PATTERN RECOGNITION OF SIGNATURE VERIFICATION USING CELLULAR AUTOMATA METHODS

[1]ADIANANDA, [2]RETANTYO WARDOYO

[1]Program Studi S2 Ilmu Komputer, FMIPA UGM, Yogyakarta

[2]Jurusan Ilmu Komputer dan Elektronika, FMIPA UGM, Yogyakarta

e-mail: **1**adi.dirza@gmail.com, 2rw@ugm.ac.id

ABSTRACT

A pattern recognition technique in the field of machine learning and can be defined as "the act of taking raw data and act upon data classification". Much research has been done on the topic of pattern recognition using a variety of methods one of which is by using the cellular automata. In this study used cellular automata method for finding and extracting characteristics of an image of the signature and realized in a pattern recognition software that can verify the authenticity of the signature image by using cellular automata method for the extraction process characteristics. In this study used data 57 respondents with 6 signatures used as a reference image. Three pieces of the original signature image and 10 pieces of counterfeit signature image is used as the test images (query). From the testing that has been done precision 88.30%, recall 65.37% and accuracy 71.31%..

Keywords: *pattern recognition, cellular automata, verification, signature*

1. INTRODUCTION

Pattern recognition for signature verification is one of the fields of pattern recognition that is developing today, where the application can be applied in various fields, especially in the field of security (security system), such as permit the withdrawal of money in the bank, validation checks, and so on. Pattern recognition technology to verify the signature included in the field of biometric technology which in this field using behavioral characteristics of human nature. Signature verification is necessary because each individual has a unique signature patterns different from each other. In general, to identify or verify the signature can be done manually, by matching the signature at the time of the transaction with a signature that is used as a legal reference. The manual system has a weakness where the examiner signature less careful in doing matching. Therefore, we need a method that is capable of analyzing the characteristics of the signatures automatically so hopefully it easier to recognize the signature of a person[20].

There are several methods used for pattern recognition (signature verification), one of which is a method based on neural networks [1]. Artificial neural networks have been used widely in the field of pattern recognition and generally show the advantages compared to other methods of learning, the nature of generalization and adaptability, as well as his trademark strength in performing non-linear mapping [2]. Nevertheless the introduction of the signature included most difficult problems in pattern recognition. This is because the examples signatures are the same person, but not identical [3].

In this study, methods of Cellular Automata (CA) was used for pattern recognition signature. Where, Stephen Wolfram stated that in such a simple program that there is universality to decipher all the phenomena that exist in nature with cellular automata models, such as looking for leaf shape, the shape of snow crystals, animal skin patterns, and other forms of artistic [4]. Cellular automata (CA) is a discrete spatiotemporal system [5], which can be used to model complex systems and dynamic. In the implementation of a cellular automata system is composed of cells, where each cell will have a state (state) specific. The state of a cell is changed or not depending on the state of the cell itself and the state of the neighboring cells that are owned by these cells [20].

Various domain issue has been investigated and the success can be solved by using CA's. In this case, digital image processing such as research done by [6], [7], [8], [10], [11], [12], [13], [14], [15]. Similarly, in the field of pattern recognition methods have been introduced CA's basic pattern recognition [16], [17], [18]. In this study is made classificatory pattern by using a simple local network of Primary or Elementary Cellular automata Cellular Automata (ECA) [19].

In this study will be used in the form of a data signature image file signatures obtained through the scanning of multiple signatures. To determine the characteristics of each image vector signatures, will be used the method of measuring the frequency of occurrence of the patterns of neighboring owned by cellular automata. As known that the state of a cell will be influenced by the neighboring cells, so that in this study will count the number of occurrences of each pattern that affects a cell neighbor when the neighbor patterns iterated on a signature image. Patterns neighbor used in this study is a model neighbor 5 Neumann [20].

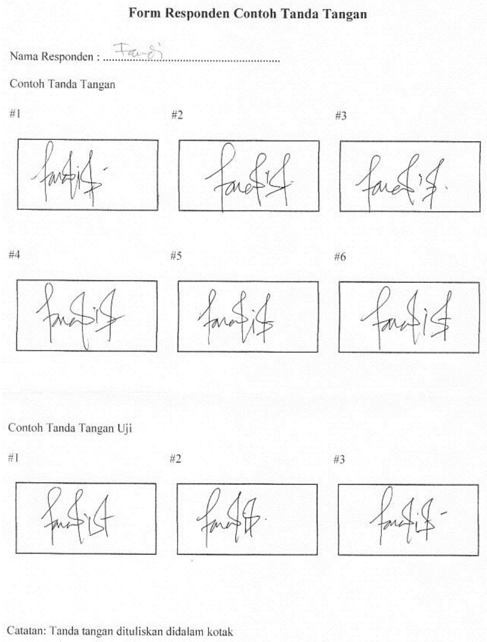
2. RESEARCH METHODS

2.1 Data Acquisition

To solve the problems in this study, the method used is divided into three phases: data acquisition, feature extraction stage and the stage of introduction pattern. Process data acquisition is the process to get the signatures of data that will be used as an input system. In this process begins with a signature sample collection from multiple users. The collection of signatures is done on a sheet of paper and wrote a predetermined signatures in the form of a rectangular box with a size of 2.5 cm x 4.5 cm. In doing writing a signature the user must write it in the rectangular box and should not be out of the box provided. After sampling the signature, then do the scan to obtain the image file. Signature image file obtained through scanning results are stored with file format * .jpg, grayscale (grayscale) 8-bit with a resolution of 100 dpi. The scans sample signature as seen in Figure 1.

Examples of image signatures that have been scanned and then cropped to obtain nine pieces of image signatures, each signature image is also saved with the file format * .jpg grayscale (grayscale) 8 bit with a size of 90 x 190 pixels. Automatic crop is presented in Figure 2. The process of separating into 9 image signature is done by using an algorithm as follows:

1. Make sure that the image file that is used grayscale (grayscale) by converting any file input image into grayscale.
2. Convert the image into a binary image.
3. Perform dilation process to thicken lines in the image.
4. Identify the contours of the object.
5. Cut the image based on the contours obtained with a size of 90 x 190 pixels and save the result.


Form Responden Contoh Tanda Tangan

Nama Responden : ....

Contoh Tanda Tangan

#1 #2 #3

#4 #5 #6

Contoh Tanda Tangan Uji

#1 #2 #3

Catatan: Tanda tangan dituliskan didalam kotak

Figure 1. The scans sample signature

Figure 2. The result of Automatic crop

Stage feature extraction is the process to obtain a set of characteristics of the image of the signature. There are several processes that can be done to get these traits. In this study the process to get traits were calculated using cellular automata. Cellular automata can be defined as a collection of cells that form the array where each cell contains the value of digital conditions. Conditions such cells are updated in parallel, while the method for renewing the cell condition is local and homogeneous. Parallel nature, local and homogeneous is a characteristic of cellular automata.

Cellular automata consist of one, two, to n dimensions. A one-dimensional cellular automata will form a line one-dimensional array, where each cell only touches two other cells, one on the left and one on the right. In this study used only one-dimensional cellular automata, the use of cellular automata for feature extraction in this study conducted by calculating the frequency of occurrence of configuration patterns of neighboring owned by cellular automata. Configures neighborhoods pattern used is a model Neumann with five neighbors.

If the signature image that measures 90 x 190 pixel are denoted as f and there are 32 patterns of cellular automata adjacency, then the image of the signature can be expressed as a column vector x dimension 32, with 32 is the number of neighborhood patterns. So it can be written in matrix form as follows:

$$x = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{Bmatrix}$$

The elemtents of x, e.g. x_1, x_2, x_3, …, x_{32} are the frequency of present of each neighborhood pattern (CA0, CA1, CA2, ..., CA31).

Feature extraction process by using cellular automata are performed using the following algorithm :

1. Ensure that the image file that is used grayscale (grayscale) by converting any file input image into grayscale.
2. Convert the images into binary images.
3. CAn = 0 ; for n = 0 ... 31
4. Run n adjacency patterns of cellular automata in the signature image, compare the pixel values of the image of the signature patterns of neighborhood.
5. If the comparison results are equal, then CAn = CAn + 1
6. Save CAn.

2.2 Pattern Recognition

Phase pattern recognition conducted in this study consists of several stages. The stages are as follows :

- Classification Process.

  In the pattern recognition classification process is a process to classify the characteristics of adjacent or similar into one. This grouping process is usually done by measuring the distance between the traits characteristic to each other. The characteristics are considered as a group if it has a small difference in distance or close to zero. In this study will be used method of measuring distance Euclidean distance. Euclidean distance calculating root of the squared differences 2 vectors. Euclidian used to calculate the distance equation :

$$d(a,b) = \sqrt{\sum_{i=1}^{n}(a_i - b_i)^2}$$

- Threshold-values Calculation Process.

  To determine the authenticity of the signatures will be used threshold value method. The threshold value is a value to be used as a reference and is used in the matching process between a query signature image with the image of a reference signature. It can be said also that the threshold value is a value that represents the characteristics possessed by the image of a reference signature. This threshold value will be calculated by finding the difference between the average distance characteristic signature image of the whole reference to the characteristics of each image signature. The calculations were made to seek NA is as follows:

  a. Average values Calculation of Reference Image.

     If m is sum of images, and x_1, x_2. … x_p are features of each images. Then, average of x are sum of all fitures ($x_{p1}+x_{p2}+\ldots+x_{pm}$) devide by sum of images (m).

     $$X_{p\_rata-rata} = \frac{1}{m}(x_{p1} + x_{p2} + \cdots + x_{pm})$$

  b. Distance (dx) of each features calculated using $X_{p\_rata-rata}$. :

     $$dx_p = jarak(X_{rata-rata}:x_{pm})$$

  c. Calculate the average distance and standard deviation of features distance (S) from averages of reference images.

     $$.dx_{rata-rata} = \frac{1}{p}(dx_1 + dx_2 + \cdots + dx_p)$$

     $$\sigma = SD(dx_1, dx_2, \ldots, dx_p)$$

     where : p is count of features and m is count of images

  d. Threshold-value Calculation (NA)

     To calculate threshold value (NA) :

$$NA = \sqrt{\sum_{t=1}^{p}\left(dx_{rata-rata,t} + \sigma_t\right)^2}$$

  where p = count of features

- Experimental Process

  Testing is part of the system used to determine and verify the signature image inserted is genuine or fake. Processed in this test also will be known how acceptance and rejection on the image of the input system.

  In the process of testing done some calculations, as for the calculations can be explained as follows:

  a. Single characteristic distance calculations with average (dx_i)

     The calculation of the distance characteristics of each single image is done in the same way at the time to calculate the distance to the calculation of the value of NA is by using equation:

     $$dx_p = jarak(X_{rata-rata}:x_{pm})$$

  b. NQ Value Calculation

     Nq is a value that represents the entire image characteristics possessed by the image of the signature test. To perform the calculation of the value of NQ in this study used the equation:

     $$Nq = \sqrt{\sum_{i=1}^{p} dx_i^{\,2}} \quad , \text{p = count of features}$$

- Inference Process

  The decision making process is the process of determining the calculation results of the testing process. This process is used to determine the authenticity of the signature image is tested by comparing the calculation results with the NQ NA. If NQ <= NA then the image of the signature is "authentic", otherwise the image of the signature which is input is "false".

2.3 Result Evaluation

The success rate of this study will be measured is the precision, recall and accuracy. Precision will be used to measure the level of acceptance of the system against the signature of the original test of each respondent. Precision measurement will be done using the equation :

$$precision = \frac{TP}{TP + FP}$$

where : TP are sum of recognized handwriting
TF are sum of unrecognized handwriting

While the recall will be used to test level system rejection of the false signatures are included. Measuring the level of recall will be used equation:

$$recall = \frac{TP}{TP + FN}$$

where : TP are sum of false-handwriting rejected
FN are sum of false-handwriting recognized

To measure the degree of accuracy of the overall system accuracy parameters will be used, which were for level measurement accuracy of the system will be used equation :

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

where : TP are sum of recognized original handwriting
FP are sum of unrecognized original handwriting
FN are sum of recognized false handwriting
TN are sum of unrecognized false handwriting

3. RESULTS AND DISCUSSION

3.1 Pattern Recognition

In this study by using cellular automata configurations of neighborhood characteristics can be generated from the image of the signature. In the extraction characteristics is done by comparing the adjacency with each pixel configuration which is owned by the image of the signature, if the same results will be created a counter that will count up.

In this study, the pattern recognition process is done in several steps before it can be concluded about the authenticity of the signatures are verified image. In Table 1 are presented the value of NA, on average, D_rata average and D_stdv for respondents noid = 2.

Table 1. Calculation of Average, D_rata average, D_stdv and NA

NoID	pattren	average	D_average	D_stdv	NA
2	CA0	1.07	1.07	0.96	6873.91
	CA1	1.45	1.45	0.89	
	CA2	1.45	1.45	1.01	
	CA3	6.71	6.71	4.62	
	CA4	0.00	0.00	0.00	
	CA5	0.50	0.50	0.00	
	CA6	1.33	1.33	1.33	
	CA7	34.29	34.29	19.06	
	CA8	1.98	1.98	1.92	
	CA9	1.60	1.60	1.40	
	CA10	4.45	4.45	3.64	
	CA11	11.50	11.50	8.93	
	CA12	0.71	0.71	0.30	
	CA13	5.02	5.02	2.71	
	CA14	2.55	2.55	1.16	
	CA15	85.02	85.02	55.25	
	CA16	1.43	1.43	0.98	
	CA17	19.45	19.45	13.81	
	CA18	1.17	1.17	0.82	
	CA19	20.83	20.83	11.25	
	CA20	0.64	0.64	0.38	
	CA21	1.33	1.33	0.94	
	CA22	4.29	4.29	2.98	
	CA23	66.90	66.90	46.96	
	CA24	6.33	6.33	4.89	
	CA25	19.86	19.86	12.95	
	CA26	12.05	12.05	8.65	
	CA27	50.55	50.55	25.45	
	CA28	33.69	33.69	20.20	
	CA29	68.40	68.40	48.72	
	CA30	84.50	84.50	54.54	
	CA31	7226.55	7226.55	5078.91	

From the results of such calculations are presented in Table 1. The value of NA is a value used as a reference for decision-making about the authenticity of the signatures to be tested. NA value is obtained by following the steps as below :

1. Calculate the average of each characteristic of 6 pieces of reference image signatures.
2. Calculate the distance of each single characteristic of the average reference image.
3. Calculate the average distance and standard deviation of step No. 2
4. Count NA

3.2 Examination

The final goal of this study was to test the feature extraction results performed using cellular automata method for verifying signatures. There are two tests to be performed, namely :

a. Examination of original handwriting

The first testing was used to determine the level of acceptance of the system in verifying the original signature of each respondent. For that in this verification test will be used three pieces of signatures for each respondent. Three signatures are used for testing an original signature. Prior to verify it with the same process as at the time of calculation and determination of the threshold value, then the signatures to be verified (Q) is extracted character, then after the obtained characteristics for each pattern were calculated within each characteristic. The results are presented in Table 2 .

Table 2. Calculation Results of NQ 3 (original signature)

NoId	The Value of The query image test		
	NQ1	NQ2	NQ3
1	232.76	873.48	2099.29
2	259.00	255.10	354.72
3	668.98	60.34	255.37
4	2326.82	215.06	715.54
5	168.61	1748.90	524.18
6	2104.30	1784.24	2963.29
7	414.13	168.27	1289.47
8	317.30	413.89	727.94
9	297.00	525.75	534.29
10	29.17	369.18	988.18

b. Examination of false handwriting

Testing the second is to test the system rejection of the false signatures are used as inputs. To perform this test was taken 10 pieces of random signature sample of 57 respondents who have registered. The testing process is carried out in general is the same as the original signature of the testing process, i.e. each sample signature taken extracted to be taken of the characteristics, then the calculated distance and the calculated value of NQ characteristics of each sample signature, as presented in Table 3.

Table 3. Calculation Results of NQ 10 (false signature)

References		NQ 10 false signature									
NoID	NA	NQ1	NQ2	NQ3	NQ4	NQ5	NQ6	NQ7	NQ8	NQ9	NQ10
1	3423.56	2099.29	4433.89	5662.04	1839.70	202.92	1641.00	1799.70	1220.72	240.60	252.57
2	6873.91	6787.16	259.00	974.09	6526.04	4513.93	6311.19	2899.92	3531.51	4522.03	4692.92
3	2396.07	498.79	7022.18	8250.33	753.72	2765.91	990.09	4381.52	3774.97	2758.38	2595.48
4	4266.59	5434.10	1101.17	2326.82	5172.78	3161.53	4959.62	1550.81	2185.02	3169.89	3341.95
5	2952.02	2155.31	4390.67	5617.56	1894.04	186.82	1665.83	1746.23	1193.88	226.96	81.53
6	1236.68	1624.64	4907.89	6136.51	1367.01	654.66	1172.81	2271.39	1689.60	667.78	536.44
7	685.01	2655.51	9184.07	10412.12	2914.59	4927.14	3135.59	6542.26	5931.63	4919.10	4752.73
8	1847.18	1499.82	8027.32	9255.54	1759.14	3770.01	1981.30	5385.60	4779.04	3763.64	3597.14
9	3409.91	783.82	5748.84	6977.18	525.75	1493.84	393.86	3110.66	2511.15	1490.68	1337.07
10	1330.24	1832.37	8362.02	9590.25	2091.39	4106.10	2321.64	5721.80	5109.80	4097.19	3934.75

3.3 Result of Examination

Based on experiments that have been done, to be able to assess the success of the proposed method in this study, the testing or verification. In the first testing is acceptance testing of the system, in this test used a test sample signature is genuine with a total of 171 pieces of the image of the signature (3 pieces of test images x 57 respondents). In the testing process is done calculating the value of NQ NQ then the value is compared with a threshold value (NA) and the results are presented in Table 4.

Table 4. Testing Results of Original Signature

the amount of the original image		
test	known	rejected
171	151	20

The testing process system rejection by using the method of taking 10 pieces image random signature from 57 respondents there, then all 10 images was tested to all respondents except himself (tested to 56 respondents), and the test results are presented in Table 5.

Table 5. Testing Results of False Signature

the amount of testing	the number of fake image		
	test	known	rejected
560	10	194	366

From the test results, it appears that for the acceptance testing of the 171 original signatures of the tested successfully identified well as 151 pieces are original signatures while 20 pieces of signatures considered false. While the fake signatures testing of the 10 signatures test tested as much as 560 times, about 366 times rejected by the system identified as fake signatures and 194 times recognized as an original signature.

To determine the level of accuracy of this system, then the calculated level of precision, recall and accuracy. The results of the calculations are presented in Table 6.

Table 6. The Accuracy of Testing Results

the result of evaluation	
PRECISION	88.30%
RECALL	65.36%
ACCURACY	71.31%

4. CONCLUSION

- Results of verification by using 171 pieces of original image with the signature obtained acceptance rate was 151 pieces of original image recognizable signature and 20 pieces of original image is not recognizable signature.
- Results of the verification of the 10 pieces of fake signature image taken at random from the 57 respondents and tested to 56 respondents obtained results fake signatures 366 times and 194 times successfully rejected the false signatures are recognized as the signature image of the original.
- The level of precision of the proposed method is 88.30%, to recognize the image of the original signatures.
- The level of recall of the proposed method is 65.36%, to reject the false signatures.
- The level of the overall system accuracy is equal to 71.31%.

5. SUGGESTION

For the improvement of the proposed method should add characteristics extracted from the image of the signature, or by combining cellular automata method with other methods.

REFERENCES

[1] Cemil, OZ., Ercal, F., Demir, Z.,(2003) Signature Recognition and Verification with ANN, Third International Conference on Electrical and Electronics Engineering, Bursa, Turkey
[2] Simon, H. (1999). Neural networks: A comprehensive foundation, New Jersey:Prentice Hall.
[3] Abbas, R., (1994), A Prototype System for off-line Signature Verificationusing Multilayered Feedforword Neural Networks. Tesis Departemen of Computer Science, RMIT.
[4] Wolfram,S., (2002), A New Kind of Science, Wolfram Media Publisher.
[5] Neumann, J.V. (1966). Theory of Self-Reproducing Automata. University of Illinois Press.
[6] Wongthanavasu, S., & Lursinsap, C. (2004). A 3D CA-based edge operator for 3D images. in Image Processing, ICIP'04. 2004 International Conference on, 2004, 235-238.
[7] Wongthanavasu, S., & Sadananda, R. (2003). A CA-based edge operator and its performance evaluation. Journal of Visual Communication and Image Representation, 14, 2, 83-96.
[8] Wongthanavasu, S., & Tangvoraphonkchai, V. (2007). Cellular Automata-Based Algorithm and its Application in Medical Image Processing. Image Processing, ICIP 2007. IEEE International Conference on, 2007, 41-44.
[9] Rosin, P.L. (2006). Training cellular automata for image processing. IEEE .ransactions on Image Processing, 15(7), 2076-2087.
[10] Anand Prakash Shukla, Swati Chauhan, Suneeta Agarwal,(2013), Training of Cellular Automata for Image Filtering, Proc. of Int. Conf. on Advances in Computer Science and Application, © Elsevier, 2013, pp. 632-639
[11] Popovici, A and Popovici, D, (2002), Cellular automata in image processing, in Proceedings of the 15th International Symposium on the Mathematical Theory of Networks and Systems.
[12] Peer M. A., et al., (2012), Investigations of Cellular Automata Game of Life Rules for Noise Filtering and Edge Detection, I.J. Information Engineering and Electronic Business, 2012, 2, pp. 22-28
[13] Priyanka Shotrya, Sanjeev Bhardwaj, (2013), Image Segmentation Using Cellular Automata: A Technical Survey, International Journal Of Engineering And Computer Science ISSN:2319-7242, Volume 2 Issue 4 April, 2013 Page No. 1268-1272
[14] R.S.RajKumar, G.Niranjana, (2013), Image Segmentation and Classification of MRI Brain Tumor Based on Cellular Automata and Neural Networks , IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 1, Issue 1, March, 2013
[15] Zijl, Botha, (2009), Feature Extraction for Image Pattern Matching with Cellular Automata, pp. 3–14., Proceedings of the Prague Stringology Conference 2009, Prague, Czech Republic, August 31 - September 2, 2009, ISBN 978-80-01-04403-2
[16] Ganguly, N., Maji, P., Sikdar, B.K., & Chaudhuri, P.P. (2002). Generalized Multiple Attractor Cellular Automata (GMACA) For Associative Memory. International Journal of Pattern Recognition and Artificial Intelligence, Special Issue: Computational Intelligence for Pattern Recognition, 16(7), 781-795.
[17] Maji, P., Ganguly, N., & Chaudhuri, P.P. (2003). Error correcting capability of cellular automata based associative memory. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 33(4), 466-480
[18] Maji, P., & Chaudhuri, P.P. (2008). Non-uniform cellular automata based associative memory: Evolutionary design and basins of attraction. Information Sciences, 178(10), 2315-2336.
[19] Wolfram, S., (1994). Cellular automata and complexity: collected papers. Addison-Wesley Pub. Co.
[20] Herulambang, Wiwiet (2016), Modeling The Effect Of Fertilization On Growth Pattern Of Brassica Rapa Using Backpropagation Neural Network, *JEECS (Journal Of Electrical Engineering And Computer Sciences) Vol 1 Number 1, 2016*, Surabaya, Indonesia.