Differential Diagnosis of Attention-Deficit / Hyperactivity Disorder and Bipolar Disorder using Steady-State Visual Evoked Potentials

Xiaoxia Li*
Jianghan University
Wuhan Hubei, 430056, China

***Abstract*—Bipolar disorder and Attention-deficit/Hyperactivity disorder (ADHD) are two prevalent disorders whose symptoms are similar. In order to reduce the misdiagnosis between bipolar disorder and ADHD, a machine learning-based system using electroencephalography (EEG) and steady state potentials (i.e., steady-state visual evoked potential [SSVEP]) was evaluated to classify ADHD, bipolar disorder and normal conditions. Indeed, this research was conducted for the first time with the aim of designing a machine learning system for EEG detection of ADHD, bipolar disorder, and normal conditions using SSVEPs. For this purpose, both linear and nonlinear dynamics of extracted SSVEPs were analyzed. Indeed, after data preprocessing, spectral analysis and recurrence quantification analysis (RQA) were applied to SSVEPs. Then, feature selection was utilized through the DISR. Finally, we utilized various machine learning techniques to classify the linear and nonlinear features extracted from SSVEPs into three classes of ADHD, bipolar disorder and normal: k-nearest neighbors (KNN), support vector machine (SVM), linear discriminant analysis (LDA) and Naïve Bayes. Experimental results showed that SVM classifier with linear kernel yielded an accuracy of 78.57% for ADHD, bipolar disorder and normal classification through the leave-one-subject-out (LOSO) cross-validation. Although this research is the first to evaluate the utilization of signal processing and machine learning approaches in SSVEP classification of these disorders, it has limitations that future studies should investigate to enhance the efficacy of proposed system.**

Keywords—Attention-deficit/Hyperactivity disorder (ADHD); bipolar disorder; electroencephalography (EEG); steady-state visual evoked potential (SSVEP); machine learning; classification

I. Introduction

Correct and accurate diagnosis of people with neuro-psychiatric illnesses like Attention-deficit/Hyperactivity disorder (ADHD), impulse control disorder, borderline personality disorder, depression, bipolar disorder, and so on has always been a challenge for experts in the fields of psychology and psychiatry [1], [2]. Since the symptoms of these disorders are very similar, it is usually difficult and time-consuming to diagnose the type of mental disorder. This diagnosis is usually made with the help of psychological tests and a specialized interview with the patient, which can be biased due to factors such as intelligence quotient (IQ), the subject's mood, and the patient's willingness to cooperate [3], [4]. Also, the experience and ability of the doctor has a high impact on the accuracy of the result. Currently, the standard method for diagnosing and distinguishing between these types of patients is to use the DSM-5, which tries to differentiate between mental disorders by setting certain criteria based on the symptoms seen in the patient [5]. Among these disorders, ADHD and bipolar disorder (especially type 2) share similar symptoms, including fast talking, competitive thoughts, less need for sleep, inattention, and high energy that manifests as high physical activity and rapid mood swings [6], [7]. It is difficult for psychiatrists to separate these two patient groups based on clinical observations, at least in the initial interview sessions [8]. The prevalence of bipolar disorder at young ages is low, so this disorder is proposed for children as a secondary diagnosis next to ADHD [9]. A previous study showed that 28.6% of patients with bipolar disorder are misdiagnosed as ADHD [10], [11]. Therefore, providing a reliable and accurate method for diagnosing patients with ADHD and bipolar disorder can provide a useful tool to the psychiatrist to increase the certainty of the doctor's diagnosis in addition to shortening the diagnosis time and starting the treatment faster.

In the meantime, one of the investigated ways to diagnose these disorders is the computational analysis of the electroencephalogram (EEG) signal [12]. EEG has emerged as a valuable instrument in the detection of psychiatric disorders, including bipolar disorder and ADHD [13], [14], [15], [16], [17], [18]. EEG measures the electrical activity of the brain and provides insights into its functioning [19]. This signal contains helpful data regarding the activity of brain cells and cognitive functions, and due to its unique properties, such as high time resolution, low cost, non-invasiveness and easy access, it has been used as a useful tool to diagnose psychiatric disorders [20]. Recent reviews supported the application of machine learning to EEG as an innovative approach to help clinicians diagnose bipolar disorder and ADHD [21], [22]. A study was conducted to compare adolescents and youths with bipolar disorder with patients with ADHD and a control group of individuals without any neurological conditions. The objective of the study was to distinguish between bipolar disorder and ADHD clients based on their VEP features. In order to achieve this, the researchers employed the 1NN technique for classification. Results showed a promising achievement with a classification accuracy of 92.85%, successfully differentiating between bipolar disorder, ADHD, and healthy subjects [23]. Another study focused on using synchronization attributes, specifically phase locking values, to differentiate between patients with bipolar disorder and schizophrenia. By employing a SVM technique, a classification accuracy of 92.45% was attained [24]. Additionally, Sadatnezhad and colleagues investigated EEGs through various nonlinear and linear methods such as auto-regressive coefficients, fractal dimensions, band power, and time-frequency approach for detecting clients with bipolar disorder and ADHD [25]. Their findings showed a classification accuracy of 86.44%. Overall, despite its clinical importance, very few researches have tried to provide an automatic system for the diagnosis of ADHD and bipolar disorder from the EEG.

Steady-state visual evoked potential (SSVEP) is an electrophysiological potential produced by the electrical activity of cerebral cortex which is extracted if a repetitive visual stimulation is delivered to a subject [26]. Previous studies have proven the high diagnostic value of this informative potential in different psychiatric disorders, including schizophrenia [27], ADHD [28] and bipolar disorder [29]. However, none of the previous studies attempted to employ SSVEP to differentiate ADHD and bipolar disorder. Therefore, this research was conducted for the first time with the aim of designing a machine learning system for EEG detection of ADHD, bipolar disorder, and normal using SSVEPs. The rest of this article is organized as follows: Section II presents the proposed methodology including the used EEG dataset, feature extraction, feature selection and classification models. Experimental findings and observation are provided in Section III. Discussion is given in Section IV. Finally, a brief conclusion is presented in Section V.

II. Methods

The designed system to automatically differentiate ADHD and bipolar disorder had different steps, including specific EEG recording protocol to elicit SSVEP, data preprocessing, feature extraction, and classification. In this section, detail of each step is explained.

A. Data Recording

In this study, EEGs were captured from 25 clients with ADHD, 27 clients with bipolar disorder and 30 healthy subjects. Patients were diagnosed by a psychiatrist using DSM-5 diagnostic criteria. None of the participants had a history of head trauma, neurological disorders, and brain stimulation, and all of them were right-handed. Table I shows baseline data of the participants. As shown, there is no significant difference between group of patients and healthy people in terms of gender and age. Study was conducted based on principles of the Declaration of Helsinki (1996) and the current Good Clinical Practice guidelines. An outline of the project was explained to the participants to signing informed consents.

Participants were equipped with a 16-channel EEG net from Electrical Geodesics Inc. During the capturing, participants were seated comfortably in an armchair with their eyes closed, ensuring minimal muscle tension or eye movement. The recording began with a two-minute period of relaxed wakefulness, followed by consecutive two-minute intervals of photic stimuli condition designed to measure SSVEP. To elicit SSVEP responses, a diode photo stimulator from Grass Technologies (model PS33-PLUS) positioned about 80 cm in front of each participant, emitted continuously modulated light stimuli at 15-Hz. Luminance of sinusoidal light emitter ranged from 300 cd/m2 at its lowest point to 800 cd/m2 at its highest point. EEG recordings were digitized and sampled at a rate of 512-Hz. Fig. 1 shows the electrode positions on the scalp. The ground and reference electrodes were also positioned in the Fpz location and the right ear, respectively.

B. Data Analysis

The EEG signals were filtered to retain frequencies ranging from 0.4 Hz to 45 Hz. To eliminate any interference caused by muscle activity and eye movements, a combination of semi-automatic techniques involving amplitude-based threshold detection and visual examination was implemented separately for each channel. This procedure was carried out using MATLAB software. Subsequently, the combined recordings of spontaneous brain activity and SSVEP were analyzed using independent component analysis (ICA), which was performed through EEGLAB plug-in in MATLAB [30]. The purpose of this analysis was to detect and eliminate artifacts related to eye movements, muscle activity, and cardiac activity. Following ICA, the individual EEG channels were visually inspected once again to remove any remaining artifacts. The recordings were then re-referenced to global average.

After removing noise and artifacts from the recorded signals, the steady state responses were extracted. It is assumed that these responses are superimposed on the background EEG as a sine wave at stimulation frequency. An example of the EEG signal recorded during 15 Hz stimulation is shown in Fig. 2. In fact, from the processing point of view, the background EEG can be considered as noise that is added with the sinusoidal response resulting from intermittent stimulation. One of the ways to remove this background EEG is to use the moving averaging method. In this method, a window whose length is an integer multiple of the stimulus period is moving over the signal at intervals of one period and divides it into segments. Then, instead of calculating the average in the trials, the average is calculated on these obtained segments. For the length of the window, five periods were considered so that both the length is long enough and the number of segments obtained is not too small. However, due to the low sampling frequency (512 Hz), it was practically impossible to move the window properly in the original signal. Because the phase difference caused by the difference between actual position of window and its correct location causes the removal or severe weakening of the steady state response. Therefore, before windowing, the signal sampling rate was increased by a factor of 4. Fig. 3 shows the result of applying this method on the signal obtained during visual stimulation.

TABLE I. Demographic Information of the Participants

Variable	ADHD group (n = 25)	Bipolar group (n = 27)	Healthy group (n = 30)	P-value
Age	20.32 ± 1.87	20.97 ± 1.79	21.01 ± 1.56	0.312
Gender	18 male, 7 female	17 male, 10 female	16 male, 14 female	0.384

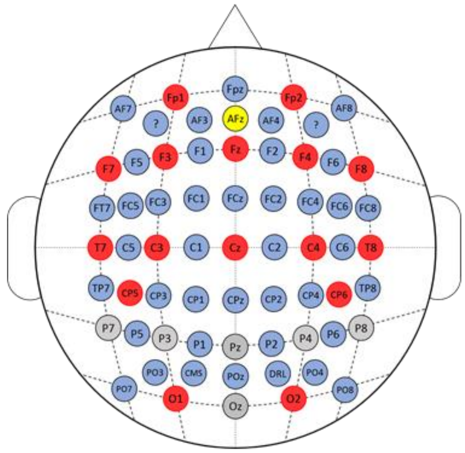

Fig. 1. Electrode placement on the scalp according to 10-20 system (red colored electrodes).

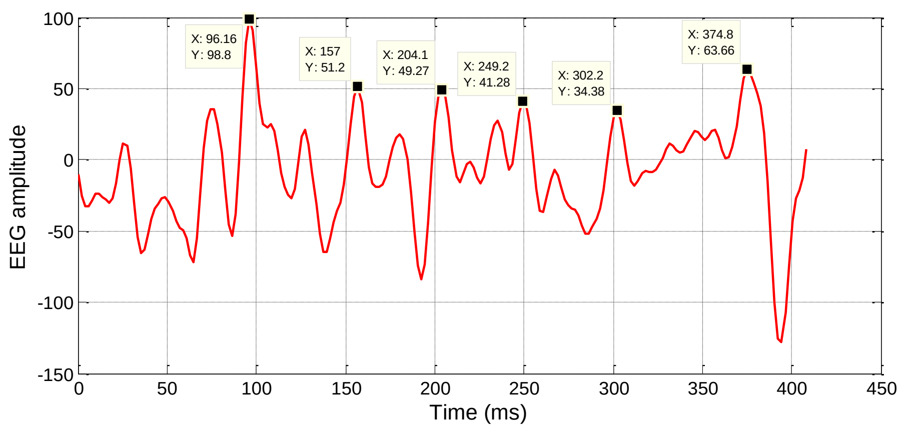

Fig. 2. An example of the recorded EEG signal and the repeating pattern in it (stimulation frequency 15 Hz).

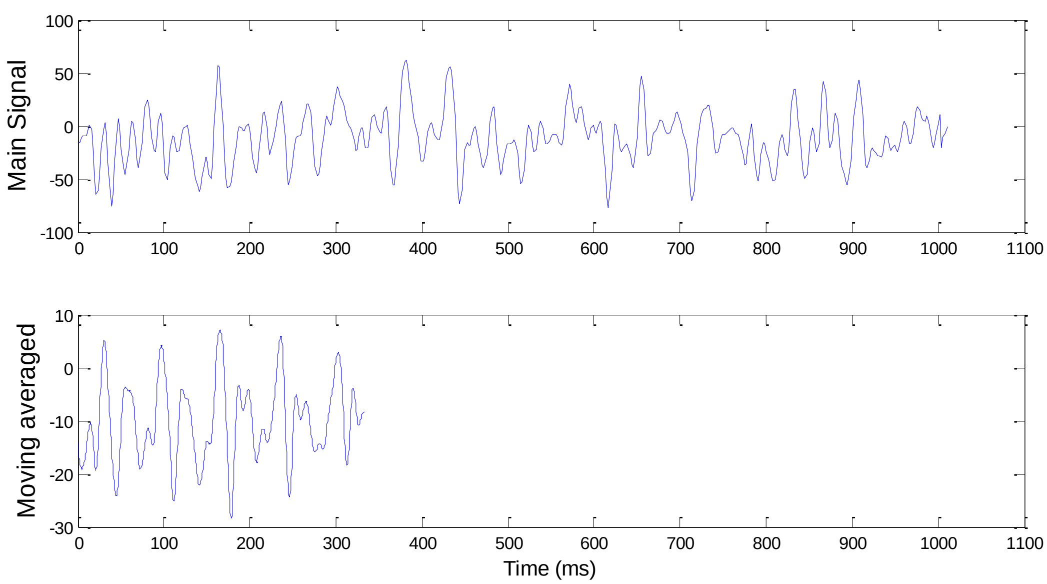

Fig. 3. Effect of moving averaging on SSVEP responses at 15 Hz.

C. Feature Extraction

In this study, feature extraction from SSVEPs was performed in two frequency and time domains. To assess spectral power, Fast Fourier Transform (FFT) was applied using a Welch's periodogram with a Hamming window. This computation resulted in a frequency resolution of 0.25-Hz. Then, amplitude and latency of SSVEPs were extracted as frequency domain features.

Afterward, for processing in the time domain, the nonlinear dynamics of SSVEPs were analyzed. This was performed through recurrence quantification analysis (RQA) in order to extract non-linear features from SSVEPs for input to classifiers. RQA is a powerful technique used in biomedical signal processing to analyze and extract valuable information from complex time series data. It provides a comprehensive approach for studying the dynamics and patterns of recurring events within a signal [31]. RQA focuses on identifying recurrent patterns, or recurrences, by measuring the similarity between different sections of the signal. By quantifying the recurrence properties, RQA enables researchers to investigate important characteristics such as the presence of regularities, irregularities, and deterministic chaos in the signal [32]. This technique plays a crucial role in various biomedical applications, including the analysis of electrocardiogram (ECG) signals, EEG recordings, and other physiological measurements [21]. With its ability to capture intricate temporal relationships and uncover hidden patterns, RQA serves as a valuable tool for understanding and interpreting complex biomedical signals, ultimately contributing to advancements in clinical applications [33]. RQA provides a recurrence plot, which can be analyzed to extract various features. To quantify structures presented in the recurrence plots, we computed and extracted multiple features:

1) Recurrence Rate (RR) = (Number of recurrent points) / (Total number of points)

2) Determinism (DET) = (Number of diagonal line structures) / (Number of recurrent points)

3) Average Diagonal Line Length (L) = (Sum of lengths of all diagonal lines) / (Number of diagonal lines)

4) Entropy (ENT) = -Σ (p × log2(p))

where, p is the probability of finding two recurrent points within a certain distance in the recurrence plot.

5) Trend = (Number of vertical line structures) / (Number of recurrent points)

6) Longest Diagonal Line (Lmax) = Maximum length among all diagonal lines

7) Divergence (DIV) = (Number of horizontal line structures) / (Number of recurrent points)

8) Trapping Time (TT) = (Average length of vertical lines) / (Average length of diagonal lines)

9) Percent Determinism = DET × 100

10) Ratio Determinism = DET / (1 - RR)

11) Average Off-Diagonal Line Length (V) = (Sum of lengths of all off-diagonal lines) / (Number of off-diagonal lines)

12) Laminarity (L) = (Number of vertical lines of length L) / (Total number of recurrent points)

13) Ratio laminarity (RL) = (L) / (RR)

14) Ratio Off-Diagonal lines (RV) = (V) / (RR)

15) Longest vertical line length (Vmax) = maximum length of vertical lines

These features provide insights into different aspects of the recurrence plot, such as the presence of recurrent patterns, diagonal line structures, vertical and horizontal line structures, entropy measures, and more.

D. Feature Selection

In this research, we extracted various features from each of the 16 channels, resulting in a feature matrix of size 16×17 for each individual. Consequently, a total of 272 features were obtained across each participant. However, it was determined that certain features were redundant and did not provide sufficient information to effectively differentiate between the three groups. Furthermore, the classification of these numerous features incurred significant computational expenses and reduced processing speed. To address these issues, we employed the double input symmetrical relevance (DISR) technique to choose optimal features. Implementation of DISR aimed to enhance the classification accuracy while optimizing computational costs. Meyer et al. [34]suggested the following measure for feature selection:

$$F = arg \max_{X_i \in X_S} \left\{ \sum_{X_j \in X_S} \frac{MI(X_{i.j}.Y)}{H(X_{i.j}.Y)} \right\} \tag{1}$$

In Eq. (1), $H(X_{i.j}.Y)$ and $MI(X_{i.j}.Y)$ denote the information entropy and mutual information, respectively.

E. Classification

In this research, we utilized various machine learning techniques to classify the linear and nonlinear features extracted from SSVEPs into three classes of ADHD, bipolar disorder and normal: k-nearest neighbors (KNN), support vector machine (SVM), linear discriminant analysis (LDA) and Naïve Bayes. In the following, each of these classifiers is briefly explained.

1) LDA. It is a supervised classification algorithm that is widely utilized in pattern recognition. LDA is a linear transformation technique that projects the data onto a lower-dimensional space, while maximizing interval between the groups. Purpose of LDA is to search for a linear integration of the input features that increases the ratio of between-group variances to within-group variances. Mathematically, this can be expressed by Eq. (2):

$$\mathrm{J(w)} = \frac{\mathrm{w^T S_B w}}{\mathrm{w^T S_W w}} \tag{2}$$

where, w indicates the weight vector, S_B indicates the between-group scatter matrix, and S_W indicates the within-group matrix. Between-group scatter matrix measures the distance between the means of the different classes, while within-group matrix measures the variance within each class. Optimal weight vector w is determined by solving generalized eigenvalue problem through Eq. (3):

$$\mathrm{S_B w} = \lambda \mathrm{S_W w} \tag{3}$$

where, λ is the eigenvalue associated with w. Once weight vector is determined, classifier can be utilized to classify new data points by projecting them onto the same lower-dimensional space and assigning them to the class with the closest mean.

2) Naïve bayes. It is a probabilistic classification approach widely utilized in machine learning and natural language processing. The algorithm works by Bayes' rule: probability of an assumption given some observed evidences is corresponding to the product of preceding probability of the assumption and likelihood of evidence for that assumption. Mathematically, this can be expressed as:

$$P(y|x_1.x_2 \dots .x_n) = \frac{P(y)P(x_1.x_2 \dots .x_n|y)}{P(x_1.x_2 \dots .x_n)} \quad (4)$$

In Eq. (4), y indicates the group label, $x_1.x_2 \dots .x_n$ indicate input attributes, P(y) is preceding probability of the group, and $P(x_1.x_2 \dots .x_n|y)$ indicates likelihood of evidence given a class. By Eq. (5), the Naïve Bayes classifier makes the hypothesis that input attributes are independent conditionally given a group label that allows the likelihood to be factorized as:

$$P(x_1.x_2 \dots .x_n|y) = \prod_{i=1}^{n} P(x_i|y) \quad (5)$$

This assumption is often called "naive" because it is rarely true in practice, but it simplifies the computation and often leads to good results. The Naive Bayes classifier can be trained by estimating the prior probabilities and the likelihoods from a labeled training set, and then using them to classify new data points by choosing the class with the largest posterior probability.

3) KNN. It is a non-parametric classification algorithm widely utilized in pattern recognition. This technique works by this concept that samples that are close in the feature space are probably to belong to the same group. Given a new data point, KNN finds K closest neighbors in training set and assigns group label that is most common among them. Mathematically, this can be expressed by Eq. (6):

$$\hat{y} = arg \max_{y_i} \sum_{i=1}^{K} [y_i = y] \quad (6)$$

$\hat{y}$ indicates predicted group label, yi indicates group label of i-th closest neighbor, and K indicates the count of neighbors. Interval between samples is typically measured using Euclidean distance defined by Eq. (7):

$$d(x_i.x_j) = \sqrt{\sum_{k=1}^{n} (x_{ik} - x_{jk})^2} \quad (7)$$

where, xik and xjk are the k-th feature value of the i-th and j-th data points, respectively. K may be determined using cross-validation. KNN algorithm is simple and easy to implement, but it may be expensive computationally for huge data and high-dimensional spaces. In this work, K = 3 was considered.

4) SVM. It is a kind of supervised technique utilized for regression and classification analyzes. SVM is especially helpful in cases where data is not distinguishable linearly, meaning that a line may not be drawn to distinguish data into various groups. Instead, SVM utilizes an approach called kernel technique to transform data into a higher dimensional space where it may be linearly separated. SVM then finds a hyperplane that best distinguishes data into different groups while increasing margin that is separation between hyperplane and nearest samples from every group. Eq. (8) defined this hyperplane:

$$W^T x + b = 0 \quad (8)$$

w indicates weight vector, x indicates input vector, and b indicates bias term. SVM allocates a new input vector to one of the two groups according to which side of decision boundary it falls on. SVM wants to search for optimum values of b and w that reduce classification errors while increasing margins. This obtains through solving following optimization problem:

$$minimize\ \frac{1}{2}\|W\|^2\ \ subject\ to\ y_i(W^T x_i + b) \geq 1\ \ for\ all\ i \quad (9)$$

In Eq. (9), ||w|| indicates Euclidean norm of weight vector, y_i indicates group label of i-th sample, and x_i indicates i-th input vector. Optimization problem may be solved through quadratic programming approaches. In this study, both linear and radial basis function (RBF) kernels were used to classify SSVEP features by SVM.

III. Results

After preprocessing, all mentioned features were calculated from SSVEPs in three groups. Fig. 4 shows an example of recurrence plots estimated from SSVEPs in an ADHD patient, a client with bipolar disorder, and a normal subject. After extracting SSVEP features by spectral analysis and RQA, as mentioned before, DISR technique was utilized to decrease feature space. Afterward, leave-one-subject-out (LOSO) technique was used to evaluate classification performance of various classifiers. This technique is a widely used cross-validation method in machine learning that involves leaving out one subject at a time from the training set to evaluate the performance of a model. This technique is particularly useful in studies with small sample sizes or highly variable data across subjects. By leaving out one subject at a time, the LOSO technique can help identify which subjects are most important for the model's performance and which ones may be less relevant. The LOSO technique may be utilized for a range of machine learning algorithms, such as neural networks, SVMs, and decision trees. Although computationally expensive, the LOSO technique remains a valuable tool for evaluating the performance of machine learning models and improving their generalization to new data. In addition, in the study, accuracy, sensitivity and specificity measures were utilized to report the classification performance of the classifiers.

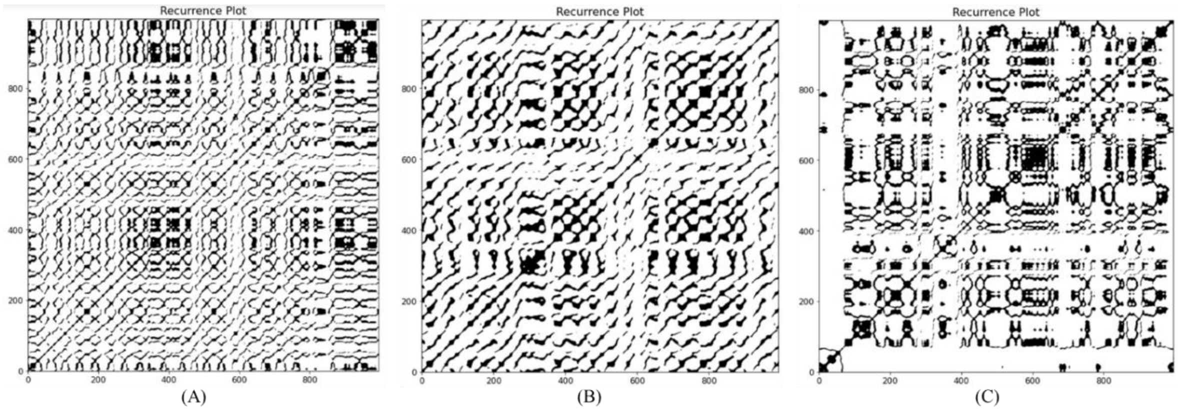

Fig. 4. An example of recurrence plots estimated from SSVEPs in (A) a healthy subject, (B) a ADHD patient, and (C) a patient with bipolar disorder.

The obtained results of the LOSO cross-validation algorithm is shown in Table II. As shown, best classification result was achieved using the selected features and SVM with linear kernel with accuracy of 78.57%, sensitivity of 79.15% and specificity of 77.94%. Naïve Bayes classifier also yielded a good accuracy of 76.20% for EEG classification of ADHD, bipolar disorder and normal groups. Furthermore, Fig. 5 shows how the accuracy percentage of the output changes with respect to the changes of the dimension of the feature space. As can be seen, it is not possible to simply determine the dimensions that are optimal for almost all classifiers. However, dimension 6 seems to be suitable for most classifiers except SVM-RBF.

TABLE II. CLASSIFICATION RESULTS FOR EEG CLASSIFICATION OF ADHD, BIPOLAR DISORDER AND NORMAL GROUPS THROUGH VARIOUS CLASSIFIERS AND SSVEP FEATURES

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)
SVM-RBF	73.81	70.30	74.08
SVM-Linear	78.57	79.15	77.94
KNN	73.81	71.29	74.55
LDA	76.19	72.41	79.88
Naïve Bayes	76.20	71.37	80.36

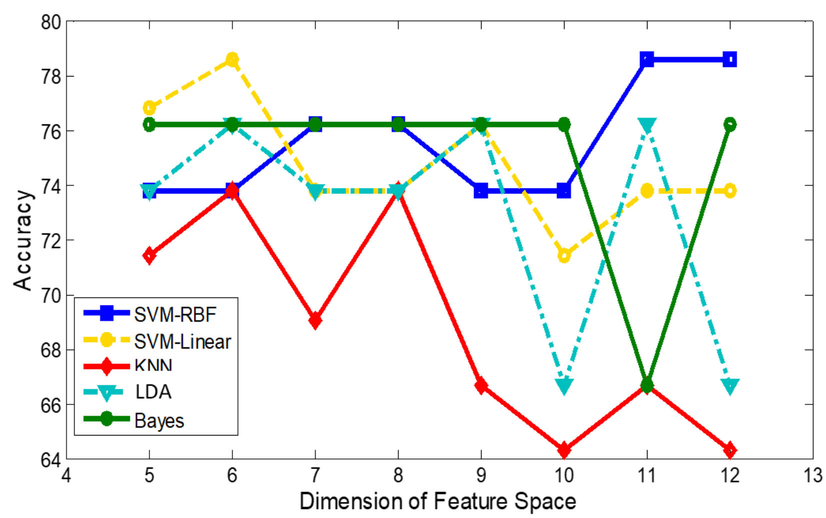

Fig. 5. Changes in output accuracy percentage versus feature dimension changes.

IV. DISCUSSION

There are psychiatric illnesses that share clinical symptoms and signs. Bipolar disorder and ADHD are two prevalent disorders whose symptoms are similar. In order to reduce the misdiagnosis between bipolar disorder and ADHD, a machine learning-based system using EEG and steady state potentials (i.e., SSVEP) was evaluated to classify ADHD, bipolar disorder and normal conditions. For this purpose, both linear and nonlinear dynamics of extracted SSVEPs were analyzed. Indeed, after data preprocessing, spectral analysis and RQA were applied to SSVEPs. Then, feature selection was utilized through the DISR. The DISR feature selection method offers several advantages. Firstly, DISR effectively identifies informative and discriminative features, eliminating redundant and uninformative ones. By doing so, it enhances the classification performance, resulting in more accurate and reliable results. Additionally, DISR optimizes the computational cost by reducing the number of features, leading to improved processing speed. Finally, SVM classifier with linear kernel yielded an accuracy of 78.57% for ADHD, bipolar disorder and normal classification through the LOSO cross-validation. While SSVEPs have many advantages, they also have limitations and shortcomings that make them unsuitable for the problem at hand. SSVEPs can sometimes suffer from low signal-to-noise ratio, especially in noisy environments. This can make it challenging to extract meaningful information from the recorded brain signals. SSVEPs are less effective in localizing brain activity compared to techniques like fMRI or EEG because they provide less spatial resolution. This means that identifying the exact brain region generating the response can be challenging. SSVEP responses can vary significantly between individuals, making it necessary to calibrate the system individually for each user. Prolonged exposure to flickering lights or screens, which are typically used to evoke SSVEPs, can lead to user fatigue or discomfort. This limits the practicality of using SSVEPs in long-term or everyday applications. Moreover, SSVEP-based systems are typically limited in the amount of information that can be reliably extracted from brain signals. This can restrict the complexity and richness of interactions that can be achieved using SSVEP interfaces.

The proposed system is less accurate compared to previous similar studies, where Nazhvani et al., Alimardani et al., and Sadatnezhad et al. reported accuracies higher than 84% for EEG classification of ADHD and bipolar disorder [23], [24], [25]. However, it should be noted that there are very few studies in the literature that have used EEG analysis to differentially diagnose these two disorders. The comparison of the present research with previous studies shows that the analysis of the resting-state or ongoing EEG signal can be a better solution for the differential diagnosis of ADHD and bipolar disorder compared to the SSVEP analysis. It should be noted that our motivation for analyzing steady-state potentials to distinguish these two disorders was previous EEG studies that reported significant differences between ADHD and bipolar disorder in terms of various EEG indices during cognitive performance [35]. Rommel et al. found that Absolute theta power may play a role as a marker of neurobiological processes in ADHD and bipolar disorder during a cued continuous performance task [36]. Furthermore, Michelini et al. reported less regulation of beta suppression in ADHD than in bipolar disorder during a cognitive task by analyzing event-related potentials (ERPs) [37]. Passarotti et al showed significant differences in the prefrontal cortex of children with ADHD and bipolar disorder during an emotional valence Stroop task [38]. However, considering that none of the previous studies have used SSVEPs as biological data to be processed, we cannot make precise comparisons between the suggested system and previous techniques. Although this research is the first to evaluate the application of signal processing and machine learning methods in SSVEP classification of these disorders, it has limitations that future studies should investigate to enhance performance of the proposed system. First, presented visual stimulation was delivered to the participants without the presence of a cognitive task, whereas previous studies often use cognitive tasks during this type of stimulation. Second, in the present study, only the stimulation frequency of 15 Hz was investigated, and other stimulation frequencies need to be tested in the future. Finally, other linear and non-linear signal processing methods should be evaluated in future studies. In addition to high comorbidity of ADHD and bipolar disorder, the closeness of the EEG patterns of the two disorders was observed in this research. Therefore, in future studies, it is better to use soft labeling methods to classify these two groups, which do not necessarily classify each subject as belonging to one group.

V. Conclusion

A new EEG classification scheme based on SSVEPs and machine learning techniques was presented in this work to distinguish ADHD from bipolar disorder. This framework exploited the linear and nonlinear properties of these cortical potentials and was tested on real-world EEG datasets from patients with ADHD and bipolar disorder. Valid performance evaluation criteria were calculated, which proved the acceptable performance of the proposed framework. However, external validation of such a framework is needed in future studies.

Acknowledgment

2023 Department of Education Philosophy and Social Science Research Special Task Project (University Student Work Brand). "Heart Chamber, Heart Matters, Heart Fitness" Growth Classroom - Building an Expressive Art Group Counseling Approach Using Internet Buzzwords as Entry Points. Project No: 22Z422.

References

[1] Duffy, G. S. Malhi, and G. A. Carlson, “The challenge of psychiatric diagnosis: Looking beyond the symptoms to the company that they keep.,” 2018.

[2] M. R. Mohammadi, A. Khaleghi, K. Shahi, and H. Zarafshan, “attention deficit hyperactivity disorder: Behavioral or Neuro-developmental Disorder? Testing the HiTOP Framework Using Machine Learning Methods,” Journal of Iranian Medical Council, vol. 6, no. 4, pp. 652–657, 2023.

[3] M. R. Mohammadi and A. Khaleghi, “Transsexualism: A different viewpoint to brain changes,” Clinical Psychopharmacology and Neuroscience, vol. 16, no. 2, p. 136, 2018.

[4] K. Allsopp, J. Read, R. Corcoran, and P. Kinderman, “Heterogeneity in psychiatric diagnostic classification,” Psychiatry Res, vol. 279, pp. 15–22, 2019.

[5] A. Khaleghi et al., “Epidemiology of psychiatric disorders in children and adolescents; in Tehran, 2017,” Asian J Psychiatr, vol. 37, pp. 146–153, 2018.

[6] A. Khaleghi, A. Sheikhani, M. R. Mohammadi, and A. M. Nasrabadi, “Evaluation of cerebral cortex function in clients with bipolar mood disorder I (BMD I) compared with BMD II using QEEG analysis,” Iran J Psychiatry, vol. 10, no. 2, p. 93, 2015.

[7] A. Khaleghi, P. M. Birgani, M. F. Fooladi, and M. R. Mohammadi, “Applicable features of electroencephalogram for ADHD diagnosis,” Research on Biomedical Engineering, vol. 36, pp. 1–11, 2020.

[8] M. R. Mohammadi, A. Khaleghi, A. M. Nasrabadi, S. Rafieivand, M. Begol, and H. Zarafshan, “EEG classification of ADHD and normal children using non-linear features and neural network,” Biomed Eng Lett, vol. 6, pp. 66–73, 2016.

[9] M. R. Mohammadi et al., “Prevalence and Socio-Demographic Factors of Bipolar Mood Disorders in Children and Adolescents: Identifying the Principal Predictors,” Iranian Rehabilitation Journal, vol. 20, no. 2, pp. 149–160, 2022.

[10] M.-R. Mohammadi et al., “Prevalence of ADHD and its comorbidities in a population-based sample,” J Atten Disord, vol. 25, no. 8, pp. 1058–1067, 2021.

[11] M. R. Mohammadi et al., “Prevalence and correlates of psychiatric disorders in a national survey of Iranian children and adolescents,” Iran J Psychiatry, vol. 14, no. 1, p. 1, 2019.

[12] A. Khaleghi, M. R. Mohammadi, K. Shahi, and A. M. Nasrabadi, “Computational neuroscience approach to psychiatry: a review on theory-driven approaches,” Clinical Psychopharmacology and Neuroscience, vol. 20, no. 1, p. 26, 2022.

[13] A. Afzali, A. Khaleghi, B. Hatef, R. Akbari Movahed, and G. Pirzad Jahromi, “Automated major depressive disorder diagnosis using a dual-input deep learning model and image generation from EEG signals,” Waves in Random and Complex Media, pp. 1–16, 2023.

[14] A. Khaleghi, M. R. Mohammadi, M. Moeini, H. Zarafshan, and M. Fadaei Fooladi, “Abnormalities of alpha activity in frontocentral region of the brain as a biomarker to diagnose adolescents with bipolar disorder,” Clin EEG Neurosci, vol. 50, no. 5, pp. 311–318, 2019.

[15] A. Khaleghi et al., “EEG classification of adolescents with type I and type II of bipolar disorder,” Australas Phys Eng Sci Med, vol. 38, pp. 551–559, 2015.

[16] M. Moeini, A. Khaleghi, N. Amiri, and Z. Niknam, “Quantitative electroencephalogram (QEEG) spectrum analysis of patients with schizoaffective disorder compared to normal subjects,” Iran J Psychiatry, vol. 9, no. 4, p. 216, 2014.

[17] M. Moeini, A. Khaleghi, and M. R. Mohammadi, “Characteristics of alpha band frequency in adolescents with bipolar II disorder: a resting-state QEEG study,” Iran J Psychiatry, vol. 10, no. 1, p. 8, 2015.

[18] M. Moeini, A. Khaleghi, M. R. Mohammadi, H. Zarafshan, R. L. Fazio, and H. Majidi, “Cortical alpha activity in schizoaffective patients,” Iran J Psychiatry, vol. 12, no. 1, p. 1, 2017.

[19] A. Khaleghi, M. R. Mohammadi, K. Shahi, and A. Motie Nasrabadi, “A neuronal population model based on cellular automata to simulate the electrical waves of the brain,” Waves in Random and Complex Media, pp. 1–20, 2021.

[20] A. Khaleghi, M. R. Mohammadi, K. Shahi, and A. M. Nasrabadi, “Possible Neuropathological Mechanisms Underlying the Increased Complexity of Brain Electrical Activity in Schizophrenia: A Computational Study,” Iran J Psychiatry, 2023.

[21] W. A. Campos-Ugaz, J. P. P. Garay, O. Rivera-Lozada, M. A. A. Diaz, D. Fuster-Guillén, and A. A. T. Arana, “An Overview of Bipolar Disorder Diagnosis Using Machine Learning Approaches: Clinical Opportunities and Challenges,” Iran J Psychiatry, 2023.

[22] H. Zarafshan, A. Khaleghi, M. R. Mohammadi, M. Moeini, and N. Malmir, “Electroencephalogram complexity analysis in children with attention-deficit/hyperactivity disorder during a visual cognitive task,” J Clin Exp Neuropsychol, vol. 38, no. 3, pp. 361–369, 2016.

[23] A. D. Nazhvani, R. Boostani, S. Afrasiabi, and K. Sadatnezhad, “Classification of ADHD and BMD patients using visual evoked potential,” Clin Neurol Neurosurg, vol. 115, no. 11, pp. 2329–2335, 2013.

[24] F. Alimardani, R. Boostani, M. Azadehdel, A. Ghanizadeh, and K. Rastegar, “Presenting a new search strategy to select synchronization values for classifying bipolar mood disorders from schizophrenic patients,” Eng Appl Artif Intell, vol. 26, no. 2, pp. 913–923, 2013.

[25] K. Sadatnezhad, R. Boostani, and A. Ghanizadeh, “Classification of BMD and ADHD patients using their EEG signals,” Expert Syst Appl, vol. 38, no. 3, pp. 1956–1963, 2011.

[26] S.-A. Mostafavi, A. Khaleghi, S. R. Vand, S. S. Alavi, and M. R. Mohammadi, “Neuro-cognitive Ramifications of Fasting and Feeding in Obese and Non-obese Cases,” Clinical Psychopharmacology and Neuroscience, vol. 16, no. 4, p. 481, 2018.

[27] A. Schielke and B. Krekelberg, “Steady state visual evoked potentials in schizophrenia: A review,” Front Neurosci, vol. 16, p. 988077, 2022.

[28] A. Khaleghi, H. Zarafshan, and M. R. Mohammadi, “Visual and auditory steady-state responses in attention-deficit/hyperactivity disorder,” Eur Arch Psychiatry Clin Neurosci, vol. 269, pp. 645–655, 2019.

[29] W. Xiao, G. Manyi, and A. Khaleghi, “Deficits in auditory and visual steady-state responses in adolescents with bipolar disorder,” J Psychiatr Res, vol. 151, pp. 368–376, 2022.

[30] M. R. Goldstein, M. J. Peterson, J. L. Sanguinetti, G. Tononi, and F. Ferrarelli, “Topographic deficits in alpha-range resting EEG activity and steady state visual evoked responses in schizophrenia,” Schizophr Res, vol. 168, no. 1–2, pp. 145–152, 2015.

[31] C. L. Webber and N. Marwan, “Recurrence quantification analysis,” Theory and Best Practices, p. 426, 2015.

[32] M. Niknazar, S. R. Mousavi, B. V. Vahdat, and M. Sayyah, “A new framework based on recurrence quantification analysis for epileptic seizure detection,” IEEE J Biomed Health Inform, vol. 17, no. 3, pp. 572–578, 2013.

[33] A. Khaleghi, M. R. Mohammadi, G. P. Jahromi, and H. Zarafshan, “New ways to manage pandemics: using technologies in the era of COVID-19: a narrative review,” Iran J Psychiatry, vol. 15, no. 3, p. 236, 2020.

[34] P. E. Meyer and G. Bontempi, “On the use of variable complementarity for feature selection in cancer classification,” in Workshops on applications of evolutionary computation, Springer, 2006, pp. 91–102.

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

[36] A.-S. Rommel et al., "Commonalities in EEG spectral power abnormalities between women with ADHD and women with bipolar disorder during rest and cognitive performance," Brain topography, vol. 29, pp. 856-866, 2016.

[37] G. Michelini et al., "Shared and disorder-specific event-related brain oscillatory markers of attentional dysfunction in ADHD and bipolar disorder," Brain topography, vol. 31, pp. 672-689, 2018.

[38] A. M. Passarotti, J. A. Sweeney, and M. N. Pavuluri, "Differential engagement of cognitive and affective neural systems in pediatric bipolar disorder and attention deficit hyperactivity disorder," Journal of the International Neuropsychological Society, vol. 16, no. 1, pp. 106-117, 2010.