Frontiers in Biomedical Technologies Vol. 8, No. 2 (2021) 115-122

ORIGINAL ARTICLE

Classification of the Children with ADHD and Healthy Children Based on the Directed Phase Transfer Entropy of EEG Signals

Ali Ekhlasi [1], Ali Motie Nasrabadi [2*], Mohammadreza Mohammadi [3]

[1] *Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran*

[2] *Department of Biomedical Engineering, Faculty of Engineering, Shahed University, Tehran, Iran*

[3] *Psychiatry and Psychology Research Center, Roozbeh Hospital, Tehran University of Medical Sciences, Tehran, Iran*

*Corresponding Author: Ali Motie Nasrabadi
Email: nasrabadi@shahed.ac.ir

Received: 24 December 2020 / Accepted: 28 February 2021

Abstract

Purpose: The present study was conducted to investigate and classify two groups of healthy children and children with Attention Deficit Hyperactivity Disorder (ADHD) by Effective Connectivity (EC) measure. Since early detection of ADHD can make the treatment process more effective, it is important to diagnose it using new methods.

Materials and Methods: For this purpose, Effective Connectivity Matrices (ECMs) were constructed based on Electroencephalography (EEG) signals of 61 children with ADHD and 60 healthy children of the same age. ECMs of each individual were obtained by the directed Phase Transfer Entropy (dPTE) between each pair of electrodes. ECMs were calculated in five frequency bands including, delta, theta, alpha, beta, and gamma. Based on ECM, an Effective Connectivity Vector (ECV) was constructed as a feature vector for the classification process. Furthermore, ECV of different frequency bands was pooled in one global ECV (gECV). Multilayer Artificial Neural Network (ANN) was used in the steps of classification and feature selection by the Genetic Algorithm (GA).

Results: The highest classification accuracy with the selected features of ECV was related to theta frequency band with 89.7%. After that, the delta frequency band had the highest accuracy with 89.2%. The results of ANN classification and GA on the gECV reported 89.1% of accuracy.

Conclusion: Our findings show that the dPTE measure, which determines effective connectivity between the brain regions, can be used to classify between ADHD and healthy groups. The results of the classification have improved compared to some studies that used the functional connectivity measures.

Keywords: Attention Deficit Hyperactivity Disorder; Electroencephalography; Phase Transfer Entropy; Effective Connectivity; Classification.

DOI: https://doi.org/10.18502/fbt.v8i2.6515

1. Introduction

Attention Deficit Hyperactivity Disorder (ADHD) is regarded as a neurodevelopmental disorder starting in childhood, although its symptoms can continue into adulthood [1]. The prevalence of ADHD in children has been estimated by 12.1 and 3.9% among boys and girls, respectively [2]. ADHD is characterized by inattention, or excessive activity, and impulsivity or their combination. Given the prevalence of ADHD and its social consequences in childhood and adulthood, its early diagnosis in children can make treatment processes and psychological interventions more effective [3, 4]. In addition to diagnosing ADHD by studying children's behaviors where they should be monitored for at least 6 months, examining children's brain function using Electroencephalogram (EEG) can also help in early the diagnosis of the disease [5].

EEG is a very popular modality for examining differences in brain function between ADHD and healthy groups due to its high time resolution and ease of data recording compared to other brain imaging modalities. One approach in processing EEG in children with ADHD is extracting features from the brain signals and applying machine learning algorithms to classify the two groups. Sadatnezhad *et al.* (2011) achieved a classification accuracy of 86.4% between the children with ADHD and healthy children by extracting fractal features and power of the EEG frequency bands [6]. Allahverdy *et al.* (2011) used non-linear features of EEG containing fractal dimension and obtained 86% of accuracy in classification [7]. Mohammadi (2016) improved classification accuracy between the two groups by 93.65% using a multilayer neural network and a combination of non-linear features [8]. Nazhvani *et al.* (2013) and Mueller *et al.* (2011) classified the healthy children and children with ADHD by extracting Evoked Related Potential (ERP) features with 92.1 and 91% of accuracy, respectively [9, 10]. Another approach in EEG processing is investigating the connectivity between brain regions that can be functional or effective. Functional connectivity provides information on the correlation or mutual dependency between recording channels, whereas effective connectivity investigates the effectiveness of various regions on each other and determines the direction and strength of information transmission. Ahmadlou (2010, 2011) reported classification accuracies of 87.5% and 95.6% for a synchronization pattern in theta and delta frequency bands by applying criteria for determining functional connectivity, such as Synchronization Likelihood (SL) and Fuzzy Synchronization Likelihood (FSL) [11, 12]. Pereda (2018) found approximately 95% of classification accuracy for a functional connectivity pattern with 280 features [13]. Recently, Kiiski (2020) used the Weighted Phase Lag Index (WPLI) measure as a neuromarker for adults with ADHD [14]. WPLI measures the phase difference between two signals; and attenuates the effect of volume conduction by weighting phase differences close to zero-phase lag. Because WPLI is not based on signal amplitude, it is insensitive to ADHD-specific spectral changes [15, 16]. Kiiski found that the elevated functional connectivity in the delta and theta bands was the best predictor for ADHD status [14].

Accordingly, in this paper, the directed Phase Transfer Entropy (dPTE) was used, which was recently introduced to determine the strength and direction of connectivity [17]. Contrary to functional connectivity measures introduced in the previous studies on EEG of the ADHD group, dPTE considers the direction of communication and measures effective connectivity. This measure uses phase time series of signals as input of Transfer Entropy (TE) function, so it has fewer limitations in interpreting connectivity of amplitude. TE is a model-free effective connectivity measure that investigates linear and non-linear couplings [18, 19]. Recent research has investigated the directional information transfer by the dPTE measurement in the brain regions on the Magnetoencephalography (MEG) in various brain disorders including, Alzheimer's and Parkinson's [20-22]. Dauwan *et al.* (2016) used the dPTE to examine the connectivity of brain regions in dementia with Lewy bodies [23] and Ahmadi *et al.* (2020) applied the dPTE method to decode covert visual attention based on EEG [24]. In more recent researches, the Phase Transfer Entropy (PTE) measure and the graph theory have been used to investigate the brain networks in brain disorders with the EEG signals [25-27].

In this research, the classification method was performed between two groups of healthy children and the children with ADHD using a dPTE-based connectivity matrix. It was hypothesized that feature extraction from Effective Connectivity Matrices (ECM) between EEG signals obtained by dPTE measure could be a reliable criterion for classifying the two groups. To the best of our knowledge, this study is the first study that classified the healthy children and those with ADHD during an attention task using dPTE values obtained from machine learning algorithms.

The rest of the paper is organized as follows: at first, in the "Methods" Section, steps of recording and preprocessing of EEG data are described. After that, the procedure of constructing an Effective Connectivity Vector (ECV) and global ECV are explained. Then, feature selection and classification methods are given. In the "Results" section, the results of the classifier are presented. In the "Discussion" section, the described results are compared with the results of the previous research, and conclusions are presented.

2. Materials and Methods

2.1. Subjects

The participants in this study were 121 children who were divided into two groups: children with ADHD (ADHD group) and healthy children (control group). The ADHD group consisted of 61 children (with a mean age of 9.62 ± 1.75 years old) whose disease was confirmed by an experienced psychiatrist according to the diagnostic and statistical manual of mental disorders –fourth edition (DSM-IV) criteria. The control group consisted of 60 children (with a mean age of 9.85 ± 1.77 years old). The medical history of the control group was evaluated to ensure the absence of psychological disorder, epileptic history, drug abuse, and head injury. All the subjects participated in the experiment voluntarily, and informed written consent was obtained from their parents for their participation in the test. This research was approved by the Institutional Review Board (IRB) and the Ethics Committee of Tehran University of Medical Sciences (TUMS).

Electroencephalogram signals were recorded by a digital device (SD-C24, Sholeh Danesh Co., Tehran, Iran) in the Psychology and Psychiatry Research Center at Roozbeh hospital (Tehran, Iran) [8]. In this task, a number of images were shown to each subject and the subject was asked to count the number of characters in each image. Correct and incorrect answers of the participants were not considered and the task was designed without rewards. During the task, EEG was recorded based on a 10-20 standard system by 19 channels with 128 Hz of sampling frequency and 16 bits of resolution. Figure 1 shows an example of these images. The duration of recording was dependent on the performance of each subject. Data of which are available now in [28].

2.2. EEG Pre-Processing

Pre-processing of brain signals was done by EEGLAB toolbox (version 14.1.1) in MATLAB 2018a software.

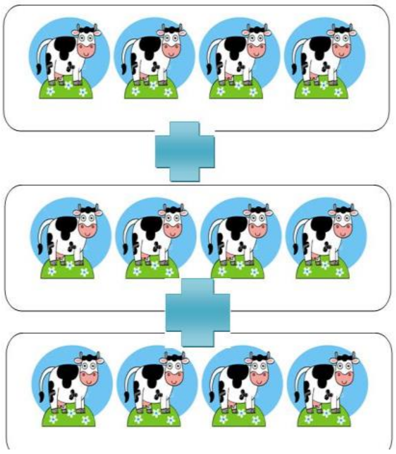

Figure 1. An example of the images shown to the children during the recording of the signals [29]

These steps were performed according to Makoto's preprocessing pipeline. A band-pass finite impulse response (FIR) filter of 1 - 48 Hz was applied to continuous EEG data. Channel interpolation was performed by adjacent channels for the channels that were incorrectly recorded. EEGs were re-referenced to the common average and were decomposed using Independent Component Analysis (ICA). Eye blinks and muscle artifacts were identified by ICA and were manually removed based on their spectra, scalp maps, and time courses. Afterward, EEG data were filtered by an offline discrete fast Fourier transform in five frequency bands, including delta (1-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13–30 Hz), and gamma (30–45 Hz). In the filtering steps, FIR filters with zero phase shift were used to avoid distorting phase response.

2.3. Measuring ECV

Effective Connectivity (EC) was investigated between all pairs of channels by dPTE in five frequency bands. The dPTE is a measure based on TE that determines the strength and direction of connectivity. TE is a model-free implementation of the Wiener's Principle of Causality and it can be expressed in terms of information theory as follows: "How much additional information does the past state of process X contain about the future observation of a value of Y given that, we already know past state of Y?" [18, 19]. For determining information transfer from X to Y, TE can be written based on conditional mutual information (or the difference between the amounts of two mutual information values) or time series probability functions as follows (Equation 1):

$$TE_{xy} = \sum P(Y_{t+\delta}|\, Y,X)\,\frac{\log P(Y_{t+\delta}|\, Y,X)}{\log P(Y_{t+\delta}|\, Y)} \qquad (1)$$

Where δ is an interaction delay between X and Y.

Lobier (2014) developed the concept of TE, called PTE, to study the transfer of information between two signals using the instantaneous phase [17]. Accordingly, instantaneous time series of the signals are extracted and applied as input to the TE function using the Hilbert transform (Equation 2).

$$PTE_{xy} = \sum P(Y_\delta)P(X)P(Y)\frac{\log P(Y_\delta | Y,X)}{\log P(Y_\delta | Y)} \quad (2)$$

Where the probabilities are obtained by building histograms of phase estimates in an epoch. For this study, dPTE, as described in the studies by Engels and Hillebrand was used [21, 22] (Equation 3):

$$dPTE_{xy} = \frac{PTE_{xy}}{PTE_{xy} + PTE_{yx}} \quad (3)$$

The dPTE indicates the direction of information transmission and determines which signal is a directional transmitter and which signal is a directional receiver of information relative to each other. The dPTE is bounded in the range of $0.5 < dPTExy \leq 1$ when information flows preferentially from a time series X to time series Y, when signal X is receiving concerning signal Y, $0 \leq dPTExy < 0.5$. If this value is equal to 0.5, it means that there is no preference in the direction and strength of information transmission between the two signals. The procedure described in [22] was implemented to determine dPTE parameters, such as interaction delay (δ), histogram calculation method, and the number of bins. For each subject in each frequency band, dPTE calculations were performed on 8-second time windows. ECMs were averaged over the windows, providing an ECM for every subject in each frequency band (see Figure 2a).

Directed connectivity matrices are usually not symmetric. The dPTE matrix is also not symmetric, although according to dPTE calculation, the two triangular parts are trivially related: dPTExy =1 – dPTEyx. As a result, both triangles express direction and strength of connectivity and do not provide more information than each other [30].

For each subject, the upper triangle of ECM was converted into a vector by placing the elements together in a row. This vector was called ECV. In this way, an ECV was obtained for each subject in each frequency band. Therefore, the ECV vector had 171 entries $\frac{(19*19)-19}{2}$, representing EC values between every pair of electrodes

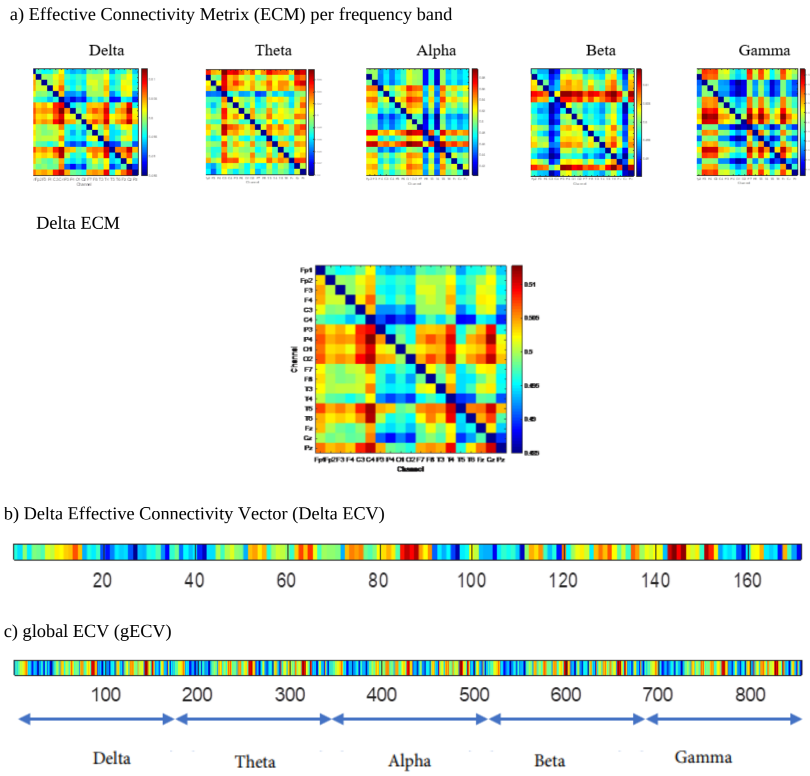

Figure 2. Example of ECV of one subject in the ADHD group. a) For every subject, and for every frequency band, an ECM was computed using dPTE. b) From every ECM, the upper triangular entries, consisting of the pair-wise connectivity between all possible combinations of electrodes were extracted. An example of an ECV in the delta band is shown in the bottom plot in. c) An ECV was computed for every frequency band and they were pooled together to obtain a single global ECV

(see Figure 2b). In addition to ECV in each frequency band, ECVs were put together for each individual to create the global ECV (gECV). Given five frequency bands, the gECV will have 855 entries (855 entries = 171*5 frequency bands) (see Figure 2c). The ECV in each frequency band and gECV were used for classification analysis.

2.4. Feature Selection and Classification

Feature selection and classification procedures were performed in two modes. In the first mode, a set of features, including ECVs of all individuals in each frequency band were used. Therefore, classification results were reported in each frequency band. In the second mode, the gECV constructed by ECV in five frequency bands was used. Feature selection is a process that reduces the number of features and selects a subset of original features. Since dPTE values were used as feature vectors in this study, the feature selection process showed how many brain connections can be more effective in distinguishing between the two groups. Features must first be normalized before performing the classification step. Thus, Z-score was used in each frequency band to normalize ECV values.

In this paper, the Genetic Algorithm (GA) was used as a feature selection algorithm and the Artificial Neural Network (ANN) was used as a classifier. GA is a common and reliable method for optimization and feature selection, inspired by biological evolution. GA uses repeated mutation and recombines parts of the best-known solutions [31]. For using GA in the feature selection, a feature vector (ECV or gECV) must be considered as a chromosome. The presence or absence of each feature in a chromosome was determined by the values of 0 and 1, respectively, in each gene. First, the population of primary chromosomes was determined and the genes on each chromosome were assigned random values of 0 and 1. A fitness function value was calculated for each chromosome. Iterative steps of the algorithm continued until the appropriate value of the fitness function was obtained by adjusting the other parameters of GA, such as mutation rate and crossover probability.

In this study, ANN was used as a fitness function in such a way that the percentage of classification accuracy by ANN was attributed to each chromosome at each step of the GA. The three-layered feed-forward Multi Layer Perceptron (MLP) with ten neurons in the hidden layer was used for fitness computation. The Levenberg-Marquardt training algorithm as a kind of backpropagation calculation was applied for training the network. Tables 1 and 2 show parameters of GA and ANN setting, respectively.

Table 1. Parameters of the GA

Parameter	Setting
Population size	100
Mutation rate	0.02
Crossover probability	0.8
Iteration number	50

Table 2. Parameters of the ANN

Parameter	Setting
Epoch number	100
Neurons number in the hidden layer	10
Training data (%)	70
Validation data (%)	15
Testing data (%)	15

3. Results

The proposed algorithm was applied in all five ECVs and gECV. The results obtained for classification using ECVs are shown in Table 3. As depicted in Table 3, accuracy for classification of the two groups was determined by ANN in two modes of all the features and the selected feature by the GA. For performing cross-validation, the K-fold algorithm was used in each classification step. Based on k-fold, the dataset was divided into ten folds randomly (k=10). Consequently, each fold was used as a test group once and nine folds were used as training data. Then, the accuracy of classifiers was calculated by averaging the

Table 3. Accuracy of ANN classification results using the k-fold method without and with feature selection and number of features selected in each ECV

ECV	ANN without Feature Selection	ANN with Feature Selection	Number of Features Selected by GA (from 172 Features)
Delta	72 ±1.21	89.2 ±1.33	86
Theta	70.2 ±1.52	89.7 ±1.14	90
Alpha	63.4 ± 1.98	83.4 ±1.32	80
Beta	68.2 ±1.35	84.7 ±1.27	72
Gamma	61.3 ±1.77	83.6 ±1.21	96

results of ten folds and the related results are illustrated in Table 3 in the form of mean and standard deviation.

In Table 3, the mean and standard deviation for the classification of test data in each ECV are given in two columns. In the first column, the feature selection process was not used and all ECV elements in each frequency band were considered as features for each subject. In the second column, a feature selection algorithm was used by the GA. The number of the selected features per frequency band was indicated in the third column. Figure 3 shows the average percentage of classification error using the GA and ANN for alpha frequency band at 50 iterations.

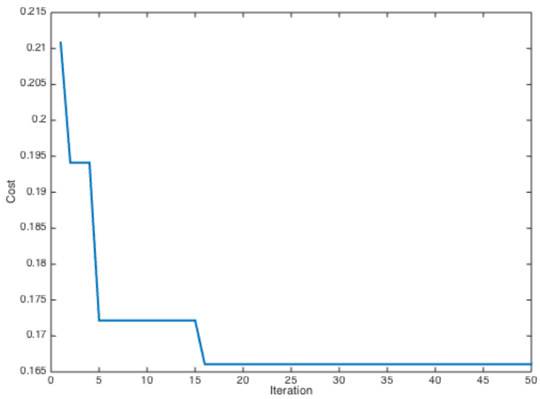

Figure 3. Percentage of classification error in alpha frequency band by GA feature selection until 50 iterations

As shown in Table 3, in all the frequency bands, in the mode where the GA was used for the feature selection step, classification was improved compared to the mode where the total features of each band were used. Approximately 20% improvement in classification was achieved for each frequency band in feature selection mode. The highest value of classification accuracy (89.7%) belongs to the ECV in the theta frequency band when the feature selection process is performed. After that, the ECV of the delta band with 89.2% classification accuracy ranked second in the classification process between frequency bands. These two percentages of classification accuracy for the delta and theta bands were obtained by selecting 86 and 90 features from 172 features by the GA, respectively. In the next step, gECV was used for classification between the two groups. Thus, the gECV of the subjects was once completely applied to ANN as input, and once again feature selection process was performed on the gECV by GA and the selected features were used for classification. The results for classification using gECV are shown in Table 4.

The results of gECV classification showed that feature selection with GA and ANN can classify both healthy and ADHD groups with 89.1% accuracy.

Table 4. Accuracy of ANN classification results using the k-fold method without and with feature selection and number of features selected in gECV

	ANN without Feature Selection	ANN with Feature Selection	Number of Features Selected by GA (from 855 features)
gECV	71±1.59	**89.1**±1.74	422

4. Discussion

In this paper, the connectivity between brain regions in the attention task was evaluated in healthy children and children with ADHD. Effective brain connectivity between the two regions was measured by dPTE. Thus, each individual had an average ECM based on the dPTE. The ECV for each subject was constructed by converting the upper triangular part of the ECM into a vector in all five frequency bands. Feature vectors were obtained by normalizing ECV values in the two groups. Classification between the two groups was performed by ANN in two modes. In the first mode, all ECV elements were used as features, and in the second mode, a feature selection process was performed on ECVs using GA. For combining the information from different bands, for every subject, ECVs of different frequency bands were pooled in one gECV. Classification steps were repeated by gECVs.

As demonstrated in Table 3, classification accuracy between the two groups based on ECVs of delta and theta bands was higher than the other frequency bands. This result is consistent with the results of the previous research that classified two groups with functional connectivity features. In Ahmadlou's (2010, 2011) studies, in which two groups were classified with SL and FSL measures, the best classification features were related to the delta and theta frequency bands [11, 12]. In Barttfeld's (2014) study, a significant difference was observed in fronto-occipital functional connectivity in the delta band between the two groups [32]. According to the results of our research and previous research, it can be concluded that more distinction between the healthy and ADHD groups in the delta and theta frequency bands can be associated with higher levels of hyperactivity, impulsivity, and depressive symptoms in the ADHD group. The highest accuracy of classification in our study with 89.7% and 89.2%, belongs to the ECV of the theta and delta bands, respectively. In addition, the classification using the gECV reported 89.1% of classification accuracy. The accuracy of the classification in our study is better than the results of some previous studies in which the two groups were classified based on the features of functional connectivity. Ahmadlou (2010) reached an accuracy of 87.50%. for the diagnosis of ADHD with FSL patterns of EEG functional connectivity [11]. The effective connectivity measures contain more information than criteria for determining functional connectivity. Since the dPTE, in addition to the strength of connectivity between regions, indicates the direction of information transfer, it can work better in classification. The results of the classification in our study were higher than some previous studies that used nonlinear features for the classification [6, 7].

As shown in Figure 3, the parameters of the GA were properly adjusted for feature selection process. Sufficient iterations to search and find the best array of features with the least amount of classification error were performed by the GA. The feature selection algorithm in the beta frequency band provided the lowest number of features (72 features) out of 172 features among the frequency bands. However, classification accuracy in this frequency band was equal to 84.7%. The gECV was constructed by putting ECV of each frequency band together. Classifying and selecting the features by gECV also produced an accuracy of 89.1%, which is less than the classification accuracy of the delta and theta bands. This may be due to large dimension of the feature vector in gECV, which has brought classification response closer to overfitting. In both modes, the number of the selected features was almost half of the total feature vector in each mode using ECVs and gECV, in the feature selection process using GA and ANN.

In conclusion, our results regarding the effective connectivity between brain regions in the healthy children and the children with ADHD during the attention task by dPTE showed that ECV produces more classification accuracy in the delta and theta frequency bands (with 89.2% and 89.7% accuracy, respectively) than higher frequency bands. A vector called gECV was constructed by the ECV of each subject in the frequency bands, which enabled feature selection and classification to classify the two groups with 89.1% of accuracy. To the best of our knowledge, this study was the first study used effective connectivity measure (dPTE) and machine learning algorithms (ANN and GA) for classification between healthy children and children with ADHD. Some of the results of the previous studies used functional connectivity measures and non-linear features were improved according to the results obtained in this paper. Our findings showed that features obtained based on dPTE can be reliable for classification between ADHD and control groups.

References

1- L. C. Fonseca, G. M. A. Tedrus, C. d. Moraes, A. d. V. Machado, M. P. d. Almeida, and D. O. F. d. Oliveira, "Epileptiform abnormalities and quantitative EEG in children with attention-deficit/hyperactivity disorder," *Arquivos de neuro-psiquiatria*, vol. 66, no. 3A, pp. 462-467, 2008.

2- R. Kessler *et al.*, "The effects of temporally secondary co-morbid mental disorders on the associations of DSM-IV ADHD with adverse outcomes in the US National Comorbidity Survey Replication Adolescent Supplement (NCS-A)," *Psychological medicine*, vol. 44, no. 8, p. 1779, 2014.

3- R. B. Tenenbaum *et al.*, "Response inhibition, response execution, and emotion regulation among children with attention-deficit/hyperactivity disorder," *Journal of abnormal child psychology*, vol. 47, no. 4, pp. 589-603, 2019.

4- F. Lenzi, S. Cortese, J. Harris, and G. Masi, "Pharmacotherapy of emotional dysregulation in adults with ADHD: a systematic review and meta-analysis," *Neuroscience & Biobehavioral Reviews*, vol. 84, pp. 359-367, 2018.

5- R. J. Barry, A. R. Clarke, and S. J. Johnstone, "A review of electrophysiology in attention-deficit/hyperactivity disorder: I. Qualitative and quantitative electroencephalography," *Clinical neurophysiology*, vol. 114, no. 2, pp. 171-183, 2003.

6- K. Sadatnezhad, R. Boostani, and A. Ghanizadeh, "Classification of BMD and ADHD patients using their EEG signals," *Expert Systems with Applications*, vol. 38, no. 3, pp. 1956-1963, 2011.

7- A. Allahverdy, A. M. Nasrabadi, and M. R. Mohammadi, "Detecting ADHD children using symbolic dynamic of nonlinear features of EEG," *in 2011 19th Iranian Conference on Electrical Engineering IEEE*, pp. 1-4, 2011.

8- M. R. Mohammadi, A. Khaleghi, A. M. Nasrabadi, S. Rafieivand, M. Begol, and H. Zarafshan, "EEG classification of ADHD and normal children using non-linear features and neural network," *Biomedical Engineering Letters*, vol. 6, no. 2, pp. 66-73, 2016.

9- A. D. Nazhvani, R. Boostani, S. Afrasiabi, and K. Sadatnezhad, "Classification of ADHD and BMD patients using visual evoked potential," *Clinical neurology and neurosurgery*, vol. 115, no. 11, pp. 2329-2335, 2013.

10- A. Mueller, G. Candrian, V. A. Grane, J. D. Kropotov, V. A. Ponomarev, and G.-M. Baschera, "Discriminating between ADHD adults and controls using independent ERP components and a support vector machine: a validation study," *Nonlinear biomedical physics*, vol. 5, no. 1, p. 5, 2011.

11- M. Ahmadlou and H. Adeli, "Wavelet-synchronization methodology: a new approach for EEG-based diagnosis of ADHD," *Clinical EEG and Neuroscience*, vol. 41, no. 1, pp. 1-10, 2010.

12- M. Ahmadlou and H. Adeli, "Functional community analysis of brain: a new approach for EEG-based investigation of the brain pathology," *Neuroimage*, vol. 58, no. 2, pp. 401-8, Sep 15 2011.

13- E. Pereda, M. García-Torres, B. Melián-Batista, S. Mañas, L. Méndez, and J. J. González, "The blessing of Dimensionality: Feature Selection outperforms functional connectivity-based feature transformation to classify ADHD subjects from EEG patterns of phase synchronisation," *PloS one*, vol. 13, no. 8, p. e0201660, 2018.

14- H. Kiiski *et al.*, "Functional EEG connectivity is a neuromarker for adult attention deficit hyperactivity disorder symptoms," *Clinical Neurophysiology*, vol. 131, no. 1, pp. 330-342, 2020.

15- S. M. Snyder and J. R. Hall, "A meta-analysis of quantitative EEG power associated with attention-deficit hyperactivity disorder," *Journal of Clinical Neurophysiology*, vol. 23, no. 5, pp. 441-456, 2006.

16- S. K. Loo and S. Makeig, "Clinical utility of EEG in attention-deficit/hyperactivity disorder: a research update," *Neurotherapeutics*, vol. 9, no. 3, pp. 569-587, 2012.

17- M. Lobier, F. Siebenhuhner, S. Palva, and J. M. Palva, "Phase transfer entropy: a novel phase-based measure for directed connectivity in networks coupled by oscillatory interactions," *Neuroimage*, vol. 85 Pt 2, pp. 853-72, Jan 15 2014.

18- T. Schreiber, "Measuring information transfer," *Phys Rev Lett*, vol. 85, no. 2, pp. 461-4, Jul 10 2000.

19- M. Wibral, R. Vicente, and M. Lindner, "Transfer entropy in neuroscience," *in Directed information measures in neuroscience: Springer*, 2014, pp. 3-36.

20- L. I. Boon, A. Hillebrand, K. T. O. Dubbelink, C. J. Stam, and H. W. Berendse, "Changes in resting-state directed connectivity in cortico-subcortical networks correlate with cognitive function in Parkinson's disease," *Clinical Neurophysiology*, vol. 128, no. 7, pp. 1319-1326, 2017.

21- M. M. A. Engels *et al.*, "Directional information flow in patients with Alzheimer's disease. A source-space resting-state MEG study," *Neuroimage Clin*, vol. 15, pp. 673-681, 2017.

22- A. Hillebrand *et al.*, "Direction of information flow in large-scale resting-state networks is frequency-dependent," *Proceedings of the National Academy of Sciences*, vol. 113, no. 14, pp. 3867-3872, 2016.

23- M. Dauwan *et al.*, "EEG-directed connectivity from posterior brain regions is decreased in dementia with Lewy bodies: a comparison with Alzheimer's disease and controls," *Neurobiology of aging*, vol. 41, pp. 122-129, 2016.

24- A. Ahmadi, S. Davoudi, M. Behroozi, and M. R. Daliri, "Decoding covert visual attention based on phase transfer entropy," *Physiology & behavior*, vol. 222, p. 112932, 2020.

25- F. Hasanzadeh, M. Mohebbi, and R. Rostami, "Graph theory analysis of directed functional brain networks in major depressive disorder based on EEG signal," J Neural Eng, vol. 17, no. 2, p. 026010, Mar 27 2020.

26- J. Wang, J. Wang, J. Sun, S. Tong, and X. Hong, "Phase Transfer Entropy between Frontal and Posterior Regions during Visual Spatial Attention," *in 2019 9th International IEEE/EMBS Conference on Neural Engineering (NER), IEEE*, pp. 550-553, 2019.

27- S. Wang *et al.*, "A Study on Resting EEG Effective Connectivity Difference before and after Neurofeedback for Children with ADHD," *Neuroscience*, 2021.

28- A. A. Ali Moti Nasrabadi, Mehdi Samavati, Mohammad Reza Mohammadi. (2020). EEG data for ADHD / Control children. Available: http://dx.doi.org/10.21227/rzfh-zn36

29- A. Allahverdy, A. K. Moghadam, M. R. Mohammadi, and A. M. Nasrabadi, "Detecting ADHD Children using the Attention Continuity as Nonlinear Feature of EEG," *Frontiers in Biomedical Technologies*, vol. 3, no. 1-2, pp. 28-33, 2016.

30- M. Demuru *et al.*, "Functional and effective whole brain connectivity using magnetoencephalography to identify monozygotic twin pairs," *Scientific reports*, vol. 7, no. 1, pp. 1-11, 2017.

31- T. T. Erguzel, S. Ozekes, O. Tan, and S. Gultekin, "Feature selection and classification of electroencephalographic signals: an artificial neural network and genetic algorithm based approach," *Clinical EEG and neuroscience*, vol. 46, no. 4, pp. 321-326, 2015.

32- P. Barttfeld *et al.*, "Functional connectivity and temporal variability of brain connections in adults with attention deficit/hyperactivity disorder and bipolar disorder," *Neuropsychobiology*, vol. 69, no. 2, pp. 65-75, 2014.