Detection of Attention Deficit and Hyperactivity Disorder by Nonlinear EEG Analysis

Farhad Nassehi, İrem Sönmez, Beril Şabanoğlu, Elifnur Yükselen, Hilal Meva Özaydın, Osman Eroğul
Biomedical Engineering
TOBB University of Economics and Technology
Ankara, Turkey
(f.nassehi,isonmez,bsabanoglu,eyukselen,hozaydin,erogul)@etu.edu.tr

***Abstract*— This study proposes to experts a fast and highly successful algorithm for the diagnosis of ADHD disorder using EEG (Electroencephalogram) signals obtained during the Attention task, reducing their dependence on subjective evaluations. Accordingly, EEG signals obtained from 61 ADHD and 60 control participants were analyzed using nonlinear features (approximate entropy, Petrosian, and Lyapunov exponent). After feature extraction, the classification process was performed using support vector machine (SVM), and K-Nearest-Neighbor (KNN), and ensemble learning. In this study t-test based and location based feature selection methods were used. We used only features that were extracted from prefrontal and frontal regions. The highest accuracy that was reached in this study was 95.8%.**

Keywords*— *Attention Deficit Hyperactivity Disorder, Electroencephalography signals, Non-linear Features, EEG, Machine Learning

I. INTRODUCTION

Attention deficit and hyperactivity disorder (ADHD) is a mental disorder manifested by inattention, excessive mobility, forgetfulness, and uncontrolled reactions. It is the most common psychiatric disorder in childhood. Recent population-based studies have shown that about 5% of children are affected by ADHD. In some cases, hyperactivity and sudden behavior predominate, while in others, signs of inattention predominate. In this study, EEG recordings were taken by showing some pictures that might be attractive to children [1].

Often diagnosed with ADHD, parents and teachers, psychologists, or psychiatrists to understand the questions and answers of honesty, which depend to a large extent in the DSM (Diagnostic and Statistical Manual of mental disorders) or the ICD (International Classification of diseases) using the criteria of judgment are made based on different versions of the diagnostic. Non-linear features are also used to distinguish children with ADHD from healthy ones.

Because linear properties (spectral, time, spatial, or frequency properties) have been used in most studies, accurate results alone cannot be obtained. Despite irregular EEG signals, quantitative measurements of chaos and nonlinear properties are appropriate descriptive tools for characterizing key information from these complex signals. Thanks to this data, children with ADHD are separated from the control group [1].

This study aims to propose a fast and highly successful decision support algorithm for the diagnosis and diagnosis of ADHD using EEG signals that will reduce their dependence on subjective tests.

II. MATERIALS & METHODS

A. Participants

Participants consisted of 61 children diagnosed with ADHD (9.62±1.75) according to DSM-IV (American Psychiatric Association Diagnostic and Statistical Manual) criteria and 60 healthy children control group (9.85 ± 1.77 years) [2].

B. Data Acquisition

These subjects were attended to EEG recording session.19 electrodes (Fz, Cz, Pz, C3, T3, C4, T4, Fp1, Fp2, F3, F4, F7, F8, P3, P4, T5, T6, O1, O2) with A1 and A2 electrodes as references on earlobes were placed on the scalp based on the 10–20 system with 128 Hz sampling frequency and 16 bits EEG resolution [2].

C. Task

Using a time series obtained from the EEG of children with ADHD while on the visual task in Task, it has been shown that the attention continuity of ADHD and control group children are different. Sampling frequency was set to 128 Hz, test times to 90 seconds. In this study, nonlinear

978-1-6654-5432-2/22/$31.00 ©2022 IEEE

property of EEG time series of ADHD and control group was extracted, then a correct classification was obtained. This task showed that the classification between ADHD and the control group was reliable [2].

D. Preprocessing

In preprocessing section, Notch filtering was used to remove 50 Hz noise, then Independent Component Analysis (ICA) was used to eye movement artifacts. Then 4th order Infinite Impulse Response (IIR) Butterworth bandpass filter in 0.5-30 Hz ranges was applied to the signal. Each signal was splatted to 1-second epochs. Mean of parameters that used in feature extraction section used as inputs in classification section.

E. Feature extraction

The approximate entropy (ApEn) method was applied to calculate the entropy of each epoch. Petrosian's fractal dimension creates a new time series by extracting successive instances of a time series. In this method, samples of a time series are extracted consecutively, and after a new time series is generated, positive and negative samples are assigned 1 and -1, respectively. In this formula, n is the number of samples in dual time series and NΔ is the number of sign changes in dual time series. The FD of the signal will be calculated:

$$D = \frac{\log_{10} n}{\log_{10} n + \log_{10}\left(\frac{n}{n + 0.4N\Delta}\right)} \tag{1}$$

Largest Lyapunov Exponent (LLE) is calculated to understand the chaotic behavior of this system thanks to the largest Lypunov base (LLE) [3]. In this formula, d_n is the distance between consecutive samples in the n_{th} time and d_0 is the sequential distance in the first time. LLE will be calculated:

$$\lambda = \frac{1}{n} ln(dn/\mathrm{d}0) \tag{2}$$

F. Classification

The purpose of classification is to learn a functional model that allows an appropriate estimation of the class label for an unknown model. In this study, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Ensemble Learning methods were used in the classification process. In the classification section, 5-fold cross-validation method was used to split data to train and test subsets.

K-Nearest Neighbors is also a machine learning method that is used in the samples which are evenly distributed.The traditional KNN is the simplest method which depends on the distance metric to determine the nearest neighbor. In the KNN algorithm, k is the important parameter. In classification with KNN algorithm, k nearest neighbors of a test example in the training set are determined. After that, the classification is made according to the class labels of the determined k nearest neighbors [3]. 5 and 7 values were selected for K parameter. In this study, the Euclidean metric of KNN classifier was used to measure the differences between the samples.

Support Vector Machine is a machine learning method that uses kernels to create a linear classification . It provides a balanced predictive performance even with limited sample sizes [4]. In this study linear kernel was used.

Ensemble learning includes consolidating numerous predictions determined by various methods with a specific end goal to create a stronger overall prediction.

G. Location and Feature selection

Two different feature reduction strategies were done in this study. Firstly, Electrode location-based feature selection strategy was done. In this section, only Frontal and Prefrontal electrodes were used for classification. An independent sample t-test was used to find the most distinctive features. LLE was shown meaningful difference (*$p<0.05$*) in all electrodes between 2 groups so LLE selected as only feature for inputs of classification algorithms.

III. RESULTS

In this section, all classification results were reported in 3 different subsections according to all features, location-based feature reduction, and t-test based feature selection method.

A. All Features

In this section, all features extracted from all participants' electrodes are used as inputs of classification algorithms. Table 1. shows the maximum, average, and standard deviation accuracies of classification algorithms.

Table 1. Accuracy results of different algorithms for 121x19x3(9196) inputs. (Subjects x Electrode x Number of Features)

Algorithms	Max.	Mean.	Standard deviation
SVM	87.5%	78.6%	8.9%
KNN (K=5)	91.66%	79.76%	11.9%
KNN (K=7)	91.66%	79.76%	11.9%
Ensemble Learning	79.16%	74.46%	4.7%

Maximum sensitivities rates also were reached in KNN algorithms for both K values with 83.1% (74.2%±8.9%). Maximum specificities rates also were reached in KNN algorithms for both K values with 100% (95.2%±4.8%).

B. Location Selection

There are several studies that reported differences in Frontal and Prefrontal regions between ADHD subjects and healthy control subjects [5,6]. Because of these reports, we selected only features from the Frontal and Prefrontal regions (Fp1, Fp2, F3, and F4). So, the number of inputs of classification algorithms reduces to 121x4x3 from 121x19x3 (Subjects x Electrodes x Number of Features). Table 2. Shows the maximum, average, and standard deviation accuracies of classification algorithms. Maximum sensitivities rates for ensemble learning that was shown maximum accuracy rate was 91.66% (80%±11.6%). Maximum specificities rates for ensemble learning was 100% (75.55%±14.45%).

C. Feature Selection

In the feature selection method, Independent sample t-test was used to select the most distinctive feature. For this purpose, feature with the most number of electrodes that showed a meaningful difference (*$p<0.05$*) between ADHD and control subjects was selected as a feature. LLE shows the meaningful difference in all electrode channels. So, only LLE values that were extracted from all channels were used as inputs of classification algorithms. And the number of inputs of classification algorithms reduces to 121x19x1 from 121x19x3 (Subjects x Electrodes x Number of Features). Table 3. Shows the maximum, average, and standard deviation accuracies of classification algorithms. Maximum sensitivities rates for ensemble learning that was shown maximum accuracy rate was 91.16% (75%±16.1%). Maximum specificities rates for ensemble learning was 92.30% (78%±14.3%).

Table 2. Accuracy results of different algorithms for 121x4x3(1452) inputs. (Subjects x Electrode x Number of Features)

Algorithms	Max.	Mean.	Standard deviation
SVM	91.66%	74.43%	17.23%
KNN (K=5)	83.33%	72.76%	10.57%
KNN (K=7)	79,16%	73.53%	5.63%
Ensemble learning	95.8%	77.73%	18.07%

Table 3. Accuracy results of different algorithms for 121x19x1(2299) inputs. (Subjects x Electrode x Number of Features)

Algorithms	Max.	Mean.	Standard deviation
SVM	79.16%	66.9%	12.7%
KNN (K=5)	76%	74.33%	1.67%
KNN (K=7)	79.16%	73.53%	5.63%
Ensemble Learning	92%	76.73%	11.27%

IV. DISCUSSION & CONCLUSION

This study aimed to propose an accurate classification algorithm to classify ADHD and control subjects. EEG signals that were recorded from 19 channels were analysed by non-linear methods such as approximate entropy, largest Lyapunov exponent (LLE) and Petrosian's fractal dimension. SVM, KNN (K=5 & K=7) and ensemble learning algorithms were used as classification methods. Firstly, these groups were classified by using all features (9196 features) including ApEn, LLE and Petrosian's values of all channels of all subjects. Maximum accuracy that was reached in this section is 91.66% for KNN algorithm for both values of K. As mentioned previously, prefrontal, and frontal regions of brain are highly related to attention abilities of subjects. So, we try to classify these groups with only 4 electrodes from the prefrontal and frontal regions (Fp1, Fp2, F3, and F4). Classification's inputs numbers was reduced to 1452 from 9196 inputs. The maximum accuracy rate that was reached by ensemble learning in this section is 85.8%. Independent sample t-test was also used to find the best distinctive feature that showed the most meaningful difference in electrodes between the two groups. LLE values show meaningful differences in all electrodes between the two groups. And we try to classify these groups only with LLE features. So, the classification's inputs number was reduced to 2299 from 9196 inputs. The maximum accuracy rate that was reached again by ensemble learning in this section is 92%. According to the results reported above, the large number of inputs of classification algorithms does not guarantee a higher accuracy rate. Especially in cognitive science, the locations of EEG electrodes are important because different regions are related to different cognitive abilities.

There are many studies that tried to classify ADHD and control subjects. Mohammadi et al. were conducted a study to classify these groups by using Multi-Layer Perceptron (MLP) classification algorithm. 4 Non-linear features were extracted from 19 EEG channels. They reached 93.65% maximum accuracy rate [1]. Khoushand et al. also conducted a study with 12 ADHD and 12 control subjects. In this study, nonlinear features and frequency band powers of EEG signals were used to classify these groups by the SVM algorithm. 83.33% Maximum accuracy rate reached [7]. Zhang et al. proposed a Convolutional Neural Networks (CNN) based framework to classify ADHD and control subjects, the Phase Lag Index (PLI) method used to calculate functional connectivity maps of EEG signals. 94.39% Maximum accuracy rate reached [8]. Dubreuil-Vall et al. also proposed a CNN-based algorithm with 88%±1.12% maximum accuracy [9]. Tor et al. tried to classify ADHD, Conduct Disorder (CD) and ADHD+CD (comorbid) subjects. This study achieved 97.88% maximum accuracy rate [10]. Allahverdy et al. conducted two different studies based on EEG nonlinear features. These studies accuracies was reported as 86% and 96.7%

when Lyapunov exponent, Katz fractal dimension, Higuchi fractal dimension, and Sevcik fractal dimension features were extracted from all frontal EEG channels [2,11].

Our study proposed an ensemble learning based method with 95.8% maximum accuracy with 3 features including ApEn, LLE, and petrosian's fractal dimensions that were extracted from only 4 channels including Fp1, Fp2, F3, and F4.

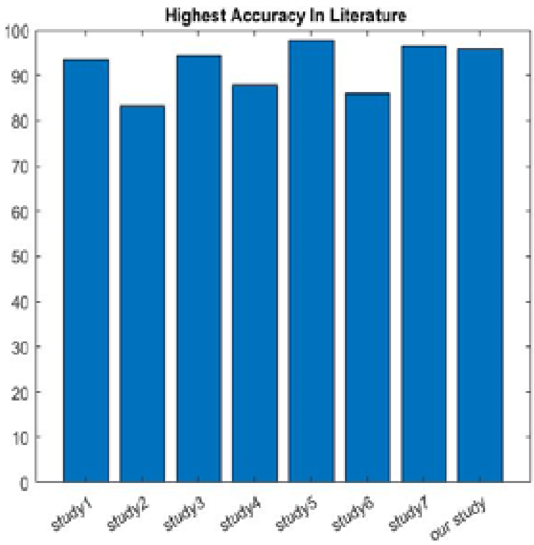

Fig 1. Highest accuracy in literature

REFERENCES

[1] Mohammadi, M. R., Khaleghi, A., Nasrabadi, A. M., Rafieivand, S., Begol, M., & Zarafshan, H. (2016). EEG classification of ADHD and normal children using non-linear features and neural network. *Biomedical Engineering Letters*, *6*(2), 66-73..

[2] Allahverdy, A., Moghadam, A. K., Mohammadi, M. R., & Nasrabadi, A. M. (2016). Detecting adhd children using the attention continuity as nonlinear feature of eeg. *Frontiers in Biomedical Technologies*, *3*(1-2), 28-33.

[3] Milošević, M., Van de Vel, A., Bonroy, B., Ceulemans, B., Lagae, L., Vanrumste, B., & Van Huffel, S. (2015). Automated detection of tonic–clonic seizures using 3-d accelerometry and surface electromyography in pediatric patients. *IEEE journal of biomedical and health informatics*, *20*(5), 1333-1341.

[4] Sun, S., & Huang, R. (2010, August). An adaptive k-nearest neighbor algorithm. In 2010 seventh international conference on fuzzy systems and knowledge discovery (Vol. 1, pp. 91-94). IEEE.

[5] Stevens, M. C., & Haney-Caron, E. (2012). Comparison of brain volume abnormalities between ADHD and conduct disorder in adolescence. *Journal of Psychiatry and Neuroscience*, *37*(6), 389-398.

[6] Altinkaynak, M., Güven, A., Nazan, D. O. L. U., İzzetoğlu, M., Demirci, E., Özmen, S., & Pektaş, F. (2018). Prefrontal brain activation in subtypes of attention deficit hyperactivity disorder: A functional near-infrared spectroscopy study. *Electrica*, *18*(2), 256-262.

[7] Khoshnoud, S., Nazari, M. A., & Shamsi, M. (2018). Functional brain dynamic analysis of ADHD and control children using nonlinear dynamical features of EEG signals. *Journal of integrative neuroscience*, *17*(1), 17-30.

[8] Zhang, Z., & Li, X. (2019, November). Use transfer learning to promote identification ADHD children with EEG recordings. In *2019 Chinese Automation Congress (CAC)* (pp. 2809-2813). IEEE.

[9] Dubreuil-Vall, L., Ruffini, G., & Camprodon, J. A. (2020). Deep learning convolutional neural networks discriminate adult ADHD from healthy individuals on the basis of event-related spectral EEG. *Frontiers in neuroscience*, *14*, 251.

[10] Tor, H. T., Ooi, C. P., Lim-Ashworth, N. S., Wei, J. K. E., Jahmunah, V., Oh, S. L., ... & Fung, D. S. S. (2021). Automated detection of conduct disorder and attention deficit hyperactivity disorder using decomposition and nonlinear techniques with EEG signals. *Computer Methods and Programs in Biomedicine*, *200*, 105941.

[11] Allahverdy, A., Nasrabadi, A. M., & Mohammadi, M. R. (2011, May). Detecting ADHD children using symbolic dynamic of nonlinear features of EEG. In *2011 19th Iranian Conference on Electrical Engineering* (pp. 1-4). IEEE.