Classification of Post-Stroke EEG Signal Using Genetic Algorithm and Recurrent Neural Networks

Ella Wahyu Guntari
Department of Informatics
Universitas Jenderal Achmad Yani
Cimahi, Indonesia
ella.guntari11@gmail.com

Esmeralda Contessa Djamal*
Department of Informatics
Universitas Jenderal Achmad Yani
Cimahi, Indonesia
***esmeralda.contessa@lecture.unjani.ac.id**

Fikri Nugraha
Department of Informatics
Universitas Jenderal Achmad Yani
Cimahi, Indonesia
fikri.nugraha@student.unjani.ac.id

Sandi Lesmana Liemanjaya Liem
Faculty of Medicine
Universitas Jenderal Achmad Yani
Cimahi, Indonesia
sandilesmana28@yahoo.com

Abstract— **Stroke is caused by a sudden burst of blood vessels in the brain, causing speech difficulties, memory loss, and also paralysis. The identification of electrical activity in the brain of post-stroke patients from EEG signals is an attempt to evaluate rehabilitation. EEG signal recording involves multiple channels with overlapping information. Therefore the importance of channel optimization is to reduce processing time and reduce the computational burden. Besides, that channel optimization can have an overfitting effect due to excessive utilization of EEG channels. This paper proposed the optimization of EEG channels for the identification of post-stroke patients using Genetic Algorithms and Recurrent Neural Networks. Data was taken from 75 subjects with a recording duration of 180 seconds in a seated state. The data was segmented and extracted using Wavelet to get the frequency of the Alpha, Theta, Mu, Delta, and Amplitude changes. The next step is the channel optimization process using Genetic Algorithms. The method applied to get a combination of channels that qualifies. Then, the EEG signal identification proceeds of the optimization of the channels used Recurrent Neural Network. The result showed that applying the Genetic Algorithm afforded 12 channels configuration with 90.00% of accuracy; meanwhile, used all channels gave a 72.22% result. Therefore, channel optimization is essential to reduce redundancy and increase recognition.**

Keywords—Post-Stroke classification; Optimization of EEG Signal; Recurrent Neural Networks; Genetic Algorithm, Wavelet.

I. INTRODUCTION

Based on data from the World Health Organization (WHO), there are 17 million new stroke cases each year. In Indonesia, more than 65% of stroke sufferers experience disabilities. Therefore medical rehabilitation becomes very important to restore post-stroke patients to be able to carry out activities of daily life. The process requires evaluation and monitoring.

One device that can identify post-stroke patients is the Electroencephalogram (EEG). EEG signals are information on electrical activity in the human brain and can be used in real-time. EEG is non-invasive, low risk to the patient, standard, and real-time output if we compared with the use of CT-Scan and MRI. Previous research, EEG signals can be identified based on several variables such as emotional [1], fatigue [2], neuropsychological [3], epilepsy [4], alcoholic's person [5], characteristics of post-stroke EEG signal [6], significant variables of EEG signals in post-stroke patients [7], classification emotions of patient post-stroke [8], and identifying ischemic stroke [9].

EEG signals have low amplitude and irregular patterns, so extraction is needed to minimize noise and obtain signal components according to the variable being reviewed. In previous research often used frequency filters. Some studies often used Fast Fourier Transform (FFT), and Wavelet transforms to extract frequency. Feature extraction using Fourier can minimize noise well, but this method does not support time information simultaneously [10]. At the same time, Wavelet changes are suitable for non-stationary signal types such as EEG signals because they can represent the time and frequency information of a signal simultaneously [10]. Some previous studies used Wavelet extraction to identify focus levels [11], find out significant variables of EEG signals in post-stroke patients [7], emotional classification of post-stroke patients [8], and classification of stroke levels [9], and find out the relationship between Alpha, Beta and Delta waves with stroke levels [12].

EEG devices have various channels of electrical signals recording in the brain. Using all of the EEG channels requires much computational time and allows overfitting because of excessive use of the EEG channel, or the channel can record the same signal. Based on previous research, the use of EEG channels can be reduced significantly without reducing accuracy [13]. However, manual channel selection does not always produce optimal accuracy. Therefore a channel selection method is needed to remove irrelevant channels or channels that record the same signal. Previous studies used Genetic Algorithms to improve correctness with EEG selection, such as classification of two movements with BCI [14], biometric identification [15], channel selection for P300 with motor imagery [16], character identification [17], and alcoholic detection [18].

Deep Learning is a part of machine learning in artificial intelligence. This method has good performance in terms of more computational computing and allows for a more complex feature extraction layer, processes large data sets, and can learn without supervision from unstructured data unlabeled. One method in Deep Learning is the Recurrent Neural Network (RNN). RNN is suitable in the signal processing or sequential data application. Sequential data processing using RNN needs to be classified information that will be processed and not processed (forgotten), so that memory usage is more realistic, such as using Long Short Term Memory (LSTM). Previous studies of LSTM RNN were used for stroke identification [19], epilepsy [4], emotion classification [20][21], sleep stages classification [22], motor imagery classification [23], neuropsychological identification [24], and motor imagery in stroke patients [25].

This study proposed a reduction in the number of EEG channels for the identification of post-stroke patients. The EEG signal from each channel was extracted first using Wavelet to obtain Alpha, Theta, Mu, and Delta waves from each data segment. The four features and amplitude changes of all channels become features that are processed by Genetic Algorithms and Recurrent Neural Networks (RNN) to classify signals in one of three classes, namely "No Stroke", "Minor Stroke" and "Moderate Stroke". Genetic Algorithms will set the optimal combination based on the value of losses throughout the iteration. We used 75 subjects consisting of stroke patients with two class conditions (moderate stroke and minor stroke), and a comparison of non-stroke.

II. Methods

Model-identification of post-stroke EEG signals with channel optimization is shown in Fig.1. The EEG signal was 120 seconds long, extracted using Wavelet so that it gets a wave according to its segment (Fig. 2). Then the training is carried out, with channel optimization using Genetic Algorithms and Recurrent Neural Networks.

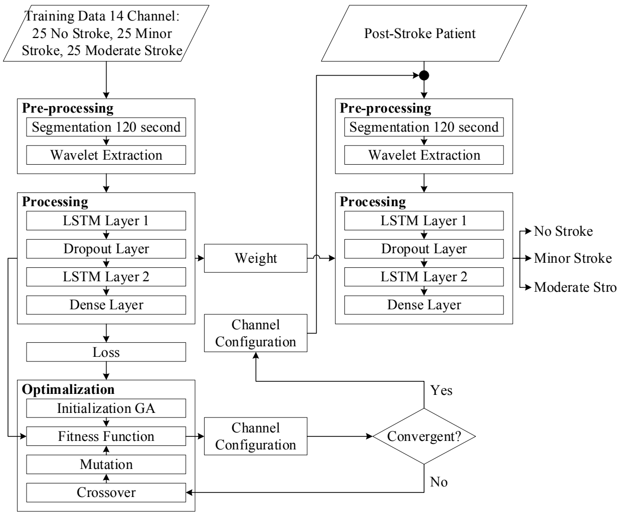

Fig. 1. Optimizing channels with Genetic Algorithm

RNN is combined with Long Short-Term Memory (LSTM) in order to overcome long-term memory. The model produced outcomes in one of three classes: Moderate Stroke, Minor Stroke, and No Stroke.

A. Data Acquisition

Previous studies have explained that stroke is caused by obstruction [26]. Obstruction can be captured by slowing down the EEG signal of post-stroke patients so that high-frequency waves such as Beta and Gamma decrease. Besides, the Alpha wave can also show indicators of stroke patients [6], as well as changes in amplitude in the EEG, which is the feature that has the most considerable influence in identifying post-stroke patients [7]. The order of segments is shown in Fig. 2.

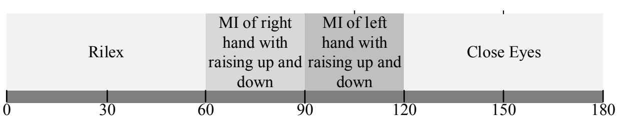

Fig. 2. EEG recording scenario

Data acquisition was of post-stroke patients at the Dustira Hospital with the Ethics Committee approval. Data retrieval lasted for two months when the patient underwent rehabilitation compared to the diagnosis of a neurologist.

Recording EEG signals using Emotiv Epoc+ 14 channels, with a sampling frequency of 128Hz for 180 seconds in a sitting position and facing the laptop to follow instructions. The first-minute containing instructions to relax, the next minute imagined raising hands and lowering hands alternately, and the last-minute containing instructions to closed eyes. Instructions given to patients are shown in Fig. 2

Instruction in Fig. 2 is intended to obtain the waves needed for the identification of post-stroke patients such as Delta, Alpha, Theta, and Mu waves. The results of the recording produced data that is used as training data and non-training data stored in a file (.csv). The amount of data on each subject when conducting one experiment is 128Hz x 180 seconds = 23,040 data points x 14 channels = 322,560 data points in a data set to practice it.

B. Wavelet Extraction

Wavelet transform has a decomposition and reconstruction process. Decomposition is extraction at a specific frequency while the reconstruction is returning to the original signal. The wavelet transformation process at the decomposition stage produces an approximation and detail signal. Approximate is a signal obtained from the convolution process of the original signal to the low-pass filter by taking data with odd indexes, and detail is a signal derived from the convolution process of the original signal to the high-pass filter by taking even indexed data.

From a sampling frequency of 128Hz, give 1-64Hz of EEG signal. So, extracted half to obtain Alpha, Beta, Theta, Gamma, and Delta waves, which will then be reduced for the identification process that is shown in Fig. 3.

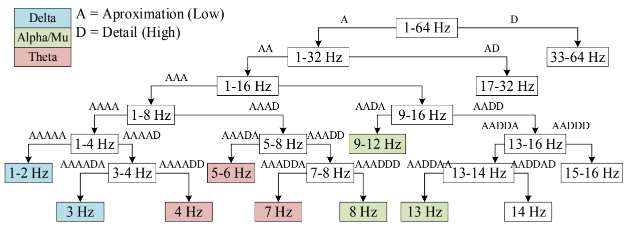

Fig. 3. Wavelet Extraction

The approximation process and details are shown in (1) and (2).

$$\text{Approximation} = ylow(k) = \Sigma n\ x(n).g(n\text{-}k) \quad (1)$$

$$\text{Detail} = yhigh(k) = \Sigma n\ x(n).h(n\text{-}k) \quad (2)$$

Where,

$x(n)$ = original signal
$g(n)$ = low-pass filter coefficient
$h(n)$ = high-pass-filter coefficient

In this study, the Wavelet used is Symlet2, which has four coefficients in each approximation and detail, as in (3) and (4).

High-pass filter coefficient

$$h_0 = \frac{1-\sqrt{3}}{4\sqrt{2}} = 0.4829,\ h_1 = \frac{3-\sqrt{3}}{4\sqrt{2}} = 0.8365,$$

$$h_2 = \frac{3+\sqrt{3}}{4\sqrt{2}} = 0.2241,\ h_3 = \frac{1+\sqrt{3}}{4\sqrt{2}} = -0.1294 \quad (3)$$

Low-pass filter coefficient

$$g_0 = \frac{1-\sqrt{3}}{4\sqrt{2}} = -0.1294,\ g_1 = -\frac{3-\sqrt{3}}{4\sqrt{2}} = -0.2241,$$

$$g_2 = \frac{3+\sqrt{3}}{4\sqrt{2}} = 0.8365,\ g_3 = -\frac{1+\sqrt{3}}{4\sqrt{2}} = -0.4829 \quad (4)$$

Record 180 seconds of EEG signal, discarded the first 30 seconds and the last 30 seconds so that it is processed 120 seconds. In the first 30 seconds, Wavelet extraction is performed to obtain Theta and Delta waves, whereas the second and third 30 seconds of extraction was carried out to obtain the Mu wave. Then in the last 30 seconds, the extraction process is carried out to obtain Alpha and Theta waves.

Wavelet extracted Delta, Theta, Alpha, and Mu waves using (1) - (4) with the hierarchy, as shown in Fig. 3. Delta has a frequency of 1-3 Hz. As Fig. 3, there are five and six steps to get each wave. Delta waves reducing the amount of data from 3,840 to 180 data points, while the Theta wave, there are four and six steps, reducing the amount of data from 3,840 to 240 data points.

The frequency ranges of the Alpha and Mu waves are the same, 8-13 Hz. Mu wave is a representation of Motor Imagery [27][28] obtained when subjects imagine raising and lowering their right hand and left hand in segments of the second 30 seconds and the third 30 seconds. However, not all channels can catch Mu wave because only the channel is in a central position in the head region or FC5 and FC6 [28]. Both waves reduced the amount of data from 3,840 to 360 data points using (3) and (4), with four and six steps as Fig. 3 shows.

C. Recurrent Neural Networks

Recurrent Neural Network (RNN) is one of the methods in Deep Learning that can use information that has been previously recorded in a long sequence or sequence so that the decisions produced are influenced by what has been learned from the past. RNN has a long-term memory problem so that Long Short Term Memory (LSTM) helps to overcome it. The architecture of the RNN variant of LSTM can be seen in Fig. 4. LSTM has also been used to identify EEG signals from stroke patients [19], classification of sleep stages [22], and identification of emotions from EEG signals [20][21].

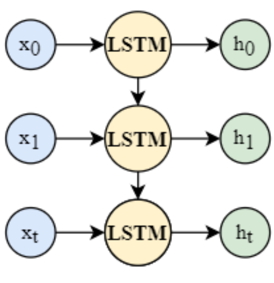

Fig. 4. Arsitektur RNN varian LSTM

LSTM has three gates. The first gate is the forget gate to determine the information that will be removed from a cell or used in a cell using the sigmoid layer. Information that is removed or used from the previous cell using (5) with the activation function using ReLU with (6).

$$f_t = \sigma(W_f . [h_{t-1}, x_t] + b_f) \quad (5)$$

$$ReLU(x) = max(0, x) \quad (6)$$

The second gate is the input gate of the sigmoid layer that will be updated—this tanh layer is created as a vector of values updated using (7) and (8).

$$i_t = \sigma(W_f \bullet [h_{t-1}, x_t] + b_i) \quad (7)$$

$$\tilde{C}_t = \tanh(W_c \bullet [h_{t-1}, x_t] + b_c) \quad (8)$$

Then cells from (5), (7), and (8) are updated using (9).

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (9)$$

Finally, the gate output updated the cell and the sigmoid layer that determines the cell to be taken as the final result using (10) and (11).

$$o_t = \sigma(W_o \bullet [h_{t-1}, x_t] + b_o) \quad (10)$$

$$h_t = o_t * \tanh(C_t) \quad (11)$$

Where c_t is an internal memory unit of combination with previous memory, index of i, f, o are input gate, forget gate, output gate. c_{t-1} is previous memory, x_t is input at each time step t at this time, s_{t-1} is the last hidden state W_f, W_t, W_c, W_o is the weight matrix, h_{t-1} previous hidden state, b_f, b_i, b_c, b_o is the bias vector, σ is a sigmoid activation function of (-1, 1), tanh is a function tangent activations of (-1, 1), and ReLU is activation functions of (0 - x).

The output gate enters the Softmax activation function, which converts the yield into a probability for each class that can be seen in (12). Where y_k is the k^{th} output data value.

$$Softmax(y_k) = \frac{e^{y_i}}{\sum_{j=1}^{n} e^{y_j}} \quad (12)$$

Next, it calculates the square of error in the output layer by using cross-entropy, which is shown in (13). Where S is the result of Softmax value, and L is the class label.

$$Losses(S, L) = -\sum_i L_i \log(S_i) \quad (13)$$

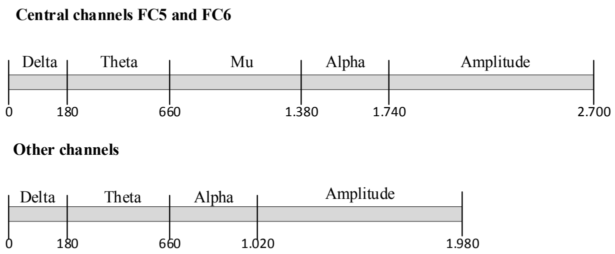

Fig. 5. The sequence of RNN features

After the EEG signal data is extracted using Wavelet, it is then identified using RNN. The order of features included in the RNN learning, as shown in Fig. 5. Based on Fig. 5, channels FC5 and FC6 produced 2,700 points from the extraction of Delta, Theta, Mu, Alpha, and amplitude waves, whereas the other 12 channels have 1,980 points of extraction without Mu waves. So the feature of the first training process is 29,160 points. The input neuron is connected to the neuron contained in the hidden layer. In the hidden layer, there is a cell that includes four steps. The first step or called first layer LSTM, which includes the activation function with ReLU from (6) to convert the negative vector to 0. In the second step, the Dropout layer is used to reduce the number of input neurons with a probability of 0.2. The third step, LSTM layer 2, using (5) through (11). The fourth step is the Dense Layer, which uses the Softmax activation function in (12). The results of this first training will produce a Losses value that will be stored as a comparison value for the results of the next training using a channel combined with Genetic Algorithms.

D. Genetic Algorithm

Genetic Algorithm is an optimization technique that starts from the beginning of an individual to produce a new individual through crossover, mutation, and elitism. The primary purpose of mutations is to prevent reaching local optimum by adding diversity to a population. Elitism is a selection process carried out by maintaining the most reliable individual from the current generation to the next generation.

It was optimizing channels for the identification of post-stroke patients by creating a chromosome architecture. The process consists of six genes that can contain six possible values of the seven odd channels of AF3, F7, F3, FC5, T7, P7, and O1, where each channel represents its partner channel. The generation of the initial population produces eight random chromosomes, which can be seen in Fig. 6.

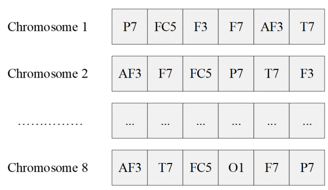

Fig. 6. Initial population

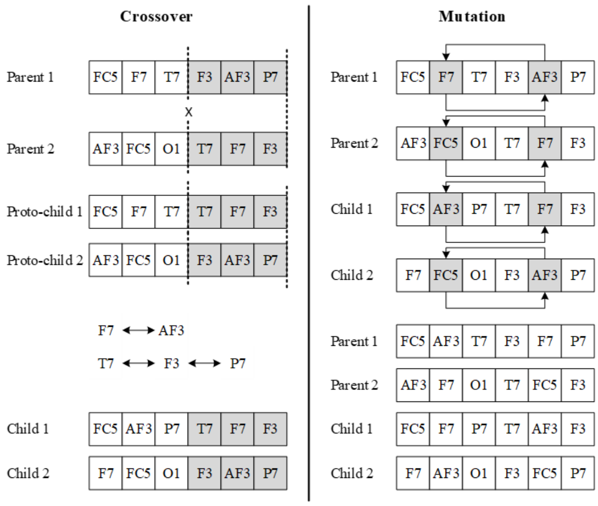

Fig. 7. Visualization of Crossover and Mutasi Operations

A combination of channels has several conditions: there cannot be the same channel in one chromosome; FC5 channels must be present in every channel, and the losses value resulting from the channel combination selection is smaller than the losses value in the first training using RNN with complete channels. In Fig. 6, chromosomes were selected by ascending of losses value training results. After obtaining the two individuals with the smallest losses values, crossover and mutation operations will be carried out, as shown in Fig. 7. The process of evolution will continue until it converges. Each channel always has a partner. So that only one pair of chromosomes is formed, as in Fig. 7. So that the data entered into the RNN training process results of optimization using the Genetic Algorithm produces 25,200 points.

III. Result and Discussion

This study used 75 data sets consisting of patients with "Moderate Stroke", "Minor Stroke," and "No Stroke" levels. The model extracted EEG signals using Wavelet, which reduces 215,040 data points to 29,160 data points from 14 channels involving Delta, Theta, Mu, Alpha, and amplitude waves. Optimization models chosen to correct weights are Adaptive Moment Estimation (Adam) and Stochastic Gradient Descent (SGD). Sixty data set as training data and 15 data sets as test data. Adam and SGD optimization models used a learning rate of 0.0001 and 0.01.

Identification involved four wave features (Delta, Theta, Alpha, and Mu) and amplitude using Genetic Algorithms and RNN. The process began with the generation of an initial population of eight chromosomes with eight genes that represent channels, as shown in Table I.

TABLE I. Initial Population

Population	Channel Combination	Losses	Accuracy (%)
Chromosome 1	P7, FC5, F3, F7, AF3, T7	0.6549	72.22
Chromosome 2	**FC5, F7, T7, F3, AF3, P7**	**0.6152**	**83.33**
Chromosome 3	AF3, F7, FC5, P7, T7, F3	0.8466	72.22
Chromosome 4	AF3, T7, P7, O1, F7, FC5	0.6664	72.22
Chromosome 5	**AF3, FC5, O1, T7, F7, F3**	**0.5184**	**83.33**
Chromosome 6	AF3, T7, FC5, P7, O1, F7	0.6233	77.78
Chromosome 7	F7, P7, AF3, T7, FC5, F3	0.7179	83.33
Chromosome 8	AF3, T7, FC5, O1, F7, P7	0.7127	66.67

Experiments of the 14 channels or initial population involved two optimization models in weight correction, namely Adam and SGD. Accuracy is shown in Fig. 8. Whereas Losses is shown in Fig. 9. Adam and SGD optimization models both have unstable graph curves in testing. However, the Adam optimization model with a learning rate of 0.0001 has an accuracy value of 72.22% higher than the Adam optimization with a learning rate of 0.01, which has an accuracy value of 61.11%, while the SGD optimization model with learning rates of 0.0001 and 0.01 has an accuracy value of 44.44% and 66.67%.

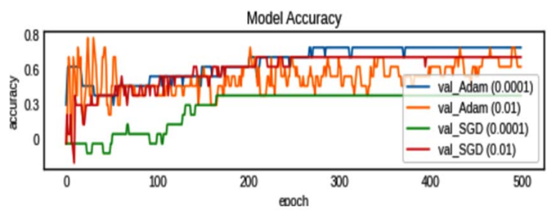

Fig. 8. Model Accuracy of Adam and SGD

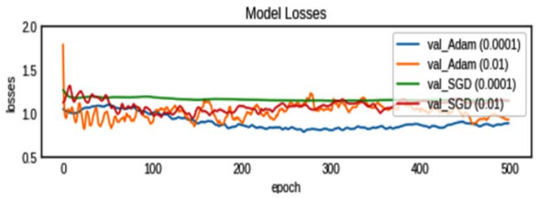

Fig. 9. Model Losses of Adam and SGD

Based on Table I, two of the best individuals were chosen, namely chromosome 2 and chromosome 5, for subsequent crossover and mutation operations, as in Fig. 7. New individual losses and accuracy values resulting from crossover and mutation operations can be seen in Table II and Table III with varying learning rates using Adam optimization model.

Table II displays the results in the form of losses and accuracy values. The use of all channels uses a learning rate configuration of 0.01, which results in an accuracy of 66.67% with 97 seconds of training time. Meanwhile, using a combination of channels produced by the Genetic Algorithm produced an accuracy of 77.78% with a training time of 65 seconds.

TABLE II. CHANNELS CONFIGURATION WITH 0.01 LEARNING RATE

Channel Combination	Epoch					
	100		300		500	
	[b]*Loss*	[a]*Acc (%)*	[b]*Loss*	[a]*Acc (%)*	[b]*Loss*	[a]*Acc (%)*
All Channels	0.892	50.00	0.925	61.11	0.784	66.67
FC5, AF3, P7, T7, F7, F3	**0.905**	**60.00**	**0.812**	**66.67**	**0.743**	**77.78**
F7, FC5, O1, F3, AF3, P7	1.053	53.33	1.007	60.00	0.951	72.22
FC5, AF3, T7, F3, F7, P7	1.110	44.44	0.993	61.11	0.965	66.67
AF3, F7, O1, T7, AF3, F3	0.978	50.00	0.988	53.33	0.968	60.00
FC5, F7, P7, T7, AF3, F3	0.950	66.67	0.847	72.22	0.826	72.22
F7, AF3, O1, F3, FC5, P7	1.124	50.00	1.007	53.33	0.984	60.00

a.Acc : Accuracy
b.Loss : Losses

TABLE III. CHANNELS CONFIGURATION WITH 0.0001 LEARNING RATE

Channel Combination	Epoch					
	100		300		500	
	[b]*Loss*	[a]*Acc (%)*	[b]*Loss*	[a]*Acc (%)*	[b]*Loss*	[a]*Acc (%)*
All Channels	0.992	55.56	0.877	72.22	0.858	72.22
FC5, AF3, P7, T7, F7, F3	**0.789**	**77.72**	**0.530**	**77.78**	**0.421**	**90.00**
F7, FC5, O1, F3, AF3, P7	0.902	66.67	0.670	66.67	0.580	77.78
FC5, AF3, T7, F3, F7, P7	0.933	50.00	0.670	72.22	0.661	72.22
AF3, F7, O1, T7, AF3, F3	0.915	55.56	0.723	66.67	0.635	72.22
FC5, F7, P7, T7, AF3, F3	0.823	77.78	0.718	72.22	0.629	77.78
F7, AF3, O1, F3, FC5, P7	0.980	50.00	0.711	61.11	0.787	66.67

a.Acc : Accuracy
b.Loss : Losses

Then a variation of learning rate was carried out at 0.0001, as shown in Table III, which displays the results in the form of losses and accuracy values from the use of the entire channel and the combination of channels from the Genetic Algorithm. Using all channels using a learning rate configuration of 0.0001 produces an accuracy of 72.22% with a training time of 80 seconds, while the use of a combination of channels produced by the Genetic Algorithm produces an accuracy of 90.00% with a training time of 50 seconds.

This experiment shows that optimizing channel usage is not only influenced by channel arrangement but is obtained from the stability of the epoch test. Increasing the epoch can increase the possibility of the learning model to correct error and accuracy. Moreover, most results were obtained at 500 epoch with a learning rate of 0.0001. The graph of accuracy results can be seen in Fig. 10 and the losses graph in Fig. 11. Table IV displays the result Adam, and SGD optimization models both have unstable graph curves in testing, but Adam optimization model has higher accuracy and lower losses.

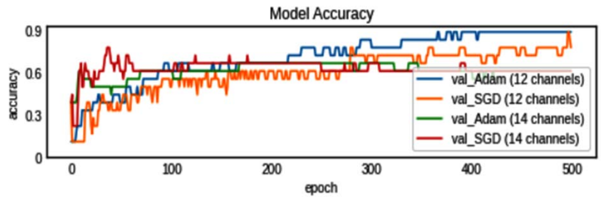

Fig. 10. Accuracy of Adam and SGD between GA and before GA

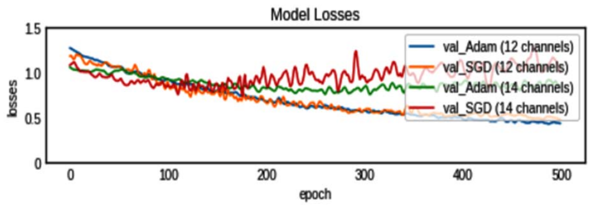

Fig. 11. Losses of Adam and SGD between GA and before GA

TABLE IV. ACCURACY

Channel Combination	SGD		Adam	
	Losses	Accuracy (%)	Losses	Accuracy (%)
All Channels	1.004	66.67	0.858	72.22
Using Genetic Algorithm	0.4817	77.78	**0.421**	**90.00**

In this study, the combination of channel FC5, AF3, P7, T7, F7, F3, F4, F8, T8, P8, AF4, and FC6 produces high accuracy values and low losses values using the Adam optimization model. Based on Fig. 10 and Fig.11, Adam optimization model converges faster. Convergence is obtained at the epoch of 400 to 500. It is because the Adam model calculates the level of learning for each parameter using the estimate of the first-moment gradient as the average of the exponential gradient.

IV. CONCLUSION

The Genetic Algorithm can optimize using channels and their sequences to identify post-stroke patients so that the algorithm can combine with Recurrent Neural Networks and Long Short-Term Memory as identification methods. The parameter test is performed to determine the effect of the sequence of features entered into the RNN learning model with identification accuracy.

The selection of EEG channels for identification of post-stroke patients using Genetic Algorithms provided an increase in accuracy from 72.22% to 90.00% and reduces computation time from 80 seconds to 50 seconds compared to the use of the whole channel. In addition to optimizing the use of the EEG channel, an optimization model is needed to improve the weight. Adam and SGD models are used to improve weights so as to reduce the value of losses in the learning process. The use of the Adam optimization model with 500 epochs results in lower losses and higher accuracy than the SGD model. So, it can be concluded that the identification of post-stroke patients can be reduced by reducing the number of channels used to get results that are faster and lighter and better results.

However, this study has not yet configured the feature sequences with the existing model, considering that the feature configuration is very influential on the use of RNN. Future studies will be continued with multivariate RNN and CNN with the Genetic Algorithm optimization model.

ACKNOWLEDGMENT

Thank you for Dustira Hospital in Cimahi-Indonesia to permit data retrieval under ethics approval.

REFERENCE

[1] E. C. Djamal and P. Lodaya, "EEG Based Emotion Monitoring Using Wavelet and Learning Vector Quantization," in *2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI 2017)*, 2017, no. September, pp. 19–21.

[2] S. Charbonnier, R. N. Roy, S. Bonnet, and A. Campagne, "EEG index for control operators' mental fatigue monitoring using interactions between brain regions," *Expert Systems with Applications*, vol. 52, pp. 91–98, 2016.

[3] E. C. Djamal, R. Indrawan, and F. Renaldi, "EEG Based Neuropsychology of Advertising Video Using Fast Fourier Transform and Support Vector Machine," *Journal of Telecommunication, Electronic & Computer Engineering (JTEC)* vol. 9, no. 3–7, pp. 105–109, 2017.

[4] I. Aliyu, Y. B. Lim, and C. G. Lim, "Epilepsy detection in EEG signal Using recurrent neural network," *ACM International Conference Proceeding Series*, pp. 50–53, 2019.

[5] J. das C. Rodrigues, P. P. R. Filho, E. P. Jr, A. K. N, and V. H. C. de Albuquerque, "Classification of EEG signals to detect alcoholism using machine learning techniques," *Pattern Recognition Letters*, 2019.

[6] N. Wang *et al.*, "Characterization of Electroencephalography of strokes based on Time-frequency analysis," *Computer-Aided Chemical Engineering*, vol. 44, pp. 2161–2166, 2018.

[7] E. C. Djamal, D. P. Gustiawan, and D. Djajasasmita, "Significant variables extraction of post-stroke EEG signal using wavelet and SOM Kohonen," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 17, no. 3, pp. 1149–1158, 2019.

[8] S. Z. Bong, K. Wan, M. Murugappan, N. M. Ibrahim, Y. Rajamanickam, and K. Mohamad, "Implementation of wavelet packet transform and non-linear analysis for emotion classification in stroke patient using brain signals," *Biomedical Signal Processing and Control*, vol. 36, no. April, pp. 102–112, 2017.

[9] O. N. Rahma, S. K. Wijaya, Prawito, and C. Badri, "Electroencephalogram analysis with extreme learning machine as a supporting tool for classifying acute ischemic stroke severity," *Proceedings - 2017 International Seminar on Sensor, Instrumentation, Measurement and Metrology: Innovation for the Advancement and Competitiveness of the Nation, ISSIMM 2017*, vol. 2017-January, pp. 180–186, 2017.

[10] M. R. Islam and M. Ahmad, "Wavelet Analysis Based Classification of Emotion from EEG Signal," *2nd International Conference on Electrical, Computer and Communication Engineering, ECCE 2019*, pp. 1–6, 2019.

[11] E. C. Djamal and H. A. Tjokronegoro, "Identification and Classification of EEG Signal Against Sound Stimulation Using Wavelet and Power Spectral," *ITB Journal of Sciences*, vol. 37, no. 1, pp. 69–92, 2005.

[12] W. R. W. Omar, R. Jailani, M. N. Taib, R. M. Isa, and Z. Sharif, "Assessment of acute ischemic stroke brainwave using Relative Power Ratio," *Proceedings - 2013 IEEE 9th International Colloquium on Signal Processing and its Applications, CSPA 2013*, pp. 310–313, 2013.

[13] M. Arvaneh, S. Member, C. Guan, K. K. Ang, and C. Quek, "Optimizing the Channel Selection and Classification Accuracy," *IEEE transactions on bio-medical engineering*, vol. 58, no. 6, pp. 1865–1873, 2012.

[14] Q. Wei, Z. Lu, K. Chen, and Y. Ma, "Channel selection for optimizing feature extraction in an electrocorticogram-based brain-computer interface," *Journal of Clinical Neurophysiology*, vol. 27, no. 5, pp. 321–327, 2010.

[15] A. Albasri, F. Abdali-Mohammadi, and A. Fathi, "EEG electrode selection for person identification thru a genetic-algorithm method," *Journal of Medical Systems*, vol. 43, no. 9, 2019.

[16] C. Y. Kee, S. G. Ponnambalam, and C. K. Loo, "Multi-objective genetic algorithm as a channel selection method for P300 and motor imagery data set," *Neurocomputing*, vol. 161, no. Mi, pp. 120–131, 2015.

[17] C. Y. Kee, R. M. Kuppan Chetty, B. H. Khoo, and S. G. Ponnambalam, "Genetic algorithm and Bayesian linear discriminant analysis based channel selection method for P300 BCI," *Communications in Computer and Information Science*, vol. 330 CCIS, pp. 226–235, 2012.

[18] S. Bavkar, B. Iyer, and S. Deosarkar, "Rapid Screening of Alcoholism: An EEG Based Optimal Channel Selection Approach," *IEEE Access*, vol. 7, pp. 99670–99682, 2019,.

[19] P. Chantamit o Pas and M. Goyal, "Long Short-Term Memory Recurrent Neural Network for Stroke Prediction," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10934 LNAI, pp. 312–323, 2018.

[20] S. Alhagry, A. Aly, and R. A., "Emotion Recognition based on EEG using LSTM Recurrent Neural Network," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 10, pp. 8–11, 2017.

[21] J. Yan, S. Chen, and S. Deng, "An EEG-based emotion recognition model with rhythm and time characteristics," *Brain Informatics*, vol. 6, no. 1, 2019.

[22] N. Michielli, U. R. Acharya, and F. Molinari, "Cascaded LSTM recurrent neural network for automated sleep stage classification using single-channel EEG signals," *Computers in Biology and Medicine*, vol. 106, no. December 2018, pp. 71–81, 2019..

[23] M. Li, M. Zhang, X. Luo, and J. Yang, "Combined long short-term memory-based network employing wavelet coefficients for MI-EEG recognition," *2016 IEEE International Conference on Mechatronics and Automation, IEEE ICMA 2016*, pp. 1971–1976, 2016.

[24] S. Roy, I. Kiral-Kornek, and S. Harrer, "ChronoNet: A Deep Recurrent Neural Network for Abnormal EEG Identification," 2018.

[25] Dawei Cheng; Ye Liu and Liqing Zhang, "Exploring motor imagery eeg patterns for stroke patients with deep neural networks," *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2561–2565, 2018.

[26] E. P. Giri, M. I. Fanany, A. M. Arymurthy, and S. K. Wijaya, "Ischemic stroke identification based on EEG and EOG using ID convolutional neural network and batch normalization," *2016 International Conference on Advanced Computer Science and Information Systems, ICACSIS 2016*, pp. 484–491, 2017.

[27] F. Singh, J. Pineda, and K. S. Cadenhead, "Association of impaired EEG mu wave suppression, negative symptoms and social functioning in biological motion processing in first episode of psychosis," *Schizophrenia Research*, vol. 130, no. 1–3, pp. 182–186, 2011, doi: 10.1016/j.schres.2011.04.004.

[28] E. C. Djamal, R. I. Ramadhan, M. I. Mandasari, and D. Djajasasmita, "Identification of post-stroke EEG signal using wavelet and convolutional neural networks," *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 5, pp. 1890–1898, 2020.