

Phase Space Reconstruction of EEG Signals for Classification of ADHD and Control Adults

Clinical EEG and Neuroscience
2020, Vol. 51(2) 102–113
© EEG and Clinical Neuroscience
Society (ECNS) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1550059419876525
journals.sagepub.com/home/eeg



Simranjit Kaur¹, Sukhwinder Singh¹ ,
Priti Arun², Damanjeet Kaur³, and Manoj Bajaj²

Abstract

Attention deficit hyperactivity disorder (ADHD) is a childhood behavioral disorder that can persist into adulthood. Electroencephalography (EEG) plays a significant role in assessing the neurophysiology of ADHD because of its ability to reveal complex brain activity. The present study proposes an EEG-based diagnosis system using the phase space reconstruction technique to classify ADHD and control adults. Electric activity is recorded for 47 ADHD and 50 control adults during the eyes-open, eyes-closed, and Continuous Performance Test (CPT) condition. Various statistical features are extracted from Euclidean distances based on phase space reconstruction of signals. The proposed system is evaluated with 2 feature selection methods (correlation-based feature selection and particle swarm optimization) and 5 machine learning methods (neural dynamic classifier, support vector machine, enhanced probabilistic neural network, k-nearest neighbor, and naive-Bayes classifier). Experimental results showed the highest testing accuracy of 93.3% under the eyes-open, 90% under the eyes-closed, and 100% under the CPT condition. This study focused on the utility of phase space reconstruction of brain signals to discriminate between ADHD and control adults.

Keywords

attention deficit hyperactivity disorder, adults, continuous performance test, EEG, phase space reconstruction

Received November 23, 2018; revised August 21, 2019; accepted August 22, 2019.

Introduction

Attention deficit hyperactivity disorder (ADHD) is a mental condition with core symptoms of hyperactivity, impulsiveness, and inattentiveness affecting 5% of children worldwide.¹ Follow-up research studies have proven the persistence of ADHD into adulthood with a prevalence between 2.5% and 4.3%.² It is found that fewer high school students with ADHD take admission in college rather than non-ADHD students. ADHD students who go to college report more depressive symptoms, task-interfering, and intrusive thoughts than their non-ADHD peers. These difficulties undermine their academic functioning in college settings, which require attention, organization, and self-management abilities. It is also found that adults with ADHD have poor driving records, low employment rate, problems in social relationships, and a high risk of drug abuse.³ The diagnosis of adult ADHD involves subjective assessment of the comprehensive history of childhood functioning and rating scales. But sometimes ADHD cases remain unidentified because of misreporting by family members and uncertainty of data. Some features of ADHD are similar to depression, which is again common in the general population. In such cases, ADHD is usually not identified.

For the past decade, electroencephalography (EEG) has emerged as a successful tool to elucidate the functioning of a human brain. Electrophysiological traits have been investigated in the literature to understand the heterogeneity of the neural profile of ADHD population. A recent meta-analysis of EEG studies for ADHD diagnosis points toward their ability to act as an objective biomarker of disease.^{4,5} Magee et al⁶ determined the power spectra of ADHD and control children during an eyes-closed condition and used a logistic regression method on independent clusters that yield classification accuracy of 87%. Ahmadi et al⁷ used the concept of synchronization

¹Department of Computer Science and Engineering, University Institute of Engineering and Technology, Panjab University, Chandigarh, India

²Department of Psychiatry, Government Medical College and Hospital, Chandigarh, India

³Department of Electrical and Electronics Engineering, University Institute of Engineering and Technology, Panjab University, Chandigarh, India

Corresponding Author:

Sukhwinder Singh, Department of Computer Science and Engineering, University Institute of Engineering and Technology, Panjab University, Chandigarh, Punjab 160014, India.

Email: sukhdalip@pu.ac.in

Full-color figures are available online at journals.sagepub.com/home/eeg

likelihood (SL) to extract the features from EEG of ADHD and control children under an eyes-closed condition. They implemented the radial basis function neural network (RBFNN) classifier, which resulted in an accuracy of 95.6% for ADHD diagnosis. Sadatnezhad et al⁸ extracted band power, fractal dimension, autoregressive model coefficients, and wavelet coefficients from signals of ADHD and bipolar mood disorder (BMD) patients under eyes-open and eyes-closed conditions. They gained an accuracy of 86.4% using linear function approximation system classifier combined with linear discriminant analysis selected features. Ahmadlou et al⁹ found that interdependencies measured by fuzzy SL are more reliable to discriminate EEGs of ADHD and control children than conventional SL during the eyes-closed condition. They reported 97.1% classification accuracy using the RBFNN with fuzzy SL-derived feature set. In another study, Ahmadlou et al¹⁰ used fuzzy SL to analyze the organization of overall and hemispheric brain networks of ADHD and non-ADHD subjects. The results of the analysis of variance test revealed that the 2 parameters, namely, clustering of connections and path length among network units, can discriminate 2 groups. Liechti et al¹¹ examined theta power, beta power and theta/beta ratio of ADHD adults, ADHD children, and matched healthy control group during an eyes-open and eyes-closed condition. The stepwise discriminant analysis yielded an overall accuracy of 53% in classifying ADHD adults and 70% for ADHD children. Poil et al¹² provided an integrated analysis of developmental effects and ADHD in children and adults using power spectra under an eyes-closed condition. ADHD adults are classified with the sensitivity of 67% and specificity of 83% using support vector machine (SVM) classifier but these features are not found suitable for ADHD children classification. Helgadottir et al¹³ examined EEG coherence features in ADHD children, ADHD adolescents, and matched healthy controls under an eyes-closed condition. They found an accuracy of 76% in age independent version using SVM classifier.

Besides the use of rest state EEG signals for ADHD classification, brain activity recorded during a cognitive task is also proven to be helpful for making a clear distinction between ADHD and normal subjects. Mueller et al¹⁴ selected independent event-related potential components of ADHD and control adults during a visual go/nogo task. Classification accuracy of 94% is achieved using SVM with a selected set of independent components. Abibullaev et al¹⁵ extracted power spectra of ADHD and control children during a focused attention task. Feature selection is performed using information theoretic based on entropy and mutual information. They obtained an accuracy of 97% using SVM classifier with the selected feature subset. Tenev et al¹⁶ made use of power spectra of ADHD and control adults under an eyes-open, eyes-closed, emotional continuous performance test, and visual continuous performance test conditions. This study obtained an accuracy of 82.3% by combining the output from 4 SVM classifiers using a logical expression of the Karnaugh map. Mohammadi et al¹⁷ computed approximate entropy, fractal dimension and Lyapunov exponent from signals of ADHD and control children during a cognitive

task. The use of a multilayer perceptron neural network yielded an accuracy of 92.3% and 93.6% for minimum redundancy maximum relevance (mRMR) and double input symmetrical relevance (DISR) method selected features, respectively.

The brain is considered a chaotic dynamical system. So, the EEG signals generated from the brain are also chaotic as their amplitude changes randomly with respect to time. The use of non-linear analysis methods for biomedical signals is widely accepted as compared with linear methods in several fields as they can reflect nonstationarity of time-series. Consider a dynamic system defined by m variables, its state at a particular moment of time can be represented by a point in m -dimensional space. This space is called phase space. Phase space reconstruction (PSR) of EEG signals is able to reflect complex brain dynamics thereby helpful in discriminating patients with neurological disorders. Sharma et al¹⁸ made use of empirical mode decomposition and obtained PSR from intrinsic mode functions of EEG signals. They achieved an accuracy of 98.76% with SVM in a classification of seizure and seizure-free brain signals using the feature set of 95% confidence area measure and interquartile range of Euclidean distances. Lee et al¹⁹ decomposed EEG signals with discrete wavelet transform and computed PSR Euclidean distance-based parameters from wavelets coefficients at different levels. Normal and epileptic EEG signals are classified with an accuracy of 98.1% using a neural network with weighted fuzzy membership function. These studies highlight the importance of reconstructed phase space in analyzing the dynamic behavior of brain signals.

As EEG signals are captured from the brain dynamical system, which is chaotic in nature, this phenomenon can be visualized by embedding the EEG time-series in delayed phase space. In this work, PSR-based features are used, which represent the all distinct states of the system and is found to be useful for chaotic time-series analysis.²⁰ Moreover, college students with ADHD remain understudied in terms of their neural profile at rest or during information processing. In the present study, an EEG-based diagnosis system is proposed for ADHD adults using PSR Euclidean distance-based features, which reflect the nonlinear dynamics of brain signals. EEG activity of university students with ADHD and control adults is recorded under 2 rest-state conditions (eyes-open and eyes-closed) and a cognitive activation condition (Continuous Performance Test [CPT]). Correlation-based feature selection (CFS) and particle swarm optimization (PSO) is employed to select the optimal feature subset for ADHD and control group discrimination. This research study has 3 main objectives: (1) to examine diagnostic utility of PSR Euclidean distance-based features under 3 different conditions and to compare the results with 2 other non-linear methods (fractal dimension and visibility graph-based features), (2) to compare the CFS-and PSO-based feature selection methods in terms of their classification performance, and (3) to compare the performance of 5 classifiers namely neural dynamic classifier (NDC), SVM, enhanced probabilistic neural network (EPNN), K-nearest neighbor (KNN), and naive-Bayes classifier (NBC). To our knowledge, this is the first study to consider both the rest-state and cognitive task-based

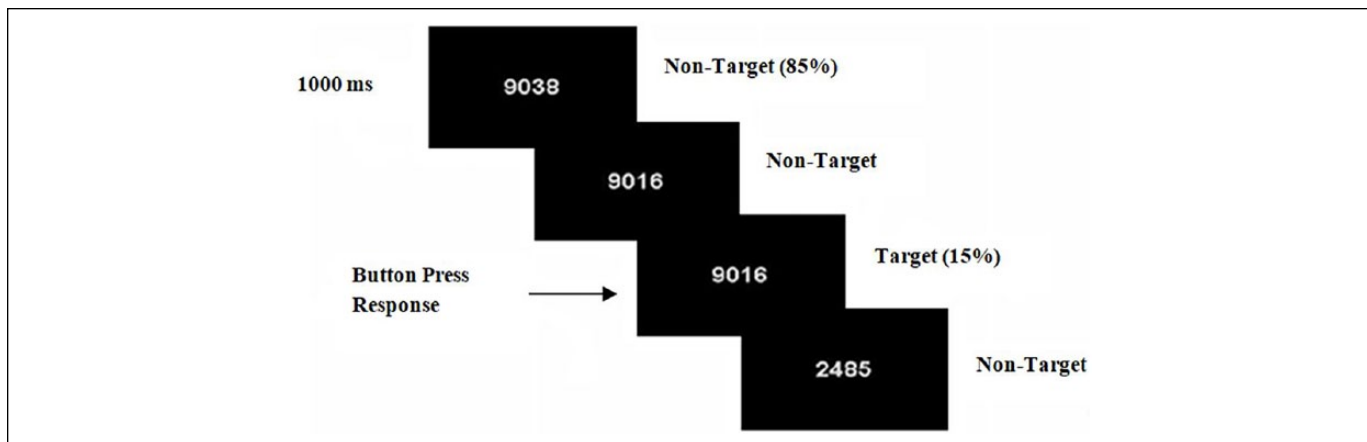


Figure 1. Visual illustration of the Continuous Performance Test.²⁴

brain signals of ADHD and control group adults for classification purpose.

Material and Methods

Participants

A total of 97 students, which include 47 adults with ADHD (8 female; mean age = 20.3, SD = 1.12) and 50 controls (9 female; mean age = 20.6, SD = 1.28) are recruited from University Institute of Engineering and Technology, Panjab University, Chandigarh, India. An initial population of 875 undergraduate university students completed the Wender Utah Rating Scale (WURS)²¹ to assess childhood ADHD symptoms and Adult ADHD Self-Report Scale (ASRS)²² to assess current symptoms of ADHD. Among 875 students, 68 scored higher than clinically significant levels on 2 rating scales. Thus, the prevalence of adult ADHD among university students is found to be 7.7%. Among selected 68, only 52 had given consent for further participation in research, which includes a clinical interview by a psychiatrist to confirm the diagnosis and EEG acquisition. ADHD diagnosis is confirmed in 48 cases after clinical assessment. ADHD group must fulfill the *Diagnostic and Statistical Manual of Mental Disorders* 5th edition (DSM-5) criteria of ADHD diagnosis. Mini International Neuropsychiatric Interview (MINI) version 6 is also filled by subjects to assess psychological comorbidities.²³ Inclusion criteria are (1) for ADHD group—a previous diagnosis of ADHD and existence of ADHD symptoms in adulthood as determined by clinical interview and ADHD rating scales and (2) for control group—age-matched adults not having ADHD or any other psychopathology. Exclusion criteria is a comorbid current psychiatric diagnosis, substance dependence except for nicotine, head injuries, a medical or neurological disorder. During a clinical assessment, all participants were asked if they are taking any kind of medication. None of the participants were taking any medication either for ADHD or any other purpose. The functional ability of ADHD participants was not much affected as they had all competed in the national engineering entrance exam and sought

admission to public engineering college, which is highly reputed in the region.

EEG Data Acquisition and Preprocessing

The EEG data are recorded for ADHD and control group adults under 3 different conditions: (1) 3-minute eyes-open, (2) 3-minute eyes-closed, and (3) 11-minute CPT. CPT evaluates person's sustained attention, that is, the ability to maintain a consistent focus on some continuous activity or stimuli. In this task, participants are required to respond to a pseudorandom stream of 4-digit numbers (eg, 2392) presented on a computer screen, each time the same number appeared in 2 consecutive trials. Trial presentation time is 750 milliseconds with an inter-trial interval of 250 milliseconds for a total of 1000 ms/trial. The task consisted of total 660 experimental trials out of which 15% are target stimuli and rest 85% are nontarget stimuli.²⁴ The task is designed using OpenSesame platform, which is an open-source, graphical experiment builder for social sciences.²⁵ The task for sustained attention is devised and conducted under the supervision of a clinical psychologist. Figure 1 depicts the visual illustration of a CPT.

Electrophysiological recordings are obtained from 19 scalp electrodes (Fp1, Fp2, F3, F4, F7, F8, C3, C4, P3, P4, T3, T4, T5, T6, O1, O2, Fz, Cz, and Pz) placed according to the international 10/20 system. Electrode impedance is maintained within a similar range (5–10 kohm). All scalp electrodes are average referenced at the time of data acquisition. The recording parameters are sample rate of 256 Hz, bandpass hardware filter of 0.1 to 70 Hz and 50-Hz notch filter. Data are exported to MATLAB (MathWorks, Natick, MA, USA) platform for further analysis.

Preprocessing of Rest-State EEG (Eyes-Open and Eyes-Closed). An automated method is used to remove electro-oculogra (EOG)- and electromyogram (EMG)-related artifactual components from rest state EEG data.²⁶ Electrocardiography (ECG)-related artifactual components are removed through independent component analysis (ICA) decomposition in EEGLAB toolbox.²⁷ Next, continuous rest-state EEG data are segmented into

2-second epochs separately for each of the 2 conditions and standard thresholding techniques are applied to reject artifacts. Epochs with extreme values, that is, exceeding $\pm 75 \mu\text{V}$ in any of the 19 channels at any time within the epoch are rejected. To detect linear drifts due to artifactual currents, data are fitted to a straight line and epoch is rejected if the slope exceeded a given threshold of $50 \mu\text{V}$ per epoch. Probability feature is employed to reject improbable epochs (represent artifact) with the help of standard deviation of a mean probability distribution. Standard deviation limit set for all channels is 5. A statistical feature kurtosis is used to reject data epochs with peaky activity value distribution and standard deviation threshold set for all channels is 5. Epochs with power spectra deviated by $+25$ or -100 dB in 20 to 40 Hz frequency window are also rejected as they might represent muscle activity. Artifact-free data of 70 epochs is selected from each participant's EEG under each of 2 rest-state conditions (eyes-open and eyes-closed).

Preprocessing of Task-Related EEG. ICA method is used to decompose raw EEG data recorded during a sustained attention task. Artifactual independent components are rejected with the help of ADJUST algorithm and through visual inspection of component topographies and power spectrum.^{28,29} Next, continuous EEG data are then segmented into stimulus-locked epochs each of duration 850 milliseconds (100 milliseconds prestimulus to 750 milliseconds poststimulus). Epochs with extreme values, that is, exceeding $\pm 75 \mu\text{V}$ in any of the 19 channels at any time within the epoch are rejected. This study has selected 300 artifact-free epoch from each participant's EEG under task condition for further processing.

Feature Extraction

Phase Space Reconstruction. The PSR of a signal presents a visual image of the transformation of the dynamical behavior of the signal over time. It is a useful method to search for patterns in the high- or low-dimensional dynamical system. A phase space diagram can be used to describe a dynamical system. This diagram provides a coordinate system where coordinates are all the variables comprising the mathematical formulation of the system. A point in the phase space represents the state of the system at any given time.³⁰⁻³² Figure 2 depicts a 2-dimensional PSR plot of EEG signal from Cz channel. Consider the time series X_i , where $i = 1, 2, \dots, N$ and N is the total number of data points. The time-delay method is widely used for PSR in which phase space is reconstructed by its trajectories according to:

$$Y_n = (X_n, X_{n+\tau}, \dots, X_{n+(m-1)\tau}) \quad (1)$$

where $n = 1, 2, \dots, N - (m-1)\tau$, m is embedding dimension and τ is a time delay.

The value of embedding dimension m is set to 2 and time lag τ is set to 1 as per literature.^{18,19} The Euclidean distance is computed between the origin (0,0) and PSR(X_i, X_{i+1}) as per equation:

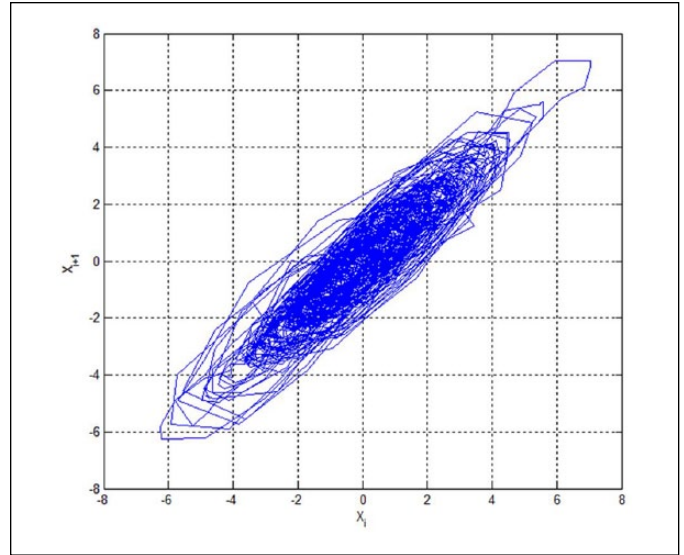


Figure 2. Two-dimensional phase space reconstruction (PSR) plot of EEG signal from Cz channel.

$$D(i) = \sqrt{(X_i)^2 + (X_{i+1})^2} \quad (2)$$

A set of 8 features, that is, maximum, minimum, mean, median, variance, skewness, kurtosis, and average power is extracted from Euclidean distance values.

Fractal Dimension. The fractal dimension (FD) of a biological signal measures its self-similarity, that is, how many times a pattern is repeated in the time-series. A more detailed description of FD and its usage in the development of EEG-based diagnosis system for various psychiatric disorders can be studied in Ahmadlou et al.³³⁻³⁵ The current study has computed Katz's and Higuchi's fractal dimension (KatzFD and HiguchiFD) to measure the nonlinearity and complexity of a signal.

Visibility Graph-Based Measures. A visibility graph (VG) converts a time-series to a graph where each node (or vertex) represents a time sample and an edge between 2 nodes shows the corresponding time samples can view each other. This graph inherits some of the properties of time-series such as periodic series results in regular graphs, random series results in random graphs and fractal series converts to scale-free networks. Previous research work has shown the robustness of VG-derived measures in the discrimination of psychiatric disorders.³⁶⁻³⁸ The current study has calculated two measures from VG of a time-series namely, power of scale-freeness in VG (PSVG) and graph index complexity (GIC).³⁸

Feature Selection

Correlation-based feature selection (CFS) is a filter-based method, which uses a heuristic function for evaluation and ranking of feature subset space.³⁹ In this method, irrelevant

features are ignored when their correlation with the class is low. Redundant features are also mean to be discarded as they are highly correlated with one or more of other features. A feature subset in which features are highly correlated with a specific class while not correlated with each other would be selected. This method is different from other feature selection methods as it uses heuristic merit for a feature subset selection instead of considering each feature independently. So, the decision of the next move is based on the option that maximizes the output of the function. CFS's feature subset evaluation function is defined as:

$$M_s = \frac{\overline{kr_{cf}}}{\sqrt{k + k(k-1)r_{ff}}} \quad (3)$$

where M_s is the heuristic "merit" of a feature subset S containing k features, $\overline{r_{cf}}$ is the mean feature-class correlation ($f \in S$), and $\overline{r_{ff}}$ is the average feature-feature intercorrelation. The numerator of Equation (3) can be thought of as providing an indication of how predictive of the class a set of features are; the denominator of how much redundancy there is among the features. The high value of M_s indicates that feature subset provides large discrimination between the 2 classes. There can be 3 heuristic search strategies: forward selection, backward elimination, and best first. Best first can start with either no features or all features. In the former, the search progresses forward through the search space adding single features; in the latter the search moves backward through the search space deleting single features. To prevent the best first search from exploring the entire feature subset search space, a stopping criterion is imposed. The search will terminate if 5 consecutive fully expanded subsets show no improvement over the current best subset. The current study has used "best first" strategy.

Particle swarm optimization (PSO) is a wrapper-based optimization technique that is inspired by the social behavior of animals such as bird flock and fish swarm.^{40,41} It is initialized by a group of random solutions and searches for an optimum solution. Each solution is represented by a particle. In every iteration, it is evaluated by a fitness function. Each particle keeps the position, velocity, and Pbest (ie, particle best, which is the best value achieved by each particle so far) value. This information is updated in each iteration. In addition to that, swarm also maintains Gbest (global best) value, which is the best value achieved by the swarm at a given time. Each particle updates its velocity which is calculated as follows:

$$V_k(l+1) = wV_k(l) + c1r1[p_k(l) - X_k(l)] + c2r2[g(l) - X_k(l)] \quad (4)$$

where w is the inertia weight, k is the index of the particle, $c1$ and $c2$ are acceleration coefficients, $r1$ and $r2$ are random values between range (0,1), $X_k(l)$ is the current position of a particle, $V_k(l)$ is the current velocity of a particle, $p_k(l)$ is the Pbest value achieved by each particle and $g(l)$ is the Gbest

value achieved by swarm at time l . In the current study, the value of w is set to 1 and the values of $c1$ and $c2$ parameters are taken as 2.⁴² Position of each particle is updated as follows:

$$X_k(l+1) = X_k(l) + V_k(l+1) \quad (5)$$

where $X_k(l)$ is the current position and $X_k(l+1)$ is the updated position. Pbest value is defined and updated as follows:

$$Pbest = \max(p_k(l)) \quad (6)$$

$$p_k(l+1) = \begin{cases} X_k(l+1), & \text{if } fit(X_k(l+1)) > fit(p_k(l)) \\ p_k(l), & \text{if } fit(X_k(l+1)) \leq fit(p_k(l)) \end{cases} \quad (7)$$

where fit is the value of fitness function. Gbest value is defined and updated as follows:

$$Gbest = g(l) \quad (8)$$

$$g(l) = \max\{fit_1(l), fit_2(l), \dots, fit_N(l)\} \quad (9)$$

where N is the number of particles in the swarm, $fit_N(l)$ is the fitness function, that is, classification accuracy achieved by N number of particles for l th iteration. Gbest is the maximum fitness value achieved until iteration l . Particle swarm optimization algorithm consists of 3 steps:

- Fitness of each particle is evaluated by a fitness function.
- Update the Pbest (particle best) and Gbest (global best) of each particle according to Equations (7) and (9).
- Update velocity and position of each particle according to Equations (4) and (5).

These three steps are repeated until the termination condition is achieved, which is the maximum number of iterations. If the termination condition is not achieved the process continues to find the best fitness value of each particle. When the termination condition has achieved, the particle having the best fitness values is selected and its corresponding feature subset is selected. The subset of features is evaluated by classification accuracy as a fitness function. Each particle is given as input to this function, and the output is computed in the form of a scalar value, which is between 0 and 1. Experiments are performed by varying the number of particles and iterations. Optimal results are obtained when the number of particles and iterations are taken as 15 and 100, respectively.

Classification

Support Vector Machine. An SVM is a supervised learning algorithm that maps the data into a higher dimensional input

Table 1. Numbers of Training and Test Sets.

Group	Training Set	Test Set	Total Set
ADHD	32	15	47
Control	35	15	50
Total	67	30	97

space and constructs an optimal separating hyperplane in this space.⁴³ It creates 2 parallel hyperplanes, one on each side of the separating hyperplane, that have the largest margin from training data. The SVM algorithm uses a set of mathematical functions, known as kernels, for handling the nonlinearly separable data.

K-Nearest Neighbor. The KNN algorithm is a nonparametric technique of pattern recognition.⁴⁴ There are K closest training examples each with labels in the input feature space. The test case is classified by assigning the label that is most frequent among the K training samples nearest to the test point.

Naive-Bayes Classifier. The NBC is a probabilistic classifier based on Bayes theorem that assumes independence between the features.⁴⁵ The NBC model uses a maximum probability algorithm to determine the class of earlier probabilities and a feature's probability distribution from a training dataset. Results are then employed with a maximized posteriori decision tree to find the specific class label for a new test instance.

Enhanced Probabilistic Neural Network. The EPNN is an enhanced version of the probabilistic neural network based on the notion of local decision circles which enable EPNN to incorporate local information and non-homogeneity existing in the training population.⁴⁶ The circle has a radius which limits the contribution of the local decision.

Neural Dynamic Classifier. The NDC is based on the concept that the probability of a new data point being correctly classified can be increased if the existing training data points are transformed into a new feature space with large margins between clusters and the close proximity of the classmates.⁴⁷ This algorithm uses a 2-phase learning approach: in the first phase, a neural dynamic optimization model is used to identify a new feature space and in the second phase, a Gaussian probability density function is employed for classification in the new feature space.

Results and Discussion

The proposed system of 2-class classification is validated on the EEG dataset of 97 adults (ADHD, 47; Control, 50) acquired under the eyes-open, eyes-closed, and CPT conditions. Table 1 presents a division of the dataset into training and test sets for classification purposes. A set of 8 PSR-based

Table 2. Classification Results Using PSR-Based Feature Set Extracted From Individual Channels Under the Eyes-Open Condition^a.

Channel	NDC	EPNN	SVM	KNN	NBC
Fp1	90.0	86.7	90.0	83.3	56.7
Fp2	76.7	60.0	73.3	66.7	63.3
F7	63.3	53.3	63.3	46.7	50.0
F3	60.0	46.7	60.0	53.3	46.7
Fz	56.7	50.0	53.3	46.7	46.7
F4	66.7	63.3	66.7	63.3	66.7
F8	66.7	63.3	63.3	60.0	63.3
T3	76.7	70.0	76.7	70.0	66.7
C3	66.7	53.3	63.3	43.3	56.7
Cz	86.7	76.7	86.7	60.0	73.3
C4	70.0	66.7	66.7	63.3	50.0
T4	76.7	73.3	76.7	56.7	53.3
T5	76.7	70.0	73.3	63.3	60.0
P3	70.0	66.7	70.0	66.7	56.7
Pz	53.3	46.7	46.7	36.7	33.3
P4	70.0	63.3	66.7	43.3	60.0
T6	56.7	53.3	56.7	50.0	43.3
O1	53.3	46.7	46.7	43.3	43.3
O2	66.7	60.0	63.3	60.0	63.3

Abbreviations: PSR, phase space reconstruction; NDC, neural dynamic classifier; EPNN, enhanced probabilistic neural network; SVM, support vector machine; KNN, K-nearest neighbor; NBC, naive-Bayes classifier.

^aValues in boldface indicate highest accuracy.

features, 2 FD features and 2 VG-based features are extracted from each of the 19 channels. Features are normalized using NORMC function of MATLAB, which normalizes the sum of the squares of the elements in each column to have a value of 1. Most discriminating features between the two groups are selected using CFS (filter-based) and PSO (wrapper-based) methods. After that, the performance of 5 machine learning methods, that is, NDC, EPNN, SVM, KNN, and NBC is compared for discrimination of EEG signals of ADHD and control group adults. The classification performance is evaluated against a test set after learning with training set using sensitivity, specificity, and accuracy parameters. Experimental results are taken using various feature sets under 3 different conditions.

Classification With a Complete Feature Set

A set of 8 PSR-based features extracted from each of the 19 channels, is fed individually as an input to the classifier. Table 2 presents testing accuracies achieved by 5 classifiers for different channels under the eyes-open condition. It is clear from the table that highest accuracy of 90% (sensitivity, 100%; specificity, 80%) is obtained by NDC and SVM classifier at Fp1 channel. Table 3 shows that NDC attained the highest accuracy of 83.3% (sensitivity, 86.7%; specificity, 80%) at F3 and F8 channel under the eyes-closed condition. Under CPT condition, the

Table 3. Classification Results Using PSR-Based Feature Set Extracted From Individual Channels Under the Eyes-Closed Condition^a.

Channel	NDC	EPNN	SVM	KNN	NBC
Fp1	73.3	63.3	70.0	60.0	56.7
Fp2	66.7	60.0	66.7	66.7	63.3
F7	66.7	63.3	63.3	53.3	63.3
F3	83.3	70.0	76.7	73.3	73.3
Fz	80.0	73.3	76.7	70.0	76.7
F4	60.0	46.7	60.0	43.3	50.0
F8	83.3	66.7	80.0	76.7	60.0
T3	66.7	63.3	66.7	63.3	60.0
C3	56.7	53.3	53.3	50.0	50.0
Cz	70.0	60.0	66.7	60.0	63.3
C4	56.7	50.0	53.3	46.7	53.3
T4	63.3	56.7	63.3	50.0	56.7
T5	66.7	53.3	66.7	66.7	56.7
P3	70.0	63.3	63.3	43.3	53.3
Pz	63.3	60.0	63.3	46.7	56.7
P4	66.7	63.3	66.7	56.7	60.0
T6	73.3	66.7	70.0	66.7	70.0
O1	70.0	63.3	66.7	50.0	63.3
O2	60.0	56.7	60.0	50.0	53.3

Abbreviations: PSR, phase space reconstruction; NDC, neural dynamic classifier; EPNN, enhanced probabilistic neural network; SVM, support vector machine; KNN, K-nearest neighbor; NBC, naive-Bayes classifier.

^aValues in boldface indicate highest accuracy.

Table 4. Classification Results Using PSR-Based Feature Set Extracted From Individual Channels Under the CPT Condition^a.

Channel	NDC	EPNN	SVM	KNN	NBC
Fp1	83.3	76.7	83.3	63.3	46.7
Fp2	76.7	70.0	73.3	73.3	70.0
F7	80.0	73.3	80.0	56.7	43.3
F3	60.0	53.3	56.7	53.3	40.0
Fz	66.7	60.0	63.3	63.3	60.0
F4	93.3	76.7	93.3	53.3	70.0
F8	56.7	50.0	50.0	36.7	43.3
T3	60.0	56.7	60.0	53.3	46.7
C3	53.3	50.0	53.3	43.3	40.0
Cz	70.0	63.3	66.7	60.0	60.0
C4	76.7	73.3	76.7	66.7	70.0
T4	63.3	53.3	53.3	50.0	60.0
T5	60.0	53.3	56.7	53.3	43.3
P3	86.7	76.7	86.7	70.0	53.3
Pz	76.7	73.3	76.7	63.3	50.0
P4	60.0	56.7	60.0	53.3	56.7
T6	60.0	53.3	56.7	46.7	56.7
O1	70.0	63.3	70.0	66.7	63.3
O2	53.3	50.0	50.0	46.7	50.0

Abbreviations: PSR, phase space reconstruction; CPT, Continuous Performance Test; NDC, neural dynamic classifier; EPNN, enhanced probabilistic neural network; SVM, support vector machine; KNN, K-nearest neighbor; NBC, naive-Bayes classifier.

^aValues in boldface indicate highest accuracy.

Table 5. Performance of 5 Classifiers for Various Feature Sets Under Different Conditions^a.

Condition	Feature Set	NDC	EPNN	SVM	KNN	NBC
Eyes-open	PSR	56.7	50.0	53.3	46.7	50.0
	FD	50.0	46.7	50.0	43.3	46.7
	VG	53.3	43.3	50.0	46.7	46.7
Eyes-closed	PSR	60.0	53.3	56.7	53.3	53.3
	FD	53.3	46.7	50.0	43.3	50.0
	VG	53.3	50.0	53.3	46.7	43.3
CPT	PSR	66.7	60.0	60.0	53.3	50.0
	FD	53.3	46.7	50.0	40.0	43.3
	VG	56.7	53.3	56.7	53.3	46.7
Joint (eyes-open, eyes-closed, and CPT)	PSR	63.3	53.3	63.3	56.7	60.0
	FD	56.7	46.7	56.7	53.3	46.7
	VG	60.0	53.3	56.7	53.3	50.0

Abbreviations: NDC, neural dynamic classifier; EPNN, enhanced probabilistic neural network; SVM, support vector machine; KNN, K-nearest neighbor; NBC, naive-Bayes classifier; PSR, phase space reconstruction; FD, fractal dimension; VG, visibility graph; CPT, Continuous Performance Test.

^aValues in boldface indicate highest accuracy.

highest accuracy of 93.3% (sensitivity, 100%; specificity, 86.7%) is achieved by NDC and SVM at F4 channel as shown in Table 4. It has been observed from the experimental results that the NDC outperforms other classifiers under 3 conditions. The NDC discovers the most effective feature spaces and finds the optimum number of features for accurate classification based on neural dynamic optimization model. After the NDC, the SVM algorithm achieved better classification accuracy than EPNN, KNN, and NBC and it works well even with few training examples. Once a hyperplane is found, most of the data other than the support vectors become redundant. This means that small changes to data cannot greatly affect the hyperplane and hence the SVMs tend to generalize very well. The KNN algorithm computes the distance and sort all the training data at each prediction, which can be slow if there are a large number of training examples. Moreover, it does not learn anything from the training data, which can result in the algorithm not generalizing well. The NBC is based on an assumption of attribute independence in the calculation of outcome probabilities, which hinders the performance of the classifier.

Table 5 exhibits a performance comparison of five classifiers for various feature sets based on PSR, FD, and VG methods under different conditions. This study has considered a total set of 152 PSR-based features computed from all channels ($19 \times 8 = 152$), 38 FD-based features ($19 \times 2 = 38$), and 38 VG-based features ($19 \times 2 = 38$). It can be analyzed from Table 5 that the NDC has attained highest accuracy of 56.7% (sensitivity, 66.7%; specificity, 46.7%) under the eyes-open condition, 60.0% (sensitivity, 60.0%; specificity, 60.0%) under the eyes-closed condition, and 66.7% (sensitivity, 76.7%; specificity, 56.7%) under the CPT condition with

Table 6. Reduced Feature Set Under 3 Conditions.

Condition	Input Feature Set	Feature Selection	Reduced Feature Set
Eyes-open	PSR features from 19 channels	CFS	Skewness_F3, Skewness_F4, Kurtosis_F8, Skewness_C3, Skewness_Cz, Skewness_T5, Kurtosis_T5, Skewness_P4
		PSO	Median_Fp2, AveragePower_Fp2, Mean_F3, Skewness_F3, Kurtosis_F3, Median_F4, Skewness_F8, Kurtosis_Cz, AveragePower_P3, Mean_T6, Maximum_O1
	FD features from 19 channels	CFS	HiguchiFD_Fp2, KatzFD_C3, KatzFD_Cz, HiguchiFD_T4
		PSO	KatzFD_Fz, KatzFD_T3, HiguchiFD_C3, HiguchiFD_Cz, KatzFD_C4, KatzFD_O1, HiguchiFD_O2
	VG features from 19 channels	CFS	GIC_F8, GIC_T3, GIC_Cz, PSVG_C4, GIC_O1
		PSO	GIC_Fp1, PSVG_T3, GIC_C3, PSVG_T5, GIC_P3, PSVG_P4, GIC_T6, PSVG_O1
Eyes-closed	PSR features from 19 channels	CFS	Skewness_Fp1, Skewness_Fp2, Skewness_F7, Skewness_F3, Skewness_Fz, Skewness_F8, Skewness_Cz, Kurtosis_Cz, Kurtosis_C4, Kurtosis_T4, Skewness_T5, Skewness_O2
		PSO	Kurtosis_Fp1, Skewness_Fp2, AveragePower_F7, Median_F3, Maximum_Fz, Variance_F8, Skewness_C3, Skewness_Cz, skewness_T5, AveragePower_P3, Skewness_P4, Median_T6
	FD features from 19 channels	CFS	HiguchiFD_C3, KatzFD_Cz
		PSO	HiguchiFD_F7, KatzFD_C3, HiguchiFD_O1
	VG features from 19 channels	CFS	PSVG_Fp2, PSVG_Pz
		PSO	PSVG_Fp2, GIC_F3, PSVG_Fz, PSVG_C4
CPT	PSR features from 19 channels	CFS	Maximum_Fp2, Minimum_Fp2, Kurtosis_Fp2, Kurtosis_P3, Kurtosis_Pz, Median_T6 and Maximum_O2
		PSO	Skewness_Fp1, Minimum_Fp2, Median_F7, Maximim_F4, Kurtosis_T3, Mean_Cz, Mean_T5, Kurtosis_P3, AveragePower_P4
	FD features from 19 channels	CFS	HiguchiFD_Fp1, HiguchiFD_Fp2, HiguchiFD_F7
		PSO	KatzFD_T4, HiguchiFD_T5, KatzFD_P3, KatzFD_O1
	VG features from 19 channels	CFS	PSVG_F8, PSVG_Cz
		PSO	PSVG_Fp1, GIC_F4, PSVG_T6, GIC_T6

Abbreviations: PSR, phase space reconstruction; CFS, correlation-based feature selection; PSO, particle swarm optimization; FD, fractal dimension; VG, visibility graph; CPT, Continuous Performance Test.

input of PSR-based feature set. The system has provided the highest accuracy of 63.3% (sensitivity, 63.3%; specificity, 63.3%) for the joint feature set obtained from 3 conditions. The PSR based feature set are able to discriminate well between ADHD and control group adults as compared with 2 other nonlinear methods (FD and VG) used in this study. This is because the state space features represent the dynamics of the system. As the system visits the unique state space without repetitions. So, the feature set extracted using PSR method best discriminated EEG signals of ADHD and control group adults.

Classification With a Reduced Feature Set

This study has applied CFS and PSO methods to find optimal or reduced feature set for ADHD and control group discrimination under 3 conditions (eyes-open, eyes-closed, and CPT). Then, the reduced feature set shown in

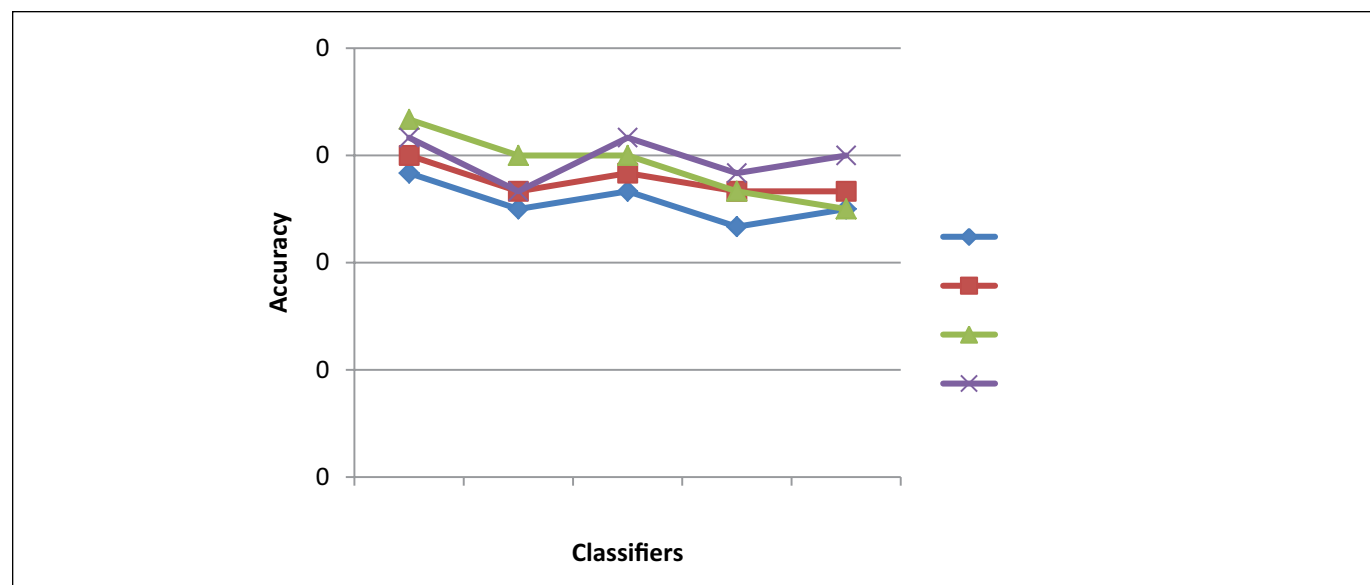
Table 6 is given as an input to each of the 5 classifiers. Each entry of the table is in form of “feature_channel.” The “feature” can be from PSR (maximum, minimum, mean, median, variance, skewness, kurtosis, and average power), FD (KatzFD and HiguchiFD) or VG (PSVG and GIC) feature set and “channel” can be Fp1, Fp2, F3, F4, F7, F8, C3, C4, P3, P4, T3, T4, T5, T6, O1, O2, Fz, Cz, or Pz. It is clear from Table 7 that the NDC and SVM has attained highest accuracy of 93.3% (sensitivity, 100%; specificity, 86.7%) under the eyes-open condition. The NDC has provided the highest accuracy of 90.0% (sensitivity, 93.3%; specificity, 86.7%) under the eyes-closed condition and 100.0% (sensitivity, 100.0%; specificity, 100.0%) under the CPT condition. The system has attained the highest accuracy of 80.0% (sensitivity, 80.0%; specificity, 80.0%) with the NDC and SVM using a reduced feature set selected from joint feature set (of 3 conditions). The use of a filter (CFS) and wrapper-based method (PSO)

Table 7. Performance of 5 Classifiers for CFS- and PSO-Selected Feature Sets Under Different Conditions^a.

Condition	Feature Set	Feature Selection	NDC	EPNN	SVM	KNN	NBC
Eyes-open	PSR	CFS	80.0	73.3	80.0	70.0	76.7
		PSO	93.3	83.3	93.3	73.3	80.0
	FD	CFS	56.7	50.0	53.3	50.0	53.3
		PSO	63.3	60.0	63.3	56.7	56.7
	VG	CFS	63.3	56.7	60.0	53.3	56.7
		PSO	73.3	66.7	66.7	60.0	63.3
Eyes-closed	PSR	CFS	80.0	73.3	76.7	73.3	73.3
		PSO	90.0	86.7	86.7	83.3	86.7
	FD	CFS	56.7	50.0	50.0	46.7	50.0
		PSO	63.3	56.7	60.0	53.3	60.0
	VG	CFS	73.3	63.3	70.0	56.7	60.0
		PSO	76.7	66.7	73.3	63.3	66.7
CPT	PSR	CFS	93.3	80.0	80.0	76.7	80.0
		PSO	100.0	86.7	96.7	86.7	93.3
	FD	CFS	56.7	50.0	53.3	43.3	46.7
		PSO	66.7	63.3	66.7	56.7	60.0
	VG	CFS	63.3	56.7	60.0	56.7	50.0
		PSO	70.0	63.3	66.7	60.0	63.3
Joint (eyes-open, eyes-closed, and CPT)	PSR	CFS	73.3	66.7	70.0	66.7	70.0
		PSO	80.0	76.7	80.0	73.3	76.7
	FD	CFS	66.7	60.0	63.3	56.7	63.3
		PSO	70.0	63.3	70.0	60.0	66.7
	VG	CFS	73.3	66.7	73.3	60.0	70.0
		PSO	80.0	73.3	76.7	66.7	73.3

Abbreviations: CFS, correlation-based feature selection; PSO, particle swarm optimization; NDC, neural dynamic classifier; EPNN, enhanced probabilistic neural network; SVM, support vector machine; KNN, K-nearest neighbor; NBC, naive-Bayes classifier; PSR, phase space reconstruction; FD, fractal dimension; VG, visibility graph; CPT, Continuous Performance Test.

^aValues in boldface indicate highest accuracy.

**Figure 3.** Performance comparison among different conditions using a complete phase space reconstruction (PSR) feature set.

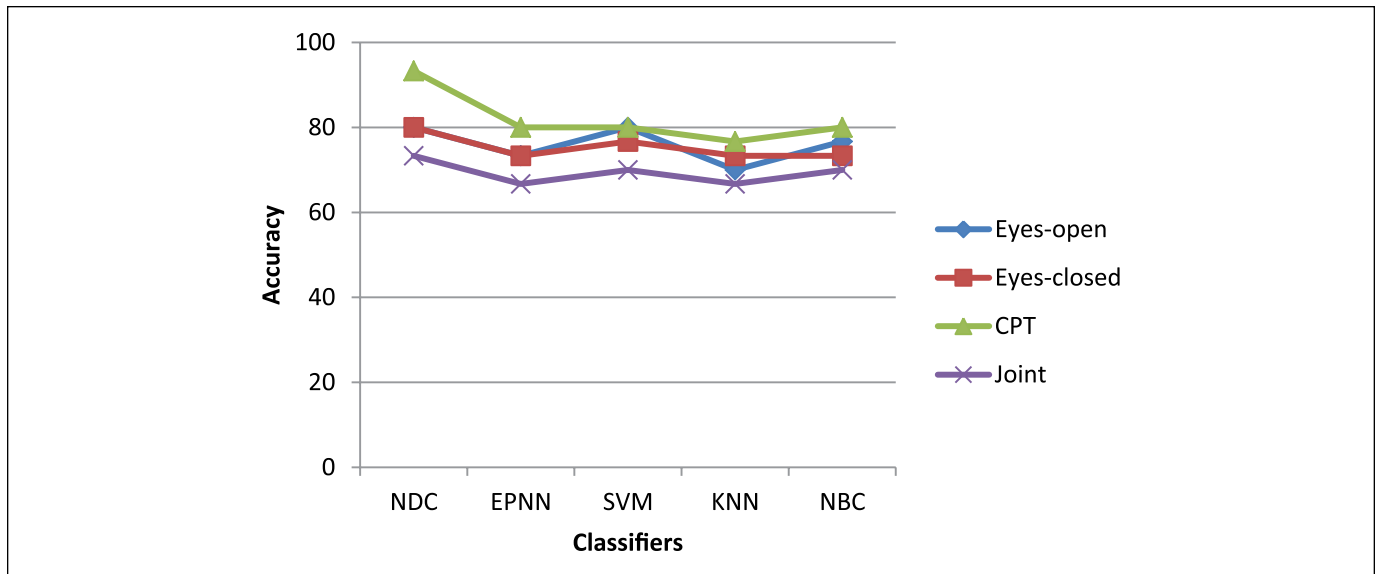


Figure 4. Performance comparison among different conditions using correlation-based feature selection (CFS)-selected phase space reconstruction (PSR) feature set.

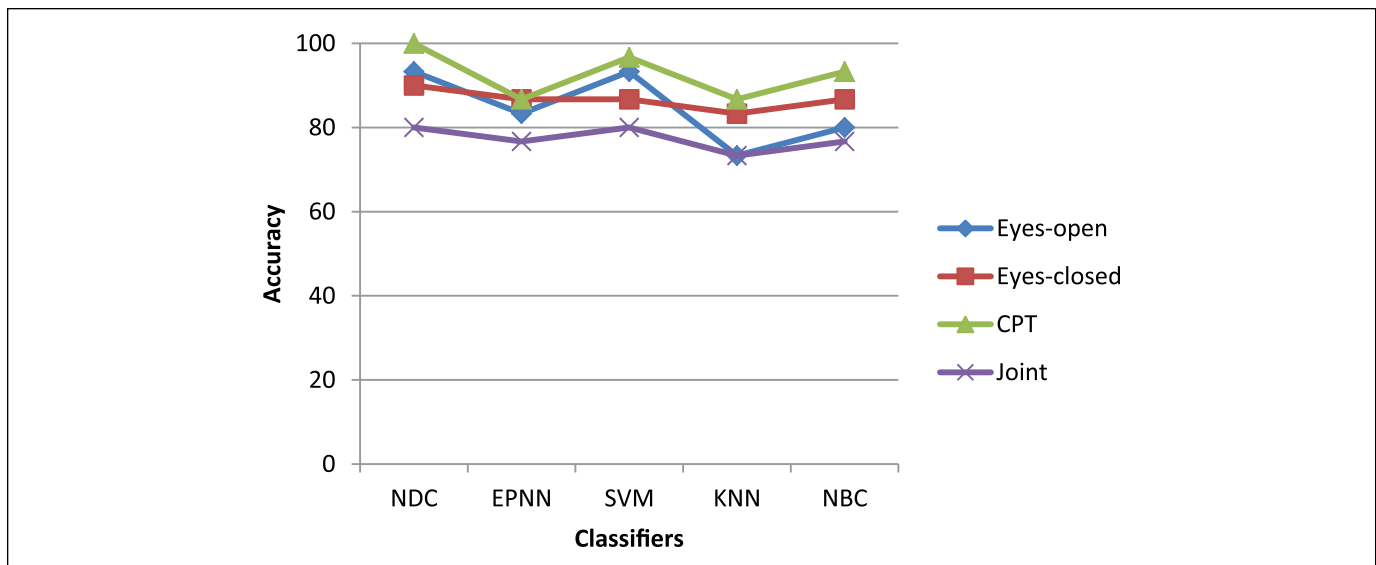


Figure 5. Performance comparison among different conditions using particle swarm optimization (PSO) selected phase space reconstruction (PSR) feature set.

for feature selection indicate that the 2 groups are best differentiated using feature set selected with the PSO method. Filter-based approach (ie, CFS) considers intrinsic properties of training set instead of cross-validation performance in selecting most discriminative feature subset. However, a wrapper-based approach (ie, PSO) uses a learning algorithm to search for optimal feature subset and provide better performance. But the wrapper method is computationally more expensive as compared with the filter method due to repeated learning steps and cross-validation.

Comparison of Classification Under Different Conditions

Figure 3 illustrates classification-based comparison using a complete PSR-based feature set extracted from EEG data recorded under the eyes-open, eyes-closed, and CPT condition. Performance of 5 classifiers for the joint feature set obtained from 3 conditions are also shown in Figure 3. The NDC and EPNN classifiers have provided the best accuracies under CPT condition whereas SVM, KNN, and NBC performed well with

the joint feature set obtained from three conditions. All the classifiers have achieved the lowest accuracies using the complete feature set extracted from the eyes-open condition. It is clear from Figure 4 that the performance by 5 classifiers remains best with *CFS-selected* PSR feature set under the CPT condition. Figure 5 depicts classification results using *PSO-selected* PSR feature set under different conditions. It can be seen that all classifiers demonstrated the highest accuracies under the CPT condition. But the classification results are less accurate for *CFS* or *PSO selected* joint PSR feature set (of 3 conditions).

It can be seen from experimental results that EEG signals captured under CPT condition are able to best discriminate between ADHD and control adults in comparison to rest state conditions (eyes-open and eyes-closed). As there is sustained attention-related deficit in ADHD adults, so their brain signals exhibited different dynamical behavior from control adults during information processing, that is, CPT condition. These differences made it possible for the proposed system to achieve the highest classification accuracy for EEG data recorded during CPT.

Conclusion

The use of EEG can play a key role in assisting clinicians in making a more accurate ADHD diagnosis. This study has explored the features obtained from PSR-based Euclidean distances and different feature selection algorithms to classify EEG signals of ADHD and control adults. The brain activity of 2 groups is recorded under eyes-open, eyes-closed, and CPT conditions. It can be concluded from experimental results that NDC gives the highest classification accuracy of 100% under CPT condition with PSO selected PSR feature set. The NDC classifier outperforms SVM, EPNN, KNN, and NBC for signal features extracted under three different conditions. This study points toward the potential of EEG signals to act as objective biomarkers of ADHD.

Author Contributions

Simranjit Kaur contributed to acquisition, analysis, and interpretation of the data. Sukhwinder Singh and Damanjeet Kaur contributed to concept and design of the study. Priti Arun and Manoj Bajaj contributed to acquisition of data for the study.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.


Ethical Approval

The Institutional Research Ethics Board of Panjab University and Government Medical College and Hospital approved this study and all participants gave informed written consent prior to the start of this study.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Sukhwinder Singh  <https://orcid.org/0000-0002-5204-6487>

References

1. Polanczyk G, de Lima MS, Horta BL, Biederman J, Rohde LA. The worldwide prevalence of ADHD: a systematic review and meta-regression analysis. *Am J Psychiatry*. 2007;164:942-948.
2. Simon V, Czobor P, Bálint S, Mészáros Á, Bitter I. Prevalence and correlates of adult attention-deficit hyperactivity disorder: meta-analysis. *Br J Psychiatry*. 2009;194:204-211.
3. Rabiner DL, Anastopoulos AD, Costello J, Hoyle RH, Swartzwelder HS. Adjustment to college in students with ADHD. *J Atten Disord*. 2008;11:689-699.
4. Lenartowicz A, Loo SK. Use of EEG to diagnose ADHD. *Curr Psychiatry Rep*. 2014;16:498.
5. Sridhar C, Bhat S, Acharya UR, Adeli H, Bairy GM. Diagnosis of attention deficit hyperactivity disorder using imaging and signal processing techniques. *Comp Biol Med*. 2017;88:93-99.
6. Magee CA, Clarke AR, Barry RJ, McCarthy R, Selikowitz M. Examining the diagnostic utility of EEG power measures in children with attention deficit/hyperactivity disorder. *Clin Neurophysiol*. 2005;116:1033-1040.
7. Ahmadi M, Adeli H. Wavelet-synchronization methodology: a new approach for EEG-based diagnosis of ADHD. *Clin EEG Neurosci*. 2010;41:1-10.
8. Sadatnezhad K, Boostani R, Ghanizadeh A. Classification of BMD and ADHD patients using their EEG signals. *Expert Syst Appl*. 2011;38:1956-1963.
9. Ahmadi M, Adeli H. Fuzzy synchronization likelihood with application to attention deficit/hyperactivity disorder. *Clin EEG Neurosci*. 2011;42:6-13.
10. Ahmadi M, Adeli H, Adeli A. Graph theoretical analysis of organization of functional brain networks in ADHD. *Clin EEG Neurosci*. 2012;43:5-13.
11. Liechti MD, Valko L, Müller UC, et al. Diagnostic value of resting electroencephalogram in attention-deficit/hyperactivity disorder across the lifespan. *Brain Topogr*. 2013;26:135-151.
12. Poil SS, Bollmann S, Ghisleni C, et al. Age dependent electroencephalographic changes in attention-deficit/hyperactivity disorder (ADHD). *Clin Neurophysiol*. 2014;125:1626-1638.
13. Helgadóttir H, Gudmundsson ÓÓ, Baldursson G, et al. Electroencephalography as a clinical tool for diagnosing and monitoring attention deficit hyperactivity disorder: a cross-sectional study. *BMJ Open*. 2015;5:e005500.
14. Mueller A, Candrian G, Grane VA, Kropotov JD, Ponomarev VA, Baschera GM. Discriminating between ADHD adults and controls using independent ERP components and a support vector machine: a validation study. *Nonlinear Biomed Phys*. 2011;5:5.
15. Abibullaev B, An J. Decision support algorithm for diagnosis of ADHD using electroencephalograms. *J Med Syst*. 2012;36:2675-2688.
16. Tenev A, Markovska-Simoska S, Kocarev L, Pop-Jordanov J, Müller A, Candrian G. Machine learning approach for classification of ADHD adults. *Int J Psychophysiol*. 2014;93:162-166.

17. Mohammadi MR, Khaleghi A, Nasrabadi AM, Rafieivand S, Begol M, Zarafshan H. EEG classification of ADHD and normal children using non-linear features and neural network. *Biomed Eng Lett*. 2016;6:66-73.
18. Sharma R, Pachori RB. Classification of epileptic seizures in EEG signals based on phase space representation of intrinsic mode functions. *Expert Syst Appl*. 2015;42:1106-1117.
19. Lee SH, Lim JS, Kim JK, Yang J, Lee Y. Classification of normal and epileptic seizure EEG signals using wavelet transform, phase-space reconstruction, and Euclidean distance. *Comput Methods Prog Biomed*. 2014;116:10-25.
20. Qiao M, Ma X, Tao H. Chaotic time series prediction based on phase space reconstruction and LSSVR model. *Paper presented at: Proceedings of the 30th Chinese Control Conference*; July 22-24, 2011; Yantai, China:3243-3247.
21. Ward MF, Wender PH, Reimherr FW. The Wender Utah Rating Scale: an aid in the retrospective diagnosis of childhood attention deficit hyperactivity disorder. *Am J Psychiatry*. 1993;150:885-890.
22. Kessler RC, Adler L, Ames M, et al. The World Health Organization Adult ADHD Self-Report Scale (ASRS): a short screening scale for use in the general population. *Psychol Med*. 2005;35:245-256.
23. Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry*. 1998;59(suppl 20):22-33.
24. Fleck DE, Eliassen JC, Durling M, et al. Functional MRI of sustained attention in bipolar mania. *Mol Psychiatry*. 2012;17:325-336.
25. Mathôt S, Schreij D, Theeuwes J. OpenSesame: an open-source, graphical experiment builder for the social sciences. *Behav Res Methods*. 2012;44:314-324.
26. Daly I, Scherer R, Billinger M, Müller-Putz G. FORCE: fully online and automated artifact removal for brain-computer interfacing. *IEEE Trans Neural Syst Rehabil Eng*. 2015;23:725-736.
27. Swartz Center for Computational Neuroscience. What is EEGLAB. <http://www.sccn.ucsd.edu/eeglab/>. Accessed August 31, 2019.
28. Chaumon M, Bishop DV, Busch NA. A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *J Neurosci Methods*. 2015;250:47-63.
29. Mognon A, Jovicich J, Bruzzone L, Buiatti M. ADJUST: an automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*. 2011;48:229-240.
30. Takens F. Detecting strange attractors in turbulence. In: Rand DA, Young LS, eds. *Dynamical Systems and Turbulence, Warwick 1980*. (Lecture Notes in Mathematics, Vol 898). Berlin, Germany: Springer; 1981:366-381.
31. Sivakumar B. A phase-space reconstruction approach to prediction of suspended sediment concentration in rivers. *J Hydrol*. 2002;258:149-162.
32. Sivakumar B, Jayawardena AW, Fernando TMKG. River flow forecasting: use of phase-space reconstruction and artificial neural networks approaches. *J Hydrol*. 2002;265:225-245.
33. Ahmadi M, Adeli H, Adeli A. Fractality and a wavelet-chaos-neural network methodology for EEG-based diagnosis of autistic spectrum disorder. *J Clin Neurophysiol*. 2010;27:328-333.
34. Ahmadi M, Adeli H, Adeli A. Fractality and a wavelet-chaos-neural network methodology for EEG-based diagnosis of Alzheimer's disease. *Alzheimer Dis Assoc Disord*. 2011;25:85-92.
35. Ahmadi M, Adeli H, Adeli A. Fractality analysis of frontal brain in major depressive disorder. *Int J Psychophysiol*. 2012;85:206-211.
36. Ahmadi M, Adeli H. Visibility graph similarity: a new measure of generalized synchronization in coupled dynamic systems. *Physica D Nonlinear Phenomena*. 2012;241:326-332.
37. Ahmadi M, Adeli H, Adeli A. Improved visibility graph fractality with application for diagnosis of autism spectrum disorder. *Physica A Stat Mech Appl*. 2012;391:4720-4726.
38. Ahmadi M, Adeli H, Adeli A. New diagnostic EEG markers of the Alzheimer's disease using visibility graph. *J Neural Transm (Vienna)*. 2010;117:1099-1109.
39. Hall MA. Correlation-Based Feature Selection for Machine Learning [thesis]. Hamilton, New Zealand: University of Waikato; 1999.
40. Shi Y, Eberhart RC. Empirical study of particle swarm optimization. In: *Proceedings of the 1999 Congress on Evolutionary Computation, CEC 99*. Vol 3. Piscataway, NJ: IEEE; 1999:1945-1950.
41. Mirjalili S, Lewis A. S-shaped versus V-shaped transfer functions for binary particle swarm optimization. *Swarm Evolutionary Comput*. 2013;9:1-14.
42. Shi Y, Eberhart RC. A modified particle swarm optimizer. In: *1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence*. Piscataway, NJ: IEEE; 1998:69-73.
43. Suykens JAK, Vandewalle J. Least squares support vector machine classifiers. *Neural Process Lett*. 1999;9:293-300.
44. Cover T, Hart P. Nearest neighbor pattern classification. *IEEE Trans Inf Theory*. 1967;13:21-27.
45. Han J, Pei J, Kamber M. *Data Mining: Concepts and Techniques*. Amsterdam, Netherlands: Elsevier; 2011.
46. Ahmadi M, Adeli H. Enhanced probabilistic neural network with local decision circles: a robust classifier. *Integr Comput Aided Eng*. 2010;17:197-210.
47. Rafiei MH, Adeli H. A new neural dynamic classification algorithm. *IEEE Trans Neural Netw Learn Syst*. 2017;28:3074-3083.