2021 3rd East Indonesia Conference on Computer and Information Technology (EIConCIT)
April 09-11. 2021, ISTTS Surabaya, Indonesia

Recognition of Indonesian Sign Language Alphabets Using Fourier Descriptor Method

2021 3rd East Indonesia Conference on Computer and Information Technology (EIConCIT) | 978-1-6654-0514-0/20/$31.00 ©2021 IEEE | DOI: 10.1109/EIConCIT50028.2021.9431883

1st Syartina Elfarika Basri
Faculty of Computer Science
Universitas Muslim Indonesia
Makassar, Indonesia
syartinaelfarika.labfik@umi.ac.id

2nd Dolly Indra
Faculty of Computer Science
Universitas Muslim Indonesia
Makassar, Indonesia
dolly.indra@umi.ac.id

3rd Herdianti Darwis
Faculty of Computer Science
Universitas Muslim Indonesia
Makassar, Indonesia
herdianti.darwis@umi.ac.id

4th A. Widya Mufila
Faculty of Computer Science
Universitas Muslim Indonesia
Makassar, Indonesia
widya.mufila@umi.ac.id

5th Lutfi Budi Ilmawan
Faculty of Computer Science
Universitas Muslim Indonesia
Makassar, Indonesia
lutfibudi.ilmawan@umi.ac.id

6th Bobby Purwanto
Faculty of Computer Science
Universitas Muslim Indonesia
Makassar, Indonesia
bobbypoerwanto@umi.ac.id

***Abstract*—Feature extraction is a process to search for special features of an object that will distinguish the characteristics between one object and another. In this study, the Fourier Descriptor method was used for extracting the feature of the Indonesian Sign Language (BISINDO) images. This research has been performed by implementing four main steps, i.e., pre-processing by converting RGB images to grayscale and binary ones, and closing operation for smoothing the images; contour detection using Moore's algorithm; feature extraction implementing Fourier descriptor with 5 kinds of coefficient i.e., 2, 5, 10, 25, and 50 to represent the feature of the images; and lastly feature recognition process by computing the image similarity using Euclidean Distance. 1820 images divided into 4 kinds i.e., standard, scaled, rotated, and translated images have been tested with 130 images of training data. Based on the test result of 130 standard images for each coefficient, the best accuracy is obtained at coefficient 25 and 50 with a similar accuracy of 96.92%. In addition, the recognition performance of Fourier descriptor and Euclidean distance reached up to above 72% in average for standard and scaled images.**

Keywords*—*BISINDO, feature extraction, Fourier descriptor, pattern recognition, Euclidean

I. Introduction

Sign language is a form of communication without making a sound but using body language and lip movements, as well as a combination of hand, arm, and body movements. In Indonesia, there are varieties of sign languages but two commonly used sign language systems are the Indonesian Sign Language which is also as known as IndoSL or *Bahasa Isyarat Indonesia* (BISINDO), and Indonesian Signed System or *Sistem Isyarat Bahasa Indonesia* (SIBI). The main difference distinguished the two is that SIBI letters can be illustrated by only one hand while BISINDO requires two hand to demonstrate some letters. As a matter of fact, an organization called *Gerakan untuk Kesejahteraan Tunarungu Indonesia* (GERKATIN) or Indonesian Association for the Welfare of the Deaf has been formed since 1966 by the deaf community and considering their needs of communication method then BISINDO was legally created and widely used in Indonesia. Vocabularies of BISINDO has grown rapidly as well and become inseparable culture for the deaf because the language was developed originally by Indonesian deaf people themselves [1].

Digital image recognition is an approach in *computer vision* that applies pattern recognition and widely developed. Regarding with BISINDO image, some previous research have been performed focusing the image contour detection using chain code [2], the usage of color tracing on the edge detection process [3], and the gesture recognition involving movements of the arms, shoulders, and hands [4].

In terms of feature extraction, Fourier descriptor is one of methods popularly used in object recognition and image processing to represent the shape of segment boundaries in images, and the advantage of this method is that it is independent on translation, rotation, and scaling in various simple or complicated shapes [5]–[10]. Considering the benefits, Fourier descriptor is specifically proposed for extracting the feature of digital image of BISINDO letters so that the study of BISINDO letters can be more deeply studied. Furthermore, as one of reliable measures for determining object similarity, Euclidean distance is used in the final process for recognizing the letters.

II. Experimental Details

This research was performed by implementing 4 main steps as covered in Fig. 1, obtaining:

- Pre-processing: Image conversion and Closing operation
- Contour detection: Moore's algorithm for edge detection.
- Feature detection: Fourier descriptor for extracting the feature of the images using 5 various coefficients.
- Feature recognition: Euclidean distance for calculating the feature similarity of the images.

A. Pre-processing

Pre-processing is the initial process to improve image quality, remove noise in the image and determine which part of the image will be processed next. Firstly, the process of converting a Red Green Blue (RGB) image is performed to a grayscale image then converting it to a binary image. Secondly, a morphological process which is the image closing operation is conducted to smooth the contours and remove small holes as well as to obtain a clearer image. On the contrary of opening operation applying erosion process followed by dilation in morphology, a closing operation is the opposite one which implements erosion after dilation. Fig. 2 illustrates the closing operation result which tends to enlarge the edge area in the image while reducing the color background in that area [11], [12].

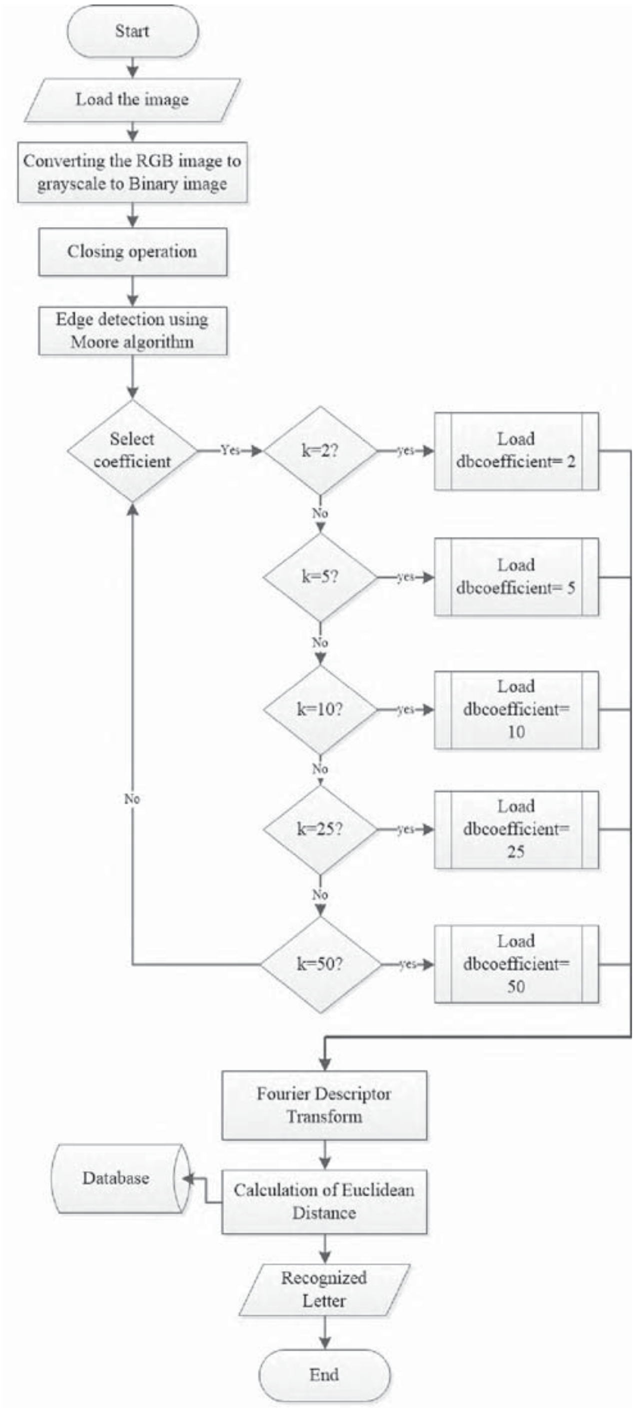

Fig. 1. Recognition process

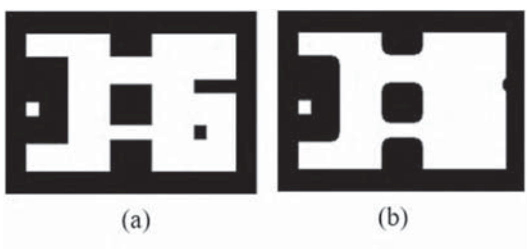

Fig. 2. (a) Original image, (b) Image closing operation result

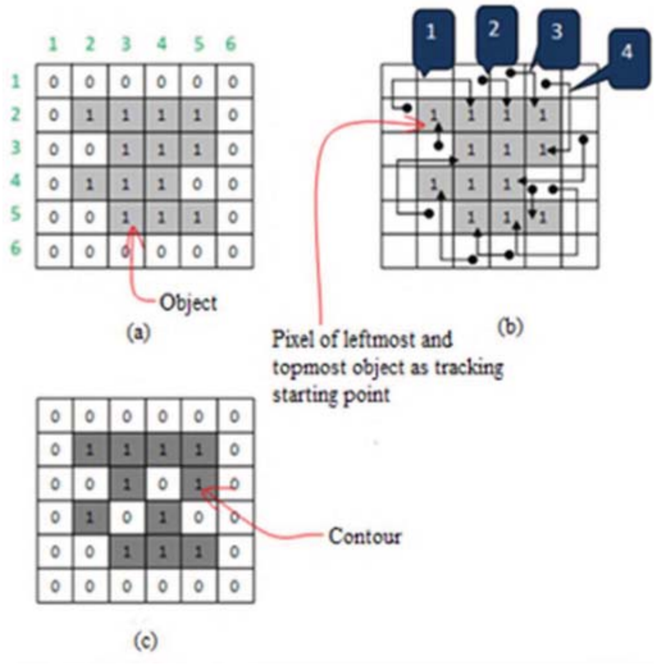

Fig. 3. Steps of Moore's algorithm

B. Contour Detection

One way to obtain internal contours that have been sorted according to pixel location is by using contour tracking of Moore's algorithm. Moore's algorithm is considered to be the most efficient algorithm in general applications [13]. Let a binary object given representing an image where each pixel has a value of 1 or 0 as in Fig. 3(a). Contour detection begins from the left with a pixel value of 1 or an object. To get a value of 1 or an object, each pixel is checked by moving clockwise as in Fig. 3(b). This is performed repeatedly so that it forms a contour as in Fig. 3(c).

C. Fourier Descriptor

Fourier transform is one of most used techniques applied in the digital signal processing and solving partial differential equations algorithms for multiplication of integers in large numbers. In this study, Fourier descriptor is used in the process of feature extraction which is a data extraction to search for special features of an object which will differentiate one object from another. Fourier descriptors itself describe the shape in terms of its spatial frequency content. Fourier descriptors are the boundary based descriptors in 2 dimensional space using the Fourier methods [7].

The basic concept for obtaining a Fourier descriptor in the process of feature extraction is that the contours of the object need to be obtained as the initial stage. Next, the pixels in the contour are transformed using Fast Fourier Transform (FFT) which refers to Discrete Fourier Transform (DFT) but in an effective algorithm computing DFT. If $\{x[n]\}_{n=0}^{N-1}$ is sequence of image pixels, DFT itself is a sequence $X[k]$ for $k = 0,1,2,\ldots,N-1$ formulated as in (1) and the inverse is written in (2).

$$X[k] = \sum_{i=0}^{N-1} x[n] W_N^{kn} \tag{1}$$

$$X[n] = \frac{1}{N}\sum_{k=0}^{N-1} X[k]W_N^{kn}, n = 0,1,2,\dots,N-1 \quad (2)$$

Note that $W_N^{kn} = e^{-\frac{2\pi jnk}{N}}$ is twiddle factor and the index starts from 1 in MATLAB.

D. Euclidean Distance

Euclidean distance is a distance calculation method used to measure the distance from 2 (two) points in Euclidean space (covering two-dimensional, three-dimensional, or even more Euclidean planes) Euclidean is a heuristic function that is obtained based on the direct distance free from obstacles such as to obtain the value of the length of the diagonal line in the triangle [14], [15]. Let x, y be two n feature values extracted from two images, $x = (x_1, x_2, \dots x_n), y = (y_1, y_2, \dots, y_n)$ then to measure the level of data similarity with Euclidean Distance is formulated as in (3).

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad (3)$$

E. Datasets

Dataset used in this study was divided into two, i.e., training data and test data. The training data used is secondary data obtained from previous research regarding the common object which is the image of 26 BISINDO letters from A to Z [2]. As for the test data, it was primary data collected by capturing several people as objects to demonstrate the BISINDO letters. The training data used was 130 images of BISINDO letters obtained from 26 original raw letters which were processed with 5 types of coefficients (2, 5, 10, 25, 50) each letter. On the other hand, the test data performed involved 4 kinds of images taken from 5 different people demonstrated the letter from A to Z with 1820 images of the letters BISINDO in total as shown in Table I.

In Table I, standard data means that the image was the original ones without any transformation and similar with training data sourced from different people. In contrast to the standard data, scaled images were image data that have been enlarged twice in size of the standard ones. As for the rotation test data, image data have been rotated 90° and 180°. Meanwhile, the translation test data was image data that has been translated by shifting the image by 15 pixels on the right of x-axis and the left of y-axis.

III. Result and Discussion

In the initial stage, pre-processing performed was image conversion. Fig. 4(a) illustrates the RGB image of letter B representing two-hands letter and letter C representing one-hand letter which were then converted into grayscale format shown in Fig. 4(b) and followed by conversion into binary image revealed in Fig. 4(c).

TABLE I. Testing Data

Transformation	Source × letters × coefficients	Total
Standard	5 × 26 × 5	650
Scale	3 × 26 × 5	390
Rotation	3 × 26 × 5	390
Translation	3 × 26 × 5	390
	Total	1820

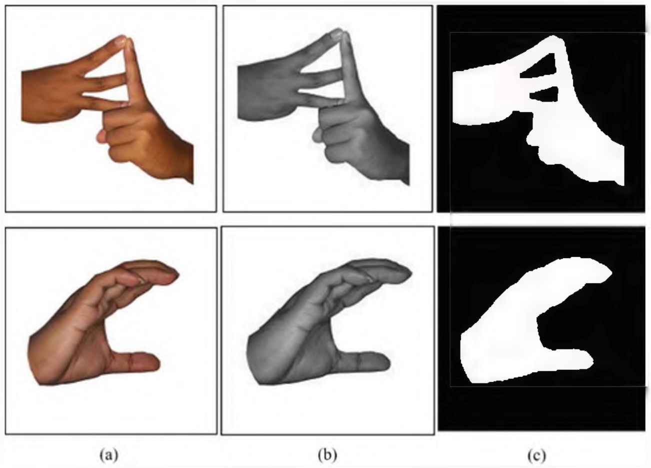

Fig. 4. Image conversion of letter B and C; (a) RGB images, (b) grayslace images, (c) Binary images

Meanwhile, closing operation was also conducted from the original binary images depicted in Fig. 5(a) to obtain clearer binary images which smooth the holes in the letters as depicted in Fig. 5 (b). In extracting the feature of the images, Fourier transform was performed with 5 kinds of coefficient (N) used i.e., 2, 5,10, 25, and 50. The feature values $X[k]$ of letter B and C for each coefficient are revealed in Fig.6 which are obtained from Fourier transform, normalized, and presented in a bar chart for each coefficient. Lastly, Euclidean distance was computed to gain the similarity of the tested data to the training data. As an example, the 5 feature values of letter A are given in Table II extracted using Fourier transform with $N = 5$.

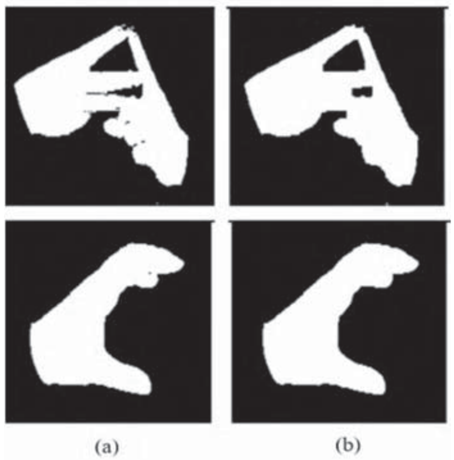

Fig. 5. Closing operation of letter B and C; (a) Original binary images, (b) Closing result.

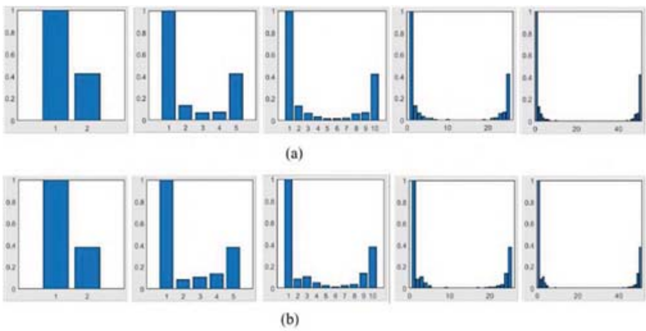

Fig. 6. Feature values of each coefficient; (a) Letter B, (b) Letter C

TABLE II. FEATURE VALUES OF LETTER B

Image	1	2	3	4	5
Training data	1	0.1505	0.0638	0.1139	0.4216
Testing data	1	0.1353	0.0707	0.0743	0.4261

The Euclidean distance is gained by using (4).

$$d(x,y) = \sqrt{(1-1)^2 + \cdots + (0.4216 - 0.4261)^2}$$
$$= \sqrt{0.0018671}$$
$$= 0.043209 \quad (4)$$

The similarity computation was repeatedly performed between the tested B ($B1$) and all the training data so that all the results are obtained as revealed in Table III. The final recognition is determined by looking for the smallest Euclidean distance obtained. A similar process was conducted up to 130 images for each coefficient so that 650 standard images in total can be successfully recognized. Table III shows that from the simulation of Fourier descriptor with coefficient $N = 5$, it was gained that the smallest distance between features of B1 and all the training data was given by the letter B which means that the image tested is recognized as letter B. Based on the experiment, the accuracy of recognition for each kind of coefficients is given in Table IV. It can be seen clearly that among 5 coefficients used in the Fourier descriptor, the higher the coefficient is, the more accurate the recognition result will be.

TABLE III. RECOGNITION OF LETTER B

Tested Letter	Training Data	Euclidean Distance	Recognition
B1	A	0.090469	B
	B	0.043209	
	C	0.089814	
	D	0.078903	
	E	0.087988	
	F	0.105701	
	G	0.108451	
	H	0.098013	
	I	0.096011	
	J	0.087975	
	K	0.132221	
	L	0.189309	
	M	0.100178	
	N	0.08932	
	O	0.149333	
	P	0.163395	
	Q	0.086296	
	R	0.168679	
	S	0.122502	
	T	0.154723	
	U	0.105979	
	V	0.123036	
	W	0.090244	
	X	0.119367	
	Y	0.069371	
	Z	0.134682	

TABLE IV. PERFORMANCE OF FOURIER DESCRIPTOR COEFFICIENT

Coefficient	Accuracy Value
2	15,38%
5	66,92%
10	94,62%
25	96,92%
50	96,92%

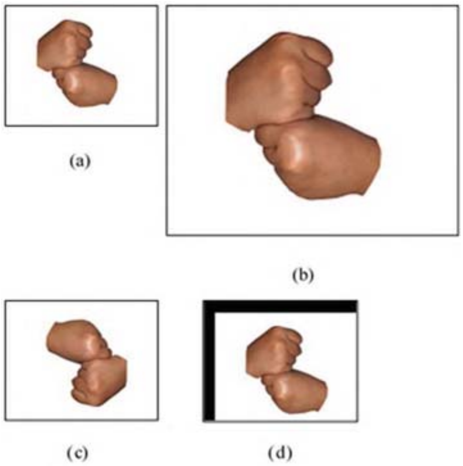

Fig. 7. (a) Standard Image, (b) Scale Image, (c) Rotation Image, (d) Translation Image.

TABLE V. AVERAGE ACCURACY

Test data	Average Accuracy
Standard	74.15%
Scale	72.30%
Rotation	57.43%
Translation	34.36%

In addition, various testing data were also conducted to the translated, scaled, and rotated images of BISINDO letters to measure the performance of Fourier descriptor and Euclidean Distance in recognizing BISINDO letter as shown in Fig.7. From the testing result of 1820 images in Table I, it is obtained the average accuracy as presented in Table V. Table V illustrates that Fourier descriptor and Euclidean distance worked powerfully for standard and scaled images reaching up to 74.15% and 72.30% of recognition accuracy in average consecutively. Meanwhile, the methods only recognized up to 57.43% and 34.36% of average accuracy for rotated and translated images, respectively.

CONCLUSION

In conclusion, Fourier descriptor can be used to extract the features of BISINDO letter images and the higher the coefficient is, the more accurate the recognition result will be. It is evidenced by 96.92% of accuracy obtained at coefficient 25 and 50. In addition, the combination of Fourier descriptor and Euclidean distance adequately recognize standard and scaled BISINDO letter images while the accuracy for translated and rotated ones are not superior yet. As for future work, it is recommended to work more on the recognition of BISINDO letter applying transformation of rotation and translation.

REFERENCES

[1] N. Palfreyman, "Sign language varieties of Indonesia a linguistic and sociolinguistic investigation," *Sign Lang. Linguist.*, vol. 20, no. 1, pp. 135–145, 2017, doi: 10.1075/sll.20.1.06pal.

[2] D. Indra, S. Madenda, and E. P. Wibowo, "Recognition of Bisindo alphabets based on chain code contour and similarity of Euclidean distance," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 7, no. 5, pp. 1644–1652, 2017, doi: 10.18517/ijaseit.7.5.2746.

[3] D. Indra, Y. Salim, Herman, D. Atmajaya, P. L. Lb, and T. Hasanuddin, "Bisindo Alphabets Edge Detection Using Color Tracing of Object Boundary," in *Proceedings - 2nd East Indonesia Conference on Computer and Information Technology: Internet of Things for Industry, EIConCIT 2018*, 2018, pp. 5–9, doi: 10.1109/EIConCIT.-2018.8878554.

[4] T. Handhika, R. I. M. Zen, D. P. Lestari, and I. Sari, "Gesture recognition for Indonesian Sign Language (BISINDO)," in *Journal of Physics: Conference Series*, 2018, vol. 1028, no. 1, p. 012173., doi: 10.1088/1742-6596/1028/1/012173.

[5] G. Zhang, Z. M. Ma, Q. Tong, Y. He, and T. Zhao, "Shape feature extraction using Fourier descriptors with brightness in content-based medical image retrieval," in *Proceedings - 2008 4th International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IIH-MSP 2008*, 2008, pp. 71–74, doi: 10.1109/IIH-MSP.2008.16.

[6] I. Guyon and A. Elisseeff, "Feature Extraction, Foundations and Applications: An introduction to feature extraction," *Stud. Fuzziness Soft Comput.*, vol. 207, pp. 1–25, 2006.

[7] F. Timm and T. Martinetz, "Statistical fourier descriptors for defect image classification," *Proc. - Int. Conf. Pattern Recognit.*, vol. 1, no. 1, pp. 4190–4193, 2010, doi: 10.1109/ICPR.2010.1018.

[8] R. Bazine, W. Huayi, and K. Boukhechba, "Fourier Descriptors Based Approach for Hyperspectral Feature Extraction," in *In 2018 International Workshop on Big Geospatial Data and Data Science (BGDDS)*, 2018, pp. 1–4.

[9] N. M. Wafi, S. N. Yaakob, N. S. Salim, M. Jusoh, A. R. A. Nazren, and M. B. Hisham, "Image analysis using new Descriptors Average Feature Optimization based on Fourier Descriptors technique," in *Proceeding - 2016 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications, ICRAMET 2016*, 2017, pp. 135–138, doi: 10.1109/ICRAMET.2016.7849599.

[10] E. Sokic and S. Konjicija, "Novel fourier descriptor based on complex coordinates shape signature," 2014, doi: 10.1109/CBMI.2014.-6849843.

[11] B. G. Batchelor and F. M. Waltz, "Morphological image processing," *Mach. Vis. Handb.*, no. November, pp. 802–870, 2012, doi: 10.1007/978-1-84996-169-1_19.

[12] K. A. M. Said, A. B. Jambek, and N. Sulaiman, "A study of image processing using morphological opening and closing processes," *Int. J. Control Theory Appl.*, vol. 9, no. 31, pp. 15–21, 2016.

[13] P. Sharma, M. Diwakar, and N. Lal, "Edge Detection using Moore Neighborhood," *Int. J. Comput. Appl.*, vol. 61, no. 3, pp. 26–30, 2013, doi: 10.5120/9910-4506.

[14] S. Meenakshi, A. Batra, and M. Sharma, "Analysis of Distance Measures in Content Based Image Retrieval," *Glob. J. Comput. Sci. Technol.* , vol. 14, no. 2, pp. 11–16, 2014.

[15] L. Wang, Y. Zhang, and J. Feng, "On the Euclidean distance of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1334–1339, 2005, doi: 10.1109/TPAMI.2005.165.