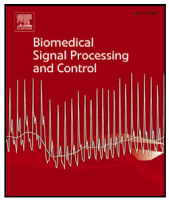
Biomedical Signal Processing and Control 81 (2023) 104512

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc



Attention deficit hyperactivity disorder recognition based on intrinsic time-scale decomposition of EEG signals

Ozlem Karabiber Cura [a,*], Sibel Kocaaslan Atli [b], Aydin Akan [c]

[a] Department of Biomedical Engineering, Izmir Katip Celebi University, Cigli, 36520, Izmir, Turkey
[b] Department of Biophysics, Faculty of Medicine, Izmir Katip Celebi University, Cigli, 36520, Izmir, Turkey
[c] Department of Electrical and Electronics Engineering, Izmir University of Economics, Balcova, 35330, Izmir, Turkey

ARTICLE INFO

Keywords:
Electroencephalography (EEG)
Attention Deficit Hyperactivity Disorder (ADHD)
Intrinsic Time-Scale Decomposition (ITD)
Machine learning
Connectivity features

ABSTRACT

Attention deficit hyperactivity disorder (ADHD), a neuro-developmental condition, is characterized by various degrees of impulsivity, hyperactivity, and inattention. Treatment of this condition and minimizing its negative impact on learning, working, forming relationships, and quality of life depends heavily on the early identification. The Electroencephalography (EEG) is a useful neuroimaging technique for understanding ADHD. This study examines the brain activity of children with ADHD by analyzing the EEG signals using the intrinsic time-scale decomposition (ITD). Different combinations of the modes, known as Proper Rotation Components (PRCs), produced by ITD, are used to extract a variety of connectivity-based features (magnitude square coherence, cross power spectral density, correlation coefficient, covariance, cohentropy coefficient, correntropy coefficient). EEG signals of 15 ADHD children and 18 age-matched health children are recorded while resting with the eyes closed. Mentioned features are calculated using different channel pairs chosen from longitudinal and transversal planes. Through various machine learning approaches and a 10-fold cross-validation method, the proposed approach is evaluated to distinguish between ADHD patients and healthy controls. Classification accuracies are obtained for the longitudinal and transverse planes, between 92.90% to 99.90% and 91.70% to 100.00%, respectively. Our results support the remarkable performance of the proposed approach, and represent a substantial advance over similar studies in terms of recognizing and classifying ADHD.

1. Introduction

Attention deficit hyperactivity disorder (ADHD) is a common neuro-developmental disease characterized by inattention, hyperactivity, and impulsivity. ADHD is a behavioral disease that is frequently diagnosed in children and adolescents. Children with ADHD exhibit these deficits early in their development and suffer significant difficulties in cognitive, social, and emotional development [1]. Approximately 5% of children in the world have ADHD [2–4]. Behavioral testing is ineffective in detecting attentional bio-markers in ADHD. Due to the lack of indicators for the disorder, diagnosis may be subjective. To overcome this struggle, brain signaling examinations are used, as well as to get at a diagnosis based on quantitative data [1].

The electroencephalography (EEG) signal which is a non-invasive neuroimaging technique is a sequence of electrical potential changes that carry information about the activities of the human brain. The use of EEG monitoring techniques is particularly well suited to pediatric research [5]. Unexpected behavior of the signals coming from the electrodes that are placed on the patient's scalp is considered a sign of ADHD. Scalp EEG data have mostly been studied in the time domain using Event-Related Potentials (ERP) or in the frequency domain using a variety of methods throughout the past several decades [6]. Different univariate and multivariate EEG signal features have commonly been evaluated utilizing linear and nonlinear techniques to identify ADHD [7].

The common approach used to examine the time-domain EEG data of ADHD patients is ERP trial averaging. Numerous research investigated morphological characteristics, such as peak amplitudes and latencies of various ERP components (P100, N100, P200, N200, P300, etc.), for the interpretation of cognitive processes in ADHD [8–12]. In one of the ERP-based studies [8], the P200 and N250 amplitudes are evaluated during target detection and it is reported that markedly reduced P200 and N250 amplitudes are observed during target identification. In another ERP-based study, morphological features namely latency and amplitude values of P300, nonlinear features called Higuchi's fractal dimension (FD), and entropy of wavelet coefficients based features in the time–frequency domain [9] are utilized. For the ADHD group, lower

* Corresponding author.
E-mail addresses: ozlem.karabiber@ikcu.edu.tr (O. Karabiber Cura), sibel.atli@ikcu.edu.tr (S. Kocaaslan Atli), akan.aydin@ieu.edu.tr (A. Akan).

https://doi.org/10.1016/j.bspc.2022.104512
Received 30 August 2022; Received in revised form 14 November 2022; Accepted 18 December 2022
Available online 22 December 2022

P300 amplitudes and longer latencies, and less complexity in the FD computation are reported. A deep learning approach using convolutional neural networks (CNN) and ERP components is presented in [11], where lower ADHD detection performance is obtained compared to previous related studies.

Numerous studies have investigated the total power, absolute, and relative power of various EEG frequency bands, including delta (≤4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (≥30 Hz), to identify ADHD [4,6,9,13]. In most of these studies, it is revealed that EEG resting-state power measurements in ADHD patients are increased in the lower frequency sub-bands; theta, and delta, and decreased in the higher frequency sub-bands; alpha, and beta. In addition to the above prominent approaches, complexity-based non-linear features have also been calculated to analyze EEG signals of ADHD patients. The following are the most significant non-linear features mentioned in the literature: entropy (approximate, sample, permutation, wavelet, fuzzy, log energy, sure, and shannon entropy) [6, 13–17], fractal dimension (Higuchi, Katz, and Petrosian) [9,15], Hurst exponent, largest Lyapunov exponent [14,15], Lempel–Ziv complexity, Kolmogorov complexity, and correlation dimension [15,18].

Using a variety of methods, EEG connectivity has also been studied in patients with ADHD in previous research. An advanced signal processing method called Bispectral analysis, which measures the quadratic phase-coupling between signal components, is one of these methods employed in the literature [6]. Another study considered the dynamic frequency warping (DFW) approach to examine the differences and similarities between intra-hemispheric or inter-hemispheric EEG channel pairs in the same frequency sub-bands [7]. They observed erroneous correlations when there is zero phase lag between the two signals. Another connectivity metric called the weighted phase lag index is introduced in a study [19], and it has been noted that EEG connectivity features do not consistently divide participants into groups. The nerve conduction network of each individual is constructed using the EEG signal by combining coherence and graph theory techniques in another study [20]. Alpha bands have been mentioned as a crucial parameter in ADHD studies when carrying out functional tasks. In a current study [21], the construction of brain networks in which the phase connectivity between the electrode pairs is calculated using the Phase Lock Value (PLV) has been proposed to investigate the EEG signals of ADHD patients. It has been reported that the highest phase connectivity disparities between ADHD and healthy groups are discovered in the anterior and posterior brain regions, and pertain to the occipital-frontal brain connectivities.

In this study, we present a new approach for the detection of ADHD using intrinsic time-scale decomposition (ITD) of EEG signals. A classification model based on longitudinal and transversal channel pairs of EEG signals recorded from ADHD patients and control subjects (CS) is proposed. Various connectivity-based features are extracted from different combinations of the ITD modes called Proper Rotation Components (PRCs). Features are classified to detect EEG segments of ADHD patients using a variety of classifiers. Several performance metrics are utilized to investigate the performance of the proposed ADHD detection method.

The novel contributions of the paper are provided in the following;

1. The classification of EEG signals of ADHD patients and control subjects is presented, using the ITD signal decomposition, and connectivity-based feature extraction.
2. Modes produced by the ITD are used to compute several features, including Magnitude Squared Coherence, Cross Power Spectral Density, Correlation Coefficient, Covariance, Correntropy Coefficient, and Cohentropy Coefficient.
3. Various combinations of the modes in accordance with the longitudinal and transversal planes are utilized for feature extraction.
4. Results reveal an impressive performance of the proposed approach and represent a significant improvement over previous studies in terms of identifying and classifying EEG signals of ADHD patients.
5. This is the first study where the ITD and connectivity features are utilized together to identify ADHD, to the best of our knowledge.

2. Materials and methods

2.1. Description of EEG dataset

In this study, 30-channel EEG data recorded at Izmir Katip «elebi University using a 1 kHz sampling frequency with the Brain vision EEG recording system is used. The data set is recorded from 2 different groups namely control subjects and ADHD patients. The first group consists of 15 ADHD patients, 8 girls, and 7 boys, with an average age of 12. The second group is the control group consists of 18 subjects, 14 girls and 4 boys, with an average age of 13. A total of 30 s of EEG data during the open-eyes resting state condition are recorded from each participant. Izmir Katip «elebi University Clinical Research Ethics Committee guaranteed ethical approval numbered 76 and dated 11.07.2019 for the collection of EEG data used in this study.

The electrode mappings are used for the examination of EEG channel connectivity. Evaluated longitudinal and transversal channel pairs are as follows: There are 12 transversal channel pairs (FP1-FP2, F7-F8, F3-F4, FC3-FC4, FT7-FT8, T3-T4, C3-C4, CP3-CP4, TP7-TP8, T5-T6, P3-P4, and O1-O2) and 20 longitudinal channel pairs (FP1-O1, F7-T5, F7-O1, FT7-TP7, FT7-O1, F3-O1, F3-P3, FC3-CP3, FC3-O1, FZ-OZ, FCZ-PZ, FP2-O2, F8-T6, F8-O2, FT8-TP8, FT8-O2, F4-O2, F4-P4, FC4-CP4 and FC4-O2). Transversal and longitudinal channel pairs are shown in Fig. 1.

In the preprocessing stage, we divided the data into 5-s segments using a Hanning window with no overlapping. Then, a 0.5–50 Hz Butterworth band-pass filter is applied to remove artifacts from the segmented EEG data. After preprocessing, filtered EEG segments are decomposed to finite number of PRCs by using ITD.

2.2. The Intrinsic Time-scale Decomposition (ITD)

The iterative intrinsic time-scale decomposition (ITD) technique which is developed for the analysis of nonlinear or non-stationary signals, divides the original signal into low-frequency (“baseline signal”, L_t) and high-frequency (“proper rotation”, H_t) components (PRCs) [22, 23]. The ITD specifically offers effective signal decomposition into “proper rotation” components, for which the instantaneous frequency and amplitude are accurately specified, together with the underlying monotonic signal trend, without the need for time-consuming and unproductive sifting or splines [22].

Let X_t be the EEG signal to be analyzed. We define an operator, $\mathcal{L}$ to remove the low-frequency component (“baseline signal”) from the signal X_t and leave behind the high-frequency component (“proper rotation”). The given signal X_t may be written as follows:

$$X_t = \mathcal{L}X_t + (1-\mathcal{L})X_t = L_t + H_t \tag{1}$$

$\mathcal{L}X_t$ denotes the baseline signal and $H_t = (1-\mathcal{L})X_t$ indicates the proper rotation component. The steps of baseline and proper rotation components extraction are detailed below [22–24].

1. A signal $X_t, t \geq 0$ and its local extremes $\tau_k, k = 1, 2, \ldots$ are assumed to exist. The notations $X(\tau_k) \equiv X_k$ and $L(\tau_k) \equiv L_k$ are introduced.
2. The L_t and H_t are provided over the interval $[0, \tau_k]$, and the signal X_t is available on $[0, \tau_k + 2]$. On the interval $(\tau_k, \tau_k + 1]$ between the two extreme points, the baseline extraction operator “$\mathcal{L}$” is defined as a piece-wise linear function as:

$$L_t = L_k + \left(\frac{L_{k+1} - L_k}{L_{k+2} - L_k}\right)(X_t - X_k), \quad t \in (\tau_k, \tau_{k+1}] \tag{2}$$

where

$$L_{k+1} = \alpha\left[\left(X_k + \frac{\tau_{k+1} - \tau_k}{\tau_{k+2} - \tau_k}\right)(X_{k+2} - X_k)\right] + (1-\alpha)X_{k+1}, \tag{3}$$

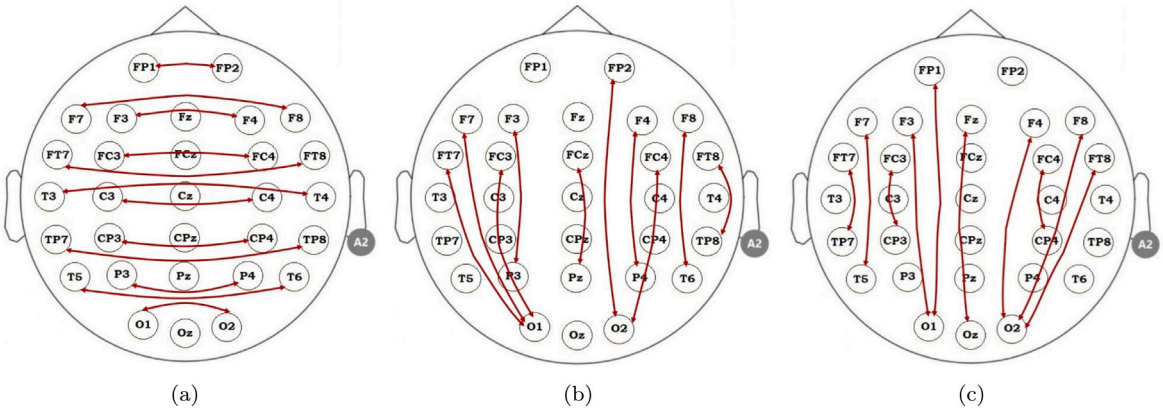


Fig. 1. Brain electrode mapping in 30-channel for; (a) Transversal channel pairs, (b)–(c) longitudinal channel pairs.

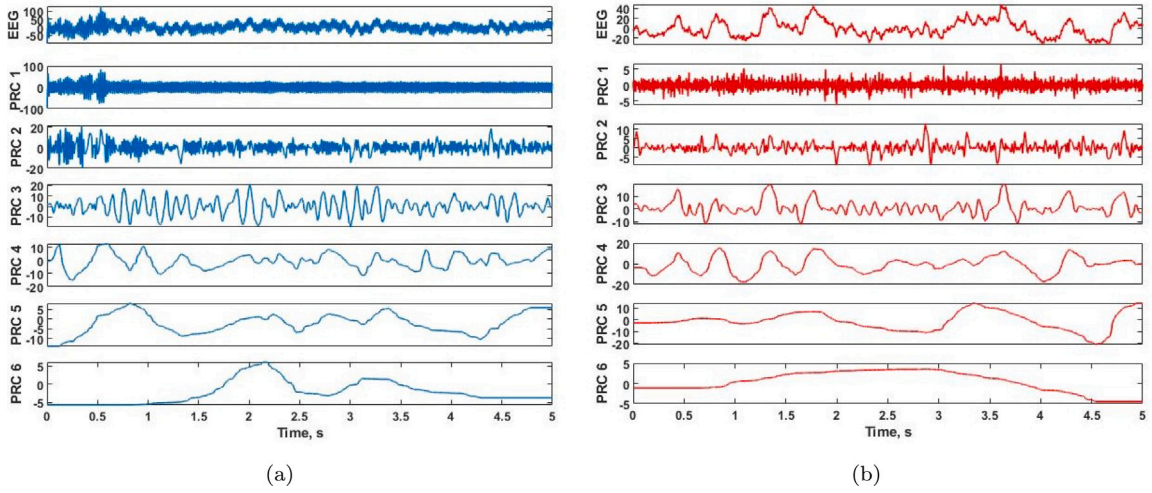


Fig. 2. PRCs obtained by the intrinsic time-scale decomposition (ITD) from a 5-s segment of (a) control subject, and (b) ADHD patient EEG signals (the first 6 components are given as examples).

$$0 < \alpha < 1, \quad \text{typically} \quad \alpha = \frac{1}{2}$$

This method of obtaining the baseline signal (L_t: low-frequency component), preserves the monotonicity of X_t between the extreme points.

3. After calculating the baseline signal as formulated in Step (2), the residual or high frequency component "PRC" is calculated as:

$$\mathcal{H}X_t = (1 - \mathcal{L})X_t = H_t = X_t - L_t \tag{4}$$

Hence, the original signal X_t can be reconstructed using the baseline L_t, and high-frequency H_t modes as follows:

$$X_t = L_t^D + \sum_{i=0}^{D} H_t^i \tag{5}$$

where the number of obtained PRCs is indicated by D.

An example of EEG signal decomposition performed using the ITD method for this study is shown in Fig. 2(a) and (b).

Different feature extraction techniques are applied to high-frequency PRCs, H_t obtained by the ITD method. In our study, various combinations of PRCs; only PRC1, PRC2, and PRC3, PRC1–PRC2, PRC1–PRC3, and PRC1–PRC2–PRC3 (shown as PRC1-to-3) are used to calculate 6 connectivity features.

2.3. Connectivity based feature extraction techniques

In this study, six connectivity features (magnitude square coherence, cross power spectral density, correlation coefficient, covariance, cohentropy coefficient, correntropy coefficient) are calculated from various combinations of PRCs in two different planes, i.e., transversal and longitudinal. These connectivity features are briefly described in the following. Note here that, for example, for a channel pair FP1-Fp2 in the transverse plane, $x[n]$ and $y[n]$ sequences in the following definitions represent each PRC obtained from channels FP1 and FP2, respectively. While calculating the features, the PRC_j of FP1 and PRC_j of FP2 channel are used ($j \in \{1,2,\ldots,D\}$).

2.3.1. Cross power spectral density

In order to analyze the cross-correlation between two-time series, the cross-power spectral density (CPSD) is employed in the frequency domain. The distribution of the covariance between two sets of data throughout the frequency range is described by the cross-spectral density [25]. It is defined as in Eq. (6)

$$P_{x,y}(\omega) = \sum_{k=-\infty}^{-\infty} R_{x,y}(k) e^{-j\omega k} \tag{6}$$

Here for the two signals $x[n]$ and $y[n]$, $R_{x,y}(k)$ indicates the cross-correlation between two signals, $P_{x,y}(\omega)$ denotes the CPSD, calculated by the discrete-time Fourier transform of the cross-correlation and is a complex function. By using the Welch estimation, the Cross Power

Spectral Density $CPSD(\omega) = \frac{1}{N}|P_{x,y}(\omega)|^2$ is calculated. The CPSD is 0 at all frequencies for signals that are not correlated [25].

2.3.2. Magnitude Squared Coherence

Magnitude-Square Coherence (MSC) is a criterion that determines how well a complex signal can be predicted using a linear model from another signal. In other words, it is a method for calculating the connection between two brain signal frequencies [26,27]. The mathematical expression of magnitude squared coherence is given in Eq. (7)

$$MSC_{x,y}(\omega) = \frac{|P_{x,y}(\omega)|^2}{P_{x,x}(\omega)P_{y,y}(\omega)} \tag{7}$$

Here, $P_{x,x}(\omega)$ and $P_{y,y}(\omega)$ denote the power spectral densities of $x[n]$ and $y[n]$, respectively, calculated by the discrete-time Fourier transforms of the auto-correlation sequences "$R_{x,x}(k)$" and "$R_{y,y}(k)$". $P_{x,y}(\omega)$ indicates the cross power spectral density of $x[n]$ and $y[n]$ calculated from "$R_{x,y}(k)$". MSC takes a value between 0 and 1 for each frequency. A value of 0 indicates linear independence, while a value of 1 indicates a perfect linear association between two signals [26,27].

2.3.3. Covariance

Information on the correlation between several random variables is available via the covariance matrix. The covariance matrix in the context of EEG signals offers details on the relationship between EEG signals recorded at different electrode positions across time. This information is crucial for determining how the activity of different brain areas has changed in relation to one another [28]. Covariance is defined as:

$$Cov_{x,y} = \frac{1}{N}\sum_{n=0}^{N-1}(x[n]-\overline{x})(y[n]-\overline{y}) \tag{8}$$

where $Cov_{x,y}$ indicates the covariance and N denotes length of the signals. $\overline{x}$ and $\overline{y}$ represent the means of two signals $x[n]$ and $y[n]$, respectively. If the signals are unrelated, the covariance will be small, but if they are similar, the covariance will be high.

2.3.4. Correlation coefficient

The most popular linear correlation coefficient is Pearson's correlation coefficient. It represents the degree to which two variables are linearly correlated and is a statistical measure of the extent to which variables vary their values with respect to one another. The linear correlation between two discrete-time signals $x[n]$ and $y[n]$ is quantified by the correlation coefficient and it is defined as follows [29–31];

$$\rho_{x,y} = \frac{Cov_{x,y}}{\sigma_x \sigma_y} \tag{9}$$

where $Cov_{x,y}$ indicates the covariance of two signal (given in Eq. (8)) and σ_x and σ_y represent the standard deviations of $x[n]$ and $y[n]$, respectively. $\rho_{x,y}$ is the correlation coefficient and it takes a value between −1 and +1. If the value is equal to 0, there is no linear relationship between the two signals. If the two signals have a positive correlation, the $\rho_{x,y}$ will be close to +1, and a strong negative correlation corresponds to a value close to −1 [29–31].

2.3.5. Correntropy coefficient

The higher-order statistical and/or nonlinear correlation between signals may be detected using the Correntropy coefficient, an extension of the correlation coefficient. The correntropy coefficient RE-Coef for signals $x[n]$ and $y[n]$ is a normalization of the centered cross-correntropy $C(x,y)$, a generalization of covariance, just like correlation is a normalization of covariance $x[n]$ and $y[n]$. Prior to evaluating the $x[n]$ and $y[n]$ signals, they must be normalized by subtracting the means and dividing by the standard deviation. Since $x[n]$ and $y[n]$ might have different dynamics, normalizing makes them dimensionless [30–32].

$$C(x,y) = \frac{1}{N}\sum_{n=0}^{N-1}K(x[n],y[n]) - \frac{1}{N^2}\sum_{n,\ell=0}^{N-1}K(x[n],y[\ell]) \tag{10}$$

Here N stands for the number of samples, while $K(.)$ indicates a symmetric positive-definite kernel function.

$$RE-Coef = \frac{C(x,y)}{\sqrt{C(x,x)}\sqrt{C(y,y)}} \tag{11}$$

where $RE-Coef$ represents the correntropy coefficient, and the kernel has a significant effect. Gaussian kernel $K(x,y) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-y)^2}{2\sigma^2}}$ is generally used in the literature. The choice of kernel width σ is critical. $\sigma = 0.9AN^{(-1/5)}$ is calculated according to Silverman's rule of thumb, where N denotes the number of samples, and A is determined by calculating the minimum of the data interquartile range scaled by 1.34 and the standard deviation of the data. The $RE-Coef$ is bounded by −1 and 1. If the two signals are independent, the $RE-Coef$ will be zero. If two variables are dependent in the same direction, $RE-Coef$ approaches 1, if they are related in opposite directions, it moves towards −1 [30–32].

2.3.6. Cohentropy coefficient

The cohentropy coefficient can be considered as a nonlinear extension of the coherence function or an adaption of the correntropy coefficient to the frequency domain.

$$CE-Coef = \frac{\langle K(X(\omega),Y(\omega))\rangle}{\sqrt{\langle K(X(\omega),X(\omega))\rangle}\sqrt{\langle K(Y(\omega),Y(\omega))\rangle}} \tag{12}$$

where $\langle . \rangle$ denotes the average and calculated over M segments of length L. Before evaluating the cohentropy coefficient, Fourier transforms of signals ($X(\omega)$ and $Y(\omega)$) must be normalized by the mean and standard deviation [30,31].

2.4. Classification and performance evaluation

Following the features extraction using the mentioned techniques, the EEG segments of ADHD patients and healthy controls are identified using a variety of machine learning algorithms, such as Decision Tree (DT), Naive Bayes (NB), Support Vector Machine (SVM), k-Nearest Neighbor (kNN), and Bagged Tree (BT). The following is a quick summary of the used classification algorithms.

2.4.1. Decision trees

The DT classification which includes structures such as root nodes, leaf nodes, and branches is a fast and high-precision method that classifies data by dividing it into various subgroups. Each internal node of a tree stands for a feature, while its branches represent feature combinations that result in classifications and its leaves stand for class labels. The samples are classified in the decision tree by navigating from root to leaf. The samples are classified in the decision tree by navigating from root to leaf [33]. The decision tree technique known as the fine tree algorithm is utilized in this study.

2.4.2. Naive Bayes

The NB Classifier is a statistical classification method based on the Bayes theorem and variables' independence and normalcy. By computing the sample's likelihood of belonging to each class in the dataset, the classification procedure is carried out. The data is assigned as a member of the class with the highest probability of membership [9]. The kernel Naive Bayes classifier is employed in this study.

2.4.3. Support vector machine

A successful approach that is commonly used in both classification and regression studies is SVM, a supervised machine learning algorithm. To classify data, one must locate the hyperplane where the best classification performance may be realized. The data is designated as an element of a different class depending on which side of the hyperplane it falls on [6,9,14,34]. In this study, a medium-Gaussian SVM is employed.

2.4.4. k-Nearest Neighbor

The kNN classification, which belongs to the pattern recognition technique class, is a non-parametric, distance-based learning model. The dataset is split into two groups, a training set, and a test set, and the learning process is carried out using the information in the training set. The distance between the sample that has to be classified and the entire training set of data is first determined. The next step is to identify the k closest neighbors with the shortest distance. Finally, the class of the new sample is chosen as the most common class among these k nearest neighbors [9]. Weighted kNN is chosen from among the nearest neighbor classifiers as a classifier in this study.

2.4.5. Ensemble classifiers: Bagged Tree

An ensemble of classifiers is a group of classifiers that classify new data by combining their separate classifications, usually by weighted or unweighted voting. The major finding is that ensembles, rather than the individual classifiers that make them up, are frequently far more accurate [35]. In this study, the Bagged Tree (BT) algorithm is used.

The following performance criteria are utilized in this study to evaluate the performance of classifiers: accuracy (ACC), sensitivity (SEN), selectivity (SPE), precision (PRE), and false discovery rate (FDR). To evaluate the effectiveness of the classifiers and achieve coherent classification accuracy, the k-fold cross-validation (CV) method has been applied. In this technique, the feature set is separated into k equal-sized subsets. In each iteration, one subset is utilized as the test data while all the other subsets $k-1$ are considered as the training data, and the procedure is repeated k times. The classification performance is then determined as the average performance for k trials [7,9,15,36].

$$
\begin{aligned}
ACC &= \frac{TP+TN}{TP+FN+FP+TN} \\
SEN &= \frac{TP}{TP+FN} \\
SPE &= \frac{TN}{FP+TN} \\
PRE &= \frac{TP}{TP+FP} \\
FDR &= \frac{FP}{TP+FP}
\end{aligned}
\tag{13}
$$

where TP and TN represent, respectively, the number of correctly identified ADHD patients and healthy controls. Additionally, FP and FN stand for the number of healthy controls and ADHD patients that are incorrectly identified.

3. Result and discussion

In this study, the classification of EEG signals of ADHD patients and control subjects using the ITD-based signal decomposition technique and connectivity-based feature extraction has been proposed. Firstly, EEG signals are decomposed into PRCs using the ITD approach, and then 6 different novel connectivity features such as Magnitude Squared Coherence, Cross Power Spectral Density, Correlation Coefficient, Covariance, Correntropy Coefficient, and Cohentropy Coefficient are calculated using different combinations of PRCs, according to the two different planes as longitudinal and transversal. In order to feature extraction both PRC1, PRC2, PRC3, and their combination PRCs1-2, PRCs1-3 and PRC1-to-3 are utilized and two different planes such as longitudinal and transversal are investigated to calculate connectivity features. To reveal the effectiveness of the ITD method, the same process is also applied to EEG signals itself without any decomposition. Afterward, a variety of classifies are utilized to distinguish EEG segments of ADHD patients and control subjects.

The box plot and mean values of calculated connectivity features for ADHD patients and control subjects are presented in Fig. 3(a) to (l) for the transversal and longitudinal planes. These values represent the average value of calculated features using the first three PCRs. There are significant differences between the connectivity features of control subjects and the ADHD patients in both transversal and longitudinal planes for MSC, CPSD, Cov, and CE-Coef features. The ADHD patient group has greater values for MSC, CPSD, and Cov, whereas the control group has greater values for RE-Coef and CE-Coef features. These values calculated for ADHD patients may be higher or lower than the control group, depending on the EEG channel pair in which the connectivity-based features are calculated. For both transversal and longitudinal planes, the MSC, CPSD, and Cov values of the ADHD group are considerably greater than those of the control group, indicating that functional connectivity is stronger in the ADHD group than in the age-matched control group.

The classification performances of the connectivity-based features calculated using the various EEG channel pairs in both longitudinal and transversal planes have been evaluated in order to more accurately reveal the efficacy of the proposed ITD-based method. The ADHD patients and control subject EEG segment classification performances of our proposed ITD-based approach are presented in Tables 1 and 2 for the longitudinal, and transversal planes, respectively. These performances are obtained using both feature sets generated utilizing single PRCs (PRC1, PRC2, and PRC3) and their combinations (PRCs1-2, PRCs1-3, and PRC1-to-3). All EEGs demonstrate that the feature set used in the classification stage is constructed using the EEG signal itself without any decomposition. Classification results obtained using connectivity features in the longitudinal plane are given in Table 1. According to the results, the ITD algorithm provides a significant improvement in terms of performance (ACC:99.06%, F1-S: 98.94%) compared to the classification conducted without using the algorithm. In the results obtained without the ITD algorithm, the ACC and F1-S values vary between 78.75% to 89.29% and 71.25% to 87.42%, respectively, while the ITD algorithm provides higher performance for all classifiers. Although all PRCs and their combinations yield higher classification performance compared to the EEG case, it is remarkable that the classification performance of a single PRC is lower compared to their combinations. Additionally, the ITD algorithm provides higher classification performances in the longitudinal plane compared to the EEG case (without decomposition), the transversal plane also yields a minor improvement in the performance (ACC:99.46%, F1-S: 99.40%, presented in Table 2) of single PRCs and their combinations compared to the longitudinal plane. However, considering only PRC1, PRC2, or PRC3, classification performances are diminished, in contrast to combinations of components compared to the EEG case (given in Table 2). Moreover, the first three PRC combinations provide the highest classification performance for both transversal and longitudinal planes and all classifiers. This proves that when the components are evaluated together, the proper rotation allows for obtaining the highest performance.

Because higher classification performances are obtained using the PRC1-to-3 combination in both transversal and longitudinal planes (given in Tables 1 and 2), classification performances in terms of ACC, SEN, SPE, PRE, and FDR are given in Fig. 4 utilizing that PRC combination. According to the classifier comparison in the transversal plane, the highest classification performance is obtained with the BT classifier that provides 99.46% ACC, 99.47% SEN, 99.47% SPE, 99.33% PRE, and 0.69% FDR, while the lowest classification performance with 97.23% ACC, 96.40% SEN, 97.90% SPE, 97.36% PRE, and 2.64% FDR is obtained using the Naive Bayes classifier. In the longitudinal plane, performance results are more varied. The SVM classifier produces the most meaningful performances (97.84% ACC, 96.37% SEN, 99.02%

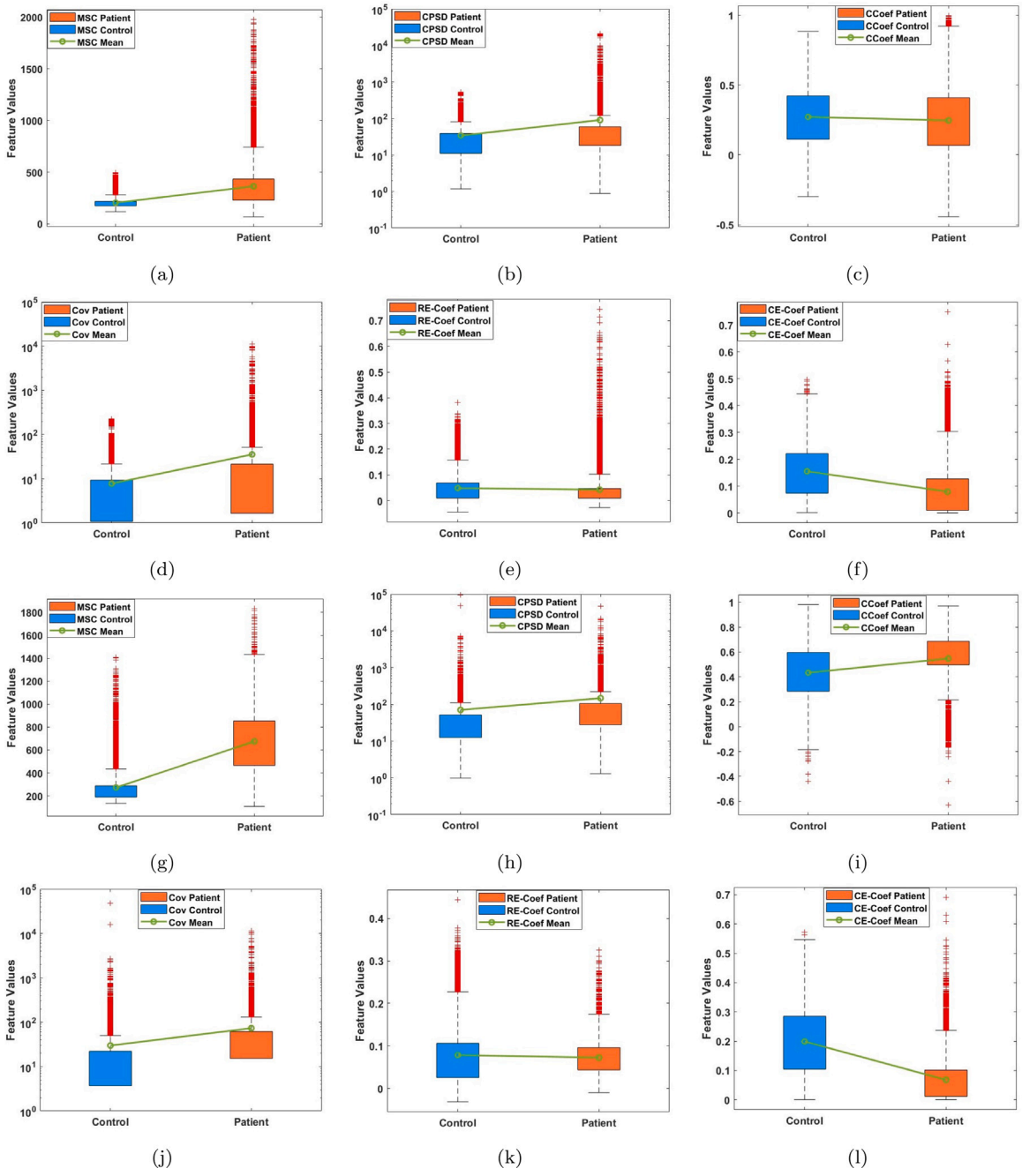


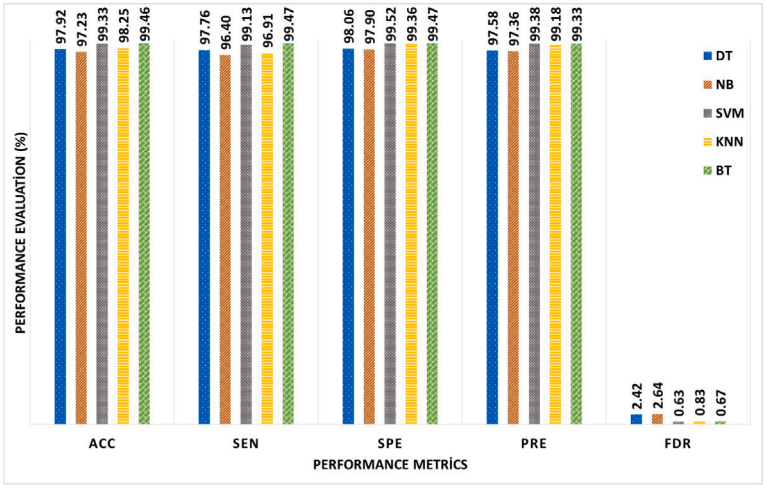
Fig. 3. Box plot of different connectivity features; on longitudinal plane (a) Magnitude Squared Coherence (b) Cross Power Spectral Density, (c) Correlation Coefficient, (d) Covariance, (e) Correntropy Coefficient, and (f) Cohentropy Coefficient and; on transversal plane (g) Magnitude Squared Coherence (h) Cross Power Spectral Density, (i) Correlation Coefficient, (j) Covariance, (k) Correntropy Coefficient, and (l) Cohentropy Coefficient for the control subjects and ADHD patients.

Table 1
The classification performances (%) of the proposed approach in the longitudinal plane. Meth.: Method, Comp.: Component.

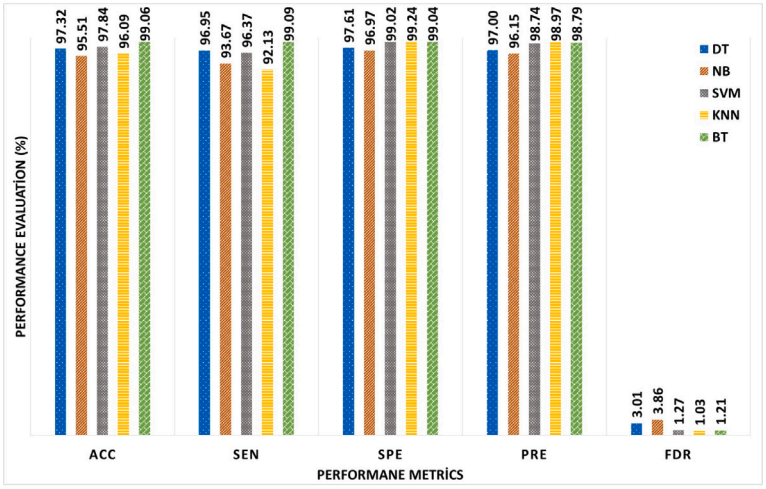
Meth.	Comp.	DT		NB		SVM		kNN		BT	
		ACC	F1-S	ACC	F1-S	ACC	F1-S	ACC	F1-S	ACC	F1-S
ITD	PRC1	94.47	93.73	89.23	88.39	92.51	91.21	94.58	93.78	96.41	95.90
	PRC2	90.79	89.63	88.06	86.20	89.02	86.79	89.97	88.12	93.60	92.66
	PRC3	90.25	89.06	89.40	87.85	89.65	87.43	89.59	87.51	93.06	92.17
	PRCs1-2	97.41	97.07	94.75	94.02	96.80	96.31	96.38	95.79	98.83	98.68
	PRCs1-3	96.16	95.66	94.50	93.82	96.00	95.40	94.90	94.03	98.08	97.85
	PRC1-to-3	97.32	96.97	95.51	94.89	97.84	97.54	96.09	95.43	99.06	98.94
EEG	all EEG	85.72	83.66	84.29	81.27	79.07	71.25	78.75	73.46	89.29	87.42

Table 2
The classification performances (%) of the proposed approach in the transversal plane. Meth.: Method, Comp.: Component.

Meth.	Comp.	DT		NB		SVM		kNN		BT	
		ACC	F1-S	ACC	F1-S	ACC	F1-S	ACC	F1-S	ACC	F1-S
ITD	PRC1	96.83	96.39	93.70	92.85	97.00	96.54	97.26	96.88	98.06	97.81
	PRC2	94.61	93.93	91.18	90.17	95.30	94.58	95.00	94.20	96.43	95.91
	PRC3	94.54	93.86	92.96	92.10	94.79	93.96	94.03	93.02	96.00	95.46
	PRCs1-2	98.11	97.86	96.03	95.52	98.88	98.73	98.19	97.94	99.28	99.19
	PRCs1-3	97.67	97.36	96.73	96.33	98.63	98.47	97.58	97.24	99.06	98.93
	PRC1-to-3	97.92	97.67	97.23	96.88	99.33	99.25	98.25	98.03	99.46	99.40
EEG	all EEG	97.01	96.61	94.21	93.87	95.27	94.57	95.24	94.54	98.03	97.78



(a)

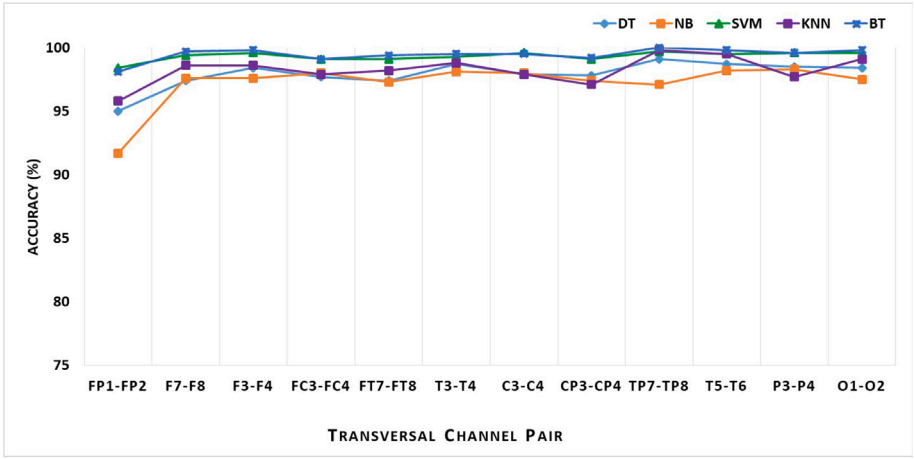


(b)

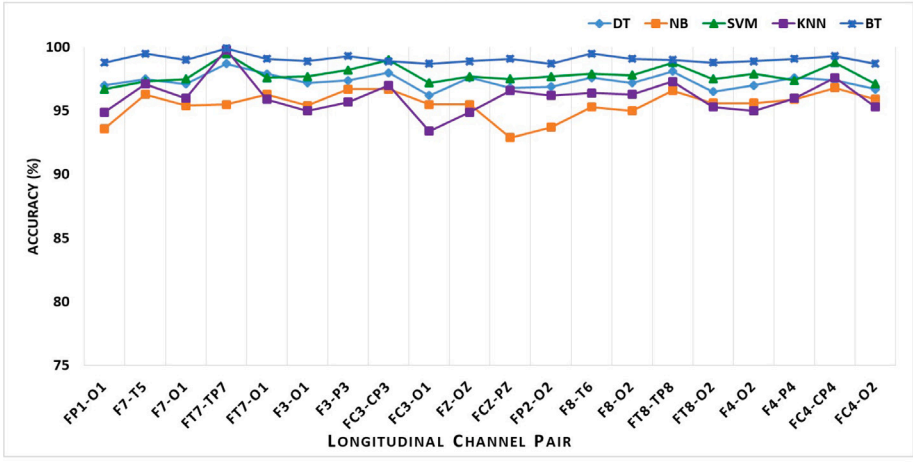
Fig. 4. The comparison of the performance of classifiers in terms of ACC, SEN, SPE, PRE, and FDR in (a) transversal and (b) longitudinal plane.

SPE, 98.74% PRE, and 1.27% FDR), while the NB classifier yields poor performance compared to the other classifiers. The observed results showed that channel pairs in the transversal plane perform better in terms of the detection of EEG segments of ADHD patients compared to channel pairs in the longitudinal plane.

Next, the classification performance of the connectivity-based features derived from the various channel pairs in both transversal and longitudinal planes has been investigated in order to more accurately evaluate the performance of the proposed approach. The classification accuracies obtained using the PRC1-to-3 combination for 12 channel pairs in the transversal plane and 20 channel pairs in the longitudinal plane are given according to different classifiers in Fig. 5(a) and (b), respectively. In the transversal plane (Fig. 5(a)), all the classifiers provide excellent channel-pair-based classification accuracies. However, in the FP1-FP2 channel pair located in the anterior region of the brain lowest accuracies for all of the classifiers are obtained. The highest accuracy values are obtained from the channel pair in the temporal and posterior regions. For the TP7-TP8 channel pair, the classification accuracies of 100%, 99.8%, 99.7%, 97.1%, and 99.1% are obtained using BT, KNN, SVM, NB, and DT classifiers, respectively. It is noticed that the channel pairs between the right and left temporal brain regions provide higher channel-pair-based classification accuracies. On the other hand, in the longitudinal plane, considerably more noticeable differences are observed between utilized classifier performances (presented in Fig. 5(b)). The classifier to be selected, as well as the channel pairs, are chosen in this plane, should be properly considered. The FT7-TP7



(a)



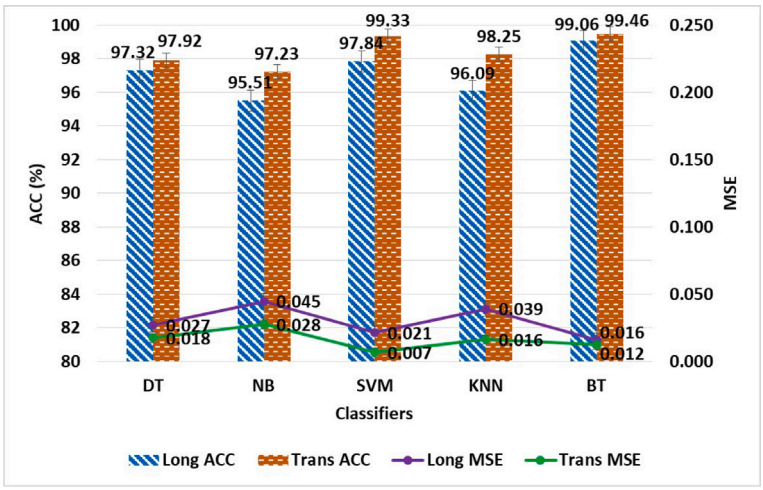
(b)

Fig. 5. The channel-pair-based classification accuracies in (a) transversal and (b) longitudinal plane.

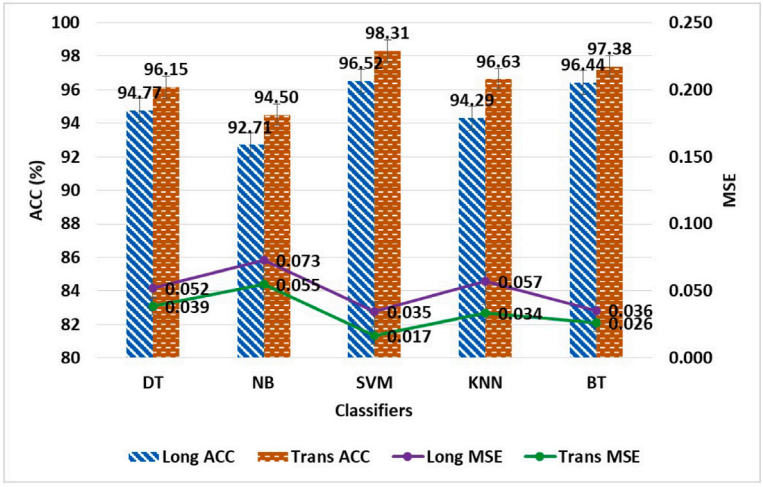
channel pair yields the best performance, followed by the FC4-CP4 channel pair, according to the average of all classifiers.

Some of the properties of all patients used for training are retained by the trained model in the k-fold CV. As a result, the effectiveness of evaluating new patients is not effectively evaluated by this technique. Therefore, in this study, the performances of the classifiers are also evaluated using the leave-one subject-out (LOSO) cross-validation technique to reveal the effectiveness of the proposed ITD-based approach to classifying unknown subjects. In this method, the test set is made up of the features of the one subject that are not included in the training set during each iteration of LOSO. With the selection of each input as test data apart from other inputs, this approach performs an impartial comparison while using other inputs as training data [9,36]. The comparison of the CV and LOSO validation-based performance results of the proposed approach in terms of ACC and Mean squared error (MSE) obtained using the PRC1-to-3 combination are given for both transversal and longitudinal planes and each classifier in Fig. 6. According to Fig. 6(a), using the 10-fold CV as the validation method yields a higher ACC ($\geq$97.23%) and lower MSE ($\leq$0.028%) for the transversal plane compared to longitudinal plane. Meanwhile, the BT yields better accuracies and lower MSE than the other classifier for both transversal and longitudinal planes. On the other hand, using the LOSO-CV as the validation method provide lower classification performance for both transversal and longitudinal planes and each classifier (presented in Fig. 6(b)). Better classification performances with 98.31% ACC and 0.017 MSE, and 96.52% ACC and 0.035 MSE are obtained using the SVM classifier for transversal and longitudinal planes, respectively. The 10-fold CV model on the transversal and longitudinal planes yielded more dramatic results compared to the LOSO-CV-based model. However, detecting the data of an unknown patient with $\leq$92.71% ACC and $\leq$0.073 MSE strikingly demonstrates the success of the proposed method. The final findings have shown the potential of the proposed ITD-based decomposition and connectivity-based feature extraction approach for a variety of validation models, chosen channel pairs, and classifiers.

The results of the current study are compared with those of previous studies of a similar nature where the connectivity and/or complexity-based features are computed to evaluate the ability of the proposed approach in classifying ADHD versus control groups. Table 3 provides a comparison of the current study with prior studies in ADHD detection. In general, the studies are formed on two bases; analyzing EEG signals recorded using a particular cognitive task or examining EEG signals recorded at a resting state. The majority of approaches in the literature, as seen in Table 3, are centered on employing cognitive tasks to distinguish between ADHD and control subjects. These studies have reported accuracy rates ranging from 81.20% to 99.75% [3,4,7,9–12,20,21,37]. Children with ADHD often have smaller brain volumes, with multi-functional frontal and parietal cortex. As a result, these children are frequently described as being easily distracted and having trouble paying attention or following directions. Also, ADHD is linked to cognitive dysfunction, particularly in the frontal and parietal cortex [9,18]. Therefore, it is expected that there will be significant differences between the EEG signals of ADHD patients and the control subjects under a specific cognitive task. However, our study using



(a)



(b)

Fig. 6. Comparison of accuracy and MSE obtained using (a) 10-fold cross-validation and (b) LOSO-based validation models for both transversal and longitudinal planes according to different classifiers.

resting state EEG signals outperformed most studies using cognitive tasks. On the other hand, there are many studies that examine the EEG signals recorded at a resting state, similar to our study [2,6,13–17,19]. Compared to the aforementioned studies, the proposed ITD and connectivity features-based study provides higher classification accuracy. In addition, there are large differences between ADHD and CS numbers in study [2], and studies [13,16] have few ADHD and CS EEG data, providing classification accuracies similar to ours. It is highly probable that these situations affected the classification's success. Only five of the aforementioned studies [6,7,19–21] proposed connectivity-based approaches. Moreover, only two of them [6,19] evaluated the resting state EEG signals in their studies. One of the connectivity-based features, bicoherence, calculated in study [6]. Although a maximum classification performance of 83.33% was reported in the proposed study, the highest performance for the bicoherence feature was 68.88. The weighted phase lag index, a different connectivity metric, is introduced in a study [19], and it has been found that EEG connectivity features do not reliably categorize participants into groups. To compare the differences and similarities between intra-hemispheric or inter-hemispheric EEG channel pairings in the same frequency sub-bands, another study [7] used the DFW method and cognitive task. They noticed false correlations when there is no phase lag between the two signals, and reported 99.17% ACC that less than that of our proposed study. Additionally, in two studies [20,21] in which the cognitive task was used, classification success was not reported. Therefore, performance comparison cannot be conducted clearly. According to the authors' knowledge, this is the first research in which the ITD approach and connectivity features have been used to identify ADHD. The classification results show that the proposed study is promising for distinguishing between ADHD patients and control subjects along different channel pairings in two different planes, longitudinal and transversal. The highest accuracies with 100.00% and 99.90% are obtained using the connectivity features of the transversal and longitudinal planes, respectively. Additionally, classification accuracies in the longitudinal and transversal planes, respectively, are ≥92.90% and ≥91.70% (given in Fig. 5).

4. Conclusion

ADHD is a prevalent disorder that affects children, and early identification is crucial for effective treatment and for reducing future problems. In the current study, an ITD-based machine learning model is proposed to help the rapid and accurate detection of ADHD by utilizing a channel-pair-based analysis of multi-channel EEG signals. The EEG data recorded from 15 ADHD patients and 18 control subjects are utilized in our experiments. Various connectivity-based features, such as Magnitude Squared Coherence, Cross Power Spectral Density, Correlation Coefficient, Covariance, Correntropy Coefficient, and Cohentropy

Table 3
Comparison between the accuracy of this method with some state-of-the-art studies in this area.

Paper	Participants	Cognitive task	Features	Accuracy (%)
[3]	46 ADHD/45 CS	Cognitive task	Dynamic connectivity tensors	99.75
[4]	53 ADHD/161 CS	Cognitive task	Subband powers of EEG	81.20
[7]	14 ADHD/19 CS	Cognitive task	DFW ERP features	99.17
[9]	23 ADHD/23 CS	Cognitive task	Higuchi's fractal dimension Entropy of wavelet coefficients Subband powers of EEG	91.30
[10]	27 ADHD/38 CS	Cognitive task	ERP features	98.40
[11]	100 ADHD/44 CS	Cognitive task	CNN	83
[12]	20 ADHD/20 CS	Cognitive task Resting state	Event-related spectrograms EEG spectrograms	88.00 66.00
[20]	16 ADHD/16 CS	Cognitive task	Coherence and graph theory	Not given
[21]	22 ADHD/22 CS	Cognitive task	Phase locking value	Not given
[37]	30 ADHD/31 CS	Cognitive task	CNN	99.06
[2]	25 ADHD/14 CS	Resting state	CNN Subband powers of EEG	99.46
[6]	50 ADHD/58 CS	Resting state	Entropy Bicoherence	84.59
[13]	12 ADHD/12 CS	Resting state	Spectral features Entropy based features Approximate entropy	99.58
[14]	12 ADHD/12 CS	Resting state	Detrended fluctuation analysis Lyapunov exponents Subband powers of EEG Entropy, correlation dimension	83.33
[15]	50 ADHD/26 CS	Resting state	Fractal dimension Lyapunov exponents	96.05
[16]	5 ADHD/5 CS	Resting state	Entropy based features	99.82
[17]	20 ADHD/20 CS	Resting state	Fuzzy entropy	98.07
[19]	38 ADHD/51 CS	Resting state	Weighted phase lag index	Not given
This study	18 ADHD/15 CS	Resting state	ITD and connectivity based features	99.46

Coefficient, are calculated using the modes obtained by the ITD. Several combinations of the modes in accordance with two different planes i.e., longitudinal and transversal are used for feature extraction. The results showed that the calculated features, especially MSC, CPSD, Cov, and CE-Coef features, are significantly different in both planes for both groups (presented in Fig. 3). Using calculated features along with the BT classifier yielded the highest ACC of 99.46% and F1-S of 99.40% for the identification of ADHD in the transverse plane. However, the EEG signals evaluated without using the ITD algorithm provide an ACC of 98.03% and F1-S of 97.78% in the same plane using the same classifier (given in Table 2). On the other hand, in the longitudinal plane, the effectiveness of the proposed ITD approach is more pronounced (given in Table 1). According to our results, the connectivity features that are derived utilizing several channel pairs in two distinct planes and the ITD modes may be a helpful and discriminative tool for the detection of ADHD.

The connectivity features of resting state EEG signals according to the various planes are explored in this study. However, when a subject focuses their attention on a situation that has been experimentally controlled, the distribution of brain resources changes. Therefore, in our future research, the use of ITD and connectivity-based feature calculation models during cognitive tasks will be investigated.

CRediT authorship contribution statement

Ozlem Karabiber Cura: Conceptualization, Investigation, Software, Methodology, Validation, Visualization, Writing – original draft. **Sibel Kocaaslan Atli:** Conceptualization, EEG data recording, Methodology, Supervision, Writing – review & editing. **Aydin Akan:** Conceptualization, Investigation, Software, Methodology, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

This work is partially funded by Izmir University of Economics, Scientific Research Projects Coordination Unit, Project no: 2022-7.

References

[1] A. Serrano-Barroso, R. Siugzdaite, J. Guerrero-Cubero, A.J. Molina-Cantero, I.M. Gomez-Gonzalez, J.C. Lopez, J.P. Vargas, Detecting attention levels in ADHD children with a video game and the measurement of brain activity with a single-channel BCI headset, Sensors 21 (9) (2021) 3221.
[2] A. Ahmadi, M. Kashefi, H. Shahrokhi, M.A. Nazari, Computer aided diagnosis system using deep convolutional neural networks for ADHD subtypes, Biomed. Signal Process. Control 63 (2021) 102227.
[3] M. Bakhtyari, S. Mirzaei, ADHD detection using dynamic connectivity patterns of EEG data and ConvLSTM with attention framework, Biomed. Signal Process. Control 76 (2022) 103708.
[4] S.J. Johnstone, L. Parrish, H. Jiang, D.-W. Zhang, V. Williams, S. Li, Aiding diagnosis of childhood attention-deficit/hyperactivity disorder of the inattentive presentation: Discriminant function analysis of multi-domain measures including EEG, Biol. Psychol. 161 (2021) 108080.
[5] R. Ho, K. Hung, Empirical mode decomposition method based on cardinal spline and its application on electroencephalogram decomposition, in: 2022 IEEE 12th Symposium on Computer Applications & Industrial Electronics, ISCAIE, IEEE, 2022, pp. 17–21.
[6] H. Chen, W. Chen, Y. Song, L. Sun, X. Li, EEG characteristics of children with attention-deficit/hyperactivity disorder, Neuroscience 406 (2019) 444–456.
[7] P. Ghaderyan, F. Moghaddam, S. Khoshnoud, M. Shamsi, New interdependence feature of EEG signals as a biomarker of timing deficits evaluated in attention-deficit/hyperactivity disorder detection, Measurement (2022) 111468.
[8] C. Mauriello, E. Pham, S. Kumar, C. Piguet, M.-P. Deiber, J.-M. Aubry, A. Dayer, C.M. Michel, N. Perroud, C. Berchio, Dysfunctional temporal stages of eye-gaze perception in adults with ADHD: a high-density EEG study, Biol. Psychol. (2022) 108351.
[9] M. Altınkaynak, N. Dolu, A. Güven, F. Pektaş, S. Özmen, E. Demirci, M. İzzetoğlu, Diagnosis of Attention Deficit Hyperactivity Disorder with combined time and frequency features, Biocybern. Biomed. Eng. 40 (3) (2020) 927–937.

[10] G. Guney, E. Kisacik, C. Kalaycıoglu, G. Saygili, Exploring the attention process differentiation of attention deficit hyperactivity disorder (ADHD) symptomatic adults using artificial intelligence on electroencephalography (EEG) signals, Turk. J. Electr. Eng. Comput. Sci. 29 (5) (2021) 2312–2325.
[11] A. Vahid, A. Bluschke, V. Roessner, S. Stober, C. Beste, Deep learning based on event-related EEG differentiates children with ADHD from healthy controls, J. Clin. Med. 8 (7) (2019) 1055.
[12] L. Dubreuil-Vall, G. Ruffini, J.A. Camprodon, Deep learning convolutional neural networks discriminate adult ADHD from healthy individuals on the basis of event-related spectral EEG, Front. Neurosci. 14 (2020) 251.
[13] M. Rezaeezadeh, S. Shamekhi, M. Shamsi, Attention Deficit Hyperactivity Disorder Diagnosis using non-linear univariate and multivariate EEG measurements: a preliminary study, Phys. Eng. Sci. Med. 43 (2) (2020) 577–592.
[14] S. Khoshnoud, M.A. Nazari, M. Shamsi, Functional brain dynamic analysis of ADHD and control children using nonlinear dynamical features of EEG signals, J. Integr. Neurosci. 17 (1) (2018) 17–30.
[15] Y.K. Boroujeni, A.A. Rastegari, H. Khodadadi, Diagnosis of attention deficit hyperactivity disorder using non-linear analysis of the EEG signal, IET Syst. Biol. 13 (5) (2019) 260–266.
[16] R. Catherine Joy, S. Thomas George, A. Albert Rajan, M. Subathra, Detection of adhd from eeg signals using different entropy measures and ann, Clin. EEG Neurosci. 53 (1) (2022) 12–23.
[17] R. Yaghoobi Karimu, S. Azadi, Diagnosing the ADHD using a mixture of expert fuzzy models, Int. J. Fuzzy Syst. 20 (4) (2018) 1282–1296.
[18] H.T. Tor, C.P. Ooi, N.S. Lim-Ashworth, J.K.E. Wei, V. Jahmunah, S.L. Oh, U.R. Acharya, D.S.S. Fung, Automated detection of conduct disorder and attention deficit hyperactivity disorder using decomposition and nonlinear techniques with EEG signals, Comput. Methods Programs Biomed. 200 (2021) 105941.
[19] H. Kiiski, L.M. Rueda-Delgado, M. Bennett, R. Knight, L. Rai, D. Roddy, K. Grogan, J. Bramham, C. Kelly, R. Whelan, Functional EEG connectivity is a neuromarker for adult attention deficit hyperactivity disorder symptoms, Clin. Neurophysiol. 131 (1) (2020) 330–342.
[20] J. Cao, Y. Li, H. Yu, X. Zhao, Y. Li, S. Wang, Investigation of brain networks in children with attention deficit/hyperactivity disorder using a graph theoretical approach, Biomed. Signal Process. Control 40 (2018) 351–358.
[21] S. Ansarinasab, S. Panahi, F. Ghassemi, D. Ghosh, S. Jafari, Synchronization stability analysis of functional brain networks in boys with ADHD during facial emotions processing, Phys. A 603 (2022) 127848.
[22] M.G. Frei, I. Osorio, Intrinsic time-scale decomposition: time–frequency–energy analysis and real-time filtering of non-stationary signals, Proc. R. Soc. A: Math. Phys. Eng. Sci. 463 (2078) (2007) 321–342.
[23] A. Voznesensky, D. Kaplun, Adaptive signal processing algorithms based on EMD and ITD, IEEE Access 7 (2019) 171313–171321.
[24] R.J. Martis, U.R. Acharya, J.H. Tan, A. Petznick, L. Tong, C.K. Chua, E.Y.K. Ng, Application of intrinsic time-scale decomposition (ITD) to EEG signals for automated seizure prediction, Int. J. Neural Syst. 23 (05) (2013) 1350023.
[25] H. Li, S. Feng, L. Ma, Z. Xu, R. Xu, T.-P. Jung, Common cross-spectral patterns of electroencephalography for reliable cognitive task identification, IEEE Access 8 (2020) 17652–17662.
[26] S. Malekpour, J.A. Gubner, W.A. Sethares, Measures of generalized magnitude-squared coherence: Differences and similarities, J. Franklin Inst. B 355 (5) (2018) 2932–2950.
[27] T. Suhail, K. Indiradevi, E. Suhara, S.A. Poovathinal, A. Ayyappan, Distinguishing cognitive states using electroencephalography local activation and functional connectivity patterns, Biomed. Signal Process. Control 77 (2022) 103742.
[28] S.S. Prabhu, N. Sinha, et al., Sleep EEG analysis utilizing inter-channel covariance matrices, Biocybern. Biomed. Eng. 40 (1) (2020) 527–545.
[29] Z. Šverko, M. Vrankić, S. Vlahinić, P. Rogelj, Complex pearson correlation coefficient for EEG connectivity analysis, Sensors 22 (4) (2022) 1477.
[30] J. Dauwels, F. Vialatte, T. Musha, A. Cichocki, A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG, NeuroImage 49 (1) (2010) 668–693.
[31] H. Bakhshayesh, S.P. Fitzgibbon, A.S. Janani, T.S. Grummett, K.J. Pope, Detecting synchrony in EEG: A comparative study of functional connectivity measures, Comput. Biol. Med. 105 (2019) 1–15.
[32] A. Gunduz, J.C. Principe, Correntropy as a novel measure for nonlinearity tests, Signal Process. 89 (1) (2009) 14–23.
[33] V.H. Kamble, M.P. Dale, Machine learning approach for longitudinal face recognition of children, in: Machine Learning for Biometrics, Elsevier, 2022, pp. 1–27.
[34] M.S. Bascil, A.Y. Tesneli, F. Temurtas, Spectral feature extraction of EEG signals and pattern recognition during mental tasks of 2-D cursor movements for BCI using SVM and ANN, Australas. Phys. Eng. Sci. Med. 39 (3) (2016) 665–676.
[35] S.K. Khare, V. Bajaj, A. Sengur, G. Sinha, Classification of mental states from rational dilation wavelet transform and bagged tree classifier using EEG signals, in: Artificial Intelligence-Based Brain-Computer Interface, Elsevier, 2022, pp. 217–235.
[36] M.A. Ozdemir, O.K. Cura, A. Akan, Epileptic EEG classification by using time-frequency images for deep learning, Int. J. Neural Syst. 31 (08) (2021) 2150026.
[37] M. Moghaddari, M.Z. Lighvan, S. Danishvar, Diagnose ADHD disorder in children using convolutional neural network based on continuous mental task EEG, Comput. Methods Programs Biomed. 197 (2020) 105738.