International Core Journal of Engineering Volume 6 Issue 10, 2020
ISSN: 2414-1895 DOI: 10.6919/ICJE.202010_6(10).0018

Classification of Two Tasks Based on Electroencephalogram in Schizophrenia People and Healthy Control Group

Rui Xu[1,a], Xiaoxia Huang[1,b]

[1]Shanghai Maritime University, Shanghai 201306, China.

[a]xurui74@stu.shmtu.edu.cn, [b]xxhuang@shmtu.edu.cn

Abstract

Schizophrenia is a chronic and severe mental disorder that affects how a person thinks, perceives and behaves. People with schizophrenia may be unable to predict impending things during cognitive activities. Electroencephalogram (EEG) is an effective and applicable way with the advantage of high temporal resolution to record brain activities during cognitive task. The aim of this paper is to try to explore the differences of two kinds of people: schizophrenia people and healthy control people under two types of cognitive task in terms of classification accuracies based on EEG. We applied two prevailing feature extraction methods: common spatial pattern (CSP) and sample-entropy to represent EEG data. Linear discrimination analysis (LDA), support vector machine (SVM), and K-nearest neighbor (KNN) were used to cope with extracted features. Further, in order to avoid dimensional disaster and improve classification accuracies, whole features, genetic algorithm (GA)-based features and reliefF-based features were respectively tested to obtain and compare the accuracies. The results showed that the average classification accuracies of two tasks in healthy control group were higher than that in schizophrenia people group. Each of the accuracy was the mean of 5-fold cross validation scores. An average classification accuracy of 86% was obtained for 32 subjects in healthy control group by GA+ LDA. In contrast, GA+LDA achieved 79% average accuracy in Schizophrenia group. The study proved that sample-entropy and CSP could effectively represent EEG signals during cognitive activities. And genetic algorithm (GA) based feature selection method actually can help to improve classification accuracies. These could provide a practical model for EEG analysis.

Keywords

Schizophrenia; Electroencephalogram (EEG); Sample-entropy; Common spatial pattern (CSP); Machine learning; genetic algorithm (GA); ReliefF.

1. Introduction

Schizophrenia is a mental disorder, which is characterized by a wide spectrum of disturbances when performing several tasks [1]. These disturbances including misperceptions and misinterpretations might result from a basic inability to make valid predictions about expected sensations and experiences [2]. If predictive mechanisms are dysfunctional ,sensations that should have been predicted but were not might take on inappropriate salience [2].

Electroencephalogram (EEG) is an effective way with high-temporal-resolution to record brain signals during cognitive activities. In [2], the EEG from two groups: diagnosed schizophrenia patients and healthy control people under three cognitive tasks, had been recorded to stimulate auditory N1 (a negative deflection in brain wave 100 milliseconds after onset of a sound), one component of the event-related-potentials (ERPs). This experiment tended to justify whether the damage of experience

ISSN: 2414-1895 DOI: 10.6919/ICJE.202010_6(10).0018

copy function can affect to generate corollary discharge in schizophrenia patients' group. The authors [2] measured N1 amplitudes under two tasks respectively to compute and analyze significance levels between two groups by analysis of variance (ANOVA). The results of ANOVA showed that N1 was suppressed in healthy control group due to normal experience copy function while it was not significant in Schizophrenia patients' group.

With the developments of feature engineering and machine learning techniques, more and more applications based on EEG have been come up and achieved great performance. Li et al. [3] adapted orthogonal empirical mode decomposition (OEMD) and common spatial pattern (CSP) to extract features and support vector machine (SVM) to do classification which achieved high accuracies during motion imaginary tasks. In [4], authors summarized 6 kinds of feature extraction methods: fracture dimension, statistics, spectral band power, higher order crossing ,signal energy and Hjorth feature in EEG-based emotion recognition. Then features extracted were evaluated and ranked by intraclass correlation coefficient (ICC) for classification. Amin et al. [5] proposed wavelet transform coping with EEG signals of epilepsy patients. Wavelet transform is well-fitted for non-stationary signals due to its scaling attribute for both time domain and spectrum. Kai-Cheng Hsu and Sung-Nien Yu [6] used sub-band nonlinear parameters and genetic algorithm to detect seizures in EEG. The EEG signals were decomposed by wavelet method to extract non-linear features for several bands. Then genetic algorithm worked with features to select the best combination. The Min-Redundancy-Max-Relevance was used to characterize suitability of a feature subset in [7] and achieved well performance during classification of emotion recognition based on EEG.

The aim of this study is to use feature extraction and machine learning techniques to explore the distinctions among different people under two tasks. We applied two effective and efficient feature extraction methods: sample-entropy and common spatial pattern (CSP). Sample-entropy is greatly fitted with unsteady EEG signals due to its non-linear trait and CSP is well-known in working with motion imagination tasks. We also used three excellent classifiers: linear discrimination analysis (LDA), support vector machine (SVM) and K-nearest neighbor (KNN) to handle with features. Moreover, features extracted from every channel might cause dimensional disaster and eliminate classification accuracies. To avoid this, we performed genetic algorithm (GA) and ReliefF technique to select best feature combination to achieve efficiently higher accuracies. The GA, a bionic algorithm, uses fitness function to evaluate every features-combination in population and selects the best from the offspring after cross-over and mutation. The ReliefF method measures features though their weights calculated.

The rest of this paper is organized as follows. Section2 covers dataset description and feature extraction, selection and classification methods in general. Section 3 displays results and discussions. Finally, the conclusion is shown in section 4.

2. Method and material

2.1 Dataset description and data preprocess

The EEG data used in this research were obtained from both healthy and Schizophrenia subjects, which are available on Kaggle (https://www.kaggle.com/broach/button-tone-sz). In this dataset, there are 81 subjects totally, 49 are the people with diagnosed (DSM-5) schizophrenia (SZ) while the others are in healthy control group (HC). And we chose 32 SZ and 32 HC depending on the qualities of data. Each subject conducted three kinds of task, task 1 and task 2 were used for classification. Task 1 was that the subject pressed the button at time 0s to trigger a 1000Hz , 80dB tone without delay between press and tone onset (Button tone) and the task 2 was that playing back tone at time 0s without pressing button (Play back tone) [2]. Fig.1 illustrates the tasks' procedures. Each task had been done for 100 times to record EEG. The EEG data were recorded through 64 channels and 8 external sites using a BioSemi ActiveTwo system [2]. EEG data were continuously sampled at 1024 Hz and referenced off-line to averaged earlobe electrodes [2]. We implemented a band-pass filter between 0.5Hz and 36 Hz to remove artifacts and direct drifts. Then, the EEG data were split into 3000-ms

ISSN: 2414-1895 DOI: 10.6919/ICJE.202010_6(10).0018

epochs and were baseline corrected at -100ms-0ms. The whole EEG process workflow is shown in Fig.2. MNE-Python [8] was used to process the data.

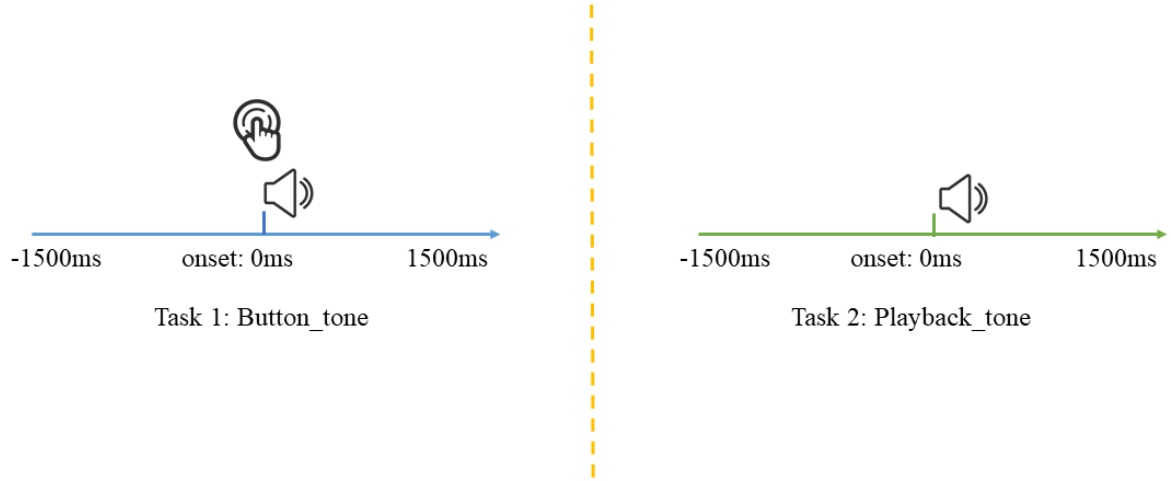

Fig1. Experimental timing scheme

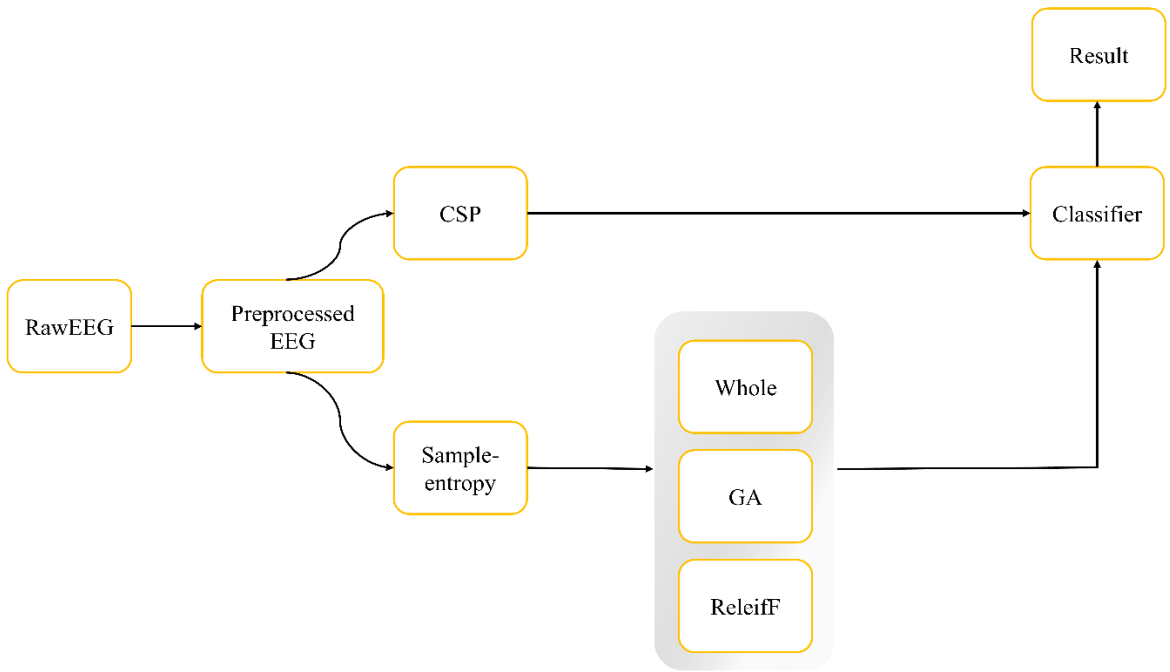

Fig.2 Block graph of EEG process

2.2 Sample-entropy Common spatial pattern (CSP)

Sample-entropy measures the regularity of the EEG signal and is independent comparing with approximate entropy [9]. Assume that x represents a N-points EEG signal, we separate x into N-(m+1)τ segments with the embedding dimension m and time delay τ (in this paper, m=3, $\tau = 1$). The r is a threshold value (in this paper, r=0.2*std(x)) to binarize the Chebyshev distance between each pair of segments. The similarity between segments can be computed as $C_i^m(r)=\frac{1}{N-(m-1)\tau}\sum_{i\neq j} \text{distance}(x_i, x_j)$, where x_i and x_j are two different segments. If the distance between two segments is larger than r, we consider the segments similar and label $\text{distance}(x_i, x_j) = 1$, if not, we label $\text{distance}(x_i, x_j) = 0$. The sample-entropy can be computed as:

ISSN: 2414-1895 DOI: 10.6919/ICJE.202010_6(10).0018

$$A(m,r)=\frac{1}{N-m+1}\sum_{i=1}^{N-m+1}\log(C_i^m(r)) \quad (1)$$

$$\text{sample-entropy}=-\log\frac{A(m+1,r)}{A(m,r)} \quad (2)$$

High value of sample-entropy implies that the signal is highly unpredictable and a low sample entropy value implies the signal is predictable [9] .

2.3 Common spatial pattern (CSP)

Common spatial pattern (CSP) is a method that uses a linear transform to project multichannel EEG data onto a low-dimensional spatial subspace with a projection matrix, W, of which each row consists of weights for each of channels [10]. The principle of the CSP algorithm is to find a set of optimal spatial filters for projection by means of diagonalization of the matrix, so as to maximize the variance difference between the two types of signals. Previous research has demonstrated deficits in preresponse motor activity in schizophrenia, as evidenced by a reduced lateralized readiness potential (LRP) [11]. Therefore, we hypothesized that the readiness potentials (RP) might be discriminative in HC group comparing with that in SZ group. Then, we cropped EEG data into two time periods which are -170ms-0ms and 0ms-170ms to apply CSP for feature extraction. For distinguishing two time periods, CSP1 represented that the data were ranging from -170ms to 0ms and CSP2 indicated that the data ranging from 0ms to 170ms were used. The detailed procedures of CSP can be seen in [3, 10].

2.4 Genetic algorithm (GA)

The genetic algorithm (GA) has been used for searching the optimum feature subset that maximizes the classification accuracies [6]. For purpose of improving accuracies efficiently, the well discriminative features were selected for further process and the insignificant features were jettisoned. In this study, we firstly set 50 as the initial number of population, each individual in population owned 65 genes. Each gene represented the feature extracted from one certain channel. The genes were set to binary code, 1 meant that the corresponding feature was available while 0 meant that the feature was excluded. When all individuals prepared, we calculated the fitness values of each individual for ranking and selecting the best parents' individuals. The fitness function is

$$f=\frac{err_1+err_2+err_3}{3} \quad (3)$$

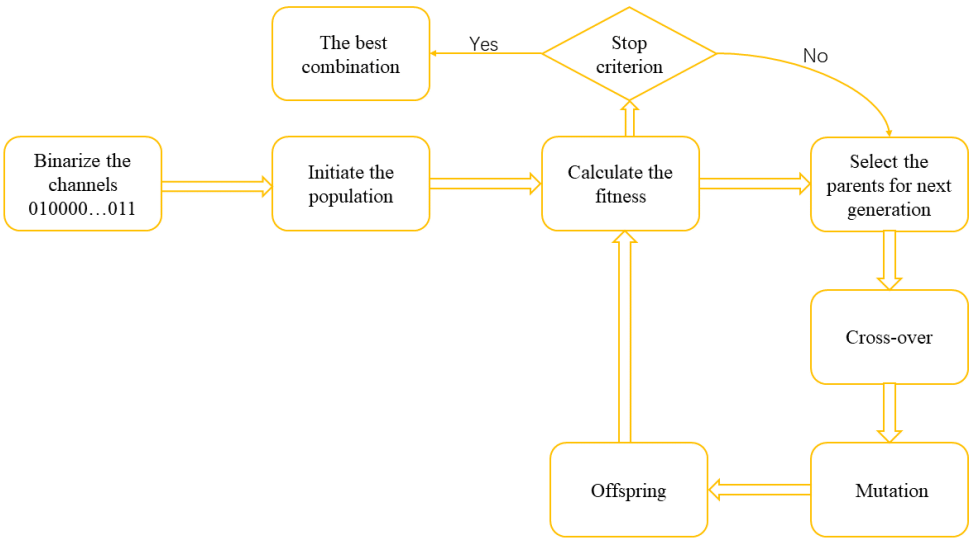

Fig.3 Architecture of genetic algorithm (GA)

ISSN: 2414-1895 DOI: 10.6919/ICJE.202010_6(10).0018

where the err_i $(i = 1, 2, 3)$ is the misclassification rates by three classifiers. The GA [12] tried to eliminate the fitness function. In this paper, we picked the top 25 individuals as parents for the next generation. When the parents produced, cross-over which meant that the two individuals had to choose their own genes to make up a new individual called offspring had to be done. The mutation followed by cross-over is a procedure that each new made-up individual has to change some genes randomly in order to avoid local optimization issue. After cross-over and mutation, the offspring were sent to fitness function to evaluate performance. The whole procedure ended until meeting the stop criterion. Fig.3 shows procedures of GA.

2.5 ReliefF

ReliefF [13, 14] is a filter method that evaluates every features by their weights. And features with higher weights mean more distinctive and typical. In this paper, we tried to select 10 features with high weights from all 65 features to reduce redundancies and improve accuracies. The main idea of ReliefF is that we chose one sample R from the sample sets. Then respectively picking K nearest neighbors from the same class and K nearest neighbors from different class by computing the Euclidean distance. Next, calculating the weights until all samples have been done. The detailed steps can be found in [13].

2.6 Classifier

In this paper, we sent features to three classifiers: linear discrimination analysis (LDA), support vector machine (SVM) and K-nearest neighbor (KNN). The aim of LDA is to create a new variable that is a combination of the original predictors. This is accomplished by maximizing the differences between the predefined groups, with respect to the new variable [15]. The SVM projected the data onto a higher dimensional space by kernel tricks and tried to separate the data by a hyper-plane in the condition of maximal margin. We adapted SVM with RBF kernel to classify the samples. K- Nearest Neighbor (KNN) classifier is a supervised learning algorithm, which categorizes a new instant based on majority votes of the nearest neighbors [16].

3. Results and discussion

We finally obtained classification results summarized in Table 1 and Table 2. These tables show each subject's accuracies by different method combinations. Each of the accuracy is the mean of 5-fold cross validation scores. The accuracy is defined as follow:

$$\mathrm{acc_i} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \tag{4}$$

$$\mathrm{ACC} = \frac{\sum_i \mathrm{acc_i}}{\text{Number of folders}} \tag{5}$$

where i means the number of folders

Table 1 Average classification accuracies (%) and standard deviation (%) by sample-entropy of 5-fold cross-validations of 2 groups

Method	Whole + LDA	Whole + SVM	Whole + KNN	GA + LDA	GA + SVM	GA + KNN	ReliefF + LDA	ReliefF +SVM	ReliefF + KNN
HC	79.3 (9.7)	78.9 (9.9)	74.2 (10.5)	84.6 (9.3)	82.0 (9.5)	80.4 (8.8)	76.1 (10.6)	76.4 (10.5)	74.4 (11.2)
SZ	76.7 (11.2)	78.7 (10.2)	73.1 (10.4)	83.1 (10.)	83.1 (8.7)	79.0 (8.8)	75.2 (10.9)	76.6 (10.8)	73.8 (11.8)

ISSN: 2414-1895 DOI: 10.6919/ICJE.202010_6(10).0018

Table 2 Average classification accuracies (%) and standard deviation (%) by csp of 5-fold cross-validations of 2 groups

Method	CSP1+LDA	CSP1+SVM	CSP1+KNN	CSP2+LDA	CSP2+SVM	CSP2+KNN
HC	73.3(14.9)	73.1(14.9)	72.4(14.7)	63.1(11.2)	63.0(10.5)	62.6(11.1)
SZ	67.9(13.9)	68.1(13.7)	67.4(13.3)	63.5(11.7)	63.2(11.4)	62.4(11.8)

3.1 Feature extraction method

In this paper, we used two feature extraction methods, one is common spatial pattern (CSP), the other is sample-entropy. CSP performed well in field of motion imagination while sample entropy can fit EEG signals well due to its non-linear attribution. We sent features to the classifiers and the results are summarized in Table 1 and Table 2. In healthy control (HC) group, the GA+SVM achieved the highest accuracy 84.6%(9.3%) comparing with other combination methods. And the CSP1+LDA, CSP1+SVM, CSP1+KNN reached 73.3%(14.9%), 73.2%(14.9%), 72.4%(14.7%) respectively. In the group of schizophrenia people, the accuracies of GA method were better than other methods and reached 83.1%(10.%), 83.1%(8.7%) and 79.0%(8.8%) with classifiers LDA, SVM and KNN. By contrast, the CSP1+classifiers achieved accuracies 67.9%(13.9%), 68.1%(13.7%), 67.4%(13.3%) as shown in Table 2.

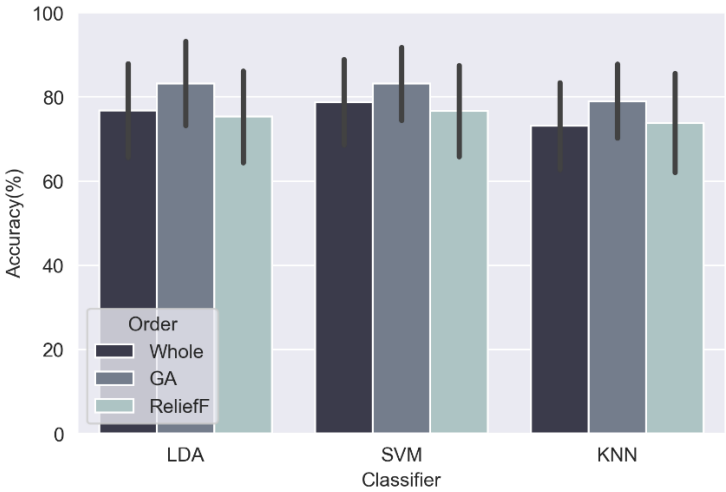

Fig.4 Comparison of average classification accuracies (%) of three-order sample-entropy features in HC group

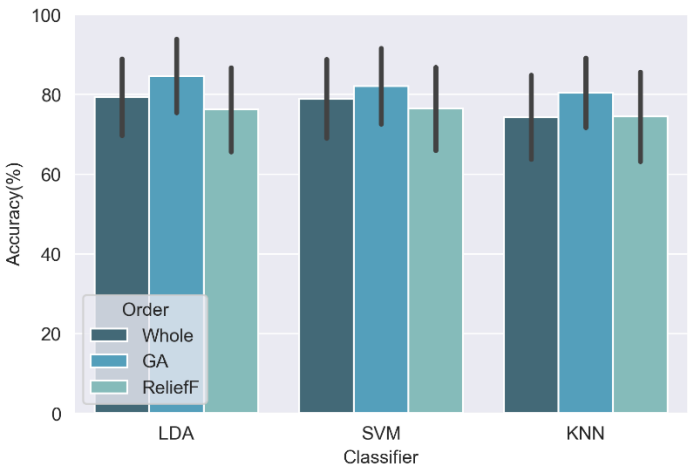

Fig.5 Comparison of average classification accuracies (%) of three-order sample-entropy features in SZ group

ISSN: 2414-1895 DOI: 10.6919/ICJE.202010_6(10).0018

3.2 Feature selection and classifier performance

In this work, we conducted three different orders: whole features, GA-based features and ReliefF-based features to test their performances. The results are shown in Table 1 and Table 2. Among all subjects, the GA-based feature selection method stood out from the others' combinations. The GA+SVM in HC group and SZ group achieved 82.0%(9.5%) and 83.1%(8.7%) accuracies respectively. And CSP1 achieved higher classification accuracies comparing with CSP2 in groups which meant that the EEG data in preparation stage were more distinguished. Three well-known classifiers were used in this work. As Fig.4 and Fig.5 displayed, the SVM achieved the best performance while LDA was slightly better than KNN.

3.3 Discussion

We supposed that the classification accuracies between task 1 and task 2 could vary due to the existence of experience copy function and N1 suppression. The results showed that the average accuracies in HC group were slightly higher than that in SZ group. In both HC and SZ groups, CSP1 achieved better performance than CSP2 did in terms of accuracies, which indicated that the EEG in preparation stage are more distinguished between two tasks. Further, the difference values of accuracies between CSP1 and CSP2 are about 10% in HC group, but about 5% in SZ group. This proved that SZ existed defects when they were in preparation stage.

4. Conclusion

We applied two feature extraction methods and three classifiers on EEG during two cognitive tasks. The results showed the classification accuracies in healthy control group were slightly better than that in patients' group. The genetic algorithm (GA) can improve accuracies effectively comparing with other feature selection methods. We also found different time stages would affect the classification accuracies by CSP, and people in SZ existed defects during preparation stage. As for limitations, 64 subjects' EEG is still small to summarize the existence of distinguishes in two groups during two tasks. In the future, more feature extraction methods and neural network techniques could be applied to test and compare.

Acknowledgements

The authors would like to thank Brain Roach and Dr. Ford for sharing their EEG dataset. Data used in this study were initially collected with funding from the National Institute of Mental Health (NIMH grant number R01MH058262).

References

[1] M. E. Shenton, C. C. Dickey, M. Frumin, and R. W. McCarley, "A review of MRI findings in schizophrenia," Schizophr. Res., vol. 49, no. 1–2, p. 1–52, 2001.

[2] J. M. Ford, V. A. Palzes, B. J. Roach, and D. H. Mathalon, "Did i do that? Abnormal predictive processes in schizophrenia when button pressing to deliver a tone," Schizophr. Bull., vol. 40, no. 4, p. 804–812, 2014.

[3] L. Mingai, G. Shuoda, Y. Jinfu, and S. Yanjun, "A novel EEG feature extraction method based on OEMD and CSP algorithm," J. Intell. Fuzzy Syst., vol. 30, no. 5, p. 2971–2983, 2016.

[4] Z. Lan, O. Sourina, L. Wang, Y. Liu, R. Scherer, and G. R. Müller-Putz, "Stable feature selection for EEG-based emotion recognition," Proc. - 2018 Int. Conf. Cyberworlds, CW 2018, p. 176–183, 2018.

[5] H. U. Amin et al., "Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques," Australas. Phys. Eng. Sci. Med., vol. 38, no. 1, p. 139–149, 2015.

[6] K. C. Hsu and S. N. Yu, "Detection of seizures in EEG using subband nonlinear parameters and genetic algorithm," Comput. Biol. Med., vol. 40, no. 10, p. 823–830, 2010.

[7] R. Jenke, A. Peer, and M. Buss, "Feature extraction and selection for emotion recognition from EEG," IEEE Trans. Affect. Comput., vol. 5, no. 3, p. 327–339, 2014.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202010_6(10).0018

[8] A. Gramfort et al., “MEG and EEG data analysis with MNE-Python,” Front. Neurosci., vol. 7, no. 7 DEC, p. 1–13, 2013.

[9] U. R. Acharya, H. Fujita, V. K. Sudarshan, S. Bhat, and J. E. W. Koh, “Application of entropies for automated diagnosis of epilepsy using EEG signals: A review,” Knowledge-Based Syst., vol. 88, pp. 85–96, 2015.

[10] M. Arvaneh, S. Member, C. Guan, K. K. Ang, and C. Quek, “Optimizing the Channel Selection and Classification Accuracy in EEG-Based BCI,” IEEE Trans. Biomed. Eng., vol. 58, no. 6, p. 1865–1873, 2012.

[11] E. S. Kappenman et al., “Response activation impairments in schizophrenia: Evidence from the lateralized readiness potential,” Psychophysiology, vol. 49, no. 1, p. 73–84, 2012.

[12] I. H. Hasan, A. R. Ramli and S. A. Ahmad, "Utilization of Genetic Algorithm for Optimal EEG Channel Selection in Brain-Computer Interface Application," 2014 4th International Conference on Artificial Intelligence with Applications in Engineering and Technology, Kota Kinabalu, 2014, p. 97-102, 2014.25.

[13] J. Zhang, M. Chen, S. Zhao, S. Hu, Z. Shi, and Y. Cao, “ReliefF-based EEG sensor selection methods for emotion recognition,” Sensors (Switzerland), vol. 16, no. 10, p. 1–15, 2016.

[14] R. Jenke, A. Peer, and M. Buss, “Feature extraction and selection for emotion recognition from EEG,” IEEE Trans. Affect. Comput., vol. 5, no. 3, p. 327–339, 2014.

[15] A. Subasi and M. I. Gursoy, “EEG signal classification using PCA, ICA, LDA and support vector machines,” Expert Syst. Appl., vol. 37, no. 12, p. 8659–8666, 2010.

[16] S. S. Poorna et al., “Classification of EEG based control using ANN and KNN-A comparison,” 2016 IEEE Int. Conf. Comput. Intell. Comput. Res. ICCIC 2016, p. 0–5, 2017.