TIPTEKNO'20 TIP TEKNOLOJİLERİ KONGRESİ
19-20 Kasım 2020
ONLINE

Firefly Algorithm Based Feature Selection for EEG Signal Classification

Ebru Ergün
Electrical and Electronics Engineering
Recep Tayyip Erdoğan University
Rize, Turkey
ebru.yavuz@erdogan.edu.tr

Onder Aydemir
Electrical and Electronics Engineering
Karadeniz Technical University
Trabzon, Turkey
onderaydemir@ktu.edu.tr

***Abstract*— Brain–computer interfaces (BCIs) recognize specific features of a person's brain signal relating to his/her intent, and output a control command that controls the outside devices or computers. BCI systems facilitate the lives of patients who cannot move any muscles but have no cognitive disorder. The high dimensions of features represent a research challenge. In recent years, especially nature inspired heuristic optimization algorithms became popular in order to eliminate unnecessary features. This paper addresses a crucial factor for effective classification of motor imaginary based EEG signals that are an optimal selection of relevant EEG features using firefly algorithm. Firefly algorithm (FA) works on the principle of directing the less shiny than the light intensity emitted by fireflies in nature towards the bright. The algorithm can adaptively select the best subset of features and improve classification accuracy. In this study, following extracted Katz Fractal Dimension based features, effective feature(s) were selected by FA. The proposed method successfully applied on open access dataset which was collected from 29 subjects. We obtained an average 76.14% classification accuracy (CA) using *k*-nearest neighbor classifier. This is 4.4% higher than the CA calculated by using all features. These results proved that used method is robust for this dataset.**

Keywords*— electroencephalography; brain computer interface; katz fractal dimension; firefly algorithm; k-nearest neighbor

I. Introduction

The main control center of the human body is the brain [1]. This control occurs when millions of nerve cells (neurons) that form the structural unit of the central nervous system communicate with each other. The electrical activity of the neurons forms the electroencephalography (EEG) signals, also known as the electrical picture of the brain [2]. EEG, which is recorded depending on electrical potential changes with electrodes placed in the skull, makes it possible to obtain and interpret the underlying information in the brain [3], [4]. Also, EEG signals form the basis of brain computer interface (BCI) systems [5]. In addition to use in different areas, BCI systems facilitate the lives of patients who cannot move any muscles but have no cognitive disorder. The basic BCI system commonly consists of pre-processing, feature extraction and classification steps. In recent years, studies have been focused on feature selection algorithms because reducing the dimensionality of features contributes to improve the classification performance in BCI approaches [6]. It can also save storage and computation time and increase comprehensibility. Particularly, nature inspired heuristic optimization algorithms became popular [7]. In this study, we applied firefly algorithm (FA) to 2-class motor imaginary (MI) based EEG signals in order to enhance the performance of a classifier and obtain a rapid and high performance BCI system without redundant features.

In literature, there are a few feature selection methods to reduce feature dimension and BCI time [6]. However, most existing feature selection methods still suffer from stagnation in local optimal and/or high computational cost [8]. Thus, an efficient global search technique is needed to better solve feature selection problems. Genetic algorithm (GA) [9], artificial bee colony [10], particle swarm optimization [11] are some of the feature selection algorithms that have been recently used in BCI studies. Hsu et al. (2010) proposed a method using subband nonlinear parameters and GA for automatic seizure detection in EEG. They used discrete wavelet transform as feature extracted method and employed as the features to train the support vector machine (SVM) with linear kernel function and radial basis function kernel function classifiers. Then, the GA was used for selecting the effective feature subset. However, the disadvantage of GA takes a long time to converge to the optimal solution [12]. In another feature selection based study, Rakshit et al. (2012) worked with MI based EEG signals. In their paper, they aimed to eliminate the redundant features of a dataset to improve the accuracy of classification. To do so, they employed artificial bee colony cluster algorithm to reduce the feature dimension. They calculated the highest classification accuracy (CA) as 64.29%. However, local search ability of this algorithm was weak [13].

In this paper, we used FA based feature selection method to select effective features because the parameters that need to be controlled are relatively few, and the precision optimization and convergence speed are high [14] compared to other algorithms. Firstly, we applied fourth order of Chebyshev type II filter (passband of 0.5 - 50 Hz) to the EEG data using the Berlin BCI toolbox [15]. Secondly, after Katz Fractal Dimension (KFD) was used for extracting features from signals, an FA algorithm based optimizing feature selection

978-1-7281-8073-1/20/$31.00 ©2020 IEEE

method was employed to identify the effective classification feature(s). Finally, we achieved average 76.14% CA from 29 subjects using *k*-nearest neighbor (*k*-NN) classifier. This result is 4.4% higher than the CA calculated using all feature(s). The results proved that the proposed method is successful for this dataset.

II. Material And Method

A. Dataset Description

Open access dataset [16] was recorded in an ordinary bright room from 29 subjects. The experimental procedure consists of three sections which is given in Figure 1. Each section has a resting period of 60 seconds (sec) before the experiment. After 2 sec visual presentation was exhibited in the middle of the screen, task period was performed for 10 sec. During the 10 sec, subjects were asked to imagine right/left hand opening-closing. Then after a 1 sec "STOP" warning, rest period was performed with a 15-17 sec. Finally, a single section was terminated with a 60 sec post-rest.

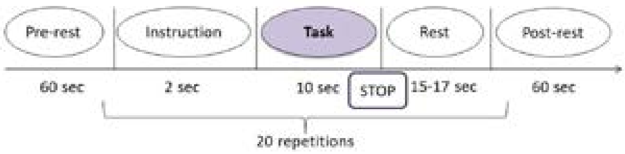

Figure 1. Experimental procedure [16]

20 trials were recorded in each section. A total of 60 trials were recorded in three sections. 30 trials are labeled as *class 0* (right hand opening-closing) while rest of them are labeled as *class 1* (left hand opening-closing). This dataset is sampled at 200 Hz and each trial consists of 2000 samples for a task period. Also, this dataset was recorded with thirty electrodes (1. AFp1, 2. AFp2, 3. AFF1h, 4. AFF2h, 5. AFF5h, 6. AFF6h, 7. F3, 8. F4, 9. F7, 10. F8, 11. FCC3h, 12. FCC4h, 13. FCC5h, 14. FCC6h, 15. T7, 16 T8, 17. Cz, 18. CCP3h, 19. CCP4h, 20. CCP5h, 21. CCP6h, 22. Pz, 23. P3, 24. P4, 25. P7, 26. P8, 27. PPO1h, 28. PPO2h, 29. POO1, 30. POO2 and Fz for ground electrode) which were placed as given in Figure 2.

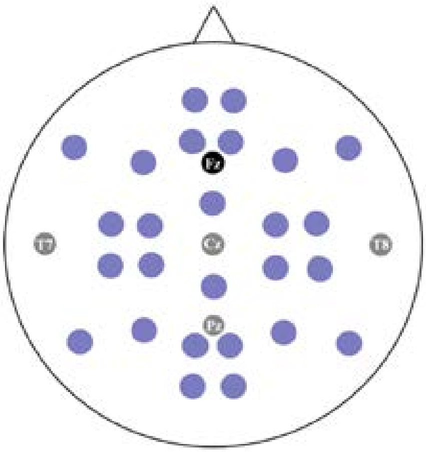

Figure 2. Electrode positions for dataset [16]

B. Proposed Method

The application procedure of the proposed method is given in Figure 3. Firstly, pre-processed EEG trials are separated by 50% randomly as training and test set. Then, KFD based feature(s) were extracted from each of these sets. During the training stage, effective feature(s) were selected with FA algorithm and the robust feature(s) in the test set were classified. Finally, the CA, sensitivity (SE), specificity (SP) and kappa coefficient (κ) were calculated for evaluating the performance of the *k*-NN.

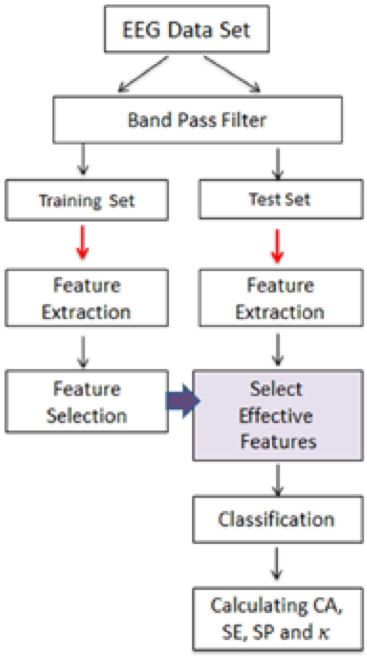

Figure 3. Flow chart of NIRS data analysis

1) Katz Fractal Dimension

We used Katz fractal dimension for extracting features. In the Katz's algorithm, Katz [17] the FD is obtained directly from the time series. P is the sum of the distance between successive data points in the signal Y ($[x_1\, y_1;\ x_2\, y_2; \dots; x_n\, y_n]$) is given in Equation 1. In this equation, n represents the size of Y.

$$P = \sum_{n=1}^{n-1} \sqrt{(x_n - x_{n+1})^2 + (y_n - y_{n+1})^2} \qquad (1)$$

The H, given in Equation 2, is the distance between the first point in the Y and the point at the farthest distance from the first point.

$$H = \max[\, \sum_{n=1}^{n} \sqrt{(x_1 - x_n)^2 + (y_1 - y_n)^2)} \qquad (2)$$

The Katz fractal dimension of the waveform is defined as in Equation 3 depending on P and H [18]. In this equation, z is the number of steps in the waveform.

$$FD = \frac{\log(z)}{\log(H/P) + \log(z)} \quad (3)$$

2) Feature Selection

In this study, firefly algorithm was used for feature(s) selection. This algorithm is herd-based Heuristic Optimization Algorithm which developed by Xin-She Yang in 2008 [19]. It was developed by considering the brightness-sensitive social behavior of fireflies. Algorithm operation consists of the following steps;

a. Random positions of fireflies are determined. Then the brightness is calculated.
b. Fireflies move towards more bright fireflies.
c. Return to stage (b) until the specified number of iterations is reached.

Brightness between x^{th} and y^{th} fireflies varies with respect to the distance and brightness decreases as the distance increases. A firefly's brightness which is seen by another firefly, is calculated as in Equation 4.

$$B(a) = B_0 e^{-\gamma a} \quad (4)$$

where B_0 is the brightness at $a = 0$ (a is the distance between any two fireflies), γ is a light absorption coefficient which controls the decrease in light intensity. x^{th} firefly is attracted by y^{th} firefly, and the movement is given in Equation 5.

$$m_x^{in+1} = B_0 e^{-\gamma (r_{xy})^2}\left(m_y^{in} - m_x^{in}\right) + \beta(RN - 0.5) \quad (5)$$

where in, β and RN mean iteration number, randomization parameter and random number generator, respectively. RN uniformly generates numbers between 0 and 1. Also, r_{xy} is the distance between the x^{th} firefly and the y^{th} firefly, which is defined by the Equation 6. In this equation, T $(i = 1 \ldots T)$ is the number of dimensions. In this study, T is 30 indicating the total number of features calculated by KFD.

$$r_{xy} = \|m_x - m_y\| = \sqrt{\sum_{i=1}^{T}(m_{xi} - m_{yi})^2} \quad (6)$$

Each firefly is moving in a direction in the searching space to find the effective feature(s) based on the accuracy of the classifier model with the selected effective features [14].

3) Classification

The k-NN algorithm stores all current states and classifies new states according to a similarity measure. This method classifies data based on similarity measure by examining the k closest neighboring distances [20]. In this study, we tested cosine as the distance measurement as given in Equation 7.

$$\cos(\theta) = \frac{f \cdot g}{||f||\,||g||} = \frac{\sum_{i=1}^{T} f_T(X_i) g_T(Y_i)}{\sqrt{\sum_{i=1}^{T} f_T(X_i)^2}\ \sqrt{\sum_{i=1}^{T} g_T(Y_i)_i^{\,2}}} \quad (7)$$

where T represents the number of features. Additionally, $f_T(X_i)$ and $g_T(Y_i)$ state the value of T^{th} feature of X_i and the value of T^{th} feature of Y_i, respectively. Also, X_i and Y_i are the components of vector X and Y, respectively. It is worthwhile mentioning that in this study, we calculated SE and SP values as correct classification percentage of the class 0 and class 1. Furthermore, the optimum k parameter was decided by random sub-sampling method.

III. Results And Discussion

In this study, KFD based features were extracted from the MI based EEG signals which were recorded from 29 subjects. Then, effective feature(s) were selected using by FA. These selected feature(s) were classified with k-NN for 100-times run using the cross validation method. CA, SE and SP values calculated for 29 people with the selected effective features are given as radar plot in Figure 4. In Figure 4, we achieved an average 76.14% CA, 76.52% SE and 75.77% SP, respectively.

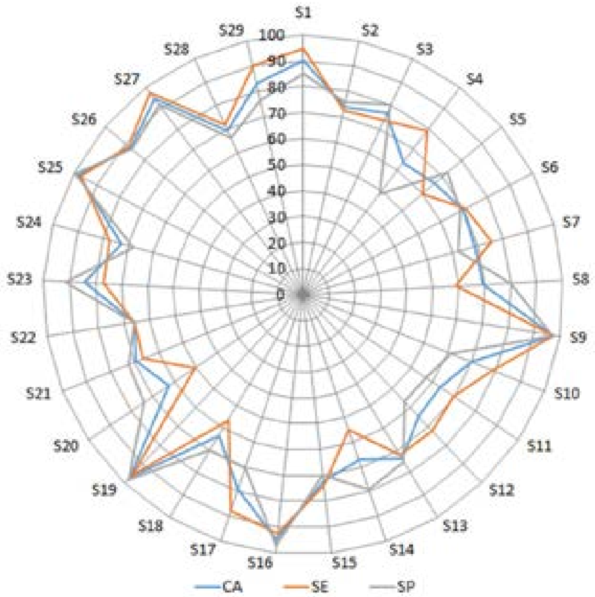

Figure 4. CAs, SEs and SPs calculated with FA

To show the effectiveness of the proposed method, we also obtained classification results using all features. In Figure 5, we showed the classification results of 29 subjects with radar plots in terms of CAs, SEs and SPs. Their average values were calculated as 71.74, 72.05% and 71.43% respectively. In these radar graphs, all subjects from 1 to 29 are expressed as $S1, S2, S3, \ldots \ldots . S29$, respectively.

Moreover, we calculated kappa coefficients as given in Figure 6 and 7 for 29 subjects with cross-validation analysis over 100-times run on the with FA (with selected effective feature(s)) and all feature. According to these values, average 0.52 and 0.43 kappa coefficients were achieved.

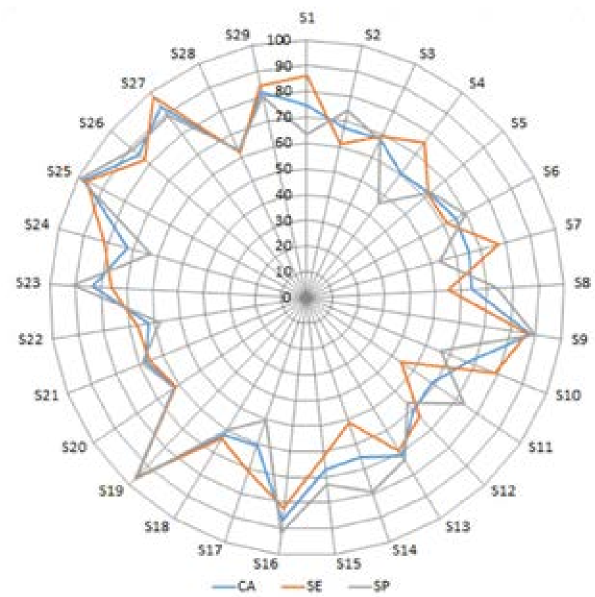

Figure 5. CAs, SEs and SPs calculated with all features

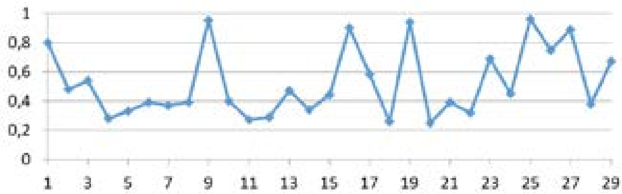

Figure 6. Kappa coefficients with FA calculated from 29 subjects

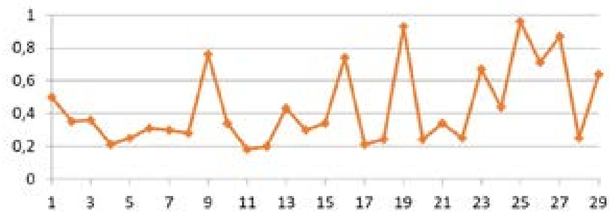

Figure 7. Kappa coefficients with all features calculated from 29 subjects

We used a FA for feature selection to choose minimal number of features and to obtain even better classification accuracy by utilizing all features. The achieved results verified that FA is an e□ective search algorithm for feature selection problems. Further, effective feature(s) selected by FA can enhance the performance of right/left hand opening-closing motor imagery BCI application.

References

[1] O. Aydemir, & E. Ergün, "A robust and subject-specific sequential forward search method for effective channel selection in brain computer interfaces", Journal of neuroscience methods, vol. 313, pp. 60-67, 2019.

[2] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy & F. Yger, "A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update", Journal of neural engineering, vol. 15(3), pp. 031005, 2018.

[3] E. Yavuz, & Ö. Aydemir, "Olfaction recognition by EEG analysis using wavelet transform features", IEEE, In 2016 International Symposium on INnovations in Intelligent SysTems and Applications (INISTA), Romania, Sinaia, pp. 1-4, August 2016.

[4] Z. Jin, G. Zhou, D. Gao, & Y. Zhang, "EEG classification using sparse Bayesian extreme learning machine for brain–computer interface", Neural Computing and Applications, pp.1-9, 2018.

[5] H. Zhiping, C. Guangming, C. Cheng, X. He, & Z. Jiacai, "A new EEG feature selection method for self-paced brain-computer interface", IEEE, In 2010 10th International Conference on Intelligent Systems Design and Applications, Cairo, Egypt, pp. 845-849, November 2010.

[6] M. Anbu, & G. A. Mala, "Feature selection using firefly algorithm in software defect prediction", Cluster Computing, pp. 1-10, 2017.

[7] G. Chandrashekar, & F. Sahin, "A survey on feature selection methods", Computers & Electrical Engineering, vol. 40(1), pp. 16-28, 2014.

[8] P. Govardhan, & P. A Prasad, "Survey on Approaches to Feature Selection. International Journal of Advanced Engineering", Management and Science (IJAEMS), pp 265-269, 2017.

[9] H. Frohlich, O. Chapelle, & B. Scholkopf, "Feature selection for support vector machines by means of genetic algorithm", IEEE, In Proceedings. 15th IEEE International Conference on Tools with Artificial Intelligence, Sacramento, CA, USA, pp. 142-148, November 2003.

[10] E. Zorarpacı, & S. A. Özel, "A hybrid approach of differential evolution and artificial bee colony for feature selection", Expert Systems with Applications, vol. 62, pp .91-103, 2016.

[11] H. H. Inbarani, A. T. Azar, & G. Jothi, "Supervised hybrid feature selection based on PSO and rough sets for medical diagnosis", Computer methods and programs in biomedicine, vol. 113(1), pp. 175-185, 2014.

[12] K. C. Hsu, & S. N. Yu, " Detection of seizures in EEG using subband nonlinear parameters and genetic algorithm", Computers in Biology and Medicine, vol. 40(10), pp. 823-830, 2010.

[13] P. Rakshit, S. Bhattacharyya, A. Konar, A. Khasnobish, D. N. Tibarewala & R. Janarthanan, Artificial bee colony based feature selection for motor imagery EEG data", In Proceedings of Seventh International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA 2012), Springer, India, pp. 127-138, 2013

[14] A. Liu, K. Chen, Q. Liu, Q. Ai, Y. Xie, & A. Chen, " Feature selection for motor imagery EEG classification based on firefly algorithm and learning automata", Sensors, vol. 17(11), pp. 2576, 2017.

[15] BBCI toolbox: It was accessed from https://github.com/bbci/bbci_public/ on 01.04.2017.

[16] J. Shin, A. von Lühmann, B. Blankertz, DW. Kim, J. Jeong, HJ. Hwang. & KR. Müller, "Open access dataset for EEG+ NIRS single-trial classification", IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 25, no. 10, pp. 1735_1745, October 2017.

[17] S. Kesić, & S. Z. Spasić, "Application of Higuchi's fractal dimension from basic to clinical neurophysiology: a review", Computer methods and programs in biomedicine, vol. 133, pp. 55-70, 2016.

[18] M. Ahmadlou, H. Adeli, & A. Adeli, "Fractality and a wavelet-chaos-neural network methodology for EEG-based diagnosis of autistic spectrum disorder", Journal of Clinical Neurophysiology, vol. 27(5), pp. 328-333, 2016.

[19] R. Diao, & Q. Shen, " Nature inspired feature selection meta-heuristics", Artificial Intelligence Review, vol. 44(3), pp. 311-340, 2015.

[20] A. Sharmila, & P. Geethanjali, " DWT based detection of epileptic seizure from EEG signals using naive Bayes and k-NN classifiers", Ieee Access, vol. 4, pp. 7716-7727, 2016.

[21] B. Blankertz, G. Curio, & K. R. Müller, "Classifying single trial EEG: Towards brain computer interfacing", In Advances in neural information processing systems, pp. 157-164, 2002.