

Research article

Functional analysis of ADHD in children using nonlinear features of EEG signals

Shiva Khoshnoud¹, Mohammad Ali Nazari¹, Mousa Shamsi^{1,*}

¹ Biomedical Engineering Faculty, Sahand University of Technology, Tabriz, Iran

² Cognitive Neuroscience Laboratory, Department of Psychology, University of Tabriz, Tabriz, Iran

*Correspondence: shamsi@sut.ac.ir (Mousa Shamsi)

<https://doi.org/10.31083/JIN-170033>

Abstract

Attention deficit hyperactivity disorder is a neurodevelopmental condition associated with varying levels of hyperactivity, inattention, and impulsivity. This study investigated brain function in children with attention deficit hyperactivity disorder using measures of nonlinear dynamics in electroencephalogram signals during rest. During eyes-closed resting, 19 channel electroencephalogram signals were recorded from 12 ADHD and 12 normal age-matched children. The multifractal singularity spectrum, the largest Lyapunov exponent, and approximate entropy were employed to quantify the chaotic nonlinear dynamics of these electroencephalogram signals. As confirmed by Wilcoxon rank sum test, the largest Lyapunov exponent over left frontal-central cortex exhibited a significant difference between attention deficit hyperactivity disorder subjects and the age-matched control groups. Further, mean approximate entropy was significantly lower in attention deficit hyperactivity disorder subjects in prefrontal cortex. The singularity spectrum was also considerably altered in attention deficit hyperactivity disorder subjects when compared to control children. Evaluation of these features was performed with two classifiers: a support vector machine and a radial basis function neural network. For better comparison, subject classification based on frequency band power was assessed using the same types of classifiers. Nonlinear features provided better discrimination between attention deficit hyperactivity disorder and control than band power features. Under four-fold cross-validation testing, the support vector machine gave 83.33% accurate classification results.

Keywords

Attention deficit hyperactivity disorder; EEG signals; largest Lyapunov exponent; approximate entropy; multifractal DFA; classification

Submitted: January 23, 2017; Accepted: May 8, 2017

1. Introduction

Attention deficit hyperactivity disorder (ADHD) refers to a heterogeneous neurodevelopmental condition or set of conditions in which children or adults exhibit varying levels of inattention, impulsivity and physical restlessness and/or hyperactivity symptoms [1]. Approximately 3–7% of school-aged children are affected by ADHD [1]. According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-V) this syndrome has three clinical subtypes: predominantly inattentive, predominantly hyperactive-impulsive, and combined [2]. Currently, diagnosis of ADHD is based on levels of symptoms listed as DSM-V diagnostic criteria, as assessed from diagnostic interviews.

Assessment of electroencephalogram (EEG) characteristics associated with ADHD, has generated substantial interest, resulting in a large number of investigations [3–16]. Most of these studies concern frequency-domain indicators, typically absolute and relative power estimates for different frequency bands or ratios of power in different frequency bands [7, 8, 10–13, 15]. Although simple to compute and visualize, these approaches are not capable of measuring the nonlinear structure of EEG brain dynamics. Methods from nonlinear dynamics and chaos theory have been used to investigate nonlinear properties of brain dynamics. EEG coherence provides useful information about the coupling of brain activity in different

regions [8]. However, the inability of coherence to characterize nonlinear interdependencies, specifically for non-stationary time series, has led to the use of nonlinear synchronization methods for analysing functional brain connectivity [3–6, 9, 14, 16].

According to Stam [17], distinct states of brain activity have different chaotic dynamics which might be quantified by nonlinear measures including entropy and Lyapunov exponents. Approximate entropy (ApEn) has proved to be particularly useful for characterizing short, noisy time series. It is capable of providing a robust, model-independent, and information-theoretic estimate of dynamical complexity [14]. A recently published study has investigated the nonlinear features of EEG signals in ADHD and normal adult subjects performing a continuous performance test [9].

Fractality is another common property of time series representing the dynamics of complex systems. Brain activity, as a complex dynamical system, exhibits a multifractal structure [18, 19]. Fractal analysis of EEG time series has proved to be useful for describing brain activity during sleep [20]. Fetterhoff *et al.* [21] show that multifractal firing patterns of hippocampal spike trains are more complex during a working memory task and significantly declined with the administration of memory impairment in rats. Zorick *et al.* [19] showed the capability of MFDFA to interfere with the recognition of changes in states of consciousness.

To the knowledge of the authors, the present study is the first

to assess multifractal spectra of EEG signals in ADHD subjects. It was predicted that multifractal analysis of EEG signal might provide useful information about the dynamics of brain electrical activity in ADHD subjects. Nonlinear EEG dynamic features were employed, including ApEn, the largest Lyapunov exponent (LLE), and multifractal spectra, to discriminate between children with ADHD and age-matched controls.

2. Materials and methods

The aim of this research was to analyse nonlinear dynamical properties of EEG signals in children with ADHD and to present a classification system based on these features. The algorithm consisted of four steps: signal preprocessing, feature extraction, feature reduction, and subject classification. Fig. 1 shows all analytic steps involved and subsequently reported here.

2.1. Dataset and pre-processing

The dataset employed in this study was EEG time series recorded at the Atieh Comprehensive Centre for Psychology and Nerve Disorders, Tehran, Iran [3]. The subject group comprised 24 right-handed children 7–12 years old (12 ADHD and 12 age-matched healthy subjects). The ADHD group included all clinical subtypes (hyperactive-impulsive, inattentive, and combined). All subjects were asked to sit comfortably in a silent room with eyes-closed. The EEG data were collected according to the 10–20 International system (Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Fz, Cz, and Pz,) with eyes-closed during rest [3]. The eyes-closed condition was preferred as the preference was to analyse ADHD brain dynamics in the absence of external stimuli. Each recording consisted of 19 channels of EEG data with a sampling frequency rate of 256 Hz and 16-bit resolution [3]. Signal pre-processing was performed in Matlab (MatlabR2013a). Artefact rejection was based on both visual inspection and computerized review. To remove out-of-band noise, EEG time series were lowpass filtered at 60 Hz and highpass filtered at a 1 Hz cut-off frequency. Line noise suppression was by notch filtering at 50 Hz. Signals were also visually analysed by an expert and any artefacts discarded. This procedure reserved about one-minute of artefact-free EEG trace out of a three-minute duration signal (15,000 time samples).

2.2. Feature extraction

Four types of feature extraction methods were selected, frequency band powers and three nonlinear features: the LLE, ApEn, and the height and width of the multifractal singularity spectrum of the EEG time series. With 19 channels for each subject, a total of 24,198 features were extracted for the analysis. The following feature extraction methods were applied to selected one minute durations of noise-free signal obtained from each channel.

2.2.1. EEG frequency band power

EEG signals were filtered with a 16th order Butterworth band-pass filter to extract four typical frequency bands: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), and beta (13–30 Hz). For each band, power was calculated by invoking Perceval's theorem and summing the signal in the time domain [22].

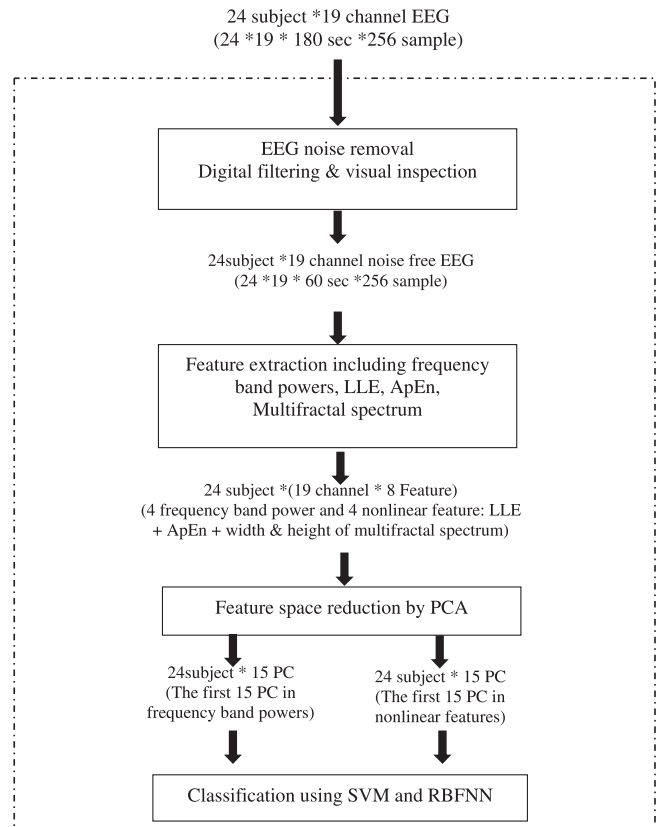


Fig. 1. Overall procedure for functional brain dynamic analysis of ADHD and control children using nonlinear dynamical features of EEG signals. First step: Signal was preprocessed to remove any noise and artefacts. Followed by calculation of LLE, ApEn, and multifractal spectra for every channel of the recorded EEG signals. In parallel with nonlinear features, frequency band power estimates were also calculated. Subsequently, the principal component coefficients of these features were computed and the 15 largest components of each feature category employed as final features. Finally, subject group classification was implemented with two classifiers: a support vector machine classifier and a radial basis function neural network.

2.2.2. Largest Lyapunov exponent

Lyapunov exponents quantify exponential divergence or convergence of nearby trajectories in a state space. Since it may indicate chaotic dynamics in a system, existence of positive LLE is of special importance [17]. Rosenstein *et al.* [23] have suggested a straightforward approach to calculate LLE [23]. The basic idea of the algorithm is that the average growth rate of inter-vector differences starting from pairs of nearest neighbours is bounded by the largest Lyapunov component. Based on Taken's theorem [24], reconstruction of trajectories from time series in a state space is given by:

$$X = (X_1, X_2, \dots, X_m)^T \quad (1)$$

where X_i is the state of the system at discrete time i , which for a N point EEG signal $x = x_1, x_2, \dots, x_N$ is calculated from:

$$X_i = (x_i, x_{i+J}, \dots, x_{i+(m-1)J}) \quad (2)$$

where J is the lag delay and m is the embedding dimension. So X would be an $M \times m$ matrix in which $M = N - (m - 1)J$. To avoid information loss, the lag delay and embedding dimension are determined

so that the reconstructed trajectories in the new state space are not folded [17]. After reconstructing the trajectories, the initial distance from the “ j th point” X_j to its nearest neighbour $X_{\hat{j}}$ is expressed as:

$$d_j(0) = \min_{X_{\hat{j}}} \|X_j - X_{\hat{j}}\| \quad (3)$$

where $d_j(0)$ is the initial distance and $\|\dots\|$ denotes the Euclidean norm. In Rosenstein’s algorithm, the growth rate of inter-vector differences starting from nearest neighbour pairs is bounded by LLE [23]. Therefore, $LLE(\lambda_1)$ of the EEG signal is calculated from:

$$\lambda_1(i) = \frac{1}{i\Delta t} \frac{1}{(M-i)} \sum_{j=1}^{M-i} \ln \frac{d_j(i)}{d_j(0)} \quad (4)$$

where Δt is the sampling period of the EEG signal and $d_j(i)$ is the distance between the j th pair of the nearest neighbours after i discrete time steps. A positive Lyapunov exponent implies that any two points that are initially very close together will diverge exponentially quickly in the direction of the positive Lyapunov exponent [23]. Since reconstructing a trajectory in a state space with the embedding dimension lower than its minimum embedding dimension, causes the trajectory to be folded, the minimum embedding dimension must be calculated first [17]. Moreover, very small lag delays may bring about almost linear trajectory reconstruction with high correlations between consecutive phase space points and very large delays may disregard the deterministic structure of the sequence [17]. According to previous investigations, the minimum embedding dimension for EEGs varies mostly in the range of 9 ± 3 [3, 4]. The minimum embedding dimension of the EEG time series employed here, for all channels and all subjects, calculated using the false nearest neighbour algorithm [25], varied in the range of (6–15), so the optimum embedding dimension employed was greater than 15. In this study, the embedding dimension was assumed to be 15 because this value was the maximum value found for minimum embedding dimension in the dataset analysed here. Another important factor was the optimum lag delay. As it gave a better result at the classification step, after testing several different values from 1 to 10, a lag delay of $J = 1$ was selected for use here.

2.2.3. Approximate entropy

The apen, as suggested by Pincus [26], is a nonlinear measure that quantifies the complexity or irregularity of a system. It measures the logarithm of the frequency at which the neighbourhoods of temporal patterns with the same length within a certain distance by one time point augmentation in phase space remain close [26]. Given the embedding dimension m , and time lag J , the state space of the N dimensional EEG signal $x = x_1, x_2, \dots, x_N$ was computed from Equation (2). ApEn was computed from the correlation integral $C^m(r)$, which represents the number of points within distance r from the i th point of the EEG time series when the signal is embedded in a phase space with embedding dimension m . The correlation integral was then calculated as:

$$C_i^m(r) = \frac{1}{(N-m+1)} \sum_{j=1}^{N-m+1} \Theta(r - \|X_i - X_j\|) \quad (5)$$

In eqn(5), Θ is the Heaviside function and X_i and X_j are vectors in the state space and r is a threshold value for distance. Finally, the ApEn is defined as [26]:

$$\text{ApEn}(m, r, N) = \frac{1}{(N-m+1)} \sum_{i=1}^{N-m+1} \log C_i^m(r) - \frac{1}{(N-m)} \sum_{i=1}^{N-m} \log C_i^{m+1}(r) \quad (6)$$

The embedding dimension can be found as described above. In some applications a value between 10 and 25 % of the data standard deviation (SD) is used for the threshold value [27]. Here, the parameter r (distance) was set to be 10 % of the SD.

2.2.4. Multifractal detrended fluctuation analysis

Detrended fluctuation analysis (DFA) is a widely-used method for exploring the fractal properties of a time series. Nonetheless, time series with a complex structure are multifractal and multifractal detrended fluctuation analysis (MFDFA) has been proposed for evaluation of their fluctuations [28]. The MFDFA procedure consists of five steps. In the first step, the structure of the EEG signal must be converted into a random walk [28]. If $x = \{x_1, x_2, \dots, x_N\}$ is an EEG signal with length N , then:

$$Y(i) = \sum_{k=1}^i [x_k - \langle x \rangle], \quad i = 1, 2, \dots, N \quad (7)$$

where $\langle x \rangle$ denotes the average amplitude of an EEG time series. Next, the profile $Y(i)$ is divided into $N_s \equiv \text{int}(N/s)$ non-overlapping segments with equal lengths s . For each segment, the root mean square (RMS) variance is calculated from Equation (8), in which $v = N_s + 1, \dots, 2N_s$ and $y_v(i)$ is the fitting polynomial in segment v :

$$F^2(s, v) = \frac{1}{s} \sum_{i=1}^s \{Y[N - (v - N_s)s + i] - y_v(i)\}^2 \quad (8)$$

To obtain the q th-order fluctuation function, the mean RMS value over all the segments is calculated from:

$$F_q(s) = \left\{ \frac{1}{2N_s} \sum_{v=1}^{2N_s} [F^2(s, v)]^{q/2} \right\}^{1/q} \quad (9)$$

Spatial and temporal variations in the scale-invariant structure of a multifractal time series require Steps two and three to be repeated at several different time scales s . Finally, the scaling behaviour of the fluctuation functions is determined by analysing log–log plots of $F_q(s)$ versus s for each value of q :

$$F_q(s) \approx s^{h(q)} \quad (10)$$

where $h(q)$ is called the q -order Hurst exponent [28]. For multifractal EEG time series, there is significant dependence of $h(q)$ on q [28]. For positive values of q , $h(q)$ describes the scaling behaviour of segments with large fluctuations. For negative values of q , $h(q)$ describes the scaling behaviour of segments with small fluctuations [28]. The q -order Hurst exponent $h(q)$, is only one of several factors used to parameterize the multifractal structure of an EEG time series. It is related to the scaling exponents $\tau(q)$ by:

$$\tau(q) = qh(q) - 1 \quad (11)$$

Thereafter, scaling exponents can be converted into the q -order singularity exponent (α) and the q -order singularity dimension ($f(\alpha)$) with the following equations to obtain a multifractal singu-

larity spectrum:

$$\alpha = \tau'(q) \quad \text{and} \quad f(\alpha) = q\alpha - \tau(q) \quad (12)$$

where $\tau'(q)$ is the derivative of $\tau(q)$. The width and shape of the multifractal spectrum are valuable factors for distinguishing different multifractal structures [28]. In this study, the width ($d\alpha$) and its height ($df(\alpha)$) were used to evaluate the multifractal spectrum of EEG signals:

$$\begin{aligned} d\alpha &= \max(\alpha) - \min(\alpha) \\ df(\alpha) &= \max f(\alpha) - \min f(\alpha) \end{aligned} \quad (13)$$

2.3. Feature reduction using principal component analysis

Principal component analysis (PCA) was used to extract features strongly correlated to each other. In this procedure, the principal component coefficients of the dataset are calculated and the k largest components are considered to give the joint features of the dataset that capture the most data variance in k dimensions [29]. Considering the feature vector X with mean μ_x and a covariance matrix of C_x , the eigenvectors and eigenvalues of the features can be calculated and sorted into an order of descending eigenvalues. In this case, transforming a data vector X gives:

$$y = A_K(X - \mu_x) \quad (14)$$

where A_K is a matrix consisting of the first k eigenvectors of the covariance matrix C_x and the components of y are then calculated as the orthogonal transformations of X . Finally, the original data vector X can be approximately reconstructed from y by:

$$X = A_K^T y + \mu_x \quad (15)$$

In this study, the principal component coefficients of previously extracted features (Frequency band powers as well as nonlinear features) were calculated and the 15 largest coefficients of each category were then used as the final feature set.

2.4. Support vector machines classifier

The support vector machine (SVM) proposed by Vapnik [30] is a supervised machine learning algorithm preferred for many classification problems. The aim of the SVM is to compute an optimal separating hyperplane to which the distance from each nearest data sample in each class is maximized. In non-separable cases, SVM offers a solution using kernel mapping. To allow separation in such cases, the data can be projected into a higher-dimensional feature space using a nonlinear function ($\phi(\cdot)$).

In general, the formulation of the hyperplane is as follows:

$$W^T \phi(\cdot) + b = 0 \quad (16)$$

in which W and b are a weight vector and a bias term, respectively. To find such an optimum hyperplane, the optimization problem is:

$$\begin{aligned} \text{Minimize } J(W) &= \frac{1}{2} \|W\|^2 \\ \text{Subject to } d_i(W^T \phi(x) + b) &\geq 1 \end{aligned} \quad (17)$$

The above-mentioned problem is solved using Lagrangian optimization theory [31]. Here, linear, polynomial, and radial basis function (RBF) kernel functions were tested. The RBF kernel led to better discrimination accuracy.

2.5. Radial basis function neural networks classifier

The radial basis function neural network (RBFNN), a supervised machine learning algorithm, is an artificial neural network that uses radial basis functions as activation functions [32]. The output of the network is a linear combination of radial basis functions of the inputs and neuron node parameters. A RBFNN typically has three layers: input, hidden with a nonlinear RBF activation function, and a linear output layer. For an input vector X , the output of the network is given by:

$$\phi(x) = \sum_{i=1}^N a_i \rho(\|X - C_i\|) \quad (18)$$

where N is the number of neurons in the hidden layer, C_i is the centre vector for neuron node i , and a_i is the weight of neuron node i in the output neuron node [32]. The most commonly used radial basis function (ρ) is Gaussian:

$$\rho(\|X - C_i\|) = \exp[-\beta \|X - C_i\|^2] \quad (19)$$

RBFNNs are typically trained by back propagation [32]. In this study, the MATLAB neural network toolbox was used to implement a RBFNN. It is important to consider the optimum value for the RBFNN spread parameter. It should be large enough that the radial basis layer neuron nodes respond to overlapping regions in the input space, but not so large that all neuron nodes respond in essentially the same manner.

3. Result

To evaluate the proposed algorithm, noise removal was first performed. The features mentioned above for the noise-free one-minute long EEG signals were then extracted. Because of the small sample size, evaluation of discrepancies in nonlinear feature category between ADHD and normal subjects was performed using the nonparametric Wilcoxon rank sum test. The most highly distinguished LLEs were the scalp channels Fp1 ($p = 0.0379$), Fp2 ($p = 0.0499$), F3 ($p = 0.0379$), C3 ($p = 0.0148$), F7 ($p = 0.0148$), and Cz ($p = 0.0148$).

The average LLE of the ADHD group was significantly higher than the control group at left frontal central scalp electrode channels. This suggests that for those channels EEG chaoticity for ADHD subjects was higher than the age-matched control group.

Similarly, the Wilcoxon rank sum test revealed significant group differences for the ApEn measure. Mean ApEn in ADHD subjects was lower than for the control group. Mean ApEn for ADHD subjects trended lower at all 19 EEG channels, reaching statistical significance at channels Fp1 ($p = 0.0351$) and Fp2 ($p = 0.0531$) (Fig. 2).

High values of ApEn indicate a more complex and irregular system and low values, a lower complexity system. Fig. 2 shows that the complexity of the EEG signals in the ADHD group was lower over the prefrontal scalp.

The height of the multifractal spectrum for ADHD subjects at most scalp channels was dissimilar to that of controls. No significant differences were observed between the widths of the multifractal spectrum in the two groups, except for channels Fp1 ($p = 0.0166$) and F7 ($p = 0.0404$).

The height of the multifractal spectrum statistical evaluation result is shown in Fig. 3. Multifractal spectrum differences between ADHD and control subjects were found in some frontal and right scalp channel signals including Fp1 ($p = 0.0304$), F4 ($p = 0.0086$),

F7 ($p = 0.0404$), F8 ($p = 0.0141$), P4 ($p = 0.0226$), T4 ($p = 0.0141$), O2 ($p = 0.0464$), Cz ($p = 0.0262$), and Pz ($p = 0.0086$). For ADHD subjects the multifractal spectrum was higher than the control group over these scalp electrodes.

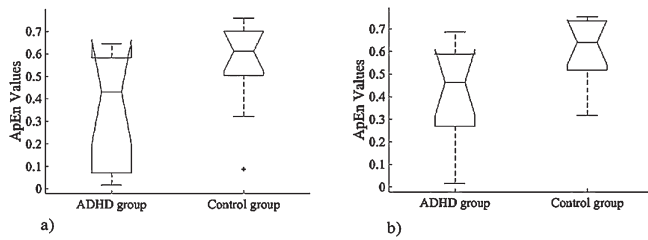


Fig. 2. Results of the Wilcoxon rank sum test for the mean ApEn feature between ADHD and control group at channels: (a) Fp1, (b) Fp2. Mean ApEn in subjects with ADHD was significantly lower than control in these two scalp electrodes.

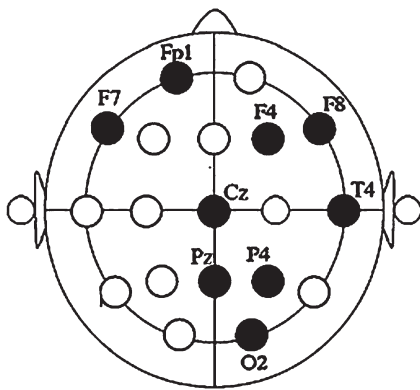


Fig. 3. Black disks indicate channels with multifractal spectrum height differences (p -value < 0.05) between ADHD and normal subjects. The multifractal spectrum for ADHD subjects was significantly higher than the control group over these scalp electrodes.

After evaluating the nonlinear features of channel signals, their ability to discriminate the two groups was considered. The first 15 principle components in each feature category (band power and nonlinear) were chosen; these 15 PCs preserved 90% of signal variance. Next, two classifiers, an SVM and an RBFNN, were constructed and the first 15 principle components in each feature category were separately analyzed. 80% of the data were used to train the classifiers; the other 20% were kept for testing. The four-fold cross validation procedure was used for evaluating the classifiers' performance. Fig. 4 summarizes the results of applying frequency band power values in the classifiers. As shown in Fig. 4, using alpha power in the SVM gave the highest accuracy (83.33%). Moreover, beta and delta power using RBFNN gave the least accurate classification discrimination (58.33%).

In Table 1, the results of using nonlinear features (LLE, ApEn, and the multifractal spectrum) are given for both classifiers. When it was used as input to the SVM classifier, the multifractal spectrum features produced the highest accuracy, 79.17%. Accuracy was lowest when ApEn was used as the classifier input for SVM (59%). Largest Lyapunov exponents with RBFNN exhibited a high accuracy (79.17%) in classifying two groups as well.

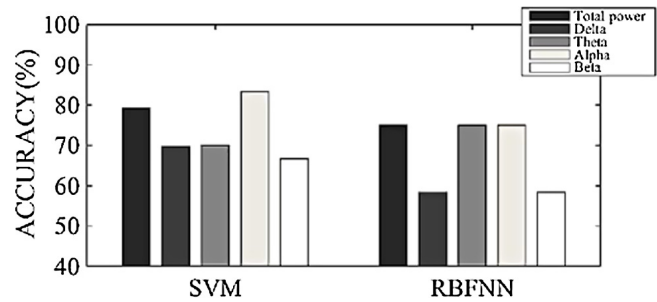


Fig. 4. Group classification using frequency band powers.

Table 1. Classification using nonlinear dynamical properties

Classifier	Nonlinear features		
	Largest Lyapunov exponent	Approximate entropy	Multi fractal spectrum
SVM	70.83	59	79.17
RBF	79.17	75	66.67

Table 2. Classification using combined features

Classifier	Band power features		Nonlinear features	
	Without PCA (%)	With PCA (%)	Without PCA (%)	With PCA (%)
SVM	79.17	79.17	83.33	83.33
RBF	83.33	70.83	70.83	66.67

Finally, classification results using combined band power and nonlinear features are shown in Table 2. SVM accuracy for frequency band power (79.17%) did not change after PCA. Accuracy of band power features with RBFNN before PCA (83.33%) declined after PCA (70.83%). The highest classification accuracy (83.33%) was achieved when nonlinear features were applied to the SVM classifier. This result remained unchanged after applying PCA.

In general, using nonlinear features in the SVM classifier provided much better classification of ADHD and controls. Findings confirmed that these nonlinear features increased the accuracy of classification compared to using band power features.

4. Discussion

In this study, resting EEG from 12 ADHD subjects and 12 controls were analysed using nonlinear analysis metrics. It was believed that nonlinear analysis would offer new tools for exploration of brain information processing. Thus, three measures of nonlinear dynamics were explored: LLEs, ApEns, and the multifractal spectrum of one-minute noise-free EEG signals extracted from 19 scalp channels. To visualize the effects of each feature set, the mean and SD of all features for the two groups are shown in Fig. 5. It is clear that the MF spectrum height contains more discriminative information.

A Wilcoxon rank sum test revealed that mean LLE for ADHD subjects was significantly higher than the control group in scalp electrode channels over left frontal and left central scalp. Even though the LLE distinguished the two groups, it could not be concluded with certainty that these results indicate impaired cortical information

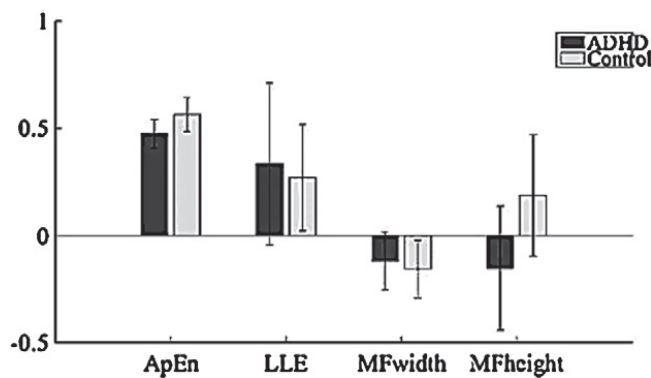


Fig. 5. Mean and SD of different features for two groups (ADHD: Black bar, Control: Gray bar).

processing in these brain regions.

ApEn was significantly lower in subjects with ADHD in scalp channels over prefrontal scalp. This suggests that the complexity of cortical dynamics in prefrontal regions of ADHD subjects may be lower than that of control subjects. Reduced EEG complexity has been associated with elevated dysfunction in some previous studies [14, 17, 33, 34]. Ikawa *et al.* [34] described region-specific correlations between neuropsychological performance and dynamical EEG complexity in subjects with attention deficit disorder. Sohn *et al.* [14] also suggested an association between reduced cognitive performance on the continuous performance test in ADHD adolescents and impaired cortical information processing, as indicated by the lower complexity of the EEG. Therefore the reduced ApEn in ADHD children reported for this study might possibly relate to their relative inability to deal with demands for flexible information processing, also noted to be of difficulty for ADHD adults.

To the authors' best knowledge, this is the first time that MFDFA has been used for discrimination of ADHD. Several recent studies suggest that changes in the multifractal structure of EEG signal reflect changes in the adaptability of underlying physiological process [19, 21]. According to the results shown in Fig. 3, the shape of the multifractal spectrum, including both its width and height differed in ADHD and healthy children. Specifically, in scalp channels over the right hemisphere the multifractal spectrum for ADHD subjects was higher than the age-matched control subjects. The elevated height and width of multifractal spectrum in children with ADHD show a more complex pattern in these children compared to healthy subjects. These results are in line with previous study of epileptic subjects which show that the degree of multifractality of EEG for patients in an epileptic seizure are much higher when compared to normal healthy people [18].

It has been suggested that symptoms of ADHD originate from dysfunction of neural systems in brain regions associated with attention and response inhibition [7]. A vast majority of previous investigation provides evidence for a deficit in the allocation of neural resources in the frontal cortices [11–13, 35]. A neuropsychological approach was previously used to examine the frontal lobe and right parietal lobe theories of attention deficit hyperactivity disorder [36]. Hence, lower ApEn at prefrontal and higher LLE at frontal scalp electrodes (Fp1, Fp2, F3, F7) might suggest a lower complexity and predictability of underlying neural systems at these scalp regions. Also the distinct multifractal spectrum shape at frontal (Fp1, F4, F7,

F8) and right parietal scalp channels might indicate impaired cortical activity at these regions compared to that of healthy subjects.

These features, as well as spectral power features, were used for classification (by use of two different classifiers: SVM and RBFNN). Classification accuracy was significantly higher when nonlinear features were used as input. In classifications using frequency band power, the highest accuracy (83.33%) was achieved using the alpha band power. For nonlinear features, using multifractal spectral features in an SVM classifier produced the highest accuracy (79.17%). This did not improve after PCA. Maximal accuracy (83.33%) was achieved using a combination of nonlinear features in an SVM classifier. These results confirmed that the nonlinear features of scalp EEG signals tested here can increase the accuracy of classification compared to using band power features.

5. Conclusion

The ability of nonlinear dynamical analysis to discriminate EEG signals of children with ADHD in comparison with age-matched healthy subjects was investigated. To do so, three nonlinear dynamical measures, LLE, ApEn, and the multifractal spectra of EEG signals were extracted from different channels of the scalp and were classified. Results showed that mean LLE, mean ApEn, and mean multifractal spectrum shape were significantly different for both groups. Using nonlinear features along with SVM yielded a high accuracy (83.3%) of ADHD identification. This was the first investigation to employ MFDFA for the discrimination of ADHD. However, small sample size was a limitation of the study. It is suggested that nonlinear measures along with large datasets may reveal more aspects of cognitive information processing deficit in children with ADHD.

This study investigated nonlinear properties of EEG signals during rest. Nevertheless, the allocation of neural resources differs when a subject directs their attention in an experimentally controlled situation. Therefore, the investigation of nonlinear measures of cognitive performance tasks should be further explored in future studies.

Acknowledgments

The authors would like to express their gratitude to Mehran Ahmadlou for sharing data with them. They also wish to thank Scott Makeig for his important collaboration and technical assistance in the preparation of this paper with the first author.

Conflict of Interest

All authors declare no conflicts of interest.

References

- [1] Sadock BJ, Kaplan HI, Sadock VA (2007) *Kaplan & Sadock's Synopsis of Psychiatry: Behavioral Sciences/Clinical Psychiatry* (10th ed.). Philadelphia, PA, Lippincott Williams & Wilkins.
- [2] Jukić V, Arbanas G (2014) *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. Jastrebarsko, Naklada Slap.
- [3] Ahmadlou M, Adeli H (2010) Wavelet-synchronization methodology: a new approach for EEG-based diagnosis of ADHD. *Clinical EEG & Neuroscience* **41**(1), 1-10.
- [4] Ahmadlou M, Adeli H (2011) Fuzzy synchronization likelihood with application to attention-deficit/hyperactivity disorder. *Clinical EEG & Neuroscience* **42**(1), 6-13.

- [5] Ahmadlou M, Adeli H (2011) Functional community analysis of brain: A new approach for EEG-based investigation of the brain pathology. *Neuroimage* **58**(2), 401-408.
- [6] Adámek V, Valeš F (2012) Graph theoretical analysis of organization of functional brain networks in ADHD. *Clinical EEG & Neuroscience* **43**(1), 5-13.
- [7] Barry RJ, Clarke AR, Johnstone SJ (2003) A review of electrophysiology in attention-deficit/hyperactivity disorder: I. Qualitative and quantitative electroencephalography. *Clinical Neurophysiology* **114**(2), 171-183.
- [8] Barry RJ, Clarke AR, McCarthy R, Selikowitz M (2002) EEG coherence in attention-deficit/hyperactivity disorder: a comparative study of two DSM-IV types. *Clinical Neurophysiology* **113**(4), 579-585.
- [9] Ghassemi F, Hassan Moradi M, Tehrani-Doost M, Abootalebi V (2012) Using non-linear features of EEG for ADHD/normal participants' classification. *Procedia-Social and Behavioral Sciences* **32**, 148-152.
- [10] Lazzaro I, Gordon E, Whitmont S, Meares R, Clarke S (2001) The modulation of late component event related potentials by pre-stimulus EEG theta activity in ADHD. *International Journal of Neuroscience* **107**(3, 4), 247-264.
- [11] Lazzaro I, Gordon E, Whitmont S, Plahn M, Li W, Clarke S, Dosen A, Meares R (1998) Quantified EEG activity in adolescent attention deficit hyperactivity disorder. *Clinical Electroencephalography* **29**(1), 37-42.
- [12] Lubar JF (1991) Discourse on the development of EEG diagnostics and biofeedback for attention-deficit/hyperactivity disorders. *Biofeedback and Self-regulation* **16**(3), 201-205.
- [13] Mann CA, Lubar JF, Zimmerman AW, Miller CA, Muenchen RA (1992) Quantitative analysis of EEG in boys with attention-deficit-hyperactivity disorder: Controlled study with clinical implications. *Pediatric Neurology* **8**(1), 30-36.
- [14] Sohn H, Kim I, Lee W, Peterson BS, Hong H, Chae JH, Hong S, Jeong J (2010) Linear and non-linear EEG analysis of adolescents with attention-deficit/hyperactivity disorder during a cognitive task. *Clinical Neurophysiology Official Journal of the International Federation of Clinical Neurophysiology* **121**(11), 1863-1870.
- [15] Swartwood JN, Swartwood MO, Lubar JF, Timmermann DL (2003) EEG differences in ADHD-combined type during baseline and cognitive tasks. *Pediatric Neurology* **28**(3), 199-204.
- [16] Tcheslavski GV, Beex AA (2006) Phase synchrony and coherence analyses of EEG as tools to discriminate between children with and without attention deficit disorder. *Biomedical Signal Processing & Control* **1**(2), 151-161.
- [17] Stam CJ (2005) Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. *Clinical Neurophysiology* **116**(10), 2266-2301.
- [18] Dutta S, Ghosh D, Samanta S, Dey S (2014) Multifractal parameters as an indication of different physiological and pathological states of the human brain. *Physica A Statistical Mechanics & Its Applications* **396**(2), 155-163.
- [19] Zorick T, Mandelkern MA (2013) Multifractal Detrended Fluctuation Analysis of Human EEG: Preliminary Investigation and Comparison with the Wavelet Transform Modulus Maxima Technique. *Plos One* **8**(7), e68360.
- [20] Weiss B, Clemens Z, Bódizs R, Vágó Z, Halász P (2009) Spatio-temporal analysis of monofractal and multifractal properties of the human sleep EEG. *Journal of Neuroscience Methods* **185**(1), 116-124.
- [21] Fetterhoff D, Kraft RA, Sandler RA, Ioan O, Sexton CA, Marmarelis VZ, Hampson RE, Deadwyler SA (2015) Distinguishing cognitive state with multifractal complexity of hippocampal interspike interval sequences. *Frontiers in Systems Neuroscience* **9**, 130.
- [22] Gradshteyn IS, Ryzhik IM (2014) *Table of Integrals, Series, and Products*. San Diego, Academic Press.
- [23] Rosenstein MT, Collins JJ, Luca CJD (1993) A practical method for calculating largest Lyapunov exponents from small data sets. *Physica D-nonlinear Phenomena* **65**(1, 2), 117-134.
- [24] Takens F (1981) Detecting strange attractors in turbulence. In: D. Rand and LS. Young (Eds) *Dynamical Systems and Turbulence* (pp. 366-381), Groningen, Rijksuniversiteit Groningen. Mathematisch Instituut.
- [25] Kennel MB, Brown R, Abarbanel HD (1992) Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical review A, Atomic, molecular, and optical physics* **45**(6), 3403-3411.
- [26] Pincus SM (1991) Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences of the United States of America* **88**(6), 2297-2301.
- [27] Sanei S, Chambers JA (2013) *EEG Signal Processing*. Somerset, Wiley.
- [28] Kantelhardt JW, Zschiegner SA, Koscielny-Bunde E, Havlin S, Bunde A, Stanley HE (2002) Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A Statistical Mechanics & Its Applications* **316**(1-4), 87-114.
- [29] Semmlow JL, Griffel B (2009) *Biosignal and Medical Image Processing*. Boca Raton, CRC Press.
- [30] Vapnik VN, Lerner AY (1963) Recognition of Patterns with help of Generalized Portraits. *Avtomat i Telemekh* **24**(6), 774-780.
- [31] Vapnik VN (1999) *The Nature of Statistic Learning Theory*. New York, Springer.
- [32] Schwenker F, Kestler HA, Palm G (2001) Three learning phases for radial-basis-function networks. *Neural Networks* **14**(4), 439-458.
- [33] Burioka N, Cornélissen G, Maegaki Y, Halberg F, Kaplan DT, Miyata M, Fukuoka Y, Endo M, Suyama H, Tomita Y (2005) Approximate entropy of the electroencephalogram in healthy awake subjects and absence epilepsy patients. *Clinical EEG & Neuroscience* **36**(3), 188-193.
- [34] Ikawa M, Nakanishi M, Furukawa T, Nakaaki S, Hori S, Yoshida S (2010) Relationship between EEG dimensional complexity and neuropsychological findings in Alzheimer's disease. *Psychiatry & Clinical Neurosciences* **54**(5), 537-541.
- [35] Clarke AR, Barry RJ, McCarthy R, Selikowitz M (2010) Electroencephalogram differences in two subtypes of attention-deficit/hyperactivity disorder. *Psychophysiology* **38**(2), 212-221.
- [36] Aman CJ, Roberts RJ, Pennington BF (1998) A neuropsychological examination of the underlying deficit in attention deficit hyperactivity disorder: frontal lobe versus right parietal lobe theories. *Developmental Psychology* **34**(5), 956-969.