Pattern Recognition of Batak Script Using Habbian Method

1st Frinto Tambunan
Faculty of Engineering and Computer Science
Universitas Potensi Utama
Medan, Indonesia
frintoaja@gmail.com

2nd Erwin Ginting
Faculty of Engineering and Computer Science
Universitas Potensi Utama
Medan, Indonesia
erwinginting82@gmail.com

3rd Edy Victor Haryanto
Faculty of Engineering and Computer Science
Universitas Potensi Utama
Medan, Indonesia
edyvictor@gmail.com

4th Muhammad Fauzi
Faculty of Engineering and Computer Science
Universitas Potensi Utama
Medan, Indonesia
fauzixx@gmail.com

***Abstract*—The introduction of patterns is an analytical work, which is to process raw data (not loading explicit structures) to find the various elements of the attributes/information associated with it. This analytical work is commonplace for human beings. By using the five senses, human beings can see, hear, and feel the condition of its environment to then identify the things that are in it, digest and understand it, so that he can do the activities or interactions that suit his wishes. Facial recognition, voice, signposts and writings, the computer performs pattern recognition by following several stages: sensing, segmentation, feature extraction, and classification/inference. One of the methods used to recognize the pattern is with the help of neural networks. In recognizing the Batak alphabet pattern, it uses one of the methods that exist in the neural network, using the Habb method. So, it is hoped that the Habb method can recognize pattern patterns or features of the Batak script.**

Keywords—Artificial neural network, Habb, Pola, Batak script

I. INTRODUCTION

The introduction of the pattern learns the use of computers to find regularity in data and to obtain important information/understanding of the regularity that has been found. In a simpler sense, the field introduction to Pattern recognition aims to provide the computer with the ability to detect the existence of objects/events in an environment and to determine the type/name of the object/event. In everyday life, pattern recognition appears in the form of a computer capable of recognizing the human face and voice, able to turn speech into text (transcribe), able to read or determine the meaning of signs/posts on the highway, able to predict the weather, predict the price of commodities, and so forth. With the ability of the computer is expected to recognize the form of Batak script that will be researched

II. RELATED WORK

Pattern Recognition Studies study the use of computers to find regularity in data and to obtain important information/understanding of the regularity that has been found [1]. In a simpler sense, the field introduction to Pattern recognition aims to provide the computer with the ability to detect the existence of objects/events in an environment and to determine the type/name of the object/event [2][24]. In everyday life, the introduction of patterns appears in the form of computers that are able to recognize human faces and voices, able to turn speech into text (transcribe), able to read or determine the meaning of signs/posts on the highway, able to predict the weather, predict the price of commodities, and so forth [3].

Pattern recognition is an analytical work, which is to process the raw data (not loading the explicit structure) to find the various elements of the attributes/information associated with it [4]. This kind of analytical work is commonplace for human beings. By using the five senses, human beings can see, hear, and feel the condition of its environment to then identify the things that are in it, digest and understand it, so that he can do the activities or interactions that suit his wishes. Facial recognition, voice, signposts, and writings, for example, are routine activities that humans always work on [5] [6]. While humans can do these things quickly and easily, the actual details of the process are not fully understood by researchers and cannot yet be emulated by the computer as a whole. The introduction of the pattern, therefore, is still an open matter that is continually investigated by many researchers [7].

The computer performs a pattern recognition by following several stages: sensing, segmentation, feature extraction, and classification/inference [8]. Sensing is the first phase of the introduction of patterns where the computer uses various types of sensors to record data from the surrounding environment. Sound sensors (microphones), image and video sensors (cameras), electric current/voltage sensors, and vibration/motion sensors are examples of sensors commonly used by computers for this purpose [9].

Segmentation aims to determine the existence of a specific target that will be the object of the introduction in the data record [2][10][25]. The camera catches, for example, can be processed at the segmentation stage to find the existence and location of human faces [9] [11]. Similar to this, the sound recording from the microphone can be segmented to find the existence and timing of speech occurrences in it. When pattern recognition is done in an uncontrolled environment, segmentation can be a fairly complex job with a variety of issues that are not easy to handle [12].

Feature Extraction is the third stage of the introduction of the pattern where the computer takes the attributes (descriptors) of a target that has been segmented [12] [13]. For human faces, for example, these descriptors can face length/width, length/width of facial components, or skin colour [14][23].

Descriptors-the descriptors that have been taken are then used by the computer to perform the classification/inference, for example, for facial identification. A good descriptor, therefore, is a discriminatory descriptor, which is what supports the more modest classification/inference [15].

Pattern recognition has been performed on different types of data. Sound recordings, stock price footage, Seismography recordings, Electrocardiographic Recordings, musical note sequences, and word sequences are examples of 1-dimensional signal-shaped data [16]. Image[22] imagery[21], images on the Internet, satellite imagery, X-ray imagery is an example of data in the form of 2-dimensional signals. The video footage is an example of a 3-dimensional signal-shaped data [4] [10]. Jam data in advance, there is also multimodal data IE data collected from a combination of a number of sensors. Each type of data has its own characteristics. The same applies to the different types of problems and types of environment. Pattern recognition technology, therefore, does not know a generic solution that works well in all situations [17] [2]. Instead, the introduction of patterns is always related to the search for an alternative solution that is most appropriate for the situation, conditions, and the typical problems that are being faced [5] [14].

III. RESEARCH METHODS

A. Artificial Neural Networks

The neural network is a processor that does a massive distribution, which has a natural tendency to save an introduction that has been experienced, in other words, JST it has the ability to be able to do the learning and detection of objects[6][18]. Fundamentally, the learning system is the process of adding knowledge to JST that is continuity so that when the knowledge is used, it will be maximized to the fullest in recognizing an object. Neurons are a fundamental part of the JST's processing.

Neural networks established generalization of mathematical models of biological neural networks. The basic form of a neuron can be seen in the following figure:

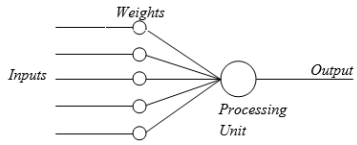

Fig. 1. The basic shape of neurons

Where:

- Input is the input used both in learning and in identifying an object.
- Weight, a load that is constantly changing every time it is given input as a learning process.
- Processing Unit is where the process of introducing an object based on the loading given.
- Output is output from the result of the introduction of an object.

B. Habb Algorithm

The Habb or Hebbian network is one of the branches of the artificially intelligent (artificial intelligence) branch [19] which, the Habb network is developed by modifying its weight and severity. The Habb network does not use the Hiden layer, which is a layer of hidden that is found in the derivative algorithm of the artificial neural network.

• Workaround steps

To complete the installation using the HABB algorithm is in each loop, the weight and bias are changed based on the multiplication of all inputs x_i directly connected to the output y unit, the weight value changes are done based on the equation

$$w_i(new) = w_i(old) + x_i y \quad (1)$$

Habb training algorithm with input s and Target unit T is as follows:

1) Initialize all weights = w_i= 0, (i=1,2,3,..,..,..,n)
2) for all input vertor and target unit t, do:
   a) Set activation for $x_i = s_i$ (i=1,2,3,..,..,..,n)
   b) Set of output Unit activation y = t
   c) fix weights by fulfilling equations

$$w_i(new) = w_i(old) + \Delta w \ (i = 1,2,3,..,..,..,n)$$

$$\Delta w = x_i y \quad (2)$$

   d) fix bias by equation b (new) = b (old) = t

Note that bias fixes are treated the same as weights.

C. Aksara Batak

As a media of writing communication, characters have significant differences in Latin letters, especially in terms of visual and technical readings. Visual differentiator, among others, is proportions, visual character, the anatomy of the letter, and geometry construction. The technical differentiator of the reading is the Batak script, which is the type of Silabik script, which describes the syllable (a-ha-ma) whereas the Latin alphabet belongs to the phonetic script type, which is the symbol of the phoneme (a-b-c). On the basis of these differences, the process of character adaptation of the letter requires a basic pattern of letters that become the intersection of the character of Batak Toba characters with the character of the Latin alphabet.

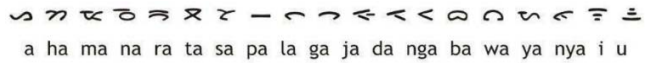

Fig. 2. Alphabetical order Aksara Batak[20]

Designing Latin letters with the character of Batak Toba characters is a challenge. Extensive cultural insights, precise data selection, and accuracy are required to process visual data to suit the purpose of design letters. Latin letters and Batak characters have a very different character that raises the question "It is quite Latin font form this Batak character?" or "How Batak is this Latin letter must be brought to the character?" The intersection between the character of the Latin alphabet and characters is the relative answer that will return to the value of the personal art of its creator and its users.

IV. RESULT AND DISCUSSION

The result of this research is by extracting the image into bipolar, which has been determined to be-1 and 1 later in the test using the specified target. Use Habb method

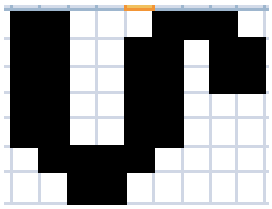

Fig. 3. First character

With the image above is obtained the following pattern for the first character.

11111-1-1 – 111111-1 - -1-1-1-1-111 – -1-1-1-1-111 – -11111-1 – 11111-1-1 – 1-1-1-1-1-1-1 – 111-1-1-1-1 – -111-1-1-1-1.

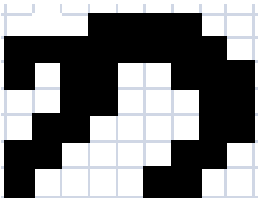

Fig. 4. Second character

With the picture above obtained pattern as follows for the second character:

-111-1-111 – -11-1-111-1 – -11111-1-1 – 1111-1-1-1 – 11-1-1-1-1-1 – 11-1-1-1-1-11 – 111-1-111 – -111111-1 – -1-1111-1-1

The extraction result from the image inserted into the table then specified the target to be achieved as shown in the table 1

TABLE I. EXTRACTION AND TARGET RESULTS

Pattern	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13
“1“	1	-1	-1	-1	1	-1	1	-1	1	-1	-1	-1	1
"2 "	-1	1	1	1	-1	1	-1	-1	-1	1	1	-1	-1

X14	X15	X16	X17	X18	X19	X20	X21	X22	X23	X24	X25	t
-1	-1	-1	1	-1	1	-1	1	-1	-1	-1	1	1
-1	1	1	-1	-1	-1	1	-1	1	1	1	-1	-1

While the weight change (ΔWI) and bias after given the pattern input 1 and 2 in table 2 as follows:

TABLE II. WEIGHT AND BIAS CHANGE VALUES

Patter n	Δw_1	Δw_2	Δw_3	Δw_4	Δw_5	Δw_6	Δw_7	Δw_8	Δw_9	Δw_{10}	Δw_{11}	Δw_{12}	Δw_{13}
“1 “	1	-1	-1	-1	1	-1	1	-1	1	-1	-1	-1	1
"2 "	-1	1	1	1	-1	1	-1	-1	-1	1	1	-1	-1

Δw_{14}	Δw_{15}	Δw_{16}	Δw_{17}	Δw_{18}	Δw_{19}	Δw_{20}	Δw_{21}	Δw_{22}	Δw_{23}	Δw_{24}	Δw_{25}	Δb
1	-1	-1	-1	1	-1	1	-1	1	-1	-1	-1	1
-1	1	1	1	-1	1	-1	-1	-1	1	1	-1	-1

And the final weight (wi) and b bias can be determined by the summation of the two weight changes above appear in table Table 3 below:

TABLE III. CHANGES IN THE WEIGHT AND BIAS

w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}	w_{11}	w_{12}	w_{13}	w_{14}
2	-2	-2	-2	2	-2	2	0	2	-2	-2	0	2	0

w_{15}	w_{16}	w_{17}	w_{18}	w_{19}	w_{20}	w_{21}	w_{22}	w_{23}	w_{24}	w_{25}	b
-2	-2	2	0	2	-2	2	-2	-2	-2	2	0

Then using the same formula as the previous calculations:

$$Net = \sum_{i=1}^{5} w_i x_i + b$$

Then the pattern 1

net = {1(2)+(-1)(-2)+(-1)(-2)+(-1)(-2)+(1)(2)}+{(-1)(-2)+1(2)+(-1)(0)+1(2)+(-1)(-2)}+{(-1)(-2)+(-1)(0)+1(2)+(-1)(0)+(-1)(-2)}+{(-1)(-2)+(1)(2)+(-1)(0)+1(2)+(-1)(-2)}+{1(2)+(-1)(-2)+(-1)(-2)+(-1)(-2)+1(2)} = 42 then *f(net)* = 1

Then the pattern 2

net = {(-1)(2)+1(-2)+1(-2)+1(-2)+(-1)(2)}+{1(-2)+(-1)(2)+(-1)(0)+(-1)(2)+1(-2)}+{1(-2)+(-1)(0)+(-1)(2)+(-1)(0)+(1)(-2)}+{1(-2)+(-1)(2)+(-1)(0)+(-1)(2)+(1)(-2)}+{(-1)(2)+(1)(-2)+(1)(-2)+1(-2)+(-1)(2)} = -42 then *f(net)* = -1

V. CONCLUSION

Conclusions from the research conducted:

1. Neural networks, especially the method Habb can work well to recognize the picture patterns in this research is a picture of the Batak script
2. Habb can do pattern recognition well by using bipolar input with-1 and 1.
3. The Batak alphabet pattern can be well recognized in the research case above Conclusions from the research conducted:

ACKNOWLEDGEMENT

The authors would like to thank Universitas Potensi Utama for their advice and financial support. The authors are also grateful to all parties involved in this research.

REFERENCES

[1] M. Susmikanti, "Pengenalan Pola Berbasis Jaringan Syaraf Tiruan Dalam Analisa Ct Scan Tumor Otak Beligna," vol. 2010, no. Snati, pp. 26–31, 2010.

[2] R. F. Ardiansyah, "Pengenalan Pola Tanda Tangan Dengan Menggunakan Metode Principal Component Analysis (PCA)," Fak. Ilmu Komput. Univ. Dian Nuswantoro, vol. 2, p. 14 pages, 2013.

[3] F. Tambunan, "Implementation of Data Mining using the Clustering Method (Case: Region of the Actors of Theft Crime by Province)," IJISTECH (International J. Inf. Syst. Technol., vol. 2, no. 2, p. 75, 2019, doi: 10.30645/ijistech.v2i2.25.

[4] V. Pebrianasari, E. Mulyanto, and D. Erlin, "Analisis pengenalan motif batik Pekalongan," Techno.COM, vol. 14, no. 4, pp. 281–290, 2015.

[5] Faradiba, "Pengenalan Pola Sinyal Suara Manusia Menggunakan Metode Back Propagation Neural Network," J. EduMatSains, vol. 2, no. 1, pp. 1–16, 2017.

[6] DAVID, "Perancangan Perangkat Lunak Pengenalan Pola Karakter Menggunakan Jaringan Syaraf Tiruan Perceptron," J. Ilm. SISFOTENIKA, vol. 1, pp. 10–19, 2011.

[7] F. Tambunan, Y. Yudi, and M. Fauzi, "Design of Artificial Neural Networks to Recognize Fingerprint Patterns," vol. 3, no. 1, pp. 58–63, 2019.

[8] N. Nurmila, A. Sugiharto, and E. A. Sarwoko, "Algoritma Back Propagation Neural Network Untuk Pengenalan Pola Karakter Huruf Jawa," J. Masy. Inform., vol. 1, no. 1, pp. 1–10, 2010, doi: 10.14710/jmasif.1.1.

[9] W. M. Sanjaya and D. Anggraeni, "Sistem Kontrol Robot Arm 5 DOF Berbasis Pengenalan Pola Suara Menggunakan Mel-Frequency Cepstrum Coefficients (MFCC) dan Adaptive Neuro-Fuzzy Inference System (ANFIS)," Wahana Fis., vol. 1, no. 2, p. 152, 2016, doi: 10.17509/wafi.v1i2.4277.

[10] Butarbutar, R. Somya, S. Kom, and M. Cs, "Perancangan Aplikasi Pembelajaran Aksara Batak Toba Menggunakan AndEngine Berbasis Android," p. 2.

[11] M. U. Musthofa, Z. K. Umma, and A. N. Handayani, "Analisis Jaringan Saraf Tiruan Model Perceptron Pada Pengenalan Pola Pulau di Indonesia," J. Ilm. Teknol. Inf. Asia, vol. 11, no. 1, p. 89, 2017, doi: 10.32815/jitika.v11i1.56.

[12] M. Syuhada, "Realisasi Pengenalan Plat Nomor Kendaraan Dengan Metode Histogram Citra Dan Jaringan Syaraf Tiruan Backpropagation," Skripsi, pp. 1–76, 2015.

[13] B. Suteja, "Penerapan Jaringan Saraf Tiruan Propagasi Balik Studi Kasus Pengenalan Jenis Kopi," J. Inform., vol. 3, no. 1, pp. 49–62, 2007.

[14] W. A. Utari, "Pengenalan pola dengan menggunakan metode backpropagation menggunakan matlab," 2018.

[15] P. Studi, I. Komputer, J. Matematika, F. Sains, D. A. N. Teknologi, and U. S. Dharma, "OSTEOPOROSIS PREDICTION THROUGH IRIS PATTERN USING BACKPROPAGATION ALGORITHM," 2008.

[16] K. Yudhistiro, "Pemanfaatan Neural Network Perceptron pada Pengenalan Pola Karakter," Smatika J., vol. 7, no. 02, pp. 21–25, 2017, doi: 10.32664/smatika.v7i02.153.

[17] R. Purba, "Tipografi Kreasi Motif Gorga Batak," Proporsi, vol. 1, no. 2, pp. 190–201, 2016.

[18] I. Ramadhani, S. H. Pratiwi, and A. N. Handayani, "Analisis Jaringan Saraf Tiruan Pengenalan Pola Huruf Hiragana dengan Model Jaringan Perceptron," J. Ilm. Teknol. Inf. Asia, vol. 11, no. 1, p. 45, 2017, doi: 10.32815/jitika.v11i1.41.

[19] Faisol, R. Amalia, Kuzairi, T. Yulianto, and M. Fariz Fadillah Mardianto, "Aplikasi Jaringan Habb dalam Pengenalan Huruf," Zeta (Math Journal), vol. 1, no. 1, pp. 10–14, 2015.

[20] F. Tambunan, "Pengenalan Aksara Batak Dengan Metode Perceptron," Jurasik (Jurnal Ris. Sist. Inf. dan Tek. Inform., vol. 4, no. 1, p. 160, 2019, doi: 10.30645/jurasik.v4i1.129.

[21] S. E. V. Haryanto, M. Y. Mashor, A. S. A. Nasir and H. Jaafar, "Malaria parasite detection with histogram color space method in Giemsa-stained blood cell images," 2017 5th International Conference on Cyber and IT Service Management (CITSM), Denpasar, 2017, pp. 1-4, doi: 10.1109/CITSM.2017.8089291.

[22] B. S. Riza, M. Y. Mashor, M. K. Osman and H. Jaafar, "Automated segmentation procedure for Ziehl-Neelsen stained tissue slide images," 2017 5th International Conference on Cyber and IT Service Management (CITSM), Denpasar, 2017, pp. 1-5, doi: 10.1109/CITSM.2017.8089292.

[23] S. E. V. Haryanto, M. Y. Mashor, A. S. A. Nasir and H. Jaafar, "A fast and accurate detection of Schizont plasmodium falciparum using channel color space segmentation method," 2017 5th International Conference on Cyber and IT Service Management (CITSM), Denpasar, 2017, pp. 1-4, doi: 10.1109/CITSM.2017.8089290.

[24] E. V. H. S., M. Y. Mashor, A. S. A. Nasir and Z. Mohamed, "Identification of Giemsa Staind of Malaria Using K-Means Clustering Segmentation Technique," 2018 6th International Conference on Cyber and IT Service Management (CITSM), Parapat, Indonesia, 2018, pp. 1-4, doi: 10.1109/CITSM.2018.8674254.

[25] B. S. Riza, M. Y. Mashor, M. K. Osman and H. Jaafar, "Segmentation for Tuberculosis Ziehl-Neelsen Stained Tissue Slide Image Using Thresholding," 2018 6th International Conference on Cyber and IT Service Management (CITSM), Parapat, Indonesia, 2018, pp. 1-3, doi: 10.1109/CITSM.2018.8674335.