

Edge Detection and Grey Level Co-Occurrence Matrix (GLCM) Algorithms for Fingerprint Identification

Edy Winarno
Faculty of Information Technology
Universitas Stikubank
Semarang, Indonesia
edywin@edu.unisbank.ac.id

Wiwien Hadikurniawati
Faculty of Information Technology
Universitas Stikubank
Semarang, Indonesia
wiwien@edu.unisbank.ac.id

Setyawan Wibisono
Faculty of Information Technology
Universitas Stikubank
Semarang, Indonesia
setyawan@edu.unisbank.ac.id

Anindita Septiari
Faculty of Engineering
Mulawarman University
Samarinda, Indonesia
anindita@unmul.ac.id

Abstract—The fingerprint identification system is a recognition process by measuring the characteristics on human fingers and then comparing them with those in the database. The purpose of this study is to create a system for fingerprints recognition using the edge detection method and Grey Level Co-occurrence Matrix (GLCM). The method used in this fingerprint recognition research is texture analysis. Pre-processing was performed with edge detection and feature extraction with GLCM method. First, the fingerprint is captured using the fingerprint scanner. Then the fingerprint image was extracted using the Grey Level Co-occurrence Matrix (GLCM) feature. The features obtained are energy, contrast, homogeneity and correlation. The final result of the fingerprint identification system is the success of displaying the image with the identity data of the fingerprint owner. Using two methods, fingerprint identification accuracy of 83% was achieved.

Keywords—*fingerprint, edge detection, GLCM, feature analysis*

I. INTRODUCTION

By using fully automated or semi-automated systems to identify individuals, biometrics provides a simple and secure solution for identity management. Fingerprint is one of the most widely studied and used forms of recognition among all biometric recognition schemes. Fingerprint identification systems are among the most widely used biometric systems. Fingerprints will be the same all our lives until we die. The size of our finger will be bigger according to growing up but the print on the finger will never change. Even if twins, the prints will not be identical on both kids. Identification of fingerprints is the technique used to determine that both sets of fingerprint images are of the same finger [1] [2]. Biometrics is a mechanism of identification, using the specific behavioral and physiological individualities of the human body as identifiers. Identification of fingerprints is a complex problem regarding pattern recognition. Designing precision algorithms for extracting and matching features is difficult [3] [4].

Fingerprint technology plays an important role in security and e-commerce applications because of its ability to identify and provide access to users via fingers. Fingerprints are a characteristic of skin. The ridges are shaped to hold securely and not slip when we grip an object. They did arrange regular patterns by themselves and have a special arrangement and combination of patterns of ridge characteristics. With some significant features, such as reliability, invariability, nondisjunction and consistency, as an effective way to verify the identity of members, fingerprint identification has wide applications in the field of information security. That's very important fingerprint feature. For this versatility, usability and its cheap, fingerprint is very common and can be used to distinguish the prints on a crime scene not only to eliminate or delete a suspect from consideration but also to provide identification, customs access and public safety organizations of our daily social lives [5] [6].

Of all the techniques for texture analysis, the most commonly used today is probably the one based on the method of the gray level co-occurrence matrix (GLCM). [7] first suggested the GLCM approach as a way of classifying images using statistical measurements of second order. The definition of a co-occurrence gray-level matrix (GLCM) was considered in the threshold selection literature. GLCM encoded spatial information has been used successfully in various applications including extraction of features, texture classification, texture description. The use of Gray Level Co-Occurrence Matrix (GLCM) is mostly used to capture remote sensing imagery with prototypes [2] [8].

There are several techniques for matching fingerprints. Many researchers in the literature suggested the different solution to fingerprint matching system problems. [9] proposed GLCM based fingerprint recognition algorithm. Using the GLCM seven significant statistical descriptors were measured with specific pixel separation distances in four orientations between pixels. Instead, these descriptors are used to create a feature set that defines the selected enhanced fingerprint image region surrounding the core

point in a unique way. The experimental results show that the Proposed Algorithm achieves good accuracy of recognition for various combinations of GLCM pixel separation distances, number of nearest neighbors and distance metrics used in KNN classifier and can be used to recognize fingerprints. Research on fingerprints has also been studied by [8], they have developed the multimodal biometric system using ear and fingerprint. Ear characteristics are identified, and horizontal matrix and vertical matrix are generated.

The Edge detection technique is of great use in the area of, computer vision, pattern recognition and image analysis. This is also significant in the field of biomedicine. There are common edge detection techniques. Other researchers have examined fingerprint recognition with edge detection techniques. [10] in their research applied edge detection to identify fingerprints. This paper presents some edge detection operators and compares their efficiency and characteristics. The results this research concluded: a. The Robert Operator has a good performance on the steep and low-noise images, but the edges identified by the Robert Operator are rough, so the edge position is not very precise, b. The Sobel Operator works well on the gray gradient and high noise images, but the position of the edges is not very precise, the edges of the image have more than one pixel, c. The Prewitt Operator works well with gray gradient and high noise on the images, but the edges are wide and have lots of intermittent points.

In this paper, we propose a new approach to the identification of fingerprints using edge detection and GLCM. We use Robinson operator for edge detection because Robinson edge operators work fairly good in comparison to other operators [11].

II. RESEARCH METHODS

A. Edge Detection

Edge detection is a two-dimensional attempt to find discernible changes to the contrast. This approach is best suited for this particular project and so most of the algorithms in this project fall within this category. A very important part of fingerprint recognition technology is the edge detection of the fingerprint images. A group of connected pixels forming a boundary between two disjunction regions called an edge i.e. it defines the sharp changes in the intensity function. Edges can be represented by the scene intensity function in the form of discontinuities. Blade is the dividing line between objects and context, so it sets the object apart from the background. The edge detection process is done through a differential approach. The local derivative operator is the basis for this derivative method. The basic idea is to use an edge enhancement algorithm to improve the local edge, then define the edge strength of the pixels, and by setting the threshold, extract the edge point collection. Because an image's edge and contour typically have arbitrary directions, some operators that have the same edge and contour detecting capacity for arbitrary directions are required [11].

Most segmentation algorithms use a mask on the pixels of the image. Every pixel has its gray level value multiplied by a mask value and its neighboring pixels, as shown in Fig. 1.

$(i-1, j-1)$	$(i-1, j)$	$(i-1, j+1)$
$(i, j-1)$	(i, j)	$(i, j+1)$
$(i+1, j-1)$	$(i+1, j)$	$(i+1, j+1)$

Fig. 1. Neighborhood of pixel (i,j)

The sum of these values is the mask response of that point. Based on the board's uniform grid architecture it was determined that several conventional segmentation algorithms would work well to find the red mine field boundaries. Many such algorithms have been produced, often varying in the masks used to decide a pixel's chance to be an edge pixel. The Robinson Operator operates similarly to the Sobel operator but uses a set of eight masks. The Robinson Edge Detection method also similar to Kirsch masks but is easier to implement, as they depend only on 0, 1 and 2 coefficients. The masks are symmetrical about their axis of direction, the axis with the zeros. It is only necessary to determine the result on four masks and it is possible to obtain the result from other four by negating the result from the first four. The masks belong as follows in Fig. 2.

r_0	r_1	r_2	r_3
$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$	$\begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -2 \end{bmatrix}$
r_4	r_5	r_6	r_7
$\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$	$\begin{bmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}$

Fig. 2. Mask of Robinson Operator

B. Grey Level Co-Occurrence Matrix (GLCM)

Feature extraction is the first step in classifying and interpreting images. This process is related to the quantization of image characteristics in a group of appropriate feature values. Texture analysis is a process for classifying and interpreting images. An image classification process based on texture analysis requires a feature extraction stage. One of the feature extraction methods is the statistical method. The second-order statistical method considers the relationship of two pixels (neighboring pixels) in the image. Second-order texture analysis requires a co-occurrence matrix for the gray image, namely GLCM. Second-order texture analysis is better at representing image textures in measurable parameters, such as contrast, correlation, homogeneity, entropy, and energy. GLCM is a matrix that represents the neighborhood relationship between pixels in the image in various orientation directions (Θ) and spatial distance (d). This approach forms a co-occurrence matrix from the image data, the next step is to determine the characteristics as a function of the matrix. Distance is expressed in pixels and orientation is expressed in degrees. This algorithm is also one of the well-known and effective algorithms for texture analysis. In general, this algorithm has four directions that are used to create a GLCM matrix, namely the angle direction (Θ) = $0^\circ, 45^\circ, 90^\circ, 135^\circ$ [9].

The co-occurrence matrix is a square matrix with the number of elements as much as the square of the number of pixel intensity levels in the image. Each point (p,q) in the angle-oriented (θ) co-occurrence matrix contains the probability of the occurrence of a pixel with a value of p neighboring a pixel with a value of q at a distance (d) and orientation directions ($\Theta = 180-\theta$). After forming the co-occurrence matrix, the second-order statistical value that represents the image can be determined. In this study, four characteristics of second-order statistical were calculated: Contrast, Correlation, Energy, Homogeneity.

a) *Contras*

Indicates the size of the spread (moment of inertia) of the image matrix elements. If it is far from the main diagonal, the contrast value will be greater. Visually, the contrast value is a measure of the variation in the degree of gray in an image area.

$$\sum_{i,j} |i - j|^2 \cdot p(i, j) \quad (1)$$

b) *Correlation*

Shows the size of the linear dependence of the gray level of the image. It can provide clues to the presence of linear structures in the image.

$$\sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i\sigma_j} \quad (2)$$

c) *Energy*

Shows the size of the homogeneity of the image.

$$\sum_{i,j} P(i, j)^2 \quad (3)$$

d) *Homogeneity*

Shows the homogeneity of the image with the same level of gray.

$$\sum_{i,j} \frac{p(i, j)}{1 + |i-j|} \quad (4)$$

C. *Euclidean Distance*

Through the feature extraction process, certain parameter values are obtained, then proceed with the calculation of the shortest distance (Euclidean distance) the value of the image feature vector. Euclidean distance values that are close to zero will refer to certain images. The value of an image feature vector that has the same feature vector value as a particular image feature vector will have a Euclidean distance value that is close to zero [12]. If the value of the input image feature vector $A_i = (A_1, A_2, \dots, A_n)$, the value of the j^{th} image feature vector is $B_j = (B_1, B_2, \dots, B_n)$, the Euclidean distance between the input image feature vector value and the feature vector value the j^{th} image is represented by:

$$D(A, B) = \sqrt{\sum_{i=0}^{n-1} \frac{|A_i - B_i|^2}{A_i}} \quad (5)$$

III. CASE STUDY

In this fingerprint recognition system, there are processes that are taken, starting from the selected data, until finally the data can be identified. All of these processes are described in the proposed method block diagram. The proposed method that used in this research as shown in Fig. 3.

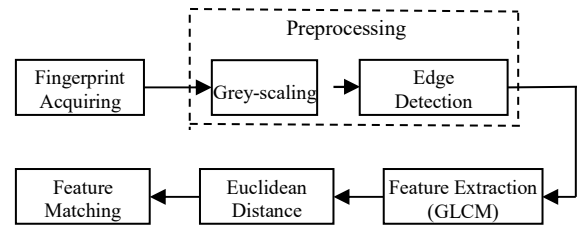


Fig. 3. The Proposed Method

The proposed method to identify fingerprints, the first step is acquiring fingerprints. The identification system starts from the training process, the user takes a fingerprint image to be identified. The first process is to change the image to a gray level. The next process is edge detection using the Robinson method. The next process is feature extraction with the GLCM algorithm. The next process is to save the image value as a database. In the testing process, the user inputs a RGB image that is converted into a gray level and the next process is edge detection using the Robinson method. The next process is feature extraction with the GLCM algorithm which will compare the image values in the database.

The image used in the training data and test data is a RGB image of the *.jpg file type. The database used is the fingerprint image of the left thumb. The fingerprint images used were taken from 12 person (fingerprints), each of which the thumb image was taken at least 3 times. That is intended to make variations in patterns and accuracy when taking fingerprints from the person concerned. The results of the training data collection can be seen in Figure 4.

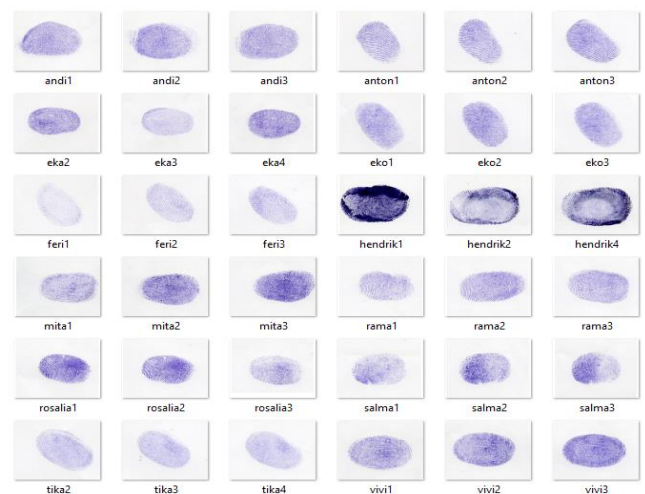


Fig. 4. Image Training Data

The RGB image in Figure 4 is then performed edge detection. The matrix calculation is performed using the convolution technique. The following is a matrix form of fingerprints. We use sample size 3x3 pixels.

239	239	244
248	248	253
248	246	248

The color image convolution technique for edge detection as follows:

$$\begin{aligned}
 \text{East} &= -1.239+0.239+1.244+2.248+0.248+ \\
 & \quad 2.253+-1.248+0.246+1.248 \\
 &= 15 \\
 \text{Northeast} &= 0.239+1.239+2.244+-1.248+0.248+ \\
 & \quad 1.253+-2.248+-1.246+0.248 \\
 &= -10 \\
 \text{North} &= 1.239+2.239+1.244+0.248+0.248+0.253+- \\
 & \quad 1.248+-2.246+-1.248 \\
 &= -27 \\
 \text{Northwest} &= 2.239+1.239+0.244+1.248+0.248+-1.253+ \\
 & \quad 0.248+-1.246+(-2.248) \\
 &= -30 \\
 \text{West} &= 1.239+0.239+-1.244+2.248+0.248+ \\
 & \quad -2.253+1.248+0.246+-1.248 \\
 &= -15 \\
 \text{Southwest} &= 0.239+-1.239+-2.244+1.248+0.248+ \\
 & \quad -1.253+2.248+1.246+0.248 \\
 &= 10 \\
 \text{South} &= -1.239+-2.239+-1.244+0.248+0.248+ \\
 & \quad 0.253+1.248+2.246+1.248 \\
 &= 31 \\
 \text{Southeast} &= -2.239+-1.239+0.244+-1.248+0.248+ \\
 & \quad 1.253+0.248+1.246+2.248 \\
 &= 33
 \end{aligned}$$

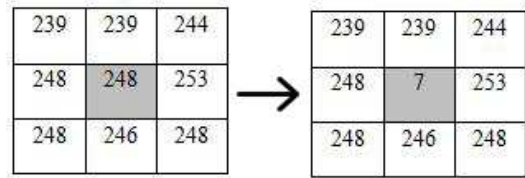


Fig. 5. Matrix Convolution Result

The result of pixel 248 is $15+-10+-27+-30+-15+10+31+33 = 7$. The combination of the matrix (3X3) resulting from the convolution is as shown in the Fig. 5.

This step is performed (convolution technique) on all pixels in the image. The fingerprint image is then feature extracted using the GLCM method. The features obtained from this feature extraction are the second angular moment, namely energy, contrast, homogeneity and correlation. Feature extraction was determined using equations (1), (2), (3) and (4). The results of the feature extraction calculation can be seen in Table I.

The values are obtained during training and the values from the training are stored automatically in a temporary storage area according to the trained image.

TABLE I. GLCM ALGORITHM TRAINING IMAGE VALUE

Data	Name	Contrast	Correlation	Energy	Homogeneity
1	Andi1	0.992071	0.529597	0.468439	0.860569
	Andi2	0.949642	0.508955	0.450153	0.858625
	Andi3	0.987702	0.513271	0.441569	0.854223
2	Anton1	0.627733	0.469008	0.605335	0.913235
	Anton2	0.664692	0.468128	0.542698	0.898672
	Anton3	0.662318	0.475664	0.564611	0.903614
3	Eka1	0.718654	0.514287	0.616318	0.907965
	Eka2	0.771452	0.535326	0.585879	0.899994
	Eka3	0.484816	0.357642	0.723399	0.9402
4	Eko1	0.511287	0.361612	0.683208	0.93084
	Eko2	0.665031	0.445448	0.561634	0.900811
	Eko3	0.598597	0.411345	0.598508	0.911698
5	Feri1	0.406067	0.333059	0.888312	0.973893
	Feri2	0.516367	0.385511	0.687107	0.932798
	Feri3	0.640613	0.463085	0.577404	0.90601
6	Mita1	1.041546	0.590252	0.475676	0.864323
	Mita2	0.789507	0.527446	0.512651	0.881477
	Mita3	1.135578	0.588443	0.451785	0.854303
7	Rama1	0.635852	0.461087	0.583961	0.907014
	Rama2	0.743249	0.50802	0.524297	0.889594
	Rama3	0.764349	0.506795	0.528183	0.889655
8	Rosa1	0.667727	0.486743	0.61945	0.91067
	Rosa2	0.885734	0.524829	0.4896	0.870908
	Rosa4	0.857233	0.520726	0.497093	0.874421
9	Hendri1	0.85864	0.567406	0.583584	0.895498
	Hendri2	0.836477	0.56125	0.586447	0.896944
	Hendri3	0.747173	0.537059	0.613858	0.907353
10	Salma1	0.663204	0.479579	0.588108	0.90741
	Salma2	0.5921	0.446693	0.632607	0.919607
	Salma3	0.704942	0.498282	0.578363	0.902554
11	Tika1	0.466611	0.341352	0.764356	0.947106
	Tika2	0.511186	0.356304	0.685812	0.929982
	Tika3	0.454351	0.340364	0.787153	0.95244
12	Vivi1	0.689946	0.453622	0.51392	0.889246
	Vivi2	0.721147	0.469642	0.503511	0.886211
	Vivi3	0.647208	0.423522	0.53386	0.89502

TABLE II. VALUE OF TEST DATA FROM GLCM ALGORITHM

Data	Name	Contrast	Correlation	Energy	Homogeneity	Prediction Result	Validation
1	Andi	0.992071	0.529597	0.468439	0.860569	Andi	valid
2	Anton	0.665969	0.454386	0.55208	0.902076	Anton	valid
3	Eka	0.807711	0.54098	0.566131	0.894147	Eka	valid
4	Eko	0.598196	0.405566	0.600041	0.911483	Eko	valid
5	Feri	0.632937	0.420786	0.592945	0.909874	Eko	invalid
6	Hendri	1.010196	0.540405	0.407437	0.847985	Hendri	valid
7	Mita	0.793925	0.517027	0.508946	0.883444	Mita	valid
8	Rama	0.987083	0.496245	0.413398	0.847721	Rama	valid
9	Rosa	0.700671	0.489852	0.623841	0.910779	Rosa	valid
10	Salma	0.654	0.441023	0.584801	0.90762	Eko	invalid
11	Tika	0.4774	0.340861	0.740724	0.941847	Tika	valid
12	Vivi	0.792983	0.459062	0.44974	0.87061	Vivi	valid

The image used in this test is a color image of the *.jpg file type. The images used were 12 persons, where 1 person took 4 samples of fingerprint images, 3 images for training data and 1 image for test data. The test value is obtained from the tested image sample. The values from the results of this test will be compared with the values of the previous training results. The results of the test data values from the GLCM algorithm can be seen in Table II.

The final result of the fingerprint identification system is the success of displaying the image with the identity data of the fingerprint owner. From the results of matching features obtained valid and invalid results. The percentage of validation from the matching feature results is

$$\frac{10}{12} \times 100\% = 83.33\%.$$

IV. CONCLUSION

The characteristics that are most efficiently used for authentication are biometric features, namely human fingerprints. A technique for fingerprint identification with the use of edge detection and feature extraction as the distinguishing characteristic analyzed in this paper. Pre-processing was performed with edge detection, fingerprint thinning and feature extraction with GLCM method. The performance of edge detection and GLCM are analyzed. From the test results, it is also found that the use of a distance of 1 pixel provides the highest recognition rate for fingerprint recognition. This is because the human fingerprint pattern shows a small co-occurrence. In addition, it can also be concluded that the factors that influence the level of image recognition are the image acquisition process, database formation and the selection of parameters for the GLCM matrix. The final result of the fingerprint identification system is the success of displaying the image with the identity data of the fingerprint owner. Using two methods, improved fingerprint identification accuracy of 83% is achieved.

REFERENCES

[1] D. Valdes-Ramirez *et al.*, "A Review of Fingerprint Feature Representations and Their Applications for Latent Fingerprint Identification: Trends and Evaluation," *IEEE Access*, vol. 7, pp. 48484–48499, 2019, doi: 10.1109/ACCESS.2019.2909497.

[2] S. P. Sandip and P. H. Zope, "Selective review of fingerprint enhancement, classification and matching techniques," *2015*

IEEE Bombay Sect. Symp. Front. Technol. Fuelling Prosper. Planet People, IBSS 2015, pp. 1–6, 2016, doi: 10.1109/IBSS.2015.7456656.

[3] S. Hemalatha, "A systematic review on Fingerprint based Biometric Authentication System," *Int. Conf. Emerg. Trends Inf. Technol. Eng. ic-ETITE 2020*, pp. 1–4, 2020, doi: 10.1109/ic-ETITE47903.2020.342.

[4] C. A. De Luna-Ortega, J. A. Ramírez-Márquez, M. Mora-González, J. C. Martínez-Romo, and C. A. López-Luévano, "Fingerprint verification using the center of mass and learning vector quantization," *Proc. - 2013 12th Mex. Int. Conf. Artif. Intell. MICAI 2013*, pp. 123–127, 2013, doi: 10.1109/MICAI.2013.21.

[5] S. Selvarani, S. Jebapriya, and R. Smeeta Mary, "Automatic identification and detection of altered fingerprints," *Proc. - 2014 Int. Conf. Intell. Comput. Appl. ICICA 2014*, pp. 239–243, 2014, doi: 10.1109/ICICA.2014.58.

[6] R. Pavithra and K. V. Suresh, "Fingerprint image identification for crime detection," *Proc. 2019 IEEE Int. Conf. Commun. Signal Process. ICCSP 2019*, pp. 797–800, 2019, doi: 10.1109/ICCSP.2019.8698014.

[7] S. Preetha and S. V. Sheela, "Selection and extraction of optimized feature set from fingerprint biometrics-a review," *Proc. 2nd Int. Conf. Green Comput. Internet Things, ICGCIoT 2018*, pp. 500–503, 2018, doi: 10.1109/ICGCIoT.2018.8753104.

[8] L. D. Bakti *et al.*, "Data extraction of the gray level Co-occurrence matrix (GLCM) Feature on the fingerprints of parents and children in Lombok Island, Indonesia," *Data Br.*, vol. 36, p. 107067, 2021, doi: 10.1016/j.dib.2021.107067.

[9] P. M. Arabi, G. Joshi, and N. Vamsha Deepa, "Performance evaluation of GLCM and pixel intensity matrix for skin texture analysis," *Perspect. Sci.*, vol. 8, pp. 203–206, 2016, doi: 10.1016/j.pisc.2016.03.018.

[10] S. Zhou and X. Lu, "Fingerprint identification and its applications in information security fields," *Proc. - 2010 Int. Conf. Inf. Sci. Manag. Eng. ISME 2010*, vol. 1, pp. 97–99, 2010, doi: 10.1109/ISME.2010.186.

[11] G. Ravivarma, K. Gavaskar, D. Malathi, K. G. Asha, B. Ashok, and S. Aarthi, "Implementation of Sobel operator based image edge detection on FPGA," *Mater. Today Proc.*, vol. 45, pp. 2401–2407, 2021, doi: 10.1016/j.matpr.2020.10.825.

[12] B. Stojanovic, O. Marques, A. Neskovic, and S. Puzovic, "Fingerprint ROI segmentation based on deep learning," *24th Telecommun. Forum, TELFOR 2016*, pp. 16–19, 2017, doi: 10.1109/TELFOR.2016.7818799.