Journal of Computer Science
https://dergipark.org.tr/en/pub/bbd

ISSN,e-ISSN: 2548-1304
Volume:IDAP-2022, pp:32-36, 2022
https://doi.org/ 10.53070/bbd.1173093
Research Paper

Analysis and Classification of Schizophrenia Using Event Related Potential Signals

Anıl Aksöz [1], Doğukan Akyüz [1], Furkan Bayır [1], Nevzat Can Yıldız [1], Fırat Orhanbulucu [*1,2], Fatma Latifoğlu [1]

[1]Department of Biomedical Engineering, Erciyes University, Kayseri, Turkey

[2]Department of Biomedical Engineering, Inonu University, Malatya, Turkey

(anilaksoz05@gmail.com, dogukanaz26@gmail.com, furkanbayir35@gmail.com, nevzatcanyldz@gmail.com, firat.orhanbulucu@inonu.edu.tr, flatifoglu@erciyes.edu.tr)

Received:Sep.08,2022 Accepted:Sep.16,2022 Published:Oct.10,2022

Abstract— Schizophrenia (SZ) is a neuropsychiatric disease that affects many people around the world and causes death if not diagnosed and treated early. One of the commonly used methods for early diagnosis is electroencephalography (EEG). The application of signal processing and machine learning methods to EEG signals can support experts and researchers who want to determine SZ disease. In this study, event-related potential (ERP) signals were obtained from the recorded EEG signals as a result of sending auditory stimuli to the SZ patient and healthy control (HC) group. P300 amplitude-latency, hjorth parameters and entropy values were calculated as features from these signals. The features obtained were evaluated with Support Vector Machines (SVM), K-Nearest Neighbor (KNN) and Artificial Neural Networks (ANN) classifiers to distinguish SZ patients from the HC group. In this study, the most successful result was obtained in the ANN classifier with an accuracy rate of 93.9%.

Keywords: *Schizophrenia, electroencephalography, event-related potential, artificial neural networks.*

1. Introduction

Although it varies from patient to patient, SZ is a serious neuropsychiatric disorder in which people generally have hallucinations and have difficulty expressing themselves [1]. According to the World Health Organization (WHO), approximately 24 million people, or 1% of the population in the world are affected by SZ [1, 2]. According to the WHO, there are many causes of SZ. SZ is a disease that can lead to death at an early age if left untreated. WHO stated that SZ is treatable if diagnosed early [2, 3]. Therefore, early diagnosis of the disease is very important. Electrophysiological and neuroimaging methods are used for the early diagnosis of neuropsychiatric and similar neurological disorders such as SZ [3, 4]. Due to the cost of brain imaging methods, EEG or ERP signals are preferred more in studies [3, 4]. EEG is a non-invasive imaging method that is more cost-effective than other imaging techniques. Thanks to the multi-channel EEG, important information about brain activitiy can be obtained. ERP is, on the other hand signals that can occur as a result of an auditory, visual or neural stimulus from EEG signals [5]. ERP signals are obtained by averaging the recorded EEG signals as a result of these target stimuli being sent many times. It has been seen in studies that the components of ERP signals provide important information about the diagnosis of diseases [4-6].

With EEG and ERP signals, early diagnosis of neurological diseases such as SZ, Parkinson's, ALS, and Alzheimer's can be made easily. In the literature, some studies on these diseases using EEG and ERP signals have been examined [3, 6-10]. In their study [3] Siuly et al. determined SZ with the Ensemble Bagged Tree (EBT) classifier with an accuracy rate of 89.59% using EEG and empirical mode decomposition (EMD) methods.In study with EEG signals obtained from healthy people and SZ patients, Devia et al. found the accuracy rate as 71% by using the Linear Discriminant Analysis (LDA) algorithm [6]. In another study on SZ detection, Zhang used EEG, and ERP signals and achieved an accuracy rate of 81.10% with the Random Forest (RF) algorithm [8]. Boostani et al. estimated SZ disease with 87.5% accuracy by extracting features from EEG signals obtained from healthy-SZ patients and applying them as input to the LDA classifier [10].

In this study, ERP signals were created from EEG signals obtained from SZ patients and the healthy control (HC) group. Some features are extracted by applying signal processing methods to these generated signals. These features are used as input in SVM, KNN, and ANN machine learning algorithms to differentiate SZ patients from the HC group. While the information about the data set and methods used in the study is given in the second part, the results and discussion are explained in the third part.

2. Materials and Methods

2.1. Data Set and EEG Recording

The data set used in the study was shared as open access and was obtained from the Kaggle website [11]. EEG signals were obtained from 49 SZ patients and 32 HC groups at the National Institute of Mental Health. The data was collected from 64 channels, but "Fz, FCz, Cz, FC3, FC4, C3, C4, CP3, CP4" channels were used in the study. The positions of the channels used in the study are given in Figure 1. EEG signals sampled at 1024 Hz were passed through a 0.5-15 Hz band-pass filter. During the recording of EEG signals, auditory stimuli were sent to the subjects. EEG signals were recorded by sending auditory stimuli to the subjects with three different tasks: pressing the button after hearing the sound, not pressing the button despite the sound, and pressing the button even though there was no sound. Detailed information about the data set and the experimental phase is available in [11, 12].

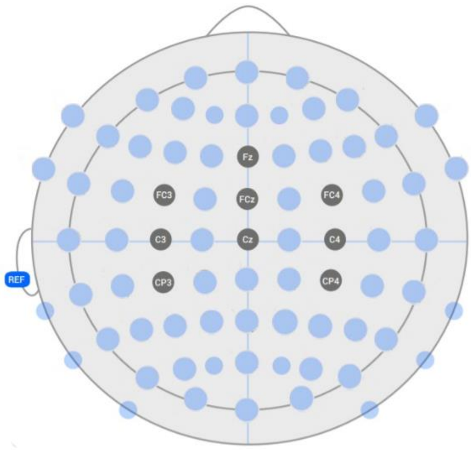

Figure 1. Electrode positions of channels used in the study

2.2. Acquisition and Pre-Processing of ERP Signals

The ERP signals were obtained from the EEG signals used in the study, and then the preprocessing steps were carried out. ERP signals are a small amplitude signal in EEG signals that can occur as a result of sending many target stimuli. ERP signals were obtained by averaging the EEG signals of the auditory target stimuli sent to the subjects at certain intervals. The generation of ERP signals is shown in Figure 2.

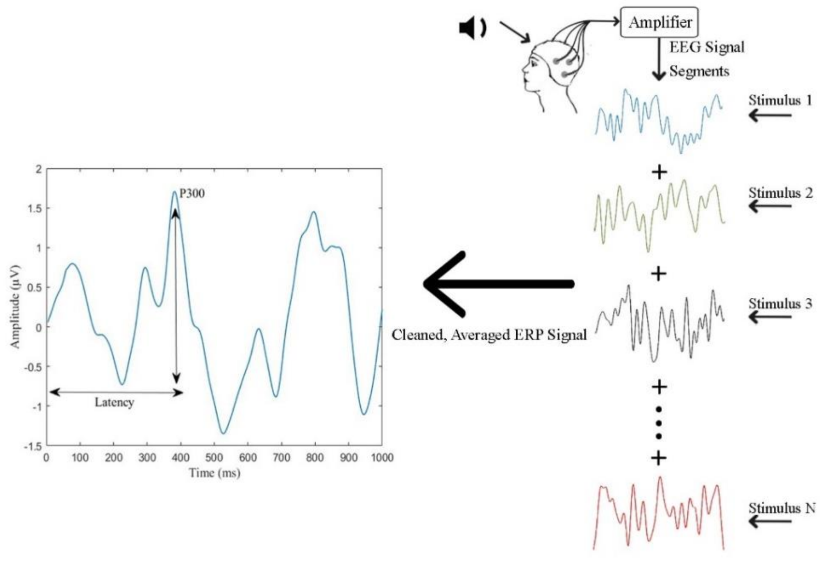

Figure 2. Stages of generating ERP signals [13]

As a pre-process in the study, after the ERP signals are obtained, the noise is eliminated by passing through the finite impulse response (FIR) filter. After the features were obtained from all nine channels, the Principal Component Analysis (PCA) process was applied and the features were made ready for the classifier. In order for the classification process to be balanced, the discrepancy in the data set was resolved by using the synthetic minority over-sampling technique (SMOTE) [14]. After the procedures, 49 SZ patients were classified as 49 HC groups.

2.3. Feature Extraction

Feature extraction is the operation applied to transform data into features. Extracted features can be evaluated by statistical interpretation or classification processes and some information can be obtained. After obtaining the ERP signals in the study, P300 amplitude and P300 latency values, which are frequently preferred in the literature, were obtained. By applying signal processing methods to ERP signals, Hjorth parameters and Entropy values are obtained from time domain. The P300 amplitude value is the maximum value seen in the ERP signal approximately 250-500 ms after the target stimulus. The time this value is observed is considered as P300 latency [15]. Hjorth parameters are a feature developed for various EEG analyses and show the activity, mobility and complexity of the signal [16]. Entropy is an indicator of the disorder in the signal. Mental changes and activations can be determined according to entropy changes [17].

2.4. Classification

Classification is a technique for estimating the relevant class using attributes as input [17]. The inputs go through the training and testing phase and a model is created according to the relationship between these inputs and the classes. The features obtained in the study were evaluated by using them as inputs in SVM, KNN, and ANN algorithms. SVM is a method that finds the situations where the distance between two classes is maximum and classifies the data by determining the boundary line on the plane. It has been shown to give successful results in electrophysiological signals [17]. The KNN algorithm is supervised learning that classifies data through a certain number of neighbors' affinity [17]. ANNs are mathematical models inspired by the information processing methods of nervous systems. Thanks to the connections established between artificial neurons, the ANN model is formed. The ANN model consists of three layers, the input layer, the hidden layer, and the output layer [18]. The ANN model in the study is a feed-forward network trained from the neural network toolbox in the Matlab program. Detailed information about the classification algorithm is explained in [18].

3. Results and Discussion

This study aimed to distinguish SZ patients and HC groups from ERP signals. For this purpose, auditory stimuli were sent to 49 SZ patients and 32 HC groups, and ERP potential signals were obtained by averaging the recorded EEG signals of the target stimuli (Figure 3). As features, P300 amplitude-latency values from ERP signals and Hjorth parameters and entropy values were obtained by applying some signal processing techniques to ERP signals. Because the obtained data did not have a balanced distribution, they were artificially reproduced and balanced using the SMOTE method. After the data became balanced, accuracy rates were obtained by evaluating with SVM, KNN, and ANN classifiers as inputs. Cross-validation (CV) was preferred as 10 in the classification process. The accuracy rates obtained in the study are given in Figure 4. Looking at Figure 4, it is seen that the most successful result was obtained with the ANN classifier with an accuracy rate of 93.9%.

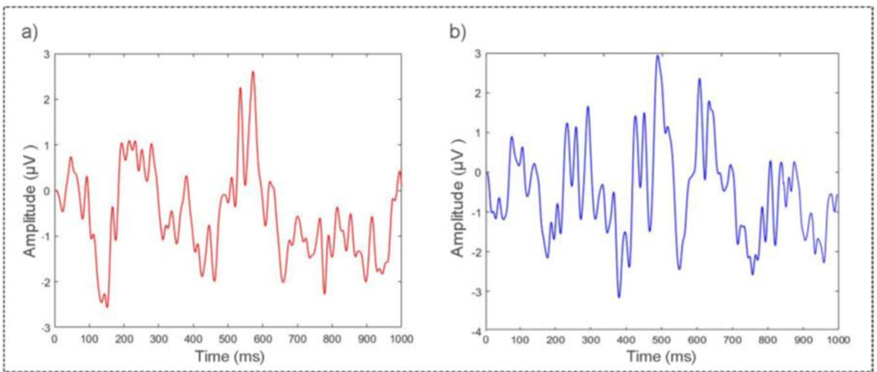

Figure 3. a) ERP signal of SZ patient b) ERP signal of HC group

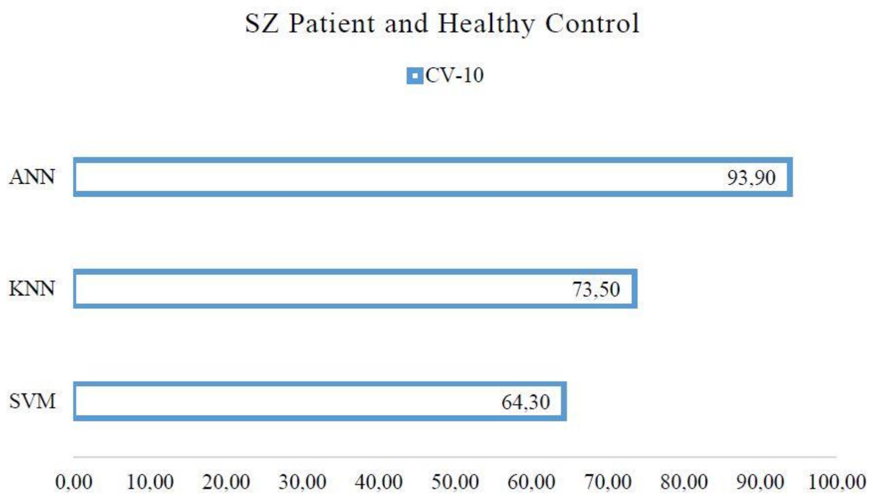

Figure 4. Accuracy rate results

Studies have shown that the ERP signal obtained as a result of auditory stimuli and the P300 amplitude-latency values of its components can be used in neuropsychological evaluations and may be a determinant for SZ disease [4, 19]. It was observed that the P300 amplitude value was decreased in SZ patients compared to healthy individuals and the P300 generation response was delayed [4]. As can be seen in the sample signals given in Figure 3, in this study, which supports the literature, P300 amplitudes were lower in SZ patients than in the HC group, while latency times occurred later. The comparison of the results of the accuracy rate obtained from this study with the studies conducted in the literature on SZ disease with EEG and ERP signals is given in Table 1. When Table 1 is examined, it is seen that the results obtained in this study compared to the studies in the literature are successful.

Table 1. Comparison of the performance of the work done with the existing studies

Study	Dataset-Method	Classifier	Accuracy (%)
Siuly et al. [3]	EEG-EMD	EBT	89.59
Devia et al. [6]	ERP- Feature extraction	LDA	71
Zhang [8]	EEG-ERP- Feature extraction	RF	81.10
Boostani et al. [20]	EEG- Feature extraction	LDA	87.5
This Study	ERP- Feature extraction	ANN	93.9

4. Conclusion

In this study, ERP signals were created from EEG signals recorded as a result of auditory stimuli from SZ patients and HC individuals. By obtaining features from the obtained ERP signals, SZ disease was tried to be predicted with machine learning algorithms. It is evaluated that the P300 amplitude-latency values obtained as features can be used in the diagnosis of the disease [4]. It is thought that the results obtained in this study may provide preliminary information about SZ disease in clinical studies in terms of P300 amplitude-latency values and may contribute to attention studies on auditory stimuli. Studies on the determination of SZ disease with machine learning methods by extracting features from EEG and ERP signals are given in Table 1. When Table 1 is examined, it is thought that evaluating the features obtained as a result of applying signal processing methods to ERP signals with the ANN classifier may contribute to the diagnosis of SZ disease.

Acknowledgment

We thank the National Institute of Mental Health (NIMH project number R01MH058262) for providing the dataset used in this study and Ford et al. for sharing it as reference [11, 12] open access.

References

1. Buettner, R., Hirschmiller, M., Schlosser, K., Rössle, M., Fernandes, M., Timm, I. J. (2019, October). High-performance exclusion of schizophrenia using a novel machine learning method on EEG data. In 2019 IEEE International Conference on E-Health Networking, Application & Services (HealthCom) (pp. 1-6). IEEE.
2. WHO. Accessed: Jul 14, 2022. [Online]. Available: https://www.who.int/mental_health/management/schizophrenia/en/
3. Siuly, S., Khare, S. K., Bajaj, V., Wang, H., & Zhang, Y. (2020). A computerized method for automatic detection of schizophrenia using EEG signals. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 28(11), 2390-2400.
4. Lapsekili, N., Uzun, Ö., Sütçigil, L., Ak, M., Yücel, M. (2011). Şizofreni Hastalarında İlk Atakta P300 Bulguları ile Nörolojik Silik İşaretler Arasındaki İlişki. Dusunen Adam: Journal of Psychiatry & Neurological Sciences, 24(3).
5. Luck, S. J. (2014). An introduction to the event-related potential technique. MIT press.
6. Orhanbulucu, F., Latifoğlu, F., Baş, A. (2020). K-Ortalamalar Kümeleme Yöntemi Kullanılarak ALS Hastalarında Dikkatin Olaya İlişkin Potansiyel Sinyalleri İle İncelenmesi. Avrupa Bilim ve Teknoloji Dergisi, 239-244.
7. Devia, C., Mayol-Troncoso, R., Parrini, J., Orellana, G., Ruiz, A., Maldonado, P. E., Egaña, J. I. (2019). EEG classification during scene free-viewing for schizophrenia detection. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 27(6), 1193-1199.
8. Zhang, L. (2019, July). EEG signals classification using machine learning for the identification and diagnosis of schizophrenia. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 4521-4524). IEEE.
9. Cavanagh, J. F., Kumar, P., Mueller, A. A., Richardson, S. P., Mueen, A. (2018). Diminished EEG habituation to novel events effectively classifies Parkinson's patients. Clinical Neurophysiology, 129(2), 409-418.
10. Boostani, R., Sadatnezhad, K., Sabeti, M. (2009). An efficient classifier to diagnose of schizophrenia based on the EEG signals. Expert Systems with Applications, 36(3), 6492-6499.
11. B. Roach, (2017). EEG data from basic sensory task in schizophrenia -button press and auditory tone event related potentials from 81 human subjects. [Online]. Available: https://www.kaggle.com/datasets/broach/button-tone-sz
12. Ford, J. M., Palzes, V. A., Roach, B. J., Mathalon, D. H. (2014). Did I do that? Abnormal predictive processes in schizophrenia when button pressing to deliver a tone. Schizophrenia bulletin, 40(4), 804-812.
13. Orhanbulucu, F., Latifoğlu, F. (2022). Detection of amyotrophic lateral sclerosis disease from event-related potentials using variational mode decomposition method. Computer Methods in Biomechanics and Biomedical Engineering, 25(8), 840-851.
14. Chawla, N. V., Bowyer, K. W., Hall, L. O., Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. Journal of artificial intelligence research, 16, 321-357.
15. Dong, S., Reder, L. M., Yao, Y., Liu, Y., Chen, F. (2015). Individual differences in working memory capacity are reflected in different ERP and EEG patterns to task difficulty. Brain research, 1616, 146-156.
16. Cecchin, T., Ranta, R., Koessler, L., Caspary, O., Vespignani, H., & Maillard, L. (2010). Seizure lateralization in scalp EEG using Hjorth parameters. Clinical neurophysiology, 121(3), 290-300.
17. Amin, H. U., Malik, A. S., Ahmad, R. F., Badruddin, N., Kamel, N., Hussain, M., Chooi, W. T. (2015). Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques. Australasian physical & engineering sciences in medicine, 38(1), 139-149.
18. Kotsiantis, S. B., Zaharakis, I. D., Pintelas, P. E. (2006). Machine learning: a review of classification and combining techniques. Artificial Intelligence Review, 26(3), 159-190.
19. Schall, U., Catts, S. V., Karayanidis, F., Ward, P. B. (1999). Auditory event-related potential indices of fronto-temporal information processing in schizophrenia syndromes: valid outcome prediction of clozapine therapy in a three-year follow-up. International Journal of Neuropsychopharmacology, 2(2), 83-93.