Biomed Eng Lett (2016) 6:66-73
DOI 10.1007/s13534-016-0218-2

ORIGINAL ARTICLE

EEG Classification of ADHD and Normal Children Using Non-linear Features and Neural Network

Mohammad Reza Mohammadi, Ali Khaleghi, Ali Moti Nasrabadi, Safa Rafieivand, Moslem Begol and Hadi Zarafshan

Received: 5 October 2015 / Revised: 10 March 2016 / Accepted: 18 March 2016
© The Korean Society of Medical & Biological Engineering and Springer 2016

Abstract
Purpose Attention-Deficit Hyperactivity Disorder (ADHD) is a neuro-developmental disorder that is characterized by hyperactivity, inattention and abrupt behaviors. This study proposes an approach for distinguishing ADHD children from normal children using their EEG signals when performing a cognitive task.
Methods In this study, 30 children with ADHD and 30 age-matched healthy children without neurological disorders underwent electroencephalography (EEG) when performing a task to stimulate their attention. Fractal dimension (FD), approximate entropy and lyapunov exponent were extracted from EEG signals as non-linear features. In order to improve the classification results, double input symmetrical relevance (DISR) and minimum Redundancy Maximum Relevance (mRMR) methods were used to select the best features as inputs to multi-layer perceptron (MLP) neural network.
Results As expected, children with ADHD had more delays and were less accurate in doing the cognitive task. Also, the extracted non-linear features revealed that non-linear indices were greater in different regions of the brain of ADHD children compared to healthy children. This could indicate a more chaotic behavior of ADHD children while performing a cognitive task. Finally, the accuracy of 92.28% and 93.65% were achieved using mRMR method and DISR method using MLP, respectively.
Conclusions The results of this study demonstrate the ability of the non-linear features to distinguish ADHD children from healthy children.

Keywords Attention-Deficit Hyperactivity Disorder (ADHD), Electroencephalogram (EEG), Non-linear features, Feature selection, Neural Network (NN)

INTRODUCTION

Attention-Deficit Hyperactivity Disorder (ADHD) is a mental disorder that is characterized by hyperactivity, inattention and abrupt behaviors [1]. Recent population based studies have shown that about 5% of children are affected by ADHD and this disorder is more prevalent in boys [1, 2]. In some cases, hyperactivity and abrupt behaviors are dominant while inattention symptoms are bold in the others [1]. Usually, ADHD symptoms begin at preschool ages, but serious problems appear in the school ages [3]. The main problem of children with ADHD is weakness in preservation and regulation of their behaviors, so they often do not show relevant responses to environmental stimulus [4-6].

Early diagnosis of this disorder is of the prime importance in preventing subsequent complications such as negative effects on children's social interactions [3]. Usually, the diagnosis of ADHD is done based on diagnostic judgment using criteria of different editions of DSM (Diagnostic and Statistical Manual of Mental Disorders) or ICD (International Classification of Diseases) [1, 7], which are highly dependent on the parents' and teachers' understanding of the psychologists

Mohammad Reza Mohammadi, Ali Khaleghi (✉), Hadi Zarafshan
Psychiatry & Psychology Research Center, Tehran University of Medical Sciences, Tehran, Iran
Tel : +98-9163464494 / Fax : +98-21-55421959
E-mail : alikhaleghi_bme84@yahoo.com

Ali Moti Nasrabadi
Biomedical Engineering Department, Shahed University, Tehran, Iran

Safa Rafieivand
Faculty of Biomedical Engineering, Amirkabir University of Technology, Tehran, Iran

Moslem Begol
Department of Biomedical Engineering, Dezful Branch, Islamic Azad University, Dezful, Iran

or psychiatrists' questions and honesty in their responses [2]. To reduce these problems, several studies attempted to propose and use more objective methods such as EEG in the diagnosis of ADHD [8-12].

Abnormalities study by EEG signals in ADHD was first performed by J. Lubar in 1973. He found that theta activity increased and also beta power dramatically reduced in ADHD [8]. In another study it was revealed that there are some indices in the EEG signals which were showing ADHD abnormality and learning disabilities [9]. Fonseca studied 30 ADHD children and 30 healthy children (as a control group) using EEG signals [13]. He utilized the absolute and relative power to detect epilepsy in ADHD children. Epileptiform activities were found in 10% (3/30) of ADHD children. Compared with controls, ADHD group had greater absolute power in delta and theta oscillations in all regions of their brain.

In a recent study, Tenev *et al*. introduced a classification model (machine learning technique) for ADHD based on EEG power spectra [10]. They assessed samples in four conditions (i.e. eyes open, eyes closed, visual continuous performance test and emotional continuous performance test). The obtained data were used for training four different SVM (Support Vector Machine) classifiers, and then the results of the classifiers were combined using logical expression. The results have shown that this approach can improve the discrimination in classification studies of ADHD. In a validation study, ERP (Event Related Potential) responses were extracted from EEG by applying a visual stimulus of go/no-go task to ADHD adults and age matched controls. Then, averaged ERP responses are decomposed using independent component analysis (ICA) to fully separate ERP and spontaneous EEG. Obtained signals were used as feature vector to SVM classifier. The obtained accuracy was 91% in the diagnosis of ADHD, which was confirmed in an independent ADHD sample with an accuracy of 94% [14].

Despite the current advances in using EEG analysis in the classification and diagnosis of ADHD, it is still controversial and it seems that we need a more accurate approach in this field [11, 12, 15, 16]. The aim of the present study was to propose a system to investigate attention in ADHD children using mathematical and engineering methods using electro-encephalogram (EEG) signal to classify and diagnose ADHD and reduce the false positives and false negatives in the diagnosis of this disorder. Also, we aimed to propose a system to reduce the dependency of ADHD diagnosis on subjective methods and to make early diagnosis of children with ADHD. For these aims, nonlinear mathematics is benefited. Nonlinear analysis of signals draws attention of many researchers in the recent years. Because of complex and chaotic nature of biomedical signals, such analysis can reveal new aspects and provide more information about the biological origin of them. Various versions of fractal dimension, approximate entropy and Lyapunov exponent are utilized in this work. These non-linear features have been used in some relevant studies for processing of EEG signals and their performances were already proved [17-20]. Thus, we prepared a new protocol during signal recording based on behavioral abnormalities in ADHD and exploited non-linear features and neural network (due to its high generalizability) to get high performance for ADHD diagnosis by their electrophysiology as an objective method.

METHODS

This research was approved by the Institutional Review Board (IRB) and Ethical Committee of Tehran University of Medical Sciences (TUMS). Parents of all participants read and signed an informed consent for participation in the experiments.

Participants

Participants were 30 children (22 boys and 8 girls, 9.62 ± 1.75 years), who were diagnosed as having ADHD by an experienced psychiatrist of child and adolescent according to DSM-IV criteria (among them, 25, 3 and 2 children were diagnosed as combined, inattentive and hyperactive subtypes, respectively); and 30 healthy controls (25 boys and 5 girls, 9.85 ± 1.77 years). The patients were referred for ADHD evaluation to the psychiatric clinic of Roozbeh hospital in Tehran, Iran, for the first time and they were drug-naïve. The control group was selected from two settings: the first and major setting was a primary school from which 25 boys were selected; and the 5 girls were selected from an all-girl's primary school. To identify possible disorders, a psychiatrist of child and adolescent evaluated the control group. Based on this evaluation, none of children in the control group had psychiatric problems. Exclusion criteria for children with ADHD and the healthy group were the history of major neurological disorders, brain injury (including epilepsy), a major medical illness, learning or verbal disability, other psychiatric disorders and use of benzodiazepine and barbiturate drugs. Furthermore, after performing the Raven Progressive Matrices Test for children, those participants who were above the medium level were entered into the study.

Experiment

Electroencephalogram signals were recorded by a digital device (SD-C24) in the Psychology and Psychiatry Research Center at Roozbeh hospital (Tehran, Iran). EEG recording was performed based on 10-20 standard by 19 channels (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 located on earlobes. Eye movement was recorded by two electrodes that were placed below and above the right eye. The experiment was carried out in a noiseless room; and to minimize the artifacts, we asked the participants to have minimal movement especially in their head area.

We designed our recording protocol according to behavior of ADHD children. Children with ADHD have difficulty in paying attention. However, if the attention stimulus is a sound then the difference between ADHD and healthy children is not significant, but if the attention stimulus is a picture then ADHD children have an obvious difference [3]. This idea led to our recording protocol that is based on cognitive visual stimulation. Therefore, in this study, during EEG recording we showed some pictures to children which could be attractive to them (like some cartoon characters), and the pictures caused a cognitive mental act in children. Seventeen images were shown to children successively and after each image, they were asked to say how many characters were in the image. The number of characters in each image was randomly selected between 5 and 16, and the size of the pictures was large enough to be easily visible and countable by children. These images were displayed on a monitor near the child. To have a continuous stimulus during the signal recording, each image was displayed immediately and uninterrupted after the child's response. Thus, the duration of EEG recording throughout this cognitive visual task was dependent on the child's performance (i.e. response speed). An example of these images is shown in Fig. 1. Finally, response time and the number of incorrect responses were recorded for each child.

Pre-processing of signals

The EEG signals were digitized at the sampling rate of 256 Hz and recorded in the frequency range of 0.1-70 Hz. We utilized a FIR (finite duration impulse response) Butterworth band-pass filter (order 7) with cut-off frequencies at 0.3 and 70 Hz and a notch filter of 50 Hz to noise and power line interference cancellation, respectively, by Matlab 8.2 (The Mathworks, Inc.). Then, an average number of 38 ± 5 epochs of 200 points (0.78125 second) without artifact were selected for each subject by two experienced neurologists. They were excluded the epochs containing noise and artifact from eye movement (detected by electrooculography), head movement, sweating or swallowing by visual inspection. Therefore, we had a cleaned EEG segment with a 30 second length on average for all individuals.

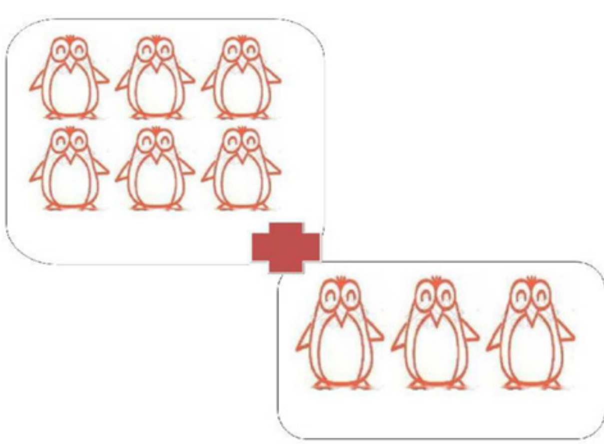

Fig. 1. An example of the images shown to the children during the EEG signals recording.

Feature extraction

In this section, we explained the implemented features of this work. The non-linear features used to characterize the chaotic pattern in the EEG signals are fractal dimension (FD) based features including Higuchi, Katz and Petrosian fractal dimensions, largest lyapunov exponent and approximate entropy. All these features are computed for each 19 channels. The features are described as follows:

Fractal Dimension (FD) is a ratio giving a statistical index of complexity in terms of details in the pattern variations with the scale [21-23]. There are various algorithms to calculate the FD. here, we have used three methods to compute it.

Katz Fractal Dimension: One way of calculating the fractal dimension is through the Katz method which was introduced in 1988. This algorithm has been proposed as an effective method in dynamic changes. In this algorithm, fractal dimension (FD) is calculated as follows [21]:

$$FD = \frac{\ln(N-1)}{\ln(N-1)-\ln\left(\frac{d}{L}\right)} \tag{1}$$

where L indicates the sum of distances between consecutive points, N is the length of data sequence and d is the diameter of data sequence, the FD calculated in this manner is dependent on the unit of measurement.

Higuchi Fractal Dimension: In this method, based on a time series x(1), x(2), …, x(N) as an input, a new time series is obtained as follows [22]:

$$x_m^k = \left\{x(m),\ x(m+k),\ x(m+2k),\ \ldots,\ x\left(m+\left\lfloor\frac{N-m}{k}\right\rfloor k\right)\right\},$$

$$\mathit{for}\ m = 1, 2, 3, \ldots, k \tag{2}$$

where m is the first sample and $\lfloor . \rfloor$ indicates the integer part of series. Length $L_m(k)$ for x_m^k is given by:

$$L_m(k) = \frac{\sum_{i=1}|x(m+ik)-x(m+(i-1)k|(N-1)}{\left\lfloor\frac{N-m}{k}\right\rfloor k} \tag{3}$$

where N is the number of samples in the time series and $\frac{N-1}{\left[\frac{N-m}{k}\right]k}$ is the normalization coefficient. For all k, total average length, L(k), is computed for k_1 to k_{max}. This method of calculating FD is more time consuming than other methods.

Petrosian Fractal Dimension: Another method to compute the FD of a signal is Petrosian which is introduced by Stoica and Moses [23]. In this method, samples of a time series are subtracted consecutively and a new time series is produced. Then, positive and negative samples are allocated to 1 and -1, respectively. Therefore, the number of sign changes in the produced time series is equal to the number of local extrema in the primary time series. The FD of signal is computed as:

$$D = \frac{\log_{10}n}{\log_{10}n + \log_{10}\left(\frac{n}{n+0.4N_{\Delta}}\right)} \tag{4}$$

where n and N_{Δ} are the number of samples and number of the sign changes in the binary time series, respectively. In this algorithm, the number of sign changes is important, while in Katz algorithm amplitude differences are important to calculate the FD. Therefore, Petrosian method is faster and more sensitive to noise.

Largest Lyapunov Exponent (LLE): Lyapunov exponent is a measure to understand the chaotic behavior of a system. If the LLE is positive, then the system has a chaotic behavior. LLE of a time series is defined based on the expansion rate of the differences between consecutive samples (not required to be consecutive). For a signal, LLE is calculated as follows [24, 25]:

$$\lambda = \frac{1}{n}\ln\left(\frac{d_n}{d_0}\right) \tag{5}$$

where d_n is the distance between consecutive samples in n^{th} time and d_0 is the consecutive distance in the initial time.

Approximate Entropy: One of the features discussed in quantifying chaos is entropy. Approximate entropy was introduced by Pincus to overcome the limitations of other measures of entropy. It is a method for determining the amount of the discipline or randomness in a signal. Approximate entropy formulation is as follows [26]:

$$ApEn(m, r_f, N) = \Phi^m(r_f) - \Phi^{m+1}(r_f) \tag{6}$$

$$\Phi^m(r_f) = \frac{1}{N-m+1}\sum_{i=1}^{N-m+1}\log C_i^m(r_f) \tag{7}$$

$$C_i^m(r_f) = \frac{number\ of\ such\ j\ that\ d[x_m(i), x_m(j)] \le r_f}{N-m+1} \tag{8}$$

$$d[x_m(i), x_m(j)] = \max_{k=1,2,\ldots,m}\left(|s(i+k-1)-(j+k-1)|\right) \tag{9}$$

$$x_m(i) = \{s(i), s(i+1), \ldots, s(i+m-1)\};\ 1 \le i \le N-m+1 \tag{10}$$

where m and r_f are positive real integers and indicate data length and filtering level, respectively. N is the considered number of samples and parameter d is the distance between vectors $x_m(i)$ and $x_m(j)$. When ApEn has a small value (nearly zero) then the signal is regular and predictable. In this study, based on [13] m = 2 and r_f = 0.2 SD were selected.

Feature selection

At first glance, it seems logical to use all the extracted features, but this will lead to inclusion of inappropriate or redundant features and reduce the classification accuracy. Relevant features are features that are informative and descriptive. Redundant features are defined as features that have the same information with one or more other features. Reduced data, reducing the feature set, improving the performance and speed and intuition are motivations to use the feature selection method [27, 28]. Here, we used two feature selection methods: Double Input Symmetrical Relevance (DISR) and minimum Redundancy Maximum Relevance (mRMR). DISR is an iterative algorithm that supposes that the mutual information of a subset S and a target variable Y has a lower bound and the lower bound is equal to an average quantity for all sub-subset [29]. In this method, this lower bound is maximized rather than the mutual information; mRMR is a two-step feature selection algorithm: in the first step a set of candidate features is selected and in the second step a compact feature subset is determined [27]. In this method, mutual information is used to classify the features. Feature selection was performed with Matlab 8.2 using FEAST toolbox [28].

Neural network and classification

Neural networks (NNs) have a good flexibility and ability to comply with almost any functions. According to their capabilities, NNs are quickly developed and have different structures and learning algorithms [30, 31]. One of the most common applications of these structures is Multi-Layer Perceptron (MLP) with back-propagation learning algorithm. MLP is used in different fields such as classification, estimation and so on [31]. These networks, in terms of architecture, have an input layer (in which the number of neurons is equal to the length of the input vector or number of features), one or more hidden layers (number of neurons in this layer can vary depending on the application and it is important to determine the design of these networks), and an output layer (number of neurons in this layer can vary depending on the

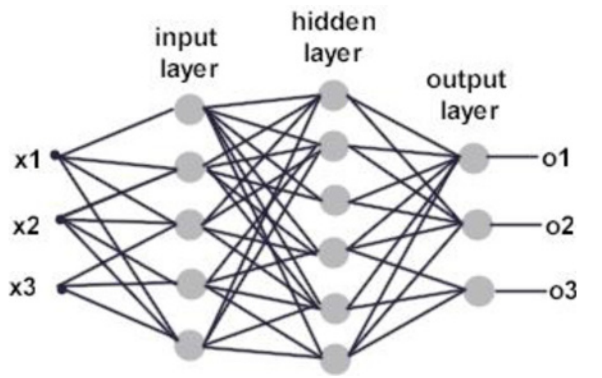

Fig. 2. General scheme of a perceptron with three input neurons, a hidden layer with six neurons and an output layer with three neurons.

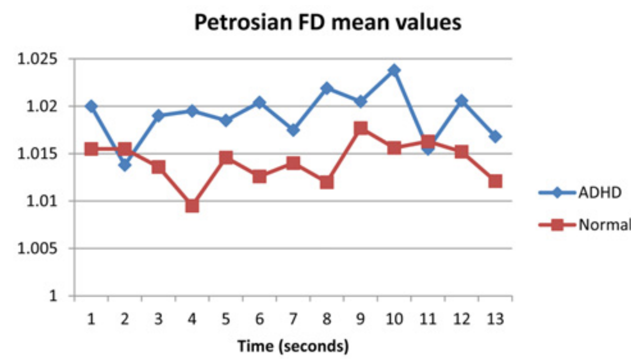

Fig. 4. An example of Petrosian fractal dimension time series in C3 channel.

application) (Fig. 2) [30].

MLP that has been used in this study has unknown parameters that are required to be carefully analyzed before training. The most important parameters are the number of hidden layers, number of neurons in each layer, learning rate and learning times on a training and test data. Here for data classification, one hidden layer with five neurons was determined [30, 31].

RESULTS

Table 1 demonstrates demographics information of both groups; as shown in this table, no significant differences exist between the patient and control groups in age, gender and IQ scores.

Fig. 3 displays the average response time and incorrect answers for ADHD and healthy children. All the mentioned features were extracted from each epoch of 200 points (equivalent to 0.78125s) and then the average values obtained for all the epochs was determined as the main used feature. Figs. 4 and 5 showed an example of Petrosian FD time series and ApEn time series in C3 channel, respectively.

After feature extraction, we selected 60 best features using DISR and mRMR methods, respectively. Thus, we had two distinct sets of features which have been used as input vectors to MLP neural network. To determine the number of

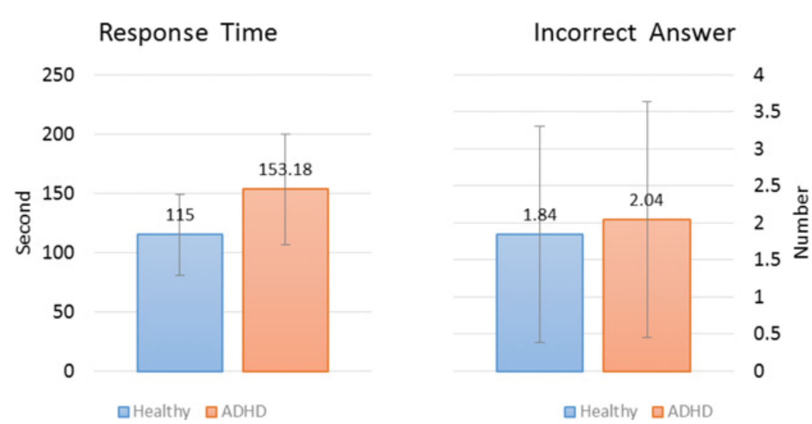

Fig 3. Average response time and incorrect answers for ADHD and healthy children.

Table 1. Demographics variables of children with ADHD and normal children

Variables	ADHD (n = 30)	Normal (n = 30)	P-value	t-statistics
Age (mean ± S.D)	9.62 ± 1.75	9.85 ± 1.77	0.818	-0.232
Gender (male, female)	m = 22 , f = 8	m = 25 , f = 5	0.754	0.317
IQ (mean ± S.D)	110.41 ± 8.68	112.27 ± 6.94	0.934	0.084
Handedness	Right (26), left (3), ambidextrous (1)	Right (27), left (3)		
Subtypes	Combined (25), inattentive (3), hyperactive (2)			

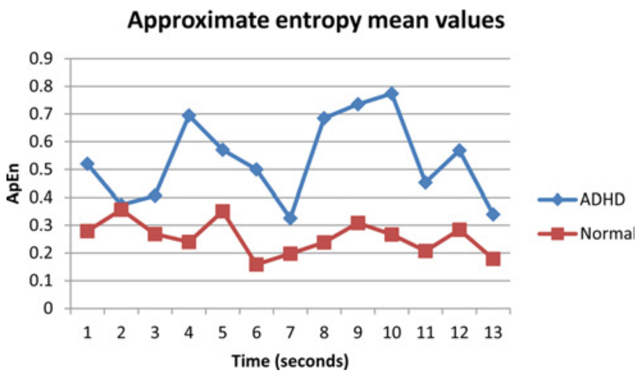

Fig. 5. An example of approximate entropy time series in C3 channel.

hidden layer neurons in the neural network structure, MLP were implemented for 3 to 8 neurons and based on the mean and variance of the error, the best results were obtained for five neurons. Network performance was evaluated based on mean square error (MSE) measure and other settings were determined with Matlab 8.2 software in the default configuration. Of the selected feature set (input vector), 70% was used to train MLP, 10% was used for validation and the remaining 20% was used to test the neural network. For each feature set, MLP was implemented 15 times and the mean and variance of error was calculated (8% was considered as a maximum acceptable variance) after each iteration. Table 2 demonstrates the results achieved by MLP neural network for both feature sets. The best performance of MLP, 93.65%, was achieved by features set obtained from DISR feature selection method and the diagnostic accuracy of MLP by features obtained from mRMR feature selection method was 92.28%.

DISCUSSION AND CONCLUSION

ADHD is a disorder that is most common in children and its early diagnosis is important to prevent complications. In this paper, we presented an off-line and semi-automatic system to detect ADHD, using non-linear features of EEG signals. This system is based on non-linear (chaotic) features and neural network. So far, most studies have used linear features (spectral, time, spatial or time-frequency features) to discriminate ADHD patients [10, 14, 16, 32, 33]. Although some of these studies achieved promising results, the recent controversial issues [15] indicate that new advanced methods are needed to analyze the EEG signals. ADHD discriminating non-linear features of EEG has already been reported but only in adults. Ghassemi *et al*. [34] evaluated the non-linear features of EEG signals in 50 normal adult subjects and 10 ADHD adult subjects. They used the lyapunov exponent, the correlation dimension and wavelet-entropy features; and they could achieve 96% accuracy in ADHD diagnosis with wavelet-entropy feature. The imbalance between the number of participants in the two groups is evident in their work. Also, in another recent study, a different non-linear approach has been reported using wavelet analysis, synchronization likelihood features, and Radial Basis Function (RBF) neural network classifier [35]. The authors obtained the accuracy of 95.6% with a variance of 0.7%.

Table 2. The results achieved by multi-layer perceptron neural network for two selected feature sets via DISR and mRMR methods in EEG classification of ADHD from healthy children.

	Accuracy (%)	
	M	V
MLP with DISR	93.65	0.64
MLP with mRMR	92.28	0.82

M = mean; V = variance.

We used two feature selection methods (DISR and mRMR) prior to classification by MLP neural network to improve the accuracy. DISR method is more computationally expensive than the mRMR method, but the accuracy obtained from DISR is obviously better. Therefore, the DISR method combining MLP neural network was able to achieve good accuracy (93.65%) in classification of ADHD and normal children.

Our results show that the non-linear features are appropriate features to analyze and characterize the EEG signals. In spite of irregular EEG signals, these noise-like fluctuations are a rule-based system. Therefore, quantitative measures of chaos and non-linear features are convenient descriptive tools to characterize the main information from these complex signals. The application of non-linear analysis to EEG has provided new information about electrophysiological abnormalities in neuropsychiatric disorders that cannot be evident by linear analysis. According to our results, the brain system of ADHD children is more complex and irregular than that of normal children due to greater values of non-linear features such as FD and approximate entropy which is depicted in Figs. 4 and 5.

Many different tasks, such as go/no-go and stop-signal tasks, determine group differences of ability to inhibit responses, but our task involved both complex cognitive visual processing and numerical operations which may engage more neurons in information processing than simple tasks of inhibition. Therefore, we put the child's brain in a high function level to perform the task in order to make more difference in electrophysiological activity of the brain.

The limitations of this study are due to lack of access to more ADHD patients, we could not conduct a validation study to validate our proposed methodology; the subjects did not undergo a cognitive assessment to have a clinical assessment and to get more information about cognitive differences between groups. In future studies, we will design this task to extract the event-related potentials (ERPs) or steady-state potentials for evaluation and classification of ADHD children.

ACKNOWLEDGMENT

We acknowledge the financial support of Tehran University of Medical Sciences through a grant from Psychiatry and Psychology Research Center (Grant No. 16534). We also wish to thanks Dr.Moeini and Mr.Parvaneh for their great advices and helps in evaluation of participants and recording of EEG.

CONFLICT OF INTEREST STATEMENTS

Mohammadi MR declares that he has no conflict of interest in relation to the work in this article. Khaleghi A declares that he has no conflict of interest in relation to the work in this article. Nasrabadi AM declares that he has no conflict of interest in relation to the work in this article. Rafieivand S declares that he has no conflict of interest in relation to the work in this article. Begol M declares that he has no conflict of interest in relation to the work in this article. Zarafshan H declares that he has no conflict of interest in relation to the work in this article.

REFERENCES

[1] American Psychiatric Association., American Psychiatric Association. DSM-5 Task Force. Diagnostic and statistical manual of mental disorders : DSM-5. 5th edn. American Psychiatric Association, Washington, D.C. 2013.
[2] Meysamie A, Fard MD, Mohammadi M-R. Prevalence of attention-deficit/hyperactivity disorder symptoms in preschool-aged Iranian children. Iran J Pediatr. 2011; 21(4):467.
[3] Jafari P, Ghanizadeh A, Akhondzadeh S, Mohammadi MR. Health-related quality of life of Iranian children with attention deficit/hyperactivity disorder. Qual Life Res. 2011; 20(1):31-6.
[4] King JA, Colla M, Brass M, Heuser I, von Cramon D. Inefficient cognitive control in adult ADHD: evidence from trial-by-trial Stroop test and cued task switching performance. Behav Brain Funct. 2007; 3:42. doi: 10.1186/1744-9081-3-42.
[5] Aboitiz F, Ossandon T, Zamorano F, Palma B, Carrasco X. Irrelevant stimulus processing in ADHD: catecholamine dynamics and attentional networks. Front Psychol. 2014: 4:183. doi: 10.3389/fpsyg.2014.00183.
[6] Mohammadi MR, Malmir N, Khaleghi A. Comparison of sensorimotor rhythm (SMR) and beta training on selective attention and symptoms in children with attention deficit/ hyperactivity disorder (ADHD): a trend report. Iran J Pediatr. 2015; 10(3):165-74.
[7] Organization WH. The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines. WHO: Geneva; 1992.
[8] Lubar JF. Discourse on the development of EEG diagnostics and biofeedback for attention-deficit/hyperactivity disorders. Biofeedback Self-Reg 1991; 16(3):201-25.
[9] Tansey MA. Brainwave signatures--an index reflective of the brain's functional neuroanatomy: further findings on the effect of EEG sensorimotor rhythm biofeedback training on the neurologic precursors of learning disabilities. Int J Psychol. 1985; 3(2):85-99.
[10] Tenev A, Markovska-Simoska S, Kocarev L, Pop-Jordanov J, Muller A, Candrian G. Machine learning approach for classification of ADHD adults. Int J Psychol. 2014; 93(1):162-6. doi:10.1016/ j.ijpsycho.2013.01.008.
[11] Poil SS, Bollmann S, Ghisleni C, O'Gorman RL, Klaver P, Ball J, Eich-Hochli D, Brandeis D, Michels L. Age dependent electroencephalographic changes in attention-deficit/hyperactivity disorder (ADHD). Clin Neurophysiol. 2014; 125(8):1626-38. doi:10.1016/j.clinph.2013.12.118.
[12] Mazaheri A, Fassbender C, Coffey-Corina S, Hartanto TA, Schweitzer JB, Mangun GR. Differential oscillatory electro-encephalogram between attention-deficit/hyperactivity disorder subtypes and typically developing adolescents. Biol Psychiat. 2014; 76(5):422-9. doi:10.1016/j.biopsych.2013.08.023
[13] Fonseca LC, Tedrus GMA, Moraes Cd, Machado AdV, Almeida MPd, Oliveira DOFd. Epileptiform abnormalities and quantitative EEG in children with attention-deficit/hyperactivity disorder. Arq Neuro-Psiquiat. 2008; 66(3A):462-7.
[14] Mueller A, Candrian G, Grane VA, Kropotov JD, Ponomarev VA, Baschera G-M. Discriminating between ADHD adults and controls using independent ERP components and a support vector machine: a validation study. Nonlinear Biomed Phys. 2011; 5(1):5.
[15] Arns M, Gordon E. Quantitative EEG (QEEG) in psychiatry: Diagnostic or prognostic use? Clin Neurophysiol. 2014; 125(8):1504-6. doi:http://dx.doi.org/10.1016/j.clinph.2014.01.014.
[16] Lenartowicz A, Loo SK. Use of EEG to diagnose ADHD. Curr Psychiat Rep. 2014; 16(11):498. doi:10.1007/s11920-014-0498-0.
[17] Zarafshan H, Khaleghi A, Mohammadi MR, Moeini M, Malmir N. Electroencephalogram complexity analysis in children with attention-deficit/hyperactivity disorder during a visual cognitive task. J Clin Exper Neuropsychol. 2015; 1-9.
[18] Khaleghi A, Sheikhani A, Mohammadi MR, Nasrabadi AM, Vand SR, Zarafshan H, Moeini M. EEG classification of adolescents with type I and type II of bipolar disorder. Australasian Phys Eng Sci Med. 2015; 38(4):551-9.
[19] Ahmadlou M, Adeli H, Adeli A. Fractality and a wavelet-chaos-neural network methodology for EEG-based diagnosis of autistic spectrum disorder. J Clin Neurophysiol. 2010; 27(5):328-33.
[20] Sadatnezhad K, Boostani R, Ghanizadeh A. Classification of BMD and ADHD patients using their EEG signals. Exp Syst Appl. 2011; 38(3):1956-63.
[21] Higuchi T. Approach to an irregular time series on the basis of the fractal theory. Physica D. 1988; 31(2):277-83.
[22] Petrosian A, Kolmogorov complexity of finite sequences and recognition of different preictal EEG patterns. Proc 8th Comp Med Sy. 1995; 212-7.
[23] Stoica P, Moses RL. Introduction to spectral analysis, vol 1. Prentice hall Upper Saddle River, NJ. 1997.
[24] Principe J, Lo P. Towards the determination of the largest Lyapunov exponent of EEG segments. In: Proceedings of the conference on measuring chaos in the human brain. World Scientific, Singapore, 1991. pp. 156-66.
[25] Röschke J, Fell J, Beckmann P. The calculation of the first positive Lyapunov exponent in sleep EEG data. Electroencephalography Clin Neurophysiol. 1993; 86(5):348-52.

[26] Acharya UR, Vinitha Sree S, Swapna G, Martis RJ, Suri JS. Automated EEG analysis of epilepsy: a review. Knowledge-Based Syst. 2013; 45:147-65.
[27] Peng H, Long F, Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE T Pattern Anal. 2005; 27(8):1226-38.
[28] Brown G, Pocock A, Zhao M-J, Luján M. Conditional likelihood maximisation: a unifying framework for information theoretic feature selection. J Mach Learn Res. 2012; 13(1):27-66.
[29] Meyer PE, Bontempi G. On the use of variable complementarity for feature selection in cancer classification. In: Applications of Evolutionary Computing. Springer. 2006. pp. 91-102.
[30] Theodoridis S, Pikrakis A, Koutroumbas K, Cavouras D. Introduction to Pattern Recognition: A Matlab Approach: A Matlab Approach. Academic Press. 2010.
[31] Nelles O. Nonlinear system identification: from classical approaches to neural networks and fuzzy models. Springer. 2001.
[32] Liechti MD, Valko L, Muller UC, Dohnert M, Drechsler R, Steinhausen HC, Brandeis D. Diagnostic value of resting electroencephalogram in attention-deficit/hyperactivity disorder across the lifespan. Brain Topogr. 2013; 26(1):135-51. doi:10.1007/s10548-012-0258-6.
[33] Alba-Sanchez F, Yanez-Suarez O, Brust-Carmona H. Assisted diagnosis of attention-deficit hyperactivity disorder through EEG bandpower clustering with self-organizing maps. Conf Proc IEEE Eng Med Biol Soc, 2010; 2010:2447-50.
[34] Ghassemi F, Hassan_Moradi M, Tehrani-Doost M, Abootalebi V. Using non-linear features of EEG for ADHD/normal participants' classification. Proc Soc Behav Sci. 2012; 32:148-52.
[35] Ahmadlou M, Adeli H. Wavelet-synchronization methodology: a new approach for EEG-based diagnosis of ADHD. Clin EEG Neurosci. 2010; 41(1):1-10.