Automatic detection of ischemic stroke based on scaling exponent electroencephalogram using extreme learning machine

H A Adhi[1], S K Wijaya[2], Prawito[2], C Badri[1] and M Rezal[1]
[1]Biomedical Engineering, Master Program, Universitas Indonesia, Jakarta, Indonesia
[2] Faculty of Mathematics and Natural Sciences, Universitas Indonesia, West Java, Indonesia

E-mail: habbi.ananto@ui.ac.id, skwijaya@sci.ui.ac.id, prawito@sci.ui.ac.id, cholidbadri@yahoo.co.id and muhammad.rezal@ui.ac.id

Abstract. Stroke is one of cerebrovascular diseases caused by the obstruction of blood flow to the brain. Stroke becomes the leading cause of death in Indonesia and the second in the world. Stroke also causes of the disability. Ischemic stroke accounts for most of all stroke cases. Obstruction of blood flow can cause tissue damage which results the electrical changes in the brain that can be observed through the electroencephalogram (EEG). In this study, we presented the results of automatic detection of ischemic stroke and normal subjects based on the scaling exponent EEG obtained through detrended fluctuation analysis (DFA) using extreme learning machine (ELM) as the classifier. The signal processing was performed with 18 channels of EEG in the range of 0-30 Hz. Scaling exponents of the subjects were used as the input for ELM to classify the ischemic stroke. The performance of detection was observed by the value of accuracy, sensitivity and specificity. The result showed, performance of the proposed method to classify the ischemic stroke was 84 % for accuracy, 82 % for sensitivity and 87 % for specificity with 120 hidden neurons and sine as the activation function of ELM.

1. Introduction

Stroke is one of cerebrovascular diseases which are caused by obstruction of blood flow to the brain [1]. The obstruction causes brain tissue damage. Based on data from WHO in 2012, stroke is the first leading cause of death in Indonesia and the second leading cause in the world [2]. Cardiovascular disease, stroke, and diabetes had become the biggest cause of disability in Indonesia [3] and in some developed countries. Ischemic stroke accounts for most of all stroke cases, up to 87% [4]. Computed Tomography (CT)-Scan was often used to support the diagnosis of stroke. While in ischemic stroke has been established that Diffused / Perfusion-Weighted (D/P-W) MRI as the gold standard in the examination of ischemic stroke [5]. D/P-W MRI is expensive and has limited availability [6]. However, CT scan or MRI, included D/P-W MRI, may take several hours to obtain examination reports while the initial aid is quite limited period for stroke patients [7].

Electroencephalography is a method to record electrical activity of the brain. Nowadays, electroencephalography widely used for evaluation of seizures and epilepsy [8]. In ischemic stroke, the role of electroencephalography as an indicator of disturbed function (locally, regionally, or diffusely) in the event of stroke has not always been duly appreciated [9] especially when CT-Scan and MRI is available. As outlined before that CT-Scan and MRI has limited availability, many research tried to

optimize electroencephalography in ischemic stroke detection. Electroencephalography offers a cheaper cost, widely available [10], continuous monitoring [11] and good temporal resolution [12]. Some of the advantages are the basis for optimizing electroencephalography to detect acute ischemic stroke [13]. Tissue damage in ischemic stroke patients causes the changes in electrical activity of brain. Those changes can be known through EEG power [14], variability [15, 16] and functional connectivity of the cortex [17]. Detrended fluctuations analysis (DFA) is one of method for analysing variability of EEG. The result of the DFA is the fluctuation of the EEG indicated by the scaling exponents.

Automatic detection is a technique to analyse signal quantitatively using a computer. Automatic detection offers great promise to assist in interpretation of EEG studies [18]. In automatic detection, there are two main considerations they are feature extraction techniques and classification techniques. Machine learning method is commonly applied in classification problem. Machine learning method based on artificial neural networks were applied in stroke classification problems such as extreme learning machine (ELM) [19] and 1D convolutional neural network (1DCNN) [20]. In this study, we presented the result of automatic detection of ischemic stroke based on the scaling exponent of EEG using ELM.

2. Material and methods

2.1. Data collection

This study was conducted and approved by the Ethics Committees of National Centre for Brain Hospital (RS PON), Jakarta. Until this study was conducted we obtain data from 31 ischemic stroke patients (age 57,6 ± 7,9 years) and 30 normal subjects (age 54,2 ± 8,1 years). The number of normal subjects selected to balance with the amount of ischemic stroke patients data. The EEG recording was conducted for 30 minutes with 32 channels and the electrode were placed with 10-20 international system. The recording was performed using Biologic System and Xltek Netlink EEG 32U with sampling frequency is 256 and 512 Hz and stored in a file in the format of edf (European Data Format).

2.2. Proposed method

The proposed method shows in figure 1. It consists of two stages namely training and testing stages. The method starts from edf file acquisition, to be read as the EEG data. Then, it is followed by the pre-processing; selecting 18 channels of EEG (FP1, F3, F7, C3, T3, P3, T5, T1, O1, FP2, F4, F8, C4, T4, P4, T6, T2 and O2) and bandpass filter the signal in the main frequencies of interest between 0-30 Hz. Channels are selected by the availability of its pair in both hemispheres. The purpose of feature extraction is to get the features of EEG, there are 36 scaling exponents obtained using DFA, two moments and two normalized moments. These features are used as input for ELM algorithm. As outlined before, system divided into two stages there are training and testing. Training stage is intended to get the optimum parameters (input weights, bias hidden neurons and output weights) for the classification. Meanwhile, testing stage is intended to validate the performance of ELM, with parameter

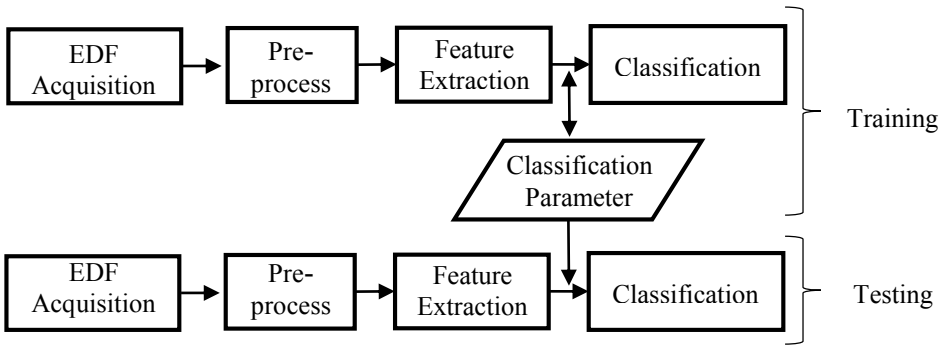

Figure 1. Block diagram of proposed method

was obtained from training stage, as a classifier. Performance of the proposed method were performed using the *K*-fold cross validation where $K = 10$. Performance of detection was evaluated based on accuracy, sensitivity and specificity.

2.3. Detrended fluctuations analysis

DFA is one of method to analyse the variability of a biomedical signal [21]. DFA is represented by two scaling exponent (α_1 and α_2) [15]. Scaling exponent is used as an indicator of EEG fluctuations that obtained from the slope of a line formed on plots of ln (*F(k)*) versus ln (*k*). In the spatial domain, the moments of the distributions of α_1 and α_2 summarize the mean and variability across the scalp[15]. Application of DFA with normalize moment follows:

i) Divide the EEG signal, *x(t),* into *B* segment with *k* data point where $1 \leq t \leq T$ and *T* is the number of data points. Floor is function to round the value of (*T/B*) to nearest integer less than or equal. *k* obtained using equation (1).

$$k = floor\ (T/B) \tag{1}$$

ii) Get the trend of signal *x(t)* on the segment *b*, $\bar{x}_b(t)$, equation (2). Linear-fit is function to get linear line with one independent and one dependent variable from a group of data point in segment *b*.

$$\bar{x}_b(t) = linear - fit\ [x(t)]\ ; \quad (b\text{-}1)\ k < t < bk \tag{2}$$

iii) Calculate the average fluctuation *x(t)* using equation (3).

$$F^2(k) = \frac{1}{kB}\sum_{t=1}^{kB}[x(t) - \bar{x}_b(t)]^2 \tag{3}$$

iv) Get the scaling exponent using equation (4), where *i* is the order of the scaling exponent, *q* is a positive integer and *N* is the number of EEG channel. Then, calculate the moment of the scaling exponent, $G^{(i)}$, using equation (5).

$$\ln(F(k)) \propto \alpha \ln(k) \tag{4}$$

$$G_q^{(i)} = \frac{1}{N}\sum_{j=1}^{N} \alpha_{i,j}^q \tag{5}$$

v) Calculate normalized moments, $M_q^{(i)}$, using equation (6).

$$M_q^{(i)} = \frac{G_q^{(i)}}{\left(G_1^{(i)}\right)^q} \tag{6}$$

2.4. Extreme learning machine

Extreme Learning Machine (ELM) was propounded for single layer hidden feedforward neural networks (SLFNs). ELM has good generalization ability, quick to perform learning and a simple structure [22]. ELM architecture consists of input layer, hidden layer and output layer. Weights input (w_{nk}) connecting *n*-th node in the input layer and the node-*k* in the hidden layer. While the weights output (β_k) connecting node-*k* in the hidden layer and output layer. For each sample of training data (x_i, t_i) where x_i is the value at the input node and t_i is the expected target of a sample. Activation function, *g(x)*, in the standard SLFNs with *K* hidden node is modelled in equation (7), where b_j is a bias in the *k*-th hidden neuron. Equation (8) is used to obtain the target value where the input weights and biases hidden neurons randomly determined. Output weights, β, modelled on the equation (10), $H^{\dagger}$ is the Moore-Penrose's generalized inverse matrix of the matrix *H*.

In artificial neural network, a parameter in the hidden layer neurons to calculate the value from input layer is called activation function. Thus *g(x)* on ELM will be recalculated using this activation function to pass the value to the output layer. Some activation functions which are often used in the artificial neural network are sigmoid, sine and radial basis. The activation function and the number of neurons in the hidden layer will affect the performance of ELM.

$$\sum_{j=1}^{K} \beta_j g\left(w_{jk} \cdot x_n + b_j\right) = t_n \tag{7}$$

$$H\beta = T \tag{8}$$

where:

$$H(w_1, \dots, w_K, b_i, \dots, b_K, x_1, \dots, x_M) = \begin{bmatrix} g(w_1 \cdot x_1 + b_1) & \cdots & g(w_K \cdot x_1 + b_K) \\ \vdots & \cdots & \vdots \\ g(w_1 \cdot x_M + b_1) & \cdots & g(w_K \cdot x_M + b_K) \end{bmatrix}; \tag{9}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1^T \\ \beta_2^T \\ \vdots \\ \beta_K^T \end{bmatrix}_{K \times l}, \boldsymbol{T} = \begin{bmatrix} t_1^T \\ t_2^T \\ \vdots \\ t_K^T \end{bmatrix}_{M \times l}$$

$$\beta = H^{\dagger} T \tag{10}$$

2.5. Evaluation Matrices

To evaluate the performance of detection, we used the value of accuracy, sensitivity and specificity [23]. Accuracy or success rate of detection is obtained using equation (11), where false positive (FP) is the number the illness subject detected as the normal subjects, false negative (FN) is the number normal subjects detected as the illness subject. Sensitivity or true positive (TP) rate is used to measure the ability of detecting the illness subjects as correctly the illness subjects. Sensitivity is obtained using equation (12). Specificity or true negative (TN) rate is used to measure the ability of detecting the normal subjects as correctly the normal subjects. Specificity was obtained using the equation (13).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{11}$$

$$Sensitivity = \frac{TP}{TP+FN} \tag{12}$$

$$Specificity = \frac{TN}{TN+FP} \tag{13}$$

3. Experiment results and discussion

3.1. Scaling exponent analysis

In this study, we used two parameters *k* on DFA, because the data has two sampling frequency. The range $3 \leq k \leq 500$ was applied to data with sampling frequency of 256 Hz and $6 \leq k \leq 1000$ was applied to data with sampling frequency of 512 Hz. Two scaling exponents are obtained from the slope of a line formed on plots of ln (*F(k)*) versus ln (*k*) from two regions. First region, to obtained α_1, located between $1 < \ln(k) < 3.5$ for data with sampling frequency 512 Hz and $1 < \ln(k) < 3$ for data with sampling frequency 256 Hz. Second region, to obtained α_2, is located between $4.45 < \ln(k) < 5.5$ for both data.

Figure 2 shows the scatter plot of the moment of scaling exponents ($G_2^{(1)}$ and $G_2^{(2)}$) in ischemic stroke patients and normal subjects. The results showed that ischemic stroke patients generally have $G_2^{(1)}$ and

$G_2^{(2)}$ higher than the normal subject, table 1, and there is correspond to the previous study in [7]. Figure 2 shows that moment of scaling exponent cannot distinguish normal subject and ischemic stroke patient clearly. In [7], normalize moment of scaling exponent can distinguish ischemic stroke patients and normal subjects clearly and shows that ischemic stroke patient have lower normalize moment than normal subject. The result of normalize moment $M_2^{(1)}$and $M_2^{(2)}$ in this study, figure 3, shows that normalize moment still cannot be distinguish between normal subject and ischemic stroke patients clearly. Normal subjects have similar average normalize moment with ischemic stroke but ischemic stroke patients have higher value of standard deviation, table 1. The result also corresponded to the other previous study [24] which generally indicates the acute stroke thalamic ischemic patient has a higher EEG complexity than the normal subjects.

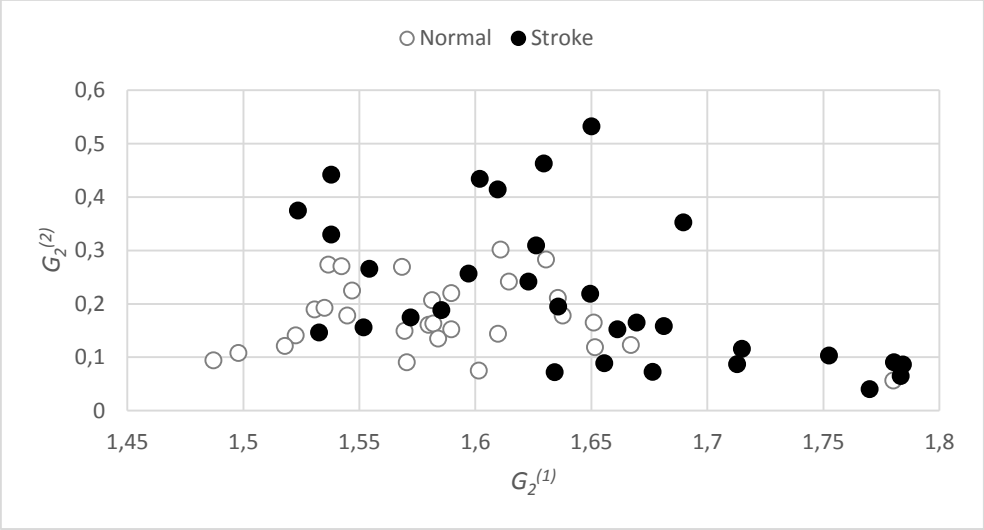

Figure 2. Moment ($G_2^{(1)}$and $G_2^{(2)}$) plot for ischemic stroke patients and normal subjects.

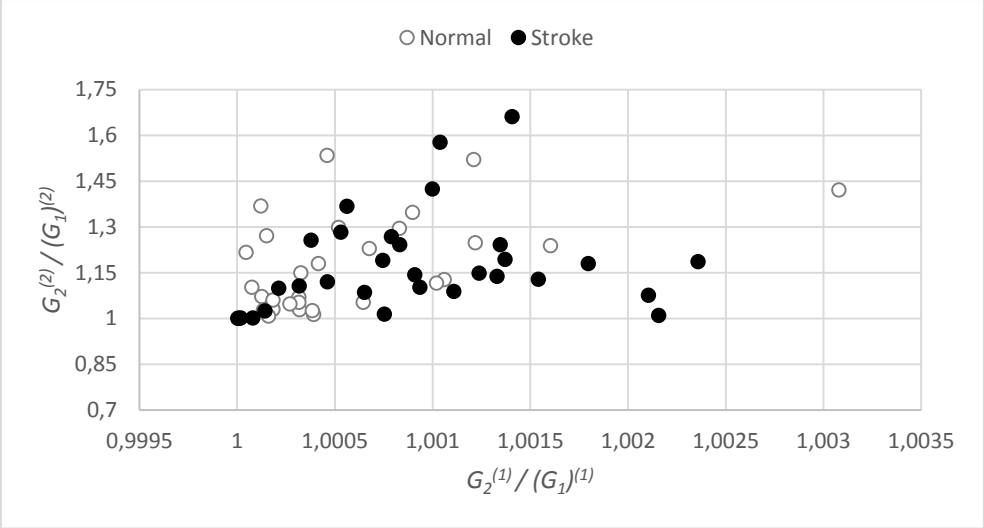

Figure 3. Normalize moment ($M_2^{(1)}$ and $M_2^{(2)}$) plot for ischemic stroke patients and normal subjects.

Table 1. Average moment and normalized moment for each scaling exponent.

Subject	$G_2^{(1)}$	$G_2^{(2)}$	$M_2^{(1)}$	$M_2^{(2)}$
Normal	1.59 ± 0.06	0.17 ± 0.07	1.00 ± 0.00	1.17 ± 0.15
Stroke	**1.64 ± 0.08**	**0.22 ± 0.14**	1.00 ± 0.00	1.17 ± 0.16

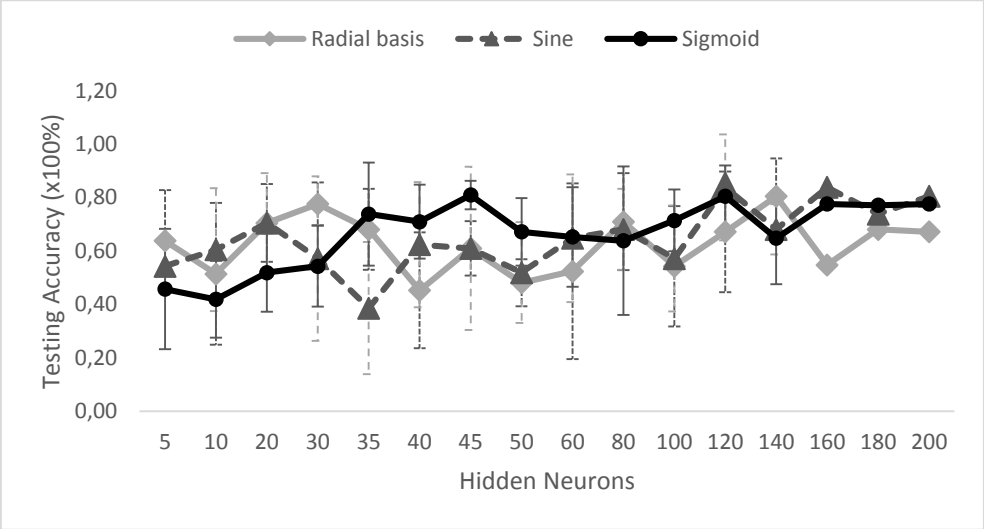

Figure 4. Accuracy of the system using scaling exponent features in testing stage.

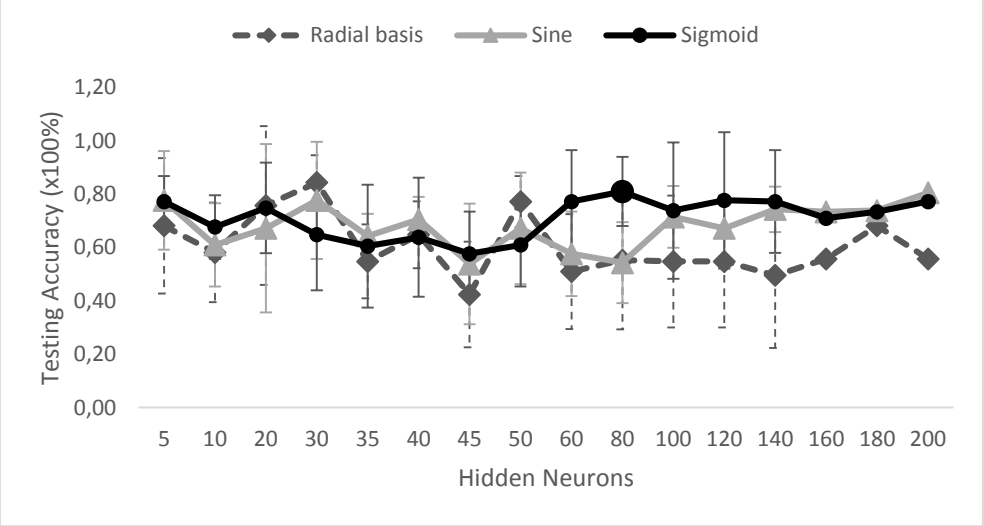

Figure 5. Accuracy of the system using combined features in testing stage.

3.2. Feature and performance analysis

In this study, we divided the features into two groups of features they are scaling exponent features and combined features. The group of scaling exponent features consists of 36 scaling exponents from 18 channels of EEG. Meanwhile, the combined features are the combination of $G_2^{(1)}$, $G_2^{(2)}$, $M_2^{(1)}$, $M_2^{(2)}$and scaling exponent features and will be consist 40 value for each subject. Experiments in those features

was conducted using three parameters of activation function (sigmoid, sine and radial basis) and the number of neurons in the hidden layer were varied. The application of some parameter settings is intended to get the optimum settings in detection. Accuracy used as main parameter to measure performance of proposed method.

The accuracy of this experiments with some parameter settings, in best case, are presented in figure 4 and 5. The results obtained in testing stage shows that the addition of neurons in hidden layer was not always followed by the increase of ELM performance. Experiments using three activation function showed, no activation function is dominant over the others. It is showed that in scaling feature using sine activation function get the best accuracy, but in combined feature the best accuracy showed with sigmoid activation function.

Table 2. Best performance for each group of features and functions activation in testing stage.

	Scaling exponent feature			Combined feature		
Parameters	Sigmoid	Sine	Radial Basis	Sigmoid	Sine	Radial Basis
Accuracy	0.81± 0.05	**0.84 ± 0.12***	0.80± 0.14	0.81 ± 0.13	0.80 ± 0.08	0.77 ± 0.10
Sensitivity	0.70± 0.18	**0.82 ± 0.17**	0.68± 0.24	0.70 ± 0.30	0.88 ± 0.16	0.68± 0.24
Specificity	0.93 ± 0.15	**0.87± 0.30**	0.93± 0.15	0.93 ± 0.15	0.73 ± 0.28	0.87 ± 0.18

*Optimum parameter settings.

Table 3. Best performance for each optimum parameter in the training stage.

	Scaling Feature			Combined Feature		
Parameters	Sigmoid	Sine	Radial Basis	Sigmoid	Sine	Radial Basis
Accuracy	1	1	1	1	1	1
Sensitivity	1	1	1	1	1	1
Specificity	1	1	1	1	1	1

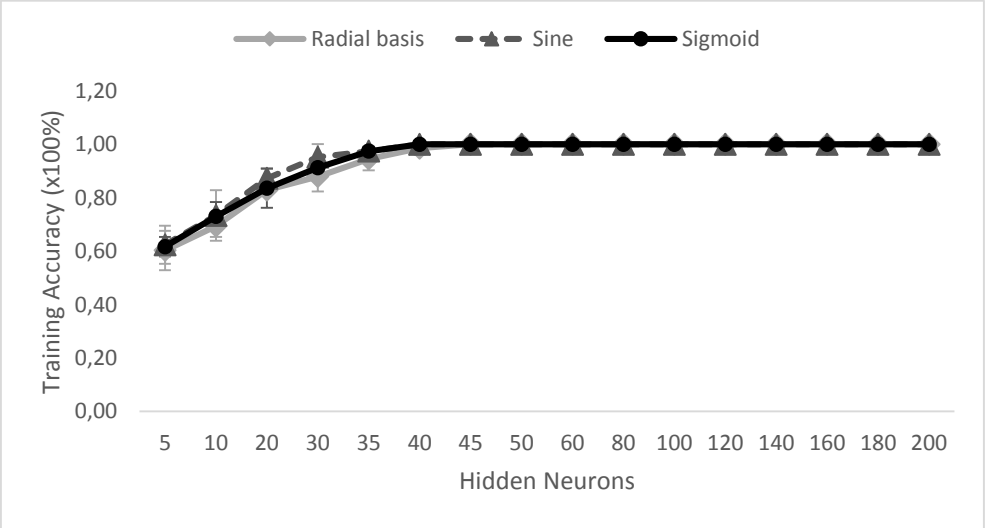

Figure 6. Accuracy using scaling exponent features in training stage.

Table 2 shows the summary of the best performance in testing stage obtained from several experiments in this study. The optimum parameters in this study was obtained with 84% as the highest accuracy using sine activation function, 120 neurons and scaling exponent features. These parameter settings gave the performance 82% for sensitivity and 87% for specificity in testing stage. The system performance indicates that the proposed method performed better to detect the normal subject than ischemic stroke patient. The reason for this situation might be because in this study there were several ischemic stroke patients in the mild ischemic stroke (NIHSS ≤ 4) thus more difficult to distinguish from normal subjects. Table 3 shows the summary of performance in the training stage by the same parameters with table 2. In the previous discussion, has been outlined that the addition of neurons is not always followed by the increase of accuracy in testing stage, but in this training stage, the addition of neurons was able to improve training accuracy as shown in figure 6. To reach 100 % accuracy in the training stage the ELM need minimum 45 neurons in the hidden layer. In the training stage, 100 % accuracy can be interpreted that the ELM was able to form a good classification model.

4. Conclusion

The results of this study shows EEG in the ischemic stroke subjects have a higher scaling exponent than the normal subjects. The best performance was achieved for the detection using sine activation function, 120 neurons and scaling exponent features as the parameters. Using these parameters, we obtained 84% for accuracy, 82% for sensitivity and 87% of specificity. In our future work, we will further explore other method of feature extraction and classifier to optimize and get better ischemic stroke detection system.

Acknowledgments

This study was supported by the Department of Education and Culture of Republic of Indonesia (0402/UN2.R12/HKP.05.00/2015) and Universitas Indonesia.

References

[1] Brass L M 1992 Stroke *Yale University School of Medicine Heart Book* ed Zaret B L, Moser M and Cohen L S (New York: Hearst Books) chapter 18 pp 215–233
[2] World Health Organization 2014 *The top 10 causes of death*
[3] World Health Organization 2012 *Indonesia: WHO statistical profile*
[4] American Stroke Association 2016 *Ischemic Strokes (Clots)*
[5] Birenbaum D, Bancroft LW and Felsberg G J 2011 Imaging in Acute Stroke *West J Emerg Med* **12** 67–76
[6] Grunwald I Q, Fassbender K and Wakhloo A K 2012 *How to set up an acute stroke service* (Berlin: Springer)
[7] Hwa R C and Ferree T C 2004 Stroke detection based on the scaling properties of human EEG *Phys A Stat Mech its Appl* **338** 246–254
[8] Acharya U R, Vinitha S S, Swapna G, Martis R J, Suri J S 2013 Automated EEG analysis of epilepsy: A review *Knowledge-Based Syst* **45** 147–165
[9] Niedermeyer E 2005 Cerebrovascular Disorders and EEG *Electroencephalography: Basic Principle, Clinical Applicationand Related fields,* vol 5 ed Niedermeyer E and Silva F L (Philadelphia: Lippincott Williams & Wilkins) chapter 17 pp 339–362
[10] Sheorajpanday R V A, Nagels G, Weeren A J T M, Surgeloose D De and Deyn P P De 2010 Additional value of quantitative EEG in acute anterior circulation syndrome of presumed ischemic origin *Clin Neurophysiol* **121** 1719–1725
[11] Schneider A L and Jordan K G 2005 Regional attenuation without delta (RAWOD): A distinctive EEG pattern that can aid in the diagnosis and management of severe acute ischemic stroke *Am J Electroneurodiagnostic Technol* **45** 102–117
[12] Rijsdijk M, Leijten F S S and Slooter A J C 2008 Continuous EEG monitoring in the Intensive Care Unit *Netherlands J Crit Care* **12** 157–162

[13] Wijaya S K, Badri C, Misbach J, Soemardi T P and Sutarmo V 2014 Optimization Electroencephalography (EEG) for Detecting Acute Ischemic Stroke: A Preliminary Study *Int. Conf. Biomed. Eng. Technol. Appl. 2014 (ICBETA 2014)* pp 165–169
[14] Foreman B and Claassen J 2012 Quantitative EEG for the detection of brain ischemia *Annual Update in Intensive Care and Emergency Medicine* ed Vincent J (Berlin: Springer) pp 746–758
[15] Ferree T C and Hwa R C 2005 Electrophysiological measures of acute cerebral ischaemia *Phys Med Biol* **50** 3927–3939
[16] Seely A J and Macklem P T 2004 Complex systems and the technology of variability analysis *Crit Care* **8** R367–R384
[17] Wu W, Jin Z, Qiu Y, Zhu Y, Li Y, Tong S 2009 Influences of bilateral ischemic stroke on the cortical synchronization *2009 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE* pp 2216–2219
[18] Hogan R E 2011 Automated EEG detection algorithms and clinical semiology in epilepsy: Importance of correlations *Epilepsy Behav* **22** S4–S6
[19] Rahma O N, Wijaya S K, Prawito and Badri C 2016 Electroencephalogram Analysis With Wavelet Transform and Neural Network As a Tool for Acute *IASTEM -The 9th Int. Conf. Medical, Biol. Pharm. Sci.* pp 108–113
[20] Giri E P, Fanany M I and Arymurthy A M 2016 Ischemic Stroke Identification Based on EEG and EOG using 1D Convolutional Neural Network and Batch Normalization
[21] Golińska A K 2012 Detrended Fluctuation Analysis (DFA) in biomedical signal processing: selected examples *Stud Logic, Gramm Rheotic* **29** 107–116
[22] Huang G-B, Zhu Q and Siew C 2006 Extreme learning machine: Theory and applications *Neurocomputing* **70** 489–501
[23] Castellino R A 2005 Computer aided detection (CAD): an overview *Cancer Imaging* **5** 17–19
[24] Liu S, Guo J, Meng J, Wang Z, Yao Y, Yang J, Qi H and Ming D 2016 Abnormal EEG Complexity and Functional Connectivity of Brain in Patients with Acute Thalamic Ischemic Stroke *Comput Math Methods Med* **2016** 1–9