

Brain Signal Classification using Genetic Algorithm for Right-Left Motion Pattern

Cahya Rahmad¹, Rudy Ariyanto², Dika Rizky Yunianto³
Information Technology, State Polytechnic of Malang
Indonesia

Abstract—Brain signals or EEG are non-stationary signals and are difficult to analyze visually. The brain signal has five waves alpha, beta, delta, gamma, and theta. The five waves have their frequency to describe the level of attention, alertness, character and external stimuli. The five waves can be used to analyze stimulation patterns when turning left and right. Giving weight to the five brain waves utilizes genetic algorithms to get one signal. Genetic algorithms can be used to find the best signal for classification. In this paper, the EEG signal will be classified to determine the right or left movement pattern. After combining the five brain waves with genetic algorithms is then classified using the Logistic Regression, Linear Discriminant Analysis, K-Neighbors Classifier, Decision Tree, Naïve Bayes Gaussian, and Support Vector Machine. From the six methods above that have the highest accuracy is 56% and SVM is a method that has better accuracy than others on this problem.

Keywords—Brain wave; EEG; genetic algorithm; classification; left right movement

I. INTRODUCTION

When the body does a job or movement, it basically coordinates with the mind. These conditions, focus on one object without being affected by other things and focus on doing the movement. Knowing someone's focus condition is not easy, one way to find out the condition of one's focus is through information on brain signals or often called an Electroencephalogram (EEG) signal.

EEG is an instrument used to record static electricity activity resulting from stimuli received by the brain. Research on the classification of EEG signals has been carried out, including classification of fatigue levels, identification of EEG signals for sound stimulation, identification of alertness, emotional conditions, attention classification, identification of epilepsy, and other studies classifying EEG signals against imagination of body movements, classification visual stimulation, classification of EEG signals with two mental conditions with an introduction of up to 83%, identification of epilepsy waves, and to recognize movements of artifacts [1][2][3].

Transforming EEG signals into a model is an effective way of analysis to classify EEG signals. An EEG signal in a person generally consists of wave components which are differentiated based on their frequency region, delta, theta, alpha, beta, and gamma [4][5][6].

TABLE I. FREQUENCY EEG

EEG	Frequency
Delta	0 - 4 Hz
Theta	4 - 8 Hz
Alpha	8 - 13 Hz
Beta	13 - 30 Hz
Gamma	30 - 100 Hz

In Table 1 it can be seen that Delta waves have a frequency of 0-4 Hz. Delta waves appear when someone is sleeping soundly. Theta waves have a frequency of 4 - 8 Hz. Theta waves appear when a person sleeps lightly and is in a happy state. Some recent research links these waves such as rapid eye movements during sleep, and hypnosis. Alfa waves have a frequency of 8-13 Hz. Alpha waves appear when a person relaxes, and eyes are closed. These waves are often used to see normal or abnormal brain functions. Beta waves have a frequency of 13-30 Hz. Beta waves appear when someone is doing activities concerning remembering such conditions as thinking. Gamma waves have a frequency of 30-100 Hz. Gamma waves are related to brain activity to integrate various stimuli [7]–[9].

The EEG signal analysis in this paper will be used to analyze for control an object for right-left movement. EEG analysis uses a genetic algorithm to combine five brain signals. the genetic algorithm is one of the heuristic methods which is a branch of an evolutionary algorithm, which is a technique for solving complex optimization problems by imitating the evolutionary process of living things. genetic algorithm proved to be suitable to be used to solve multi-objective problems. genetic algorithm develops along with the rapid development of information technology. This algorithm is widely used in the fields of physics, biology, economics, sociology, and others who often face optimization problems with complex or even difficult mathematical models. by using the genetic algorithm, one brain signal from the five brain signals is obtained. one signal will be used for classification in the next process.

II. METHODS

In this paper, several stages are carried out. These stages can be seen in figure 1.

In Figure 2 it is explained that the method proposed. Where each of the five brain waves signals will be weighted by the genetic algorithm method. The result a signal that will be used for classification.

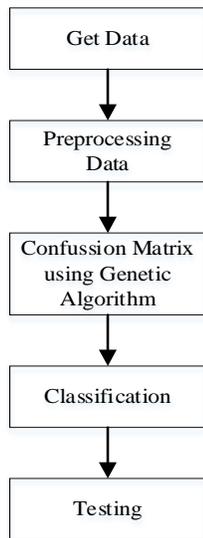


Fig. 1. Process Diagram Flow.

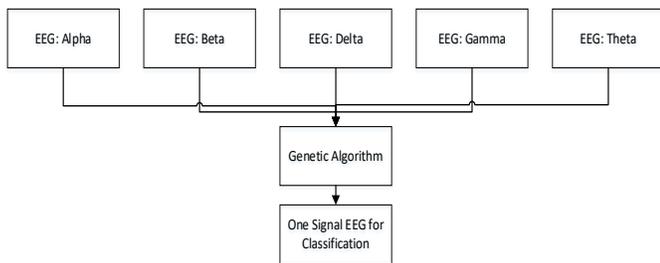


Fig. 2. Proposed Method.

A. Data

The data is taken using a brain signal reader. The tool is MindWave Mobile Brainwave Starter Kit from NeuroSky. The device can directly read the five brain waves. The use of this tool utilizes Bluetooth technology, and this tool is compatible with operating systems such as Windows, Mac, Android or Linux. The device can be seen in figure 3.

Brainwave data is taken from 10 people where each person is recorded 10 times. 5 times the brainwave recording for the condition of the right moves and 5 times the brainwave recording for the condition of the left movement. So the total data captured is 100 data.



Fig. 3. Brainwave Reader.

Recording of brain wave is done when the subject controls the right or left of the remote control car toy and recording is done for one minute. The data took 15 seconds in the middle for the next analysis. Example and visualization of data can be seen at table 2 and figure 4.

TABLE II. EXAMPLE RAW DATA

ms	1	2	...	15	CLASS
ALPHA	69987,5	49024,5	...	335993	LEFT MOVEMENTS
BETA	6527,5	9185	...	8822	
GAMMA	1409	1220	...	7123	
THETA	1532876	856975	...	716820	
DELTA	2	2	...	2	

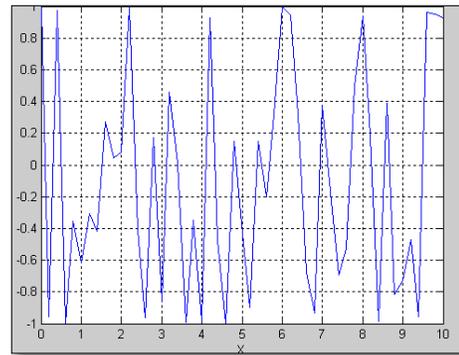


Fig. 4. Visualization Data.

B. Preprocessing

At this stage, the data will be normalized using the min-max method. This normalization is done to adjust the signal magnitude. All waves will be mapped with a range between 0 and 1. Min-max normalization calculations are carried out following equation 1.

$$y = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{1}$$

C. Genetic Algorithm

A genetic algorithm is a search heuristic that is inspired by Charles Darwin’s theory of natural evolution. This algorithm reflects the process of natural selection where the fittest individuals are selected for reproduction to produce offspring of the next generation.

The process of natural selection starts with the selection of fittest individuals from a population. They produce offspring which inherit the characteristics of the parents and will be added to the next generation. If parents have better fitness, their offspring will be better than parents and have a better chance of surviving. This process keeps on iterating, and at the end, a generation with the fittest individuals will be found[10].

This notion can be applied to a search problem. We consider a set of solutions for a problem and select the set of best ones out of them. Five phases are considered in a genetic algorithm is the initial population, fitness function, selection, crossover, and mutation.

The initial population is processed begins with a set of individuals which is called a Population. Each is a solution to the problem you want to solve. On this EEG problem at equation 2, it is known that y is the new value of combining five signals and x is the signal captured by the reader. while w is the weight sought for each type of wave where the sum of w must be 1 like equation 3. In the process of the genetic algorithm we will find the best w weights to be multiplied by each of the types of waves that have been taken on average to determine that weight [10], [11].

$$y = w1.alpha(x) + w2.beta(x) + w3.delta(x) + w4.gamma(x) + w5.theta(x) \quad (2)$$

$$w1 + w2 + w3 + w4 + w5 = 1 \quad (3)$$

An individual is characterized by a set of parameters (variables) known as Genes. Genes are joined into a string to form a Chromosome (solution). In a genetic algorithm, the set of genes of an individual is represented using a string, regarding an alphabet. Usually, binary values are used (string of 1s and 0s). We say that we encode the genes in a chromosome [10].

The fitness function determines how to fit an individual is (the ability of an individual to compete with other individuals). It gives a fitness score to each. The probability that an individual will be selected for reproduction is based on its fitness score [12].

The idea of the selection phase is to select the fittest individuals and let them pass their genes to the next generation. Two pairs of individuals (parents) are selected based on their fitness scores. Individuals with high fitness have more chance to be selected for reproduction [12].

Crossover is the most significant phase in a genetic algorithm. For each pair of parents to be mated, a crossover point is chosen at random from within the genes. Offspring are created by exchanging the genes of parents among themselves until the crossover point is reached. The new offspring are added to the population [11].

In particular new offspring formed, some of their genes can be subjected to a mutation with a low random probability. This implies that some of the bits in the bit string can be flipped. The mutation occurs to maintain diversity within the population and prevent premature convergence [10].

The algorithm terminates if the population has converged (does not produce offspring which are significantly different from the previous generation). Then it is said that the genetic algorithm has provided a set of solutions to our problem [10].

D. Classification

From the process of genetic algorithm each data produces a signal that can be classified. In this paper uses six classification methods to get the best classification.

- Logistic Regression (LR)

In the statistical model with two categories, with response variables contain elements of "success" or "failure". This binary data is the simplest form of data category. The most

frequently used model for data two the category is binary logistic regression [13].

Logistic Regression (LR) is used to measure the functional relationship between one dependent variable from a qualitative type of dichotomous with independent variables of type quantitative and qualitative. It is somewhat similar to multiple linear regression. It is usually appropriate for models where dependent variables are of the qualitative type of dichotomous. Model parameters are estimated using the maximum-likelihood method [13]. Form a logistic regression model with variables i predictors are as follows equation (4)

$$\pi(x) = \frac{e^{g(x)}}{1+e^{g(x)}} \quad (4)$$

By using a logit transformation from $\pi(x)$, then the logit function regression model can be defined as following equation (5).

$$g(x) = \beta_0 + \beta_1x_1 + \beta_2x_2 \dots + \beta_ix_i \quad (5)$$

The logit form $g(x)$ is a logit model, a linear function in its parameters, and is within the distance between $-\infty$ to $+\infty$ depends on variable X .

- Linear Discriminant Analysis (LDA)

LDA performs linear analysis has its representation (vectors base) from dimensionless EEG vector space high, depending on the statistical point of view. By projecting an EEG vector into its base vector, representation will be obtained the feature of the wave per unit time [14].

Similarity measurements will then be made between EEG representations with data testing. Representations in this method are considered as a linear transformation from the original EEG vector to in a projection space (base vectors)[14].

- K-Neighbours Classifier (KNN)

K-nearest neighbor algorithm is a classification technique a very popular one introduced by Fix and Hodges (1951), which have been proven to be a good simple algorithm. KNN is one method used in classification by using a supervised algorithm (Chan et al. 2010). The purpose of this algorithm is to classify new objects based on the distance of an object to be classified to sample data. Classifier only uses the distance function from new data to training data.

K-Nearest Neighbor is to find the closest distance between data will be evaluated with neighbor K closest neighbors in training data. Training data is projected into space many dimensions, where each dimension represent features of data. This space is divided into parts based on the classification of training data. A point in this space is marked class c , if class c is a classification that is most commonly found in neighboring fruit closest to that point [15]. Near or far neighbors are usually calculated based on distance Euclidean with the following equation (6).

$$d_i = \sqrt{\sum_{i=1}^p (x_{2i} - x_{1i})^2} \quad (6)$$

When x_1 is data sample, x_2 is data testing, and i is variable data. d is a symbol of distance and p is a symbol of data dimension.

In the learning phase, this algorithm only does vector-vector storage features and classification of learning data. In the classification phase, the features are the same calculated for test data (which is the classification not known). The distance from the new vector this is against the entire vector of learning data counted, and a number of the most fruits close taken. The new point is classification predicted included in the classification most of these points. The best k value for this algorithm depends on the data [16].

Generally, the value of k is height will reduce the effect of noise on classification, but make boundaries between each classification becomes more blurred. Value k good can be selected by optimization parameters, for example by using cross-validation. Where special cases classification is predicted based on data the closest learning (with words another, $k = 1$) is called the nearest neighbor algorithm [16].

The accuracy of the KNN algorithm is very influenced by the presence or absence of features irrelevant, or if feature weight it is not equivalent to its relevance against classification. When is the amount of data approaching infinity, this algorithm guarantees error rate of no more than twice Bayes error rate (minimum level for distribution certain data)[15].

- Decision Tree Classifier (CART)

Decision tree is one of the most popular data mining classification techniques. The appropriate decision tree is used for cases that have the following characteristics [17]:

- 1) Data or examples are expressed by pairs of attributes and values.
- 2) Labels or output data are usually discrete values.
- 3) Data has a missing value.

Entropy theory is adopted to choose the right attribute breakdown for the C4.5 algorithm, stating the average number information needed to classify samples.

- Naïve Byaes Gaussian (NB)

Naïve Bayes is an algorithm in data mining technique that applies the Bayes theory in classification. Bayes' decision theorem is a fundamental statistical approach within pattern recognition (pattern recognition). Naive Bayes based on the assumption of simplifying that value attributes conditionally free if given output value. In other words, given the output value, the probability of observing together is a product of individual probabilities. In equation 7 $P(c|x)$ is the posterior probability of class (target) given predictor (attribute) and $P(c)$ is the prior probability of class. $P(x|c)$ is the likelihood which is the probability of predictor given class and $P(x)$ is the prior probability of predictor [18][19].

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)} \quad (7)$$

- Support Vector Machine (SVM)

The support vector machine (SVM) was first introduced by Vapnik in 1992 when presented at Annual Workshop on Computational Learning Theory. The basic principle of SVM is a linear classifier, namely case classifications that are linearly separated [20].

Support Vector Machine (SVM) is a method classification that works by searching hyperplane with the biggest margin. The hyperplane is the inter-class data dividing line. The margin is the distance between the hyperplane and the closest data in each class. The closest data with a hyperplane in each of these classes which are called support vector. SVM is a method used for the classification of two classes (binary classification). In its development, several a method proposed for SVM to be used for multi-class classification by combining some binary classifier. Method ever proposed is the One-against-one method. As for the One-against-one method, it will be constructed by a number of $k(k-1) / 2$ SVM classification models with each model being trained to use data from two different classes [21][12].

Therefore, for data on classes i and j , SVM completes binary classification problems for The decision function for the function above is taken through voting, if the result is *sign* states that data x is in class i , then vote value for class i plus one. Furthermore, class predictions from data x are classes with values highest vote. Otherwise, the vote value for class j plus one. In general, the stages of document classification with multi-class SVM this research can describe as the following stages. First of all, from the results of the pre-process and the decision, then generated k -groups of data terms in the representation vector space model. Then each group was trained to use the SVM Multi-class with the One-against-one method. Therefore the document classification model will be obtained on each group [20], [21].

E. Testing

At this stage, testing of the method will be done by dividing the data into training and testing. Each type of wave on the EEG signal will be classification after weighting of the genetic algorithm process done. The testing process is done by allocating data testing at 10%, 15%, 20%, 25%, and 30%.

III. RESULTS

The implementation of the proposed method is run in the Python programming language. Python uses various types of libraries, sklearn, matplotlib and pandas.

The results of implementation and testing can be seen in Table 3. From the table it can be seen that the method has the highest average accuracy is SVM with a total average of 48%. And the method which has the smallest average is KNN method which is not at odds with the decision tree method with the average accuracy of 38% and 39%.

In figure 5 it can be seen that the SVM method has an average above other methods. of all methods the smaller the training data, the accuracy will decrease.

TABLE III. TESTING RESULTS

Data Testing	Acuration (%)					
	LR	LDA	KNN	CART	NB	SVM
10%	57	52	30	27	40	40
15%	50	53	32	35	50	40
20%	40	45	50	50	45	55
25%	50	35	45	50	50	55
30%	35	25	35	35	40	50
AVERAGE	46	42	38	39	45	48

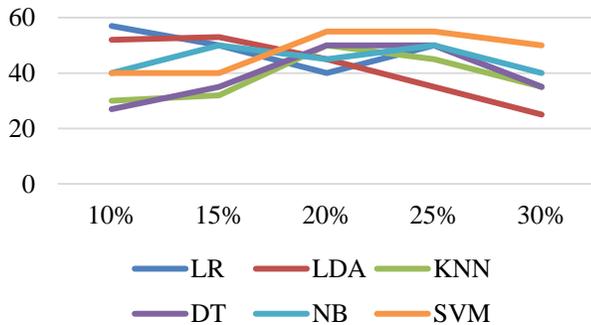


Fig. 5. Diagram Result.

IV. CONCLUSION

The brain signal has five different frequency waves. The five signals will be combined using Genetic Algorithm into one signal that can be used for classification. The signal classification is intended to determine the pattern of right-left movement thoughts. The results of the classification with the highest accuracy were 56% using the Logistic Regression method. However, the highest average accuracy is owned by SVM with 48%. This research can be further developed by further shortening the time unit in retrieving signals and changing the weight values for each type of wave to have better accuracy.

ACKNOWLEDGMENT

We are pleased to acknowledge the contribution of our colleagues (Afif, Yusron, Rendy, Yudi, Yesaya, Marcelina, Fikri, Mr. Erfan and Mr. Indrazno) and our institute Polytechnic State of Malang.

REFERENCES

[1] G. Townsend, B. Graimann, and G. Pfurtscheller, "Continuous EEG classification during motor imagery - Simulation of an asynchronous BCI," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 12, no. 2, pp. 258–265, 2004.

[2] M. Witkowski, M. Cortese, M. Cempini, J. Mellinger, N. Vitiello, and S. R. Soekadar, "Enhancing brain-machine interface (BMI) control of a hand exoskeleton using electrooculography (EOG)," *J. Neuroeng. Rehabil.*, vol. 11, no. 1, pp. 1–6, 2014.

[3] V. J. Shetty, M. Laxman, M. S. Magdum, S. B. Sidnal, and S. Paraddi, "Head Motion Controlled Wheel Chair using," vol. 2, no. 6, pp. 2393–2395, 2015.

[4] E. Haselsteiner and G. Pfurtscheller, "Using time-dependent neural networks for EEG classification," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 4, pp. 457–463, 2000.

[5] D. Wang, D. Miao, and C. Xie, "Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection," *Expert Syst. Appl.*, vol. 38, no. 11, pp. 14314–14320, 2011.

[6] I. Sturm, S. Lapuschkin, W. Samek, and K. R. Müller, "Interpretable deep neural networks for single-trial EEG classification," *J. Neurosci. Methods*, vol. 274, pp. 141–145, 2016.

[7] Y. Zhang, G. Zhou, J. Jin, Q. Zhao, X. Wang, and A. Cichocki, "Sparse Bayesian Classification of EEG for Brain-Computer Interface," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 27, no. 11, pp. 2256–2267, 2016.

[8] L. Hellstro m-Westas, I. Rosen, L. S. de Vries, and G. Greisen, "Amplitude-integrated EEG Classification and Interpretation in Preterm and Term Infants," *Neoreviews*, vol. 7, no. 2, pp. e76–e87, 2006.

[9] A. Arif and E. C. Djamal, "Klasifikasi Tingkat Kelelahan Berdasarkan Sinyal Electroencephalogram (EEG) Menggunakan Jaringan Syaraf Tiruan Backpropagation," in *SNIJA Unjani*, 2014, p. 2014.

[10] R. A. Welikala et al., "2015 - Genetic algorithm based feature selection combined with dual classification for the automated detection of proliferative diabetic retinopathy," 2015.

[11] V. Jain, A. Kumar, S. Kumar, and C. Chandra, "Weight restrictions in Data Envelopment Analysis: A comprehensive Genetic Algorithm based approach for incorporating value judgments," *Expert Syst. Appl.*, vol. 42, no. 3, pp. 1503–1512, 2015.

[12] P. Ghamisi and J. A. Benediktsson, "Feature selection based on hybridization of genetic algorithm and particle swarm optimization," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 2, pp. 309–313, 2015.

[13] D. Liu, T. Li, and D. Liang, "Incorporating logistic regression to decision-theoretic rough sets for classifications," *Int. J. Approx. Reason.*, vol. 55, no. 1 PART 2, pp. 197–210, 2014.

[14] E. Neto, F. Biessmann, H. Aurlien, H. Nordby, and T. Eichele, "Regularized linear discriminant analysis of EEG features in dementia patients," *Front. Aging Neurosci.*, vol. 8, no. NOV, pp. 1–10, 2016.

[15] S. Zhang, X. Li, M. Zong, X. Zhu, and D. Cheng, "Learning k for kNN Classification," *ACM Trans. Intell. Syst. Technol.*, vol. 8, no. 3, pp. 1–19, 2017.

[16] M. Havaei, P. M. Jodoin, and H. Larochelle, "Efficient interactive brain tumor segmentation as within-brain kNN classification," *Proc. - Int. Conf. Pattern Recognit.*, pp. 556–561, 2014.

[17] E. Gokgoz, A. Subasi, and B. S. Processing, "E. Gokgoz, A. Subasi, Comparison of decision tree algorithms for EMG signal classification," *Biomed. Signal Process. Control*, vol. 18, pp. 138–144, 2015.

[18] A. Choi, N. Tavabi, and A. Darwiche, "Structured Features in Naive Bayes Classification," *Proc. 30th Conf. Artif. Intell. (AAAI 2016)*, pp. 3233–3240, 2016.

[19] J. Wu, S. Pan, X. Zhu, Z. Cai, P. Zhang, and C. Zhang, "Self Adaptive Attribute Weighting for Naive Bayes Classification," *Expert Syst. Appl.*, vol. 42, no. 3, pp. 1487–1502, 2015.

[20] J. Cervantes, F. García Lamont, A. López-Chau, L. Rodríguez Mazahua, and J. Sergio Ruíz, "Data selection based on decision tree for SVM classification on large data sets," *Appl. Soft Comput. J.*, vol. 37, no. December, pp. 787–798, 2015.

[21] F. Riaz, A. Hassan, S. Rehman, I. K. Niazi, and K. Dremstrup, "EMD-based temporal and spectral features for the classification of EEG signals using supervised learning," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 1, pp. 28–35, 2016.