Journal of Theoretical and Applied Information Technology
31st January 2017. Vol.95. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

COMBINING DECISION TREE AND BACK PROPAGATION GENETIC ALGORITHM NEURAL NETWORK FOR RECOGNIZING WORD GESTURES IN INDONESIAN SIGN LANGUAGE USING KINECT

WIJAYANTI NURUL KHOTIMAH, YOHANES ADITYA SUSANTO, NANIK SUCIATI
Informatics Department
Institut Teknologi Sepuluh Nopember Surabaya (ITS)
Indonesia
E-mail: wijayanti@if.its.ac.id; yohanes.aditya94@gmail.com; nanik@its-sby.edu

ABSTRACT

Sign language is a media for speech and/or hearing problem's people to communicate. Different kind of sign languages exist in the world such as Indonesian Sign Language (ISL), American Sign Language (ASL), Chinese Sign Language (CSL), British Sign Language (BSL), Brazilian Sign Language (BSL), and France Sign Language (FSL). In Indonesia, the used of ISL was less extensive because not all people understand it. People that do not have understanding on ISL cannot translate it. Therefore an ISL translation system is required. Many researches about sign language translation system had been done for FSL, BSL, FSL, and CSL. However, research on ISL is still limited and still need development. Therefore we proposed a new system for recognizing ISL word gestures. In this research we captured user skeleton by using Kinect. From those skeletons only nine skeletons were used as feature by computing their vector value, angle value, and distance value. Totally 28 features were extracted. Then the combination of Decision Tree and Back Propagation Neural Network (BPGANN) was implemented for classifier. For experiment, eight ISL vocabularies were tested by two people. The recognition accuracy of this system, although evaluated with small vocabulary, presents very promising result with value 96%.

Keywords: *Indonesian Sign Language Recognition, Decision Tree, BPGANN, Kinect*

1. INTRODUCTION

Sign language is a media for deaf people to communicate with others. Visual movements of their body are very helpful for them to make their message easily understood by their communication partner. If they communicate with lip movement, their message will be less understood. It becomes reason to develop Indonesian Sign Language [1].

Different kind of sign languages exist in the world. Each has its own gestures and vocabulary. For example are Indonesian Sign Language (ISL), American Sign Language (ASL), Chinese Sign Language (CSL), British Sign Language (BSL), Brazilian Sign Language (BSL), and France Sign Language (FSL) [2].

Indonesia has two kind of sign language system: BISINDO *(Berkenalan dengan Sistem Isyarat Indonesia*) and SIBI (Sistem *Isyarat Bahasa Indonesia*). BISINDO was developed by deaf people themselves through GERKATIN (*Gerakan Kesejahteraan Tuna Rungu Indonesia*). While SIBI was developed by normal people, non-deaf people [3]. SIBI becomes Indonesian official sign language. They consist of finger position and hand movement to represent Indonesian vocabulary. Gestures in SIBI had been arranged systematically and they followed convention [1].

However, the used of SIBI was less extensive. Only those who master at SIBI can communicate with deaf people. Therefore, the communication scope of deaf people was very limited. Hence, a tool that can be used to translate sign language, SIBI, was required.

SIBI sign language automatic translation had been done before [3] [4]. However, the translated sign language in those researches limited on alphabet sign language, A-Z, and numerical sign language, 0-9. Yet in reality, as a communication

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

tool, a person need not only alphabet and numerical sign language but also term sign language.

In 2015, a research about French Sign Language (FSL) had been done by Zbakh et al [5]. In this research, Zbakh et al aimed to build an application that able to easily find the meaning of the sign in French. In the same year, Velasco et al created a real time translator for Portuguese Sign Language [6]. Two devices, Microsoft Kinect and 5DT Sensor Gloves were used to gather data about hand motion and shape. In 2014, Moreira Almeida et al presented a methodology for feature extraction in Brazilian Sign Language (BSL). They investigated relation between extracted feature and structural elements in BSL. Seven vision-based features were obtained from the RGB-D image [7].

In Indonesia, translation tool application for Indonesian word sign language had been done by Rakun et al [1]. In their research, the detected sign are *air* (water), *bibi* (aunt), *lampu* (lamp), *paman* (uncle), *panjang* (long), *ajar* (teach), *biru* (blue), *matahari* (sun), *kuning* (yellow), and *pendek* (sort). These vocabularies are limited and need to be explored.

The research on Indonesian Sign Language is still limited and still need development. One of the development is in the classification process. In classification process, Back Propagation Genetic Algorithm Neural Network (BPGANN) method is one of Neural Network method that robust for classification. However sometimes overfitting problem that decrease classification accuracy occurs when the number of class is many. In this research, we try to solve the overfitting problem by grouping the class using rule. The rule was created by using Decision Tree algorithm and each group of the class will be processed by using BPGANN.

In this research, we proposed the combination of Decision Tree and Back Propagation Genetic Algorithm Neural Network for recognizing word gestures of Indonesian Sign Language. This research used Kinect to capture user gestures. The organization of this paper are as follows. In section II and III, we present the references and methods that needed for developing the application. Section IV discusses about the experiment. The last, in section V conveys the conclusion of this proposed research.

2. REFERENCES

2.1 Indonesian Sign Language

Indonesian Sign Language (ISL) was also known as *Sistem Isyarat Bahasa Indonesia* (SIBI). Generally, its gesture can be divided in to two, alphabet gestures and word gestures. Alphabet gestures in SIBI refer to ASL. The main component of it is finger movement. They can be shown in Figure. 1. Usually, they were used to spell names or words that are not exist in vocabulary dictionary.

Different to alphabet gestures, the main component of word gesture is hand movement instead of finger movement. In practice, word gestures are more widely used than alphabet gestures. Word gestures are also have more sign than alphabet gestures [2].

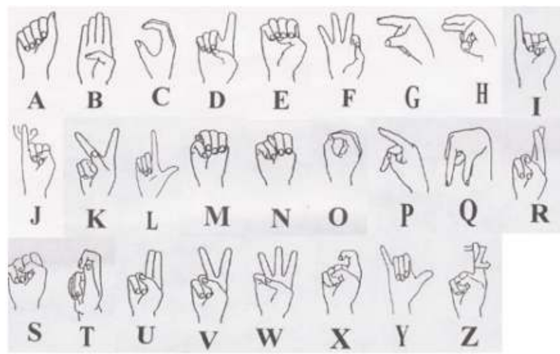

Figure. 1 Alphabet Sign Language in SIBI

In ISL, word gestures are divided into two kinds: (a) the basic word (basic gestures) as shown in Figure. 2 and (b) additional gestures (prefix, suffix) as shown in Figure. 3.

Figure. 2 Basic Gestures Example in SIBI

Figure. 3 Additional Gestures Example in SIBI

Journal of Theoretical and Applied Information Technology
31st January 2017. Vol.95. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2.2 Decision Tree

Decision tree is data mining technique that broadly uses for classification. Its model is like a tree structure with root node, parent node, child node and leaves node. The root node, parent node, and child node contain split and splitting attributes. They represent test on an atribute. Edges that coming out from those nodes represent consequences of the test. Leaf node is a node that associated directly with class label.

Generally, decision tree algorithm consists of three steps:

1. For a given dataset S, select a target attribute to split data.
2. Determine a splitting criteria to generate a partition in which each data belong to a single class. Choose best split to create a node.
3. Repeat those two steps iteratively until any stopping criteria fulfilled.

ID3 (Iterative Dichotomiser 3) and C4.5 are commonly used decision tree algorithm. They are different in the process of selecting target attribute. ID3 uses information gain as splitting criteria. The attribute with the highest information gain was used as target attribute to split data. The process to find target attribute was happen continually until the stopping criteria reached or the information gain close to zero.

C4.5 is improvement of ID3. It can handle continues attribute and missing value attribute. Instead of using information gain as splitting criteria, C4.5 used gain ratio as splitting criteria which is less bias [8].

The example of decision tree result was shown in Figure. 4. A, B, and C are feature while D1, D2, D3, and D4 are label. Rules can be generated from that tree to classify data [9]. The example rules generated from that tree was shown in Table 1.

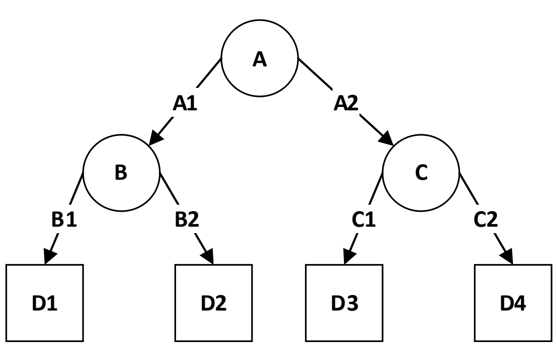

Figure. 1 Decision Tree Result Example

Table 1 Rule Example

Number	Rule
1	$A1 \wedge B1 \rightarrow D1$
2	$A1 \wedge B2 \rightarrow D2$
3	$A2 \wedge C1 \rightarrow D3$
4	$A2 \wedge C2 \rightarrow D4$

2.3 Back Propagation Genetic Algorithm Neural Network (BPGANN)

BPGANN is combination between Back Propagation Neural Network (BPNN) and Genetic Algorithm (GA). In this algorithm, the GA was used as feature selector and BPNN was used to compute the fitness value of each chromosome in GA [10]. The process of BPGANN was shown in Figure. 5.

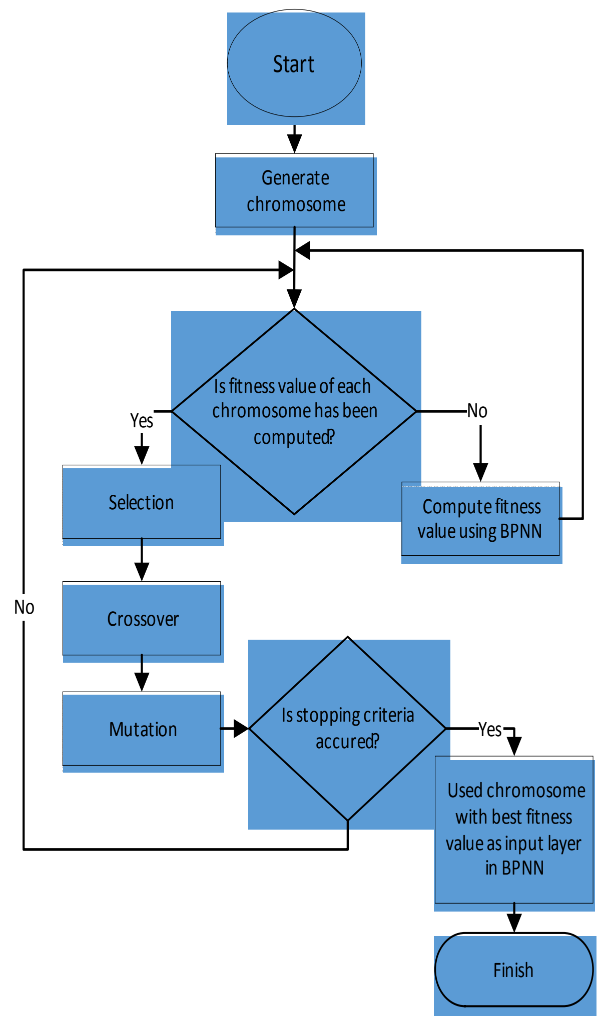

Figure. 5 The process of BPGANN

Journal of Theoretical and Applied Information Technology
31st January 2017. Vol.95. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The first time, n chromosomes were generated randomly, n is the number of population. Each chromosome was consist of x gens where x is the number of feature. The value of each gen is either 0 or 1. To compute the fitness value of each chromosome, the value of each gen was used as input layer node in BPNN. Each gen represented one feature, for example the first gen represented the first feature. If the first gen has value 1 means that the first feature will be used in the BPNN training and vice versa. Then BPNN training process was done and its accuracy was used as fitness value of the chromosome. The selection, crossover, and mutation process were done normally.

3. METHODS

The system architecture of this proposed research was shown in Figure. 6.

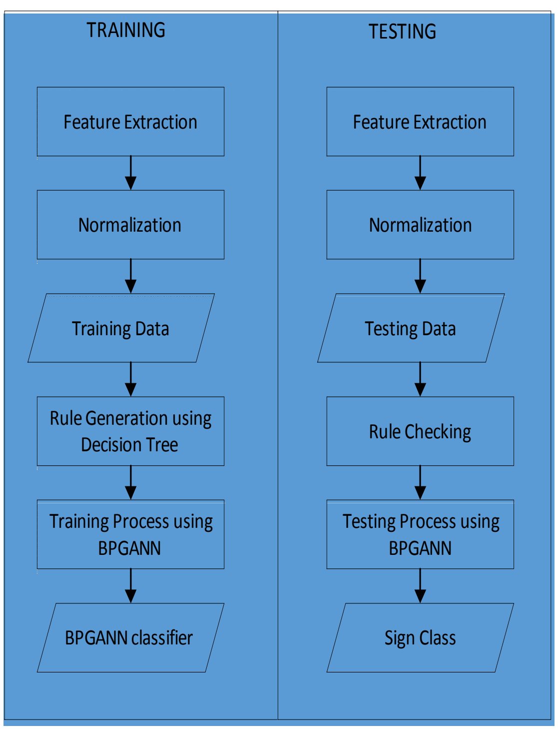

Figure. 2 System Architecture

3.1 Feature Extraction

In this research, the features used to define sign language are from skeleton joints that captured by Kinect. There are 28 skeletons joint that can be captured by Kinect as shown in Figure. 7. However, in this research we used nine skeleton joints to be processed for feature extraction. Those nine skeleton joints are:

1. Right hand (HL)
2. Right wrist (WL)
3. Right elbow (EL)
4. Right shoulder (SL)
5. Center of shoulders (SC)
6. Left shoulder (SR)
7. Left elbow (ER)
8. Left wrist (WR)
9. Left hand (HR)

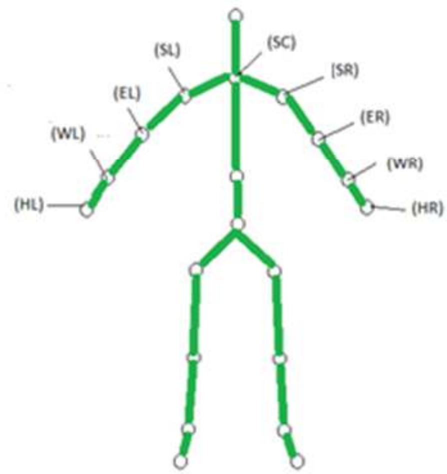

Figure. 3 Skeleton Joints

Table 2 Features from Skeleton Joints

Vector3 (x,y,z)	Angle (rad)	Distance (float)
ER → SR	∠ SC - SR – ER	
WR → ER	∠ SR - ER - WR	
HR → WR	∠ ER - WR - HR	
EL → SL	∠ SC - SL – EL	HR - HL
WL → EL	∠ SL - EL – WL	
HL → WL	∠ EL - WL – HL	
HL → HR		

From these 9 skeletons, we extract features based on vector, angle, and distance as shown in Table 2. (Totally 28 features (7x3 features from vector value, 6 features from angle, and one feature from distance) were extracted.

3.2 Feature Normalization

Feature normalization, feature scaling process, was implemented by using (1). This process was done in order to make these features in the certain range therefore features are more proportional. The maximum and minimum value of each feature was obtained from the maximum and minimum value of training data. The minimum and maximum value

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

for vector features are minus one (-1) and positive one (+1). The minimum value of angle features is zero and maximum value is 3.14.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (1)$$

Where:

X' : feature value after normalization
X : feature value before normalization
X_{min}: the minimum value of a feature
X_{max}: the maximum value of a feature

3.3 Rule Generation using Decision Tree

Rules were used to minimize the effect of overfitting in Neural Network. In this research, normalized training data were classified using Decision Tree. WEKA open source application was used to generate the tree classifier. The nodes in root and level one of tree classifier were taken as rules.

In this research, the tree classifier yielded from WEKA was shown in Figure. 8. Only root node and node in level one that were used as rules like in Figure. 9. The rules that generated from the Figure. 9 was shown in Table 3. Then, these rules were implemented in training and testing process.

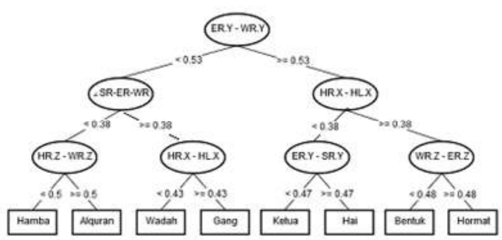

Figure. 8 Decision Tree Result

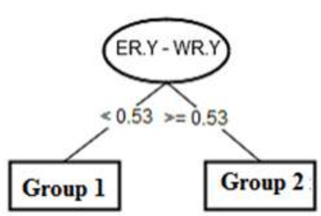

Figure. 9 The Root Node And The First Level Node

Table 3 The Rule In This Proposed Research

No	Rule	Member
1	$ER.Y - WR.Y < 0.53 \rightarrow Group\ 1$	*Hamba, Alqur'an, Wadah, Gang*
2	$ER.Y - WR.Y \geq 0.53 \rightarrow Group\ 2$	*Ketua, Hai, Bentuk, Hormat*

3.4 Training and Testing using Neural Network

After the rules created. The training data was grouped based on the rules. Each group was trained by using BPGANN and produced one classifier. In this research, the rules categorized training data into two groups therefore the classifier produced using BPGANN was two classifiers. The process of training was shown in Figure. 6.

In testing, before classified by using classifier, each testing data was tested by rules. The result of rules will determine whether the testing data will be classified using Classifier 1 or 2. The testing process was shown in Figure. 10.

Figure. 10 Testing Process

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4. EXPERIMENTS

For experiments, we used 8 ISL vocabularies: *alqur'an*, *bentuk*, *gang*, *hai*, *hamba*, *hormat*, *ketua*, and *wadah*. The gestures of these sign language were shown in Figure. 11.

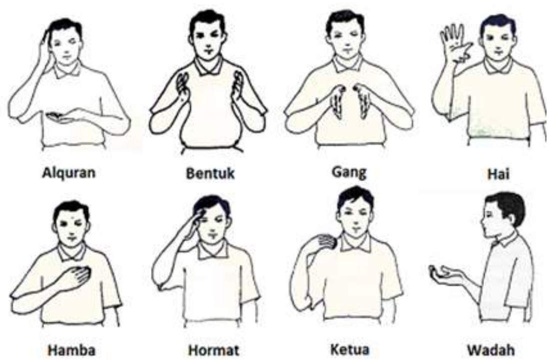

Figure. 4 Experiments gesture

For training process, the dataset were collected from two people. The number of dataset were 240 for 8 sign languages with 30 gestures for each sign language. Testing was done in real time by these two people. Each person performs five gestures for each sign language. In total, each sign language was tested ten times. The result of the testing was shown in confusion matrix in Table 4. From that table, we can see that the accuracy of this system is reach 96%. The misclassified gestures only occurs in 'Hai' and 'Hamba' gestures.

However, the ISL vocabularies that we used in this experiment only vocabularies with static gesture. A further research is required to recognize ISL vocabularies with dynamic gestures so that a full ISL recognition system can be developed. In this system, the training and testing were done by the same people. If the testing is done by different people, the accuracy may decrease. To overcome this problem, the training data should be enriched with data from many different people.

5. CONCLUSION

This research aims to contribute to the improvement of communication system of people with speech and/or hearing disabilities. The recognition module to translate ISL (Indonesian Sign Language) to text, although evaluated with small vocabulary, presents very promising result with accuracy 96%. However, word that classified in this experiment only word with static gestures, a further research is needed in order to cover not only word with static gestures but also word with dinamic gestures. It is also necessary to enrich the training data with data from many different people so that the application can be used for all people.

ACKNOWLEDGMENT

This research was supported by PNBP research grant for "Penelitian Pemula" in 2016 from Institut Teknologi Sepuluh Nopember Surabaya.

Table 3 The Confusion Matrix from Experiment

		Class Label							
System' Prediction		*Alqur'an*	*Bentuk*	*Gang*	*Hai*	*Hamba*	*Hormat*	*Ketua*	*Wadah*
	Alquran	10							
	Bentuk		10						
	Gang			10					
	Hai				8				
	Hamba					9			
	Hormat				2		10		
	Ketua							10	
	Wadah					1			10

REFRENCES:

[1] E. Rakun, M. Andriani, I. W. Wiprayoga, K. Danniswara, and A. Tjandra, "Combining depth image and skeleton data from Kinect for recognizing words in the sign system for Indonesian language (SIBI [Sistem Isyarat Bahasa Indonesia])," in 2013 International Conference on Advanced Computer Science and Information Systems (ICACSIS), 2013, pp. 387–392.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[2] Sutarman, M. A. Majid, and J. M. Zain, "A Review on the Development of Indonesian Sign Language Recognition System," J. Comput. Sci., vol. 9, no. 11, pp. 1496–1505, Nov. 2013.

[3] N. Sugianto and F. Samopa, "Analisa Manfaat Dan Penerimaan Terhadap Implementasi Bahasa Isyarat Indonesia Pada Latar Belakang Komplek Menggunakan Kinect Dan Jaringan Syaraf Tiruan," JUISI, vol. 01, no. 01, Feb. 2015.

[4] W. N. Khotimah, R. A. Saputra, N. Suciati, and R. R. Hariadi, "Comparison between Back Propagation Neural Network and Genetic Algorithm Back Propagation Neural Network for Sign Language Recognition," presented at the The International Conference on Science and Technology 2015 (ICST2015), Yogyakarta.

[5] M. Zbakh, Z. Haddad, and J. Lopez Krahe, "An online reversed French Sign Language dictionary based on a learning approach for signs classification," Pattern Recognit. Lett., vol. 67, Part 1, pp. 28–38, Dec. 2015.

[6] C. Velasco, G. Weber, J. Barroso, Y. Mohamad, H. Paredes, P. Escudeiro, N. Escudeiro, R. Reis, J. Lopes, M. Norberto, A. B. Baltasar, M. Barbosa, and J. Bidarra, "Virtual Sign – A Real Time Bidirectional Translator of Portuguese Sign Language," in Procedia Computer Science, 2015, vol. 67, pp. 252–262.

[7] S. G. Moreira Almeida, F. G. Guimarães, and J. Arturo Ramírez, "Feature extraction in Brazilian Sign Language Recognition based on phonological structure and using RGB-D sensors," Expert Syst. Appl., vol. 41, no. 16, pp. 7259–7271, Nov. 2014.

[8] D. Kaur, R. Bedi, and S. K. Gupta, "Review Of Decision Tree Data Mining Algorithms: ID3 And C4.5," presented at the International Conference on Information Technology and Computer Science, 2015.

[9] D.-A. CHIANG, W. CHEN, Y.-F. WANG, and L.-J. HWANG, "Rule generation using decision tree," J. Inf. Sci. Eng., vol. 17, no. 2001, pp. 325–339.

[10] S. Wang, S. Yin, and M. Jiang, "Hybrid Neural Network Based on GA-BP for Personal Credit Scoring," in 2008 Fourth International Conference on Natural Computation, 2008, vol. 3, pp. 209–214.