

## Identification of fingerprint image with Minkowski distance algorithm approach

Wisnu Catur Rachmad Tulloh, Dian Eka Wijayanti\*

Universitas Ahmad Dahlan, Jl. Jend. Ahmad Yani, Tamanan, Banguntapan, Bantul, DIY 55711, Indonesia

\*Corresponding E-mail: [dian@math.uad.ac.id](mailto:dian@math.uad.ac.id)

### ARTICLE INFO

### ABSTRACT

#### Article History

Received 14 December 2023

Revised 29 December 2023

Accepted 29 December 2023

#### Keywords

Data analysis

Fingerprint identification

Minkowski distance

System security

#### How to cite this article:

Tulloh, W. C. R., & Wijayanti, D. E.

(2023). Identification of

fingerprint image with

Minkowski distance algorithm

approach. *Bulletin of Applied*

*Mathematics and Mathematics*

*Education*, 3(2), 69-78.

In the digital era, fingerprint identification plays a critical role in information technology administration. Various studies have been conducted to improve the fingerprint identification process, but there are still cases of identification failures that are fatal. This research discusses fingerprint identification with the Minkowski distance method. The data of fingerprint are taken from Mathematics students and the Kaggle site. Data analysis includes the steps of image retrieval, dimensioning, conversion to gray scale, pattern matching, and accuracy measurement. Results show an improvement in data accuracy with a structured approach to data capture and preprocessing. Results from primary data obtained an accuracy of 56.67% while from secondary data obtained an accuracy of 93%.

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



## Introduction

The rapid growth in technology and information systems in the digital age is undergoing significant changes. Easy access to information through the internet has become a primary need, but the increasing use of digital technology also brings the risk of increased threats to security. Therefore, it is important to prioritize digital security protection that includes aspects of personal data protection, network security, and information system security (Gani, n.d.; Mayamin & Usuluddin, 2023; Nuriadin et al., 2021; Nuriadin, 2021).

Fingerprint sensors, as part of management information systems, play a role in the management of personal information. In the digital age, technological advancements impact on the vulnerability of information security. In the digital financial sector, this poses a threat of loss of assets or leakage of customer data, requiring simple and valid identification. Fingerprint sensors have several advantages, such as simplifying the identification process and data storage, while increasing the security of personal information (Emelda, n.d.; Rizaldi et al., n.d.; Nurliza & Widodo, 2018).

Fingerprint identification can be used to improve security and prevent data leakage. A study conducted by Satria et al. (2017) showed that the fingerprint system can prevent personal data

leakage up to 99.9%. In addition, another study conducted by Arifandi (2023) showed that a door security system using fingerprints based on Arduino Uno ATmega328 and SMS Gateway can identify residents of the house and provide solutions to homes left by their owners. In the study, the test results show access to open the door using fingerprints from outside can run well and homeowners can easily get information when someone breaks into the door by force (Arifandi, 2019; Aziz, n.d.; Gusri et al., n.d.; Hartono et al., 2023; Hermawan et al., 2023).

Various methods can be used in fingerprint identification, one of which is the Minkowski distance. The Minkowski distance method acts as an important metric for vector spaces, serving as a norm in the space, encompassing the generalized forms of Euclidean and Manhattan distances. The use of Minkowski distance in the fingerprinting approach contributes to the formation of a unique signature, ensuring a distinctive and non-copyable identification. This methodology, which is an integral part of vector space analysis, highlights the precision and complexity essential for robust applications, making it a key element in a collection of advanced techniques in the domain of data science and pattern recognition (Nurliza & Widodo, 2018; Pradana, 2017; Safwandi & Muthmainnah, n.d.).

Research that adopts the Minkowski distance method to perform distance analysis between data points in vector space can result in a deeper understanding of the structure and distribution of the data, open up the potential for discovering relevant patterns, and improve the accuracy of the analysis in the context in question. The Minkowski distance method is a very suitable method for translating distance analysis between data points, because it provides flexibility and generalization in measuring distance. The steps of this research involve calculating distance using the Minkowski distance formula to describe the relationships and patterns that exist in the data (Khairunnisa et al., n.d.; Nishom, 2019; Safwandi & Muthmainnah, n.d.; Thant & Aye, 2020).

## **Method**

This research adopting the Minkowski distance method to perform distance analysis between data points in vector space can lead to a deeper understanding of the structure and distribution of the data, open up the potential for discovering relevant patterns, and improve the accuracy of the analysis in the context in question. The Minkowski distance method is a very suitable method for translating distance analysis between data points, as it provides flexibility and generality in measuring distance. The steps of this research involved calculating distances using the Minkowski distance formula to illustrate the relationships and patterns present in the data.

## **Data collection methods**

### *Experiment research*

In the data collection process, primary data was used by randomly collecting 30 fingerprints from students of the mathematics study program at Ahmad Dahlan University. In addition, an additional 6 reference data were also obtained to be included in the calculation of the Minkowski distance method. This method requires reference data as a reference in measuring the distance between primary data. Thus, the fingerprint data collected from these students will be analyzed using the Minkowski distance method by considering the 6 references data that have been previously determined. This aims to enable proper comparison between the primary data and the reference data, thus facilitating an accurate analysis process in the context of the Minkowski distance method.

### *Literature study*

The data used in this research is secondary data obtained from the Kaggle platform. The data

retrieval process was carried out with respect to the distance in which the photos were taken and involved cropping the data. This approach allows for a more structured selection and customization of data from such secondary sources to support proper analysis in a pre-defined method.

### **Techniques for analyzing data**

Data analysis is an essential component of the research framework. The reliability of resolving the issues in focus and the conclusions that result from a study have a substantial dependence on the integrity and completeness of the data analysis conducted. The analysis methods and procedures applied in this study included the following steps.

#### *Fingerprint image capture*

At this stage, 36 images are taken, consisting of 6 fingerprint reference templates (guide1.jpg - guide6.jpg) and 30 input templates to be tested (finger.1.jpg - finger.30.jpg). Subsequently, these images are stored for further processing.

#### *The process of customizing the image dimensions to 4 × 4 pixels*

This step is designed to normalize the input image to be analyzed.

#### *Image conversion from RGB color mode to grayscale*

This is necessary so that the image can be processed using the chosen method.

#### *Pattern matching*

In this phase, processing is performed between the reference sample template and the input template that has been taken. Each template, be it the reference image or the cropped input image, is analyzed comparatively one by one. To assess the similarity between the two templates, the Minkowski distance method is used. This method calculates the distance value between the two templates being compared. An input template will be classified as similar to the reference template if the distance value between the two is equal to, or less than 30.

#### *Accuracy measurement*

In this step, the accuracy of the Minkowski distance method will be evaluated to determine whether or not this method is suitable in the template matching process (See Figure 1).

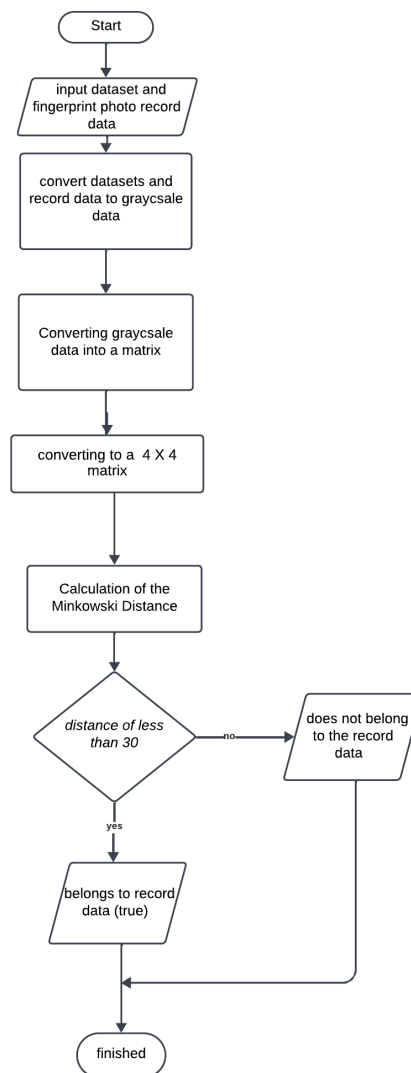
### **Minkowski distance method and accuracy technique**

The Minkowski Distance method has a formula that integrates the characteristics of Euclidean Distance and Manhattan Distance with  $p$  as a parameter. When the value of  $p = 1$ , it has the same formula as Manhattan Distance, while if  $p = 2$ , the formula tends to be Euclidean Distance. The general formula of Minkowski Distance is,

$$d(x, y) = \left( \sum_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}} \quad (1)$$

The calculation process to evaluate the accuracy of a method or procedure that has been implemented (See Figure 1). This accuracy analysis aims to measure the extent to which the results obtained are in accordance with the expected standards or meet predetermined criteria. This stage makes it possible to identify the advantages and disadvantages of a method or technique used in the context of data processing or analysis being carried out. The general formula for accuracy is,

$$Accuracy(\%) = \frac{\sum \text{many successes}}{\sum \text{total number of data}} \times 100\%$$

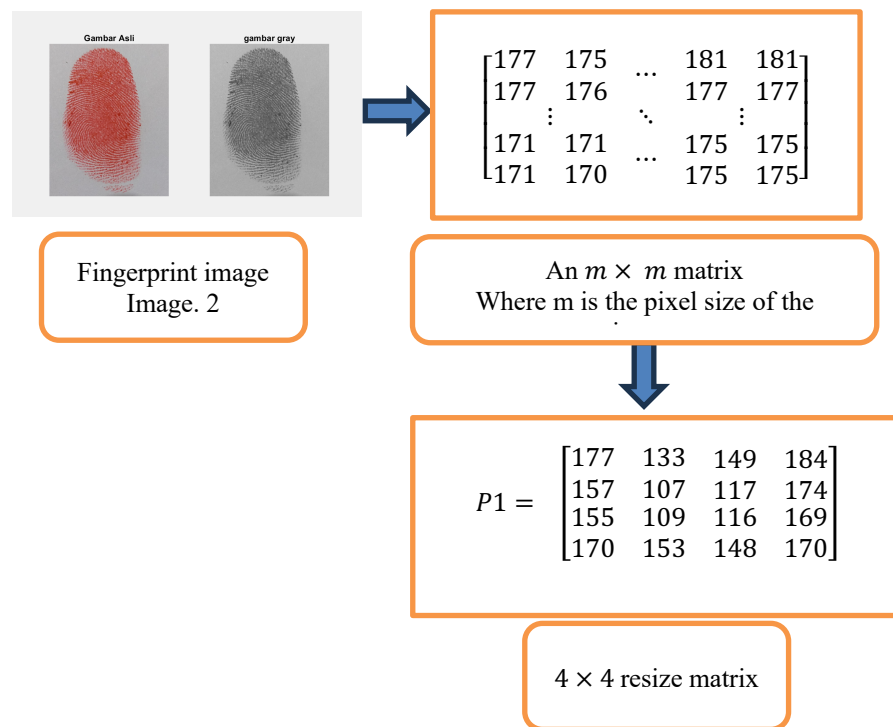


**Figure 1.** Minkowski distance method flowchart

## Results and discussion

The data that has been taken is then converted into a matrix in the RGB color scheme, then converted into a grayscale matrix. Furthermore, from the initial matrix that has a size of  $m \times m$ , dimensional adjustments are made to a  $4 \times 4$  matrix (See Figure 2). This step aims to minimize the dimensions of the data so that it can be processed more efficiently and in accordance with the predetermined analysis needs.

Matrix P1 refers to the matrix derived from guideline 1. Using the established definition of 1, the Minkowski distance will be calculated for the matrix. This process is part of the evaluation of the similarity or difference between the reference matrix and other matrices within the framework of the predefined method.



**Figure 2.** Processing fingerprint data into  $4 \times 4$  matrix

Manual calculation as follows. Comparison of respondent 1 with guideline 1 results:

$$R1 = \begin{bmatrix} 177 & 144 & 132 & 172 \\ 148 & 91 & 92 & 144 \\ 127 & 83 & 91 & 140 \\ 146 & 114 & 129 & 161 \end{bmatrix}$$

$$P1 = \begin{bmatrix} 177 & 133 & 149 & 184 \\ 157 & 107 & 117 & 174 \\ 155 & 109 & 116 & 169 \\ 170 & 153 & 148 & 170 \end{bmatrix}$$

$$p = 3$$

$$d(x, y) = (|177 - 177|^3 + |133 - 144|^3 + |149 - 132|^3 + \dots + |170 - 161|^3)^{\frac{1}{3}}$$

$$d(x, y) = 59.9717459564981$$

With a similar approach, calculations were made using Guideline 1 to Guideline 6 against the responses from Respondent 1 to Respondent 30. This step was carried out to analyze and evaluate the suitability and similarity between each guideline and the respective responses generated by the respondents involved in this study and the following Table 1 was obtained.

**Table 1.** Calculation on data from respondents based on Guideline 1 to 6

Number	Name	Guidelines 1	Guidelines 2	Guidelines 3	Guidelines 4	Guidelines 5	Guidelines 6	Classification Results	Success
1	Respondent1	59,9717	68,1699	37,9984	82,7347	50,2466	61,2808	Does not belong to guideline data	works
2	Respondent2	80,7384	89,8481	78,3769	90,3676	69,093	63,3747	Does not belong to guideline data	works
3	Respondent3	65,6447	71,2493	91,7409	46,9907	63,0258	42,7467	Does not belong to guideline data	works
4	Respondent4	71,4142	69,0025	76,1025	64,2724	60,185	54,4752	Does not belong to guideline data	works
5	Respondent5	96,3032	92,7073	128,6952	65,8542	104,6089	93,5718	Does not belong to guideline data	works
6	Respondent6	37,9037	32,4189	58,3379	36,4415	53,6541	50,5340	Does not belong to guideline data	works
7	Respondent7	50,2726	38,0567	56,3271	44,8584	46,5402	46,7214	Does not belong to guideline data	works
8	Respondent8	34,2930	46,8687	56,3271	55,9647	35,2389	35,4968	Does not belong to guideline data	works
9	Respondent9	43,4746	56,2209	34,6287	73,2616	45,2215	54,5415	Does not belong to guideline data	works
10	Respondent10	31,7185	33,2918	59,3669	23,3059	34,8326	32,6972	Belonging to the guideline data	Failed
11	Respondent11	23,4329	29,3833	30,1338	49,452	36,4162	45,4339	Belonging to the guideline data	Failed
12	Respondent12	35,6585	40,1306	46,5394	38,8048	23,3915	26,167	Belonging to the guideline data	Failed
13	Respondent13	101,2387	109,2219	98,0335	113,9386	23,3915	26,167	Belonging to the guideline data	Failed
14	Respondent14	33,0653	42,1691	30,5394	57,085	31,9056	40,4521	Does not belong to guideline data	works
15	Respondent15	42,2666	41,9367	68,8864	14,8639	40,2352	32,2301	Belonging to the guideline data	Failed
16	Respondent16	45,1987	51,9454	47,6481	55,3524	29,1376	24,744	Does not belong to guideline data	Failed
17	Respondent17	43,6056	45,9337	36,3201	51,611	22,3551	29,9599	Belonging to the guideline data	Failed
18	Respondent18	39,5631	48,1102	56,5746	44,2941	37,4866	32,3082	Does not belong to guideline data	works
19	Respondent19	29,9803	48,1102	56,5746	44,2941	37,4866	32,3082	Belonging to the guideline data	Failed
20	Respondent20	37,7057	34,6576	50,2614	35,9368	20,8415	22,1171	Belonging to the guideline data	Failed
21	Respondent21	57,8054	58,2673	85,166	24,6242	52,5613	39,5156	Belonging to the guideline data	Failed
22	Respondent22	33,0214	38,0454	42,7767	42,3725	25,4195	33,7773	Belonging to the guideline data	Failed
23	Respondent23	41,4974	39,9628	66,8592	15,9634	38,21804	27,63034	Belonging to the guideline data	Failed
24	Respondent24	51,0921	51,1571	79,7653	23,55	47,6627	33,2665	Belonging to the guideline data	Failed
25	Respondent25	46,7772	45,2199	53,8587	43,587	29,7007	25,993	Belonging to the guideline data	works
26	Respondent26	46,7772	45,2199	53,8587	43,5870	29,7007	23,3426	Belonging to the guideline data	works
27	Respondent27	45,9274	50,9755	41,5662	59,5204	27,3462	30,4041	Belonging to the guideline data	works
28	Respondent28	38,5918	42,7663	66,9229	20,397	37,2382	22,8509	Belonging to the guideline data	works
29	Respondent29	37,8877	26,0487	61,2362	21,6237	42,0035	39,996	Belonging to the guideline data	works
30	Respondent30	40,7947	43,4481	65,8272	26,0928	35,2201	22,6753	Belonging to the guideline data	works

From the previously tested data, the accuracy and success of the tested image matching can be evaluated using the following methods.

$$Accuracy(\%) = \frac{\sum \text{many successes,}}{\sum \text{total number of data}} \times 100\%$$

$$Accuracy(\%) = \frac{17}{30} \times 100\%$$

$$Accuracy(\%) = 56,67\%$$

Then, we continue to analyze the following steps (See Figure 3). Matrix P1 refers to the matrix derived from guideline 1. Using the established definition of 1, the Minkowski distance will be calculated for the matrix. This process is part of the evaluation of the similarity or difference between the reference matrix and other matrices within the framework of the predefined method.

The manual calculation is as follows. Comparison of respondent1 with guideline 1.

$$R1 = \begin{bmatrix} 223 & 197 & 202 & 189 \\ 164 & 114 & 109 & 122 \\ 140 & 110 & 116 & 115 \\ 141 & 109 & 116 & 124 \end{bmatrix}$$

$$P1 = \begin{bmatrix} 230 & 210 & 210 & 191 \\ 215 & 119 & 107 & 170 \\ 194 & 119 & 112 & 170 \\ 160 & 94 & 108 & 149 \end{bmatrix}$$

$$p = 3$$

$$d(x, y) = (|230 - 223|^3 + |210 - 197|^3 + |210 - 202|^3 + \dots + |149 - 149|^3)^{\frac{1}{3}}$$

$$d(x, y) = 84.22309366304167$$

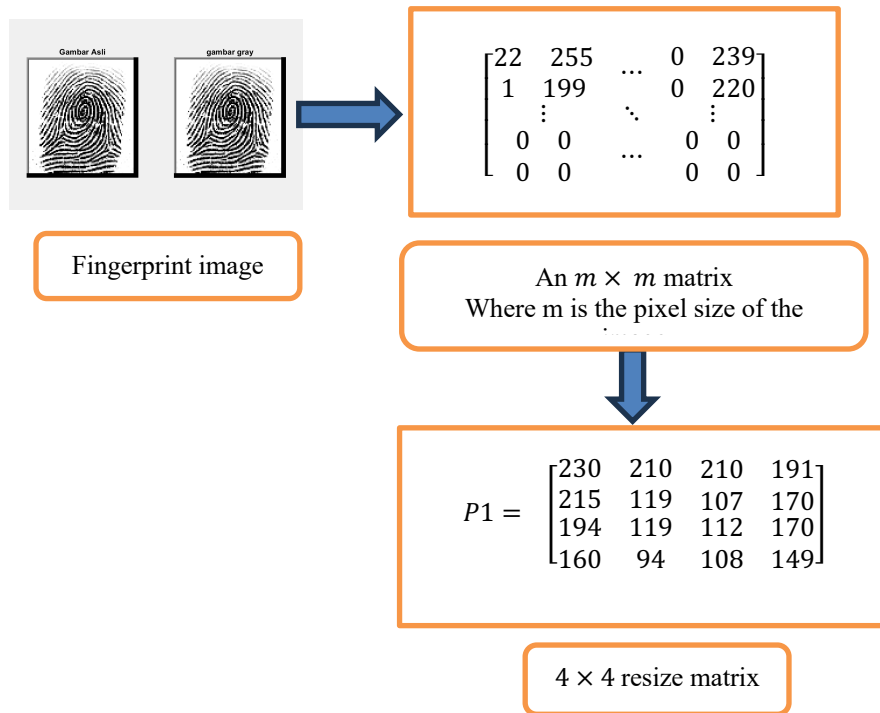


Figure 3. Further processing fingerprint data into 4 x 4 matrix

Table 2. Calculation on data from respondents based on Guideline 1 to 6

Number	Name	Guidelines 1	Guidelines 2	Guidelines 3	Guidelines 4	Guidelines 5	Guidelines 6	Classification Results	Success
1	Respondent1	84,22	73,57	88,12	67,43	78,46	98,34	Does not belong to guideline data	works
2	Respondent2	136,44	125,84	126,2	57,19	87,45	126,21	Does not belong to guideline data	works
3	Respondent3	56,96	56,24	45,5	118,57	95,86	109,39	Does not belong to guideline data	works
4	Respondent4	112,55	100,48	110,05	61,38	63,97	54,96	Does not belong to guideline data	works
5	Respondent5	82,75	65,96	56,08	89,55	71,95	104,89	Does not belong to guideline data	works
6	Respondent6	116,84	109,9	126,32	58,41	86,31	87,93	Does not belong to guideline data	works
7	Respondent7	104,3	101,75	87,09	89,55	103,86	128,97	Does not belong to guideline data	works
8	Respondent8	110,49	97,07	104,28	40,29	61,62	72,4	Does not belong to guideline data	works
9	Respondent9	112,23	120,2	104,83	196,64	176,45	175,44	Does not belong to guideline data	Failed
10	Respondent10	77,23	60,22	53,01	122,95	88,68	117,33	Does not belong to guideline data	works
11	Respondent11	117,87	109,58	97,66	84,73	94,93	137,41	Does not belong to guideline data	works
12	Respondent12	156,49	149,11	164,64	73,13	114,06	132,93	Does not belong to guideline data	works
13	Respondent13	87,7	93,45	82,32	70,12	80,18	83,35	Does not belong to guideline data	works
14	Respondent14	101,07	95,41	94,14	71,78	82,34	63,91	Does not belong to guideline data	works
15	Respondent15	150,15	140,17	154,36	58,43	99,89	100,22	Does not belong to guideline data	works
16	Respondent16	118,51	112,3	120,47	144,86	130,75	104,91	Does not belong to guideline data	works
17	Respondent17	92,22	91,22	88,97	68,91	70,87	99,97	Does not belong to guideline data	works
18	Respondent18	84,01	86,58	81,35	71,84	89,03	108,71	Does not belong to guideline data	works
19	Respondent19	87,14	62,28	71,4	94,48	22,25	68,25	Belonging to the guideline data	works
20	Respondent20	65,68	47,94	68,85	90,81	48,82	83,8	Does not belong to guideline data	works
21	Respondent21	112,01	94,76	110,85	79,91	58,23	55,23	Does not belong to guideline data	works
22	Respondent22	115,32	107,19	111,48	38,59	79,1	82,58	Does not belong to guideline data	works
23	Respondent23	154,11	140,82	137,74	75,4	108,51	156,37	Does not belong to guideline data	works
24	Respondent24	104,02	98,42	91,38	75,17	91,59	131,26	Does not belong to guideline data	works
25	Respondent25	137,45	124,63	128,49	39,47	95,98	130,64	Does not belong to guideline data	works
26	Respondent26	71,03	71,34	81,29	70,24	69,37	77,69	Does not belong to guideline data	works
27	Respondent27	120,82	101,72	120,31	58,49	54,13	56,14	Does not belong to guideline data	works
28	Respondent28	118,48	100,99	115,37	69,15	58,04	46,33	Does not belong to guideline data	works
29	Respondent29	99,31	100,26	89,67	104,67	120,56	126,62	Does not belong to guideline data	Failed
30	Respondent30	61,51	66,25	56,47	88,02	70,92	75,87	Does not belong to guideline data	works

With a similar approach, calculations were made using Guideline 1 to Guideline 6 against the responses from Respondent 1 to Respondent 30. This step was carried out to analyze and evaluate the suitability and similarity between each guideline and the respective responses generated by the respondents involved in this study and Table 2 was obtained.

From the previously tested data, the accuracy and success of the tested image matching can be evaluated using the following methods:

$$Accuracy(\%) = \frac{\sum \text{many successes,}}{\sum \text{total number of data}} \times 100\%$$

$$Accuracy(\%) = \frac{28}{30} \times 100\%$$

$$Accuracy(\%) = 93\%$$

From the results of the first data processing, an accuracy rate of 56.67% was obtained. While in the second data processing, the accuracy rate increased to 93%. This significant difference is due to the different approaches in data collection. In the first data collection, there was no consideration of the data collection distance or cropping process on the data set. In contrast, in the second data collection, a more structured approach was used with data collection at a uniform distance and the application of a cropping process, which resulted in a significant increase in data accuracy.

## Conclusion

The results of using the Minkowski Distance method on the first data showed an accuracy rate of 56.67%, which is significant but worth noting for further improvement. The use of the second data that considered the capture distance and applied cropping resulted in an increase in accuracy to 93%, confirming the importance of a structured approach to data capture. A suggestion for future research is to consider other factors that affect accuracy, such as variations in image capture angle or more complex preprocessing techniques. Further evaluation of these factors can help improve the accuracy of the method in fingerprint image recognition.

Suggestions for further research include several things that can be optimized. Firstly, it is recommended to consider the use of additional data to improve a greater level of accuracy in fingerprint recognition. Second, it is recommended to consider an approach to minimize the fingerprint image in certain parts of the diagram to obtain a more optimized matrix. Thirdly, it is recommended to further normalize the matrix to produce smaller values, simplifying the calculation process. Finally, it is important to pay attention to the pixel size when photographing fingerprint data, considering it as a key factor in image capture for more accurate analysis. With these suggestions in mind, future research is expected to make significant progress in the development of more reliable and efficient fingerprint recognition methods.

## References

- Arifandi, M. (2019). *Prototype Sistem Keamanan Pintu Menggunakan Sidik Jari (Fingerprint) Berbasis Arduino Uno ATmega328 dan SMS Gateway*.  
<https://www.researchgate.net/publication/340165694>
- Aziz, N. (n.d.). *Pemanfaatan teknologi internet dalam pendidikan*.
- Emelda. (n.d.). *Pengaruh Absensi Fingerprint dan Sanksi Hukuman Terhadap Disiplin Kerja Pegawai Pada Dinas Perdagangan Provinsi Sumatera Selatan*. *STIE Rahmanyah Sekayu*.
- Gani, A. G. (n.d.). *Pengenalan teknologi internet serta dampaknya*.



- Gusri, M., Dan, A., & Kadarisman, Y. (n.d.). *Penggunaan Teknologi Internet Dikalangan Mahasiswa Fakultas Ilmu Sosial Dan Ilmu Politik Universitas Riau*.
- Hartono, B., Eniyati, S., & Hadiono, K. (2023). Perbandingan Metode Perhitungan Jarak pada Nilai Centroid dan Pengelompokan Data Menggunakan K-Means Clustering. *Jurnal Sistem Komputer Dan Informatika (JSON)*, 4(3), 503. <https://doi.org/10.30865/json.v4i3.6021>
- Hermawan, D., Ullah, A., & Faizal, A. (2023). *Rancang Bangun Keamanan Kotak Amal dengan Akses Fingerprint Menggunakan ESP32-Cam dan Telegram Berbasis IOT*. 7(3), 1013–1021. <https://doi.org/10.30865/mib.v7i3.6252>
- Kasus Cybercrime Dengan Studi Kasus Hacker Bjorka Terhadap Pembocoran Data Zaki Rizaldi, A. M., Dwi Putra, R., Ul Hosnah, A., & Zaki Rizaldi, M. (n.d.). *Analisis Kasus Cybercrime Dengan Studi Kasus Hacker Bjorka Terhadap Pembocoran Data*. <http://jurnal.um-tapsel.ac.id/index.php/justitia>
- Khairunnisa, A., Putri Efendy, M., Zamri, M., & Tusakdiyah, S. H. (n.d.). *Aplikasi Pendeteksi Gejala Penyakit Jantung Menggunakan Metode Minkowski Distance Dengan Citra Matlab (Vol. 3, Issue 1)*.
- Mayamin, M., & Usuluddin, L. (2023). Pengaruh Absensi Sidik Jari (Finger Print) terhadap Disiplin Kerja Pegawai. *Remik*, 7(1), 602–609. <https://doi.org/10.33395/remik.v7i1.12091>
- Nishom, M. (2019). Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square. *Jurnal Informatika: Jurnal Pengembangan IT*, 4(1), 20–24. <https://doi.org/10.30591/jpit.v4i1.1253>
- Nuriadin, A., Dyan Nofia Harumike, Y., Tana Sanggamu, D. (2021). Sejarah perkembangan dan implikasi internet pada media massa dan kehidupan masyarakat. *SELASAR KPI: Referensi Media Komunikasi Dan Dakwah*, 1(1). <https://ejournal.iainu-kebumen.ac.id/index.php/selasar/index>
- Nuriadin, A. Y. D. N. (2021). *Sejarah Internet di Indonesia*.
- Nurliza, N. N., & Widodo, E. et al. (2018). *Penerapan euclidean distance pada pengenalan pola citra sidik jarl*.
- Pradana, I. H. (2017). *Klasifikasi citra sidik jari berdasarkan enam tipe pattern menggunakan metode euclidean distance*.
- Safwandi, & Muthmainnah. (n.d.). *Sistem Pendeteksi Terjemahan Kifayatul Muhtadi Ke Dalam Bahasa Indonesia Menggunakan Metode Minkowski Distance*.
- Thant, A. A., & Aye, S. M. (2020). Euclidean, Manhattan and Minkowski Distance Methods For Clustering Algorithms. *International Journal of Scientific Research in Science, Engineering and Technology*, 553–559. <https://doi.org/10.32628/ijrsrset2073118>

This page is intentionally left blank.