewitnesses. even if only briefly observed. Developing a system for searching facial images using sketches by artists is immensely helpful in identifying criminal suspects and enables law enforcement to pinpoint individuals or groups under suspicion. Overall. out of 400 face images. 328 are correctly matched. and 72 are unmatched. The overall precision for the entire dataset is 82%. In This research employs two methods for creating a criminal face recognition system. namely. segmentation and Content-Based Image Retrieval (CBIR).

General Terms
Computer vision. image processing., image segmentation

Keywords
Face recognition. CBIR. Criminal, Sketch Images

1. INTRODUCTION
The development of face recognition techniques. as stated by [1]. is quite challenging due to the complexity. multidimensionality. and dynamic nature of the human face. which changes in response to environmental shifts and an individual's mental state. Therefore. automatic face recognition systems pose a challenge for experts. Changes in facial conditions. such as identity alterations and variations caused by lighting and different facial image angles. present a challenge in representing faces accurately for face recognition purposes [1]. [2]. The creation of an application or tool is both urgent and crucial to aid law enforcement. such as determining photos of faces suspected of criminal activities. An automated tool that can provide or display suspected faces based on desired queries is essential. However. in many cases. the desired photos are unavailable or not present in the police database. To overcome this. law enforcement usually seeks the assistance of skilled artists. especially facial sketch artists. to create sketches of criminal suspects based on eyewitness accounts. even if only briefly observed [3]–[5]. Searching for a facial image using a sketch by an artist is immensely helpful in identifying criminal suspects and allows law enforcement to pinpoint individuals or groups under suspicion. Additionally. such sketches are expected to assist artists and eyewitnesses in refining sketches based on retrieved facial images. However. a remaining challenge is how to transform these sketches into realistic facial images or photos of individuals. This research aims to explore how to identify criminal suspects based on sketches and how to generate realistic facial photos or images based on these sketches. The objective is to develop an application or tool system that can be used automatically for recognizing the faces of criminal suspects based on sketches created by an artist or someone trained in sketching faces.

According to [6] .an image is a representation. resemblance. or imitation of an object. As a result of data recording systems. images can be optical. such as photos. analog. like video signals on a television monitor. or digital. directly storable on a storage medium. Digital images are those that can be processed by computers. a term highly prevalent in the present era. Various electronic devices. such as scanners. digital cameras. digital microscopes. and fingerprint readers. produce digital images and are widely used for photo editing and other purposes. For instance. Adobe Photoshop and GIMP (GNU Image Manipulation Program) provide various features for manipulating digital images. JPEG is a standard compression method for photographic images. The term JPEG stands for the Joint Photographic Experts Group committee. which established the format's standards in 1992. later recognized as ISO/IEC 10918-1 in 1994. JPEG is a graphic image format highly useful for creating high-quality photographic images in small file sizes. This graphic file format has been accepted by the International Telecommunication Union Telecommunication Standardization Sector (ITU-T) and the International Organization for Standardization (ISO). [7]. [8]. Content-Based Image Retrieval (CBIR) is a technique that uses visual content to search for images in large-scale databases. It is a well-known technique widely used in image searches based on visual content such as color. texture. shape. and other features extracted in vector form. CBIR is also referred to as a computer vision application that addresses image search issues on a large scale,[9]. [10].

In this research. Euclidean Distance calculation will be utilized. According to [11], the most frequently used method for calculating the similarity between two vectors is Euclidean Distance. The equation for this method is as follows:

$$d(A,B) = \sqrt{\sum_{j=1}^{n}(H_j^A - H_j^B)^2} \tag{1}$$

Where A: Vector A. B: Vector B .d(A.B): Euclidean distance between vector A and vector B .n: Number of vector elements. and H: Vector elements. This method used in the research to calculate the accuracy of similarity between two vectors: one from the segmented database using Multi Otsu Thresholding and the other from facial sketches as input.

Image segmentation involves separating one object from another in an image or separating objects from the background in an image. With segmentation. individual objects in an image can be extracted for further processing. There are two main approaches to image segmentation: edge based. dividing the image based on discontinuities between sub-regions. and region-based. dividing the image based on regions. The segmentation process in this research involves edge detection using the Prewitt and Sobel methods [12]. [13]. According to [7], the Otsu method is highly popular among all thresholding methods and is the best method for obtaining automatic threshold values. Thresholding converts a grayscale image into a binary image based on the threshold value (T). distinguishing between object and background areas. If the pixel value is greater than the threshold. it is set to 1; otherwise. if it is less than the threshold. it is set to 0. Equation 2. generally, describes the thresholding process.

$$g(x,y) = \begin{cases} 1 \ if \ f(x,y) \geq T \\ 0 \ if \ f(x,y) < T \end{cases} \tag{2}$$

Where T represents the threshold value. According to [8], Otsu's thresholding concept was first introduced by Nobuyuki Otsu (1979) for binarizing images based on histogram shape automatically. It assumes that the image contains two basic classes with a bimodal histogram (foreground and background). Moreover the goal of the Otsu method is to automatically divide the grayscale image histogram into two distinct regions without user assistance. The approach involves discriminant analysis. determining a variable. and maximizing that variable to separate objects from the background.

2. METHOD

This research employs an experimental research design. comprising stages such as Face Image Data Collection. Face Recognition with Sketches. Face Image Engineering. Similarity Determination. Method Experimentation and Testing. and Evaluation and Validation of Results. In this study. approximately 7.800 face images with various poses. positions. and brightness levels will be used. The database includes normal face images (facing directly or 90 degrees to the front) and abnormal face images with obstacles. such as faces wearing glasses or covered with masks or veils. The first step in face recognition with sketches is face detection. which determines if a human face is present in the image and its position. The output of face detection is expected to be the face region or area. To enhance the accuracy of face detection. face alignment is performed to standardize the face region. Additionally. pre-processing for face detection involves determining the Region of Interest (ROI).

2.1 Face Image Engineering

This research involves three stages in matching face images. The first stage is searching for topological similarities between the query face and the database face. acting as a filter. The second stage utilizes information to improve potential face images. and the final stage involves matching calculation methods to determine similarity between the sketch query and face images in the database. Before determining similarity. DC extraction will be performed. starting from pixel decomposition to grayscale image formation. The similarity determination will use Euclidean distance calculation. Two methods will be employed to determine face similarity from face sketches. The first method utilizes segmentation. specifically, Multi Otsu Thresholding segmentation. to obtain objects within the image or divide the image into regions. with each object or region having similarity attributes in the provided database for use as reference data in the similarity process. The second method uses CBIR for face similarity. determining similarity based on training data. In face similarity. Euclidean distance calculation will be applied. Two methods used to calculate face similarity in this study: Multi Otsu Thresholding segmentation and CBIR. both utilizing Euclidean distance calculation for similarity determination.

2.2 Multi Otsu Thresholding Segmentation

According to [14] image segmentation in this research is used to obtain objects within the image or divide the image into regions. with each object or region having similarity attributes. The Otsu method is used to determine the threshold value. The Otsu method automatically obtains the threshold value by converting a grayscale digital image into a black and white image based on the comparison of threshold values with the pixel color values in the digital image.

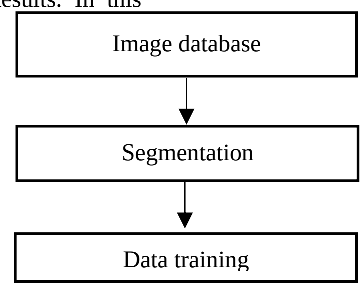

Fig. 1. Segmentation process into training data Face Similarity with CBIR

After generating face images based on sketches. CBIR is used to determine face similarity. Euclidean Distance calculation is used for feature extraction. which will be utilized for face classification and recognition. The process involves segmentation and then measuring similarity using Euclidean Distance. Moreover [16] the above process involves segmenting the database images for use as reference or training data. The next step is similarity determination in CBIR to identify image similarity. The system for Multi Otsu Thresholding segmentation is illustrated in Fig.6. The Multi Otsu Thresholding segmentation process involves reading criminal face images from the database as input. After reading the database. default values are applied to standardize the input image. Subsequently. the program generates the image into three regions using Multi Otsu Thresholding. separating three objects with different colors. This results in a vector pattern.

and the vectorized image can be used as reference data for similarity with face sketches to achieve maximum accuracy. After the generation step. the segmented image is saved for use as reference data in the similarity process. and the program concludes, [15].

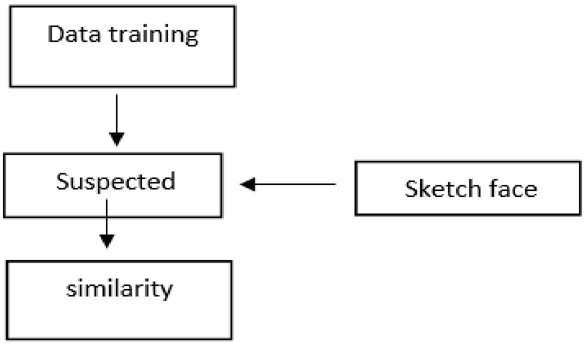

Fig. 2. Similarity process with CBIR

The matching similarity process involves finding the retrieved image based on facial similarity within the segmented database using Multi Otsu Thresholding. The testing of matching similarity results involves Euclidean Distance calculation. aiming to find the average precision and recall time values.

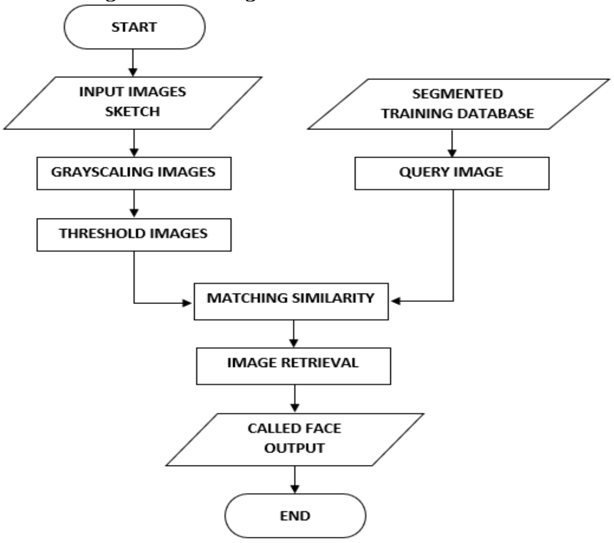

Fig. 3. Overall Matching Similarity Process with CBIR

Whilst [17] and [18] indicated that in the Multi Otsu Thresholding and CBIR segmentation process. matching similarity is performed. In Multi Otsu Thresholding. the stage involves generating face images in the database into three vector parts after the initial grayscale database process. In this process. the previously processed DC coefficient is binarized and divided into three parts based on the histogram. The foreground consists of pixels with large values. while the background consists of pixels with smaller value.

3. RESUTS AND DISCUSSION

3.1 Matching Similarity Results

The threshold value in the face sketch will undergo matching similarity by creating two variables. each for the segmented database and the face sketch. Euclidean Distance calculation is then performed. The matching similarity process using Euclidean Distance is illustrated in Figure 4.

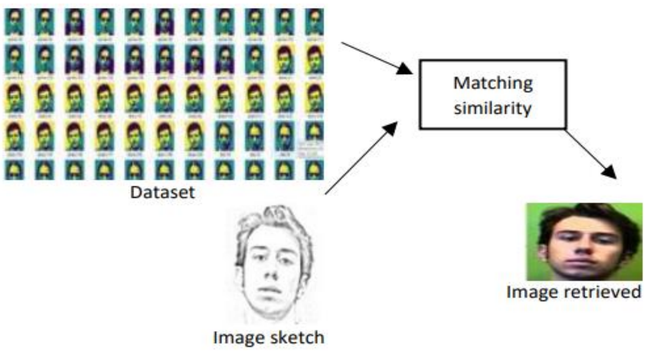

Fig. 4. Matching Similarity Process with Euclidean Distance

3.2 Euclidean Distance Calculation Results

To calculate the precision of the built CBIR model. testing is conducted on a dataset of 1080 images with 20 queries on artificial images. including 50 different face sample images. Each face sample contains 20 face queries. The test calculates the number of matching and non-matching images based on sketches and the database segmented using Multi-Otsu Thresholding. The results of the Euclidean Distance calculation are shown in Table 1. Table 1 shows the results of a face recognition system evaluation with various individuals and their corresponding metrics. Precision is a measure of the accuracy of the positive predictions made by the system. The precision values for individual names range from 0.65 to 1. On average. the precision is 82%. indicating that around 82% of the positively identified images are correct. Time represents the time taken for the face recognition process. The time values for individual names range from 10.65 to 13.04 seconds. The average time is 11.51 seconds. Overall Evaluation. the system generally performs well in terms of precision. with most individuals having high precision values (above 0.75).

The time taken for recognition varies. but the average time is 11.51 seconds. It seems that some individuals have 100% precision. suggesting that their images are always correctly identified by the system. Recall is not explicitly mentioned in the table. but the average time (11.51 seconds) may be considered as a measure of system efficiency. Individual Analysis. like "casano." "sega." "jhon." and "jaflem" have perfect precision. indicating that the system correctly identifies all their images.

Table 1. CBIR Euclidean Distance Testing Results

No	Name	Number of face images	Similar images	Unsimilar images	Precision	Time (second)
1	Felix	20	16	4	0.80	11.72
2	Baner	20	12	8	0.60	10.88
3	Casano	20	20	0	1.00	10.68
4	Jonson	20	13	7	0.65	11.48
5	Smith	20	18	2	0.90	12.42
6	Sega	20	20	0	1.00	11.01
7	Akatsi	20	16	4	0.80	10.78
8	Ambraw	20	15	5	0.75	10.86
9	Choi	20	18	2	0.90	11.21
10	Chris	20	16	4	0.80	12.29
11	Dioan	20	17	3	0.85	10.65
12	Jhon	20	20	0	1.00	10.62
13	Jaflem	20	20	0	1.00	11.21
14	Jame	20	17	3	0.85	11.36
15	Darda	20	16	4	0.80	11.83
16	Dgemen	20	15	5	0.75	11.99
17	Hsgrim	20	16	4	0.8	12.44
18	Kristo	20	14	6	0.7	12.63
19	Melky	20	16	4	0.8	11.25
20	Stave	20	13	7	0.65	13.04

Individuals like "baner" and "stave" have lower precision. suggesting that the system may have more false positives for them. The number of matched and unmatched images for each person is provided. Overall. out of 400 face images. 328 are correctly matched. and 72 are unmatched. The overall precision for the entire dataset is 82%. In summary. the system appears to have a good overall performance with high precision and reasonable processing times. However. individual variations in precision and time may require further investigation. Additionally. recall is not explicitly provided in the table. so, it's essential to consider the system's ability to correctly identify all relevant instances of a face.

In the Euclidean Distance computation results above. 400 queries were used with 50 different faces. each with 20 image samples. The average precision value is 82%. and the average recall time is 11.51 seconds. The results show the number of matching and non-matching images based on sketches and the database segmented using Multi-Otsu Thresholding. The precision value is calculated as the percentage of matching images out of the total queries. Figure 6 shows recognized Face Images

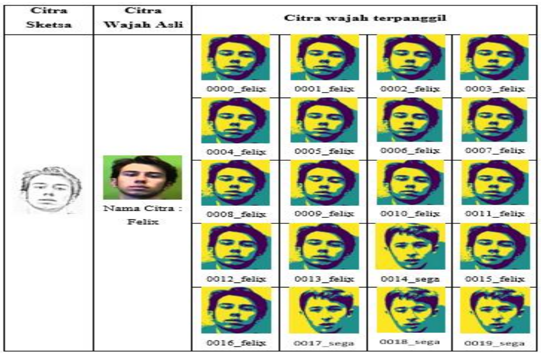

Fig. 6. The recognized face images for the name "Felix" are shown below:

The graph below depicts the results of the CBIR calculation test. including image precision. average recall time. and the number of recognized and unrecognized images illustrated on figure 7.

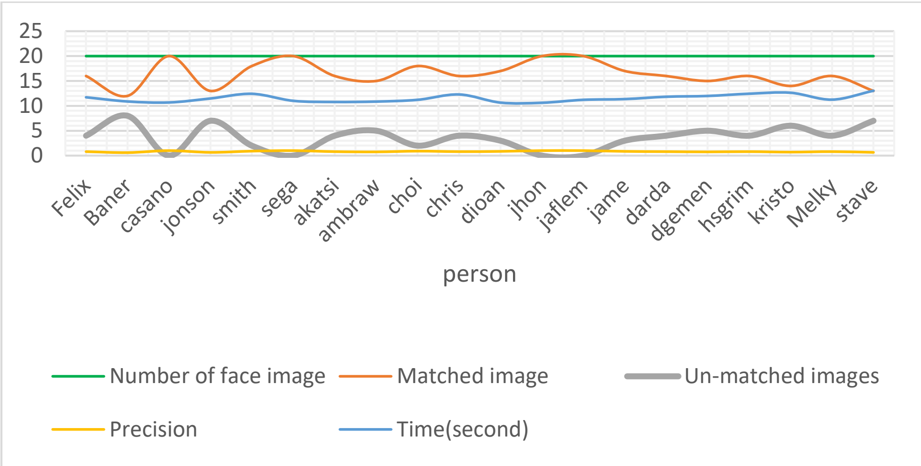

Fig 7. Graph of CBIR Euclidean Distance Calculation Test

3. CONCLUSION

Based on the conducted experiments using the segmentation process and Euclidean distance matching with the available image dataset. the following conclusions were drawn: After the preprocessing stage. which includes DC extraction by converting the images to grayscale. resolution normalization was applied to the database images. resulting in a resolution of 1800x2400. The segmentation process using Multi Otsu Thresholding produced well-segmented images. Objects and backgrounds were effectively separated. forming clear vector images and resulting in high-quality segmented images. The matching process using the Euclidean distance method achieved an average precision of 82% over 20 trials with different face sketches and a dataset of 1.080 images. The precision ranged from a minimum of 60% to a maximum of 100% in the testing of image precision using 20 different experimental face sketches.

4. CONFLICT OF INTEREST

The author declares no conflict of interest.

5. ACKNOWLEDGMENT

Our thanks to the Ministry of research and Technology, the Republic of Indonesia who support with funding through Hibah Kompetive Nasional. Also thanks to Research Center of Institute Informatics and Business Darmajaya who help finishing this work on time.

6. REFERENCES

[1] O. R. Shahin, R. Ayedi, A. Rayan, R. M. A. El-Aziz, and A. I. Taloba, “Human Face Recognition from Part of a Facial Image based on Image Stitching,” International Journal of Advanced Computer Science and Applications, vol. 12, no. 12, 2021, doi: 10.14569/IJACSA.2021.0121260.

[2] Umar Aditiawarman, Dimas Erlangga, Teddy Mantoro, and Lutfil Khakim, “Face Recognition of Indonesia’s Top Government Officials Using Deep Convolutional Neural Network,” Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), vol. 7, no. 1, 2023, doi: 10.29207/resti.v7i1.4437.

[3] S. Setumin and S. A. Suandi, “Cascaded Static and Dynamic Local Feature Extractions for Face Sketch to Photo Matching,” IEEE Access, vol. 7, 2019, doi: 10.1109/ACCESS.2019.2897599.

[4] N. Lakshmi and M. P. Arakeri, “A novel sketch based face recognition in unconstrained video for criminal investigation,” International Journal of Electrical and Computer Engineering, vol. 13, no. 2, 2023, doi: 10.11591/ijece.v13i2.pp1499-1509.

[5] C. Benavides Alvarez, J. Villegas Cortez, G. Roman Alonso, and C. Aviles Cruz, “Face Classification by Local Texture Analisys through CBIR and SURF Points,” IEEE Latin America Transactions, vol. 14, no. 5, 2016, doi: 10.1109/TLA.2016.7530440.

[6] V. Tyagi, Understanding Digital Image Processing. CRC Press, 2018. doi: 10.1201/9781315123905.

[7] G. K. Wallace, “The JPEG Still Picture Compression Standard,” Commun ACM, vol. 34, no. 4, 1991, doi:

10.1145/103085.103089.

[8] G. K. Wallace, "The JPEG still picture compression standard," IEEE Transactions on Consumer Electronics, vol. 38, no. 1, 1992, doi: 10.1109/30.125072.

[9] M. Fachrurrozi et al., "Real-time multi-object face recognition using content based image retrieval (CBIR)," International Journal of Electrical and Computer Engineering, vol. 8, no. 5, 2018, doi: 10.11591/ijece.v8i5.pp.2812-2817.

[10] M. Fachrurrozi et al., "Real-time Multi-object Face Recognition Using Content Based Image Retrieval (CBIR)," International Journal of Electrical and Computer Engineering (IJECE), vol. 8, no. 5, 2018, doi: 10.11591/ijece.v8i5.pp2812-2817.

[11] M. I. Wahyudi, E. W. Wibowo, and S. Sopiullah, "Web-Based Face Recognition using Line Edge Detection and Euclidean Distance Method," Edumatic: Jurnal Pendidikan Informatika, vol. 6, no. 1, 2022, doi: 10.29408/edumatic.v6i1.5525.

[12] H. Yasin and S. Hayat, "DeepSegment: Segmentation of motion capture data using deep convolutional neural network," Image Vis Comput, vol. 109, 2021, doi: 10.1016/j.imavis.2021.104147.

[13] Z. Wang, K. Wang, F. Yang, S. Pan, and Y. Han, "Image segmentation of overlapping leaves based on Chan–Vese model and Sobel operator," Information Processing in Agriculture, vol. 5, no. 1, 2018, doi: 10.1016/j.inpa.2017.09.005.

[14] M. H. J. Vala and A. Baxi, "A review on Otsu image segmentation algorithm," International Journal of Advanced Research in Computer Engineering and Technology, vol. 2, no. 2, 2013.

[15] S. Sikandar, R. Mahum, and A. M. Alsalman, "A Novel Hybrid Approach for a Content-Based Image Retrieval Using Feature Fusion," Applied Sciences (Switzerland), vol. 13, no. 7, 2023, doi: 10.3390/app13074581.

[16] S. Maji and S. Bose, "CBIR Using Features Derived by Deep Learning," ACM/IMS Transactions on Data Science, vol. 2, no. 3, 2021, doi: 10.1145/3470568.

[17] Z. Y. Tan, S. N. Basah, H. Yazid, and M. J. A. Safar, "Performance analysis of Otsu thresholding for sign language segmentation," Multimed Tools Appl, vol. 80, no. 14, 2021, doi: 10.1007/s11042-021-10688-4.

[18] A. M. Alhassan, "Enhanced Fuzzy Elephant Herding Optimization-Based OTSU Segmentation and Deep Learning for Alzheimer's Disease Diagnosis," Mathematics, vol. 10, no. 8, 2022, doi: 10.3390/math10081259.