

Original Research

# Utilizing Phase Locking Value to Determine Neurofeedback Treatment Responsiveness in Attention Deficit Hyperactivity Disorder

Mohammad Reza Yousefi<sup>1,2</sup>, Nikoo Khanahmadi<sup>1,2</sup>, Amin Dehghani<sup>3,\*</sup><sup>1</sup>Department of Electrical Engineering - Najafabad Branch, Islamic Azad University, 8514143131 Najafabad, Iran<sup>2</sup>Smart Microgrid Research Center - Najafabad Branch, Islamic Azad University, 8514143131 Najafabad, Iran<sup>3</sup>Department of Psychological and Brain Sciences, Dartmouth College, Hanover, Grafton County, NH 03755, USA\*Correspondence: [Amin.Dehghani@dartmouth.edu](mailto:Amin.Dehghani@dartmouth.edu) (Amin Dehghani)

Academic Editor: Gernot Riedel

Submitted: 22 December 2023 Revised: 13 March 2024 Accepted: 17 April 2024 Published: 21 June 2024

## Abstract

**Background:** Neurofeedback is a non-invasive brain training technique used to enhance and treat hyperactivity disorder by altering the patterns of brain activity. Nonetheless, the extent of enhancement by neurofeedback varies among individuals/patients and many of them are irresponsive to this treatment technique. Therefore, several studies have been conducted to predict the effectiveness of neurofeedback training including the theta/beta protocol with a specific emphasize on slow cortical potential (SCP) before initiating treatment, as well as examining SCP criteria according to age and sex criteria in diverse populations. While some of these studies failed to make accurate predictions, others have demonstrated low success rates. This study explores functional connections within various brain lobes across different frequency bands of electroencephalogram (EEG) signals and the value of phase locking is used to predict the potential effectiveness of neurofeedback treatment before its initiation. **Methods:** This study utilized EEG data from the Mendelian database. In this database, EEG signals were recorded during neurofeedback sessions involving 60 hyperactive students aged 7–14 years, irrespective of sex. These students were categorized into treatable and non-treatable. The proposed method includes a five-step algorithm. Initially, the data underwent preprocessing to reduce noise using a multi-stage filtering process. The second step involved extracting alpha and beta frequency bands from the preprocessed EEG signals, with a particular emphasis on the EEG recorded from sessions 10 to 20 of neurofeedback therapy. In the third step, the method assessed the disparity in brain signals between the two groups by evaluating functional relationships in different brain lobes using the phase lock value, a crucial data characteristic. The fourth step focused on reducing the feature space and identifying the most effective and optimal electrodes for neurofeedback treatment. Two methods, the probability index ( $p$ -value) via a  $t$ -test and the genetic algorithm, were employed. These methods showed that the optimal electrodes were in the frontal lobe and central cerebral cortex, notably channels C3, FZ, F4, CZ, C4, and F3, as they exhibited significant differences between the two groups. Finally, in the fifth step, machine learning classifiers were applied, and the results were combined to generate treatable and non-treatable labels for each dataset. **Results:** Among the classifiers, the support vector machine and the boosting method demonstrated the highest accuracy when combined. Consequently, the proposed algorithm successfully predicted the treatability of individuals with hyperactivity in a short time and with limited data, achieving an accuracy of 90.6% in the neurofeedback method. Additionally, it effectively identified key electrodes in neurofeedback treatment, reducing their number from 32 to 6. **Conclusions:** This study introduces an algorithm with a 90.6% accuracy for predicting neurofeedback treatment outcomes in hyperactivity disorder, significantly enhancing treatment efficiency by identifying optimal electrodes and reducing their number from 32 to 6. The proposed method enables the prediction of patient responsiveness to neurofeedback therapy without the need for numerous sessions, thus conserving time and financial resources.

**Keywords:** neurofeedback; hyperactivity disorder; electroencephalogram; Mendelian database; phase lock value; genetic algorithm; support vector machine

## 1. Introduction

Attention deficit hyperactivity disorder (ADHD) is a behavioral disorder primarily observed in children [1,2]. This disorder can be identified by parents or educators who notice atypical behaviors such as restlessness, challenges in sustaining concentration and attention, and heightened impulsivity, commonly associated with attention deficit disorder (ADD), particularly in educational settings where children may display unusual behavior or face difficulties with certain academic subjects. Alternatively, a therapist may

diagnose this disorder through psychological evaluations, with early intervention in childhood often leading to more successful management. Statistical data reveal that ADHD impacts both sexes, albeit with a prevalence rate three times higher in males compared with females. The incidence of this disorder is estimated at 2.7% among those under 18 years of age and 4.3% among adults [3,4]. This psychological disorder presents with diverse symptoms across individuals, with patient behavior varying according to the condition's severity. Generally, there are three primary subtypes of attention deficit hyperactivity disorder, namely (a)



ADD, (b) hyperactivity and impulsivity, and (c) combined presentation (which includes symptoms of ADD, hyperactivity, and impulsivity) [5].

The temporal lobe, integral for auditory-language functions, often shows abnormalities in children diagnosed with ADHD. Moreover, the parietal lobe is notably significant in ADHD, given its association with the disorder. Regarding visual processing, the occipital lobe's importance is underscored, with studies on children with ADHD revealing a 9% reduction in the volume of both white and gray matter in the left posterior occipital region [6]. ADHD affects the limbic area, marking it as one of the regions impacted in the brain. Additionally, research on brain electrical activity has shown reduced activity in the frontal lobe to specific visual stimuli among individuals with this disorder [6–11]. Several studies have consistently identified several key anatomical differences in the brains of individuals with ADHD compared with those in healthy people. Among the most significant findings are reduced overall brain volume and specific alterations in areas related to executive functions, impulse control, and emotional regulation. Notably, the prefrontal cortex shows reduced volume or cortical thickness in ADHD [12–15]. Other reported differences include changes in the size or shape of the corpus callosum, which facilitates communication between the brain's hemispheres, and potential alterations in limbic system structures like the amygdala and hippocampus [16–18].

Previous research indicates that cellular irregularities in fetal development are typically observed between the fifth and seventh months, suggesting that learning disorders may originate from minor brain abnormalities present from birth. Among the regions impacted by these abnormalities is the brain's temporal surface, which plays a role in language functions and is typically more developed in the left hemisphere in individuals without these defects. However, individuals with this disorder frequently display symmetrical sizes in both hemispheres [7]. Furthermore, electroencephalogram (EEG) activity patterns in children with this disorder are unconventional, marked by an increased occurrence of short amplitude brain signals in the frontal brain regions, indicating delayed development in the brain structures vital for attention and information organization. Additionally, several brain regions across the left hemisphere, anterior frontal lobe, parietal-temporal, and occipital-temporal regions show reduced activity in these children. Specific auditory centers for recognizing certain sounds are absent in these children, leading to partial word comprehension. While the exact cause of hyperactivity remains unclear, genetic factors are recognized as a primary contributor to this disorder, as is the case in many other conditions. Diagnosing this disorder presents challenges, as children often display behaviors such as restlessness, inattention, or hyperactivity, which can be exacerbated by stress and anxiety [19]. A confirmed diagnosis of ADHD mandates the persistent presence of symptoms for a mini-

mum of 6 months. Diagnosing ADHD in children under 6 years old poses challenges due to their ongoing growth and developmental changes. The criteria for an ADHD diagnosis include the manifestation of symptoms across multiple environments, a broad spectrum of behavioral symptoms, the emergence of symptoms before the age of 7 years, and the duration of symptoms extending beyond 6 months [19].

Numerous techniques and strategies have been used to treat this disorder, among which neurofeedback is emerging as a highly promising treatment approach. Neurofeedback is a non-invasive brain training technique; in this technique the brain activity and functions are fed back to the patient/participant, aiming to direct changes in a specific direction [20–36]. Neurofeedback typically utilizes auditory or visual cues to which the patient reacts, receiving either positive or negative feedback. This technique aids the brain in adjusting its activities and functions by leveraging insights from brainwave patterns. Essentially, it communicates the brain's real-time state during an activity, enabling the patient to learn to regulate their brainwaves for improved task performance [22,23,25,37–40]. In the EEG neurofeedback, this feedback loop operates solely through monitoring brainwave patterns, without introducing any electrical currents into the brain. The electrical activity of the brain is simply transmitted to a computer for analysis [41].

Previous studies have demonstrated that prompt brain stimulation can enhance synaptic development and inhibit synaptic degeneration. In contrast, the absence of stimulation results in a progressive decrease in neural functions, hindered neuronal development, and ultimately, synaptic atrophy. This process can lead to the emergence of diverse neurological disorders [42]. Consequently, neurofeedback emerges as a training approach for improving attention and concentration. By training the brain, it helps to prevent early synaptic deterioration, thus averting associated neurological complications [42]. Neurofeedback involves a training process that enables the brain to self-regulate. Through repeated practice across multiple sessions, the brain becomes adept at adopting new wave patterns for everyday functions. A critical element of neurofeedback therapy is its focus on altering the intensity and amplitude of brainwaves, rather than their frequency. For instance, if the therapist observes an excessive alpha wave frequency in the frontal area, the goal is to reduce the intensity of this frequency. For instance, alpha wave intensity might decrease from 12 microvolts to 7 millivolts. To ensure optimal treatment outcomes and maintain the effects, a sufficient number of sessions must be undertaken. A standard treatment regimen includes 40 to 45 sessions, each lasting between 30 and 45 minutes, though the exact duration can vary depending on the individual. Sessions are usually held 2 to 3 times per week. It is critical to acknowledge that the success of neurofeedback treatment is not immediate and must be evaluated over several sessions. The assessment of treatment efficacy

requires completion of numerous sessions within the designated 30 to 45-session framework. Through monitoring an individual's brain signals, improvements can be measured against the initial baseline established at the start of training. If noticeable improvement is lacking, it may be advisable to cease the treatment. Brainwave frequencies, including delta (less than 4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–100 Hz) bands, reflect different physiological states and functions in the brain.

### 1.1 Reviews of Relevant Literature

As a broad overview, there are seven distinct categories of neurofeedback methods utilized for the treatment of various disorders including frequency/power neurofeedback, slow cortical potential neurofeedback (SCP-NF), low-energy neurofeedback system (LENS), hem encephalographic neurofeedback (HEG), life Z source (LZS) neurofeedback, low resolution electromagnetic tomography (LORETA), and functional magnetic resonance imaging (fMRI).

Neurofeedback protocols predominantly focus on manipulating specific brainwave frequencies—alpha, beta, delta, theta, and gamma—either in isolation or through combinations like the alpha/beta ratio, alpha/theta, and beta/theta, among others. However, the protocols most commonly utilized involve alpha, beta, and theta waves, and particularly the alpha/beta ratio [43].

The aim of neurofeedback training is to correct irregular EEG patterns, with the goal of improving individuals' cognitive and behavioral performance. An appropriately tailored neurofeedback protocol can effectively adjust imbalances in the alpha and beta brainwave bands, aligning them closer to typical levels [44,45]. In various perceptual-cognitive states, information transmission relies on the oscillations of brain neurons. The analysis of these oscillations and the interplay between different brain regions can yield valuable insights into how the brain responds under diverse circumstances.

Previous studies have shown that in diagnosing and treating hyperactivity with neurofeedback, individuals often participate in several training sessions to assess their responsiveness or lack thereof to treatment. The results of these studies have demonstrated varying degrees of precision in their findings [6,46].

Neurofeedback training essentially follows two primary directions: one focuses on lower frequencies (alpha or theta) to promote relaxation, while the other centers on higher frequencies (low beta, beta, and gamma) to enhance activation, organization, and the ability to resist distractions [47]. In the first method, participants typically close their eyes, while in the second method, they keep their eyes open. Generally, the first method is less suitable for children, whereas both children and adults can undergo the second method.

Neurofeedback has been successfully applied in treating a variety of conditions and mental health disorders. Research indicates that individuals with ADHD show slower brainwave activity in the theta range and decreased beta activity compared with those without ADHD. In treating ADHD, the aim is to lower theta band brain activity and boost beta band activity (or reduce the theta/beta ratio) at a designated electrode site [48]. This treatment has demonstrated efficacy in diminishing hyperactivity, bolstering concentration, improving academic outcomes, raising parental satisfaction regarding their children's behavior, and augmenting indicators of sustained attention [49].

Based on studies conducted to predict the treatability of hyperactivity patients through neurofeedback, various methods have been explored, as outlined in [50]. In the first approach, researchers investigated the feasibility of employing brain electromagnetic tomography to identify treatable conditions before initiating neurofeedback treatment.

In another study [51], a combined approach involving theta/beta training was used with a particular focus on the slow potentials of the cerebral cortex during sessions 10–28. Additionally, other studies delved into the method of assessing slow potentials of the cerebral cortex, considering age and sex criteria, particularly within the initial six training sessions [52,53].

### 1.2 Main Works and Novelty

As previously mentioned, numerous studies have delved into the application of neurofeedback as a therapeutic approach for managing hyperactivity. However, accurately forecasting the responsiveness of hyperactivity patients to neurofeedback treatment remains a challenging endeavor, lacking an effective method with a notable level of precision.

The primary aim of this article is to increase the accuracy in predicting treatment responsiveness through neurofeedback, surpassing existing methods before initiating therapy. This is achieved by quantifying the communication between various brain lobes. Additionally, the application of a genetic algorithm optimizes the feature selection process, effectively narrowing down the feature space. This innovation leads to the use of a more efficient electrode setup in neurofeedback treatments, significantly reducing the number of required brain channels from 32 to just 6. The genetic algorithm serves a crucial role in enhancing classification accuracy and identifying the optimal features. Each feature within the selected space is binary-coded, with '1' signifying its presence (indicating treatability) and '0' indicating its absence (indicating intractability).

The principal contributions of this study can be summarized as follows:

(1) Improving the accuracy of outcomes relative to previous methods.

(2) Identifying the influential brain lobes in neurofeedback treatment through the application of the phase locking value and statistical index.

(3) Employing a genetic algorithm to reduce the feature space and, subsequently, pinpointing the most effective EEG channels for neurofeedback treatment.

(4) Streamlining the number of brain signal recording channels from 32 to 6, focusing on the most impactful ones.

(5) Pioneering the prediction of treatability or non-treatability in children with hyperactivity based on an analysis of the alpha and beta frequency bands of the EEG signal, as well as the assessment of functional connectivity within the brain.

### 1.3 Main Structure of This Paper

The rest of this article is structured as follows: the second section details the proposed methodology, breaking it down into six pivotal stages including preprocessing, processing, feature extraction, reduction of the feature space, classification, and ultimately assessing treatability or non-treatability. The third section focuses on materials and methods, elaborating on participant characteristics, preprocessing techniques, and data processing strategies. The fourth section conducts a detailed analysis of the results, utilizing a variety of metrics to interpret the findings. The article concludes with a fifth section that provides conclusions and outlines recommendations for future research.

## 2. Proposed Method

In this section, we describe the six fundamental steps of our proposed method, as illustrated in Fig. 1. These steps include data collection, preprocessing, processing, feature extraction, feature space reduction, and classification. At the conclusion of this section, we present the neurofeedback results and accuracy values, which validate the efficacy of our proposed approach.

In the initial step, we utilized data obtained from brain signal recordings during 10–20 sessions of neurofeedback over the treatment period. The data were preprocessed previously; however, to enhance data quality, we meticulously removed signal noise arising from artifacts.

The second step involved extracting the alpha and beta frequency bands from the brain signal using a Butterworth pass filter. In the third step, we focused on feature extraction, specifically identifying functional connectivity via the phase lock value. This method not only detects the presence of connections between different brain lobes but also quantifies their strength using a statistical index derived from the  $t$ -test. This process helped to select the optimal channels.

Subsequently, the fourth step revolved around the selection of the most suitable electrodes, accomplished through the implementation of a genetic algorithm. This optimization technique efficiently reduced the number of electrodes from 32 to a more manageable 6. The electrodes identified by the genetic algorithm were then com-

pared with those selected by the statistical test, resulting in the identification of 6 common electrodes deemed optimal.

The feature matrix, a key component, was generated in the fifth step. This matrix comprised 1200 rows (representing the number of test repetitions) and 36 columns (representing the number of features).

In the final step, classification, a combination of different classification methods was employed. The feature matrix from the previous step served as input for the classifiers. The output was a label matrix consisting of 1200 rows (reflecting the number of test repetitions) and one column (representing the number of labels). These labels distinguished between individuals as either treatable or incurable. To assess the effectiveness of our classifiers, we employed accuracy indices, and the classifiers were evaluated using the k-Fold cross-validation method, with ‘k’ set to 5.

## 3. Materials and Methods

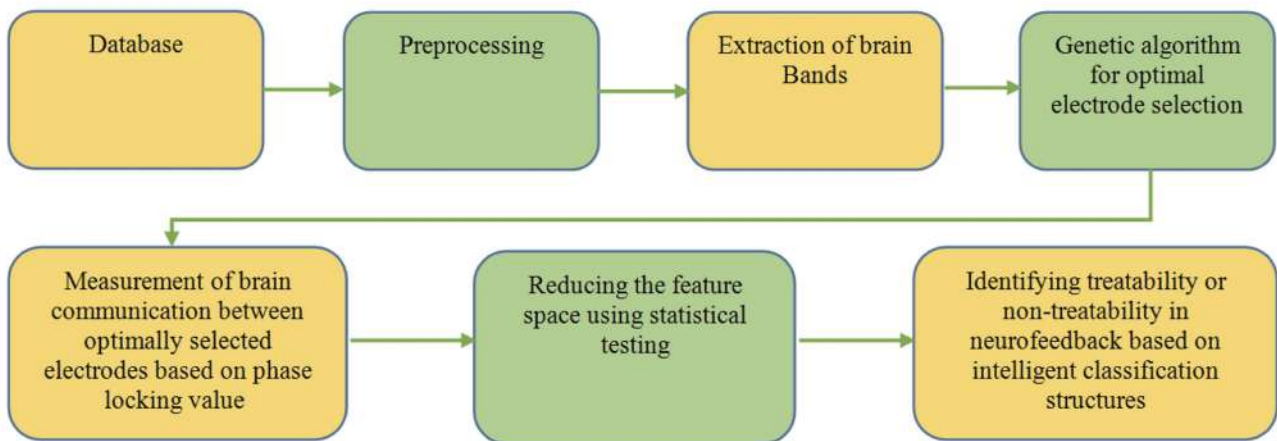
In this section, we discuss the participant characteristics, as well as the preprocessing and data processing procedures.

### 3.1 Mendelian Database

We utilized a dataset from the Mendelian database (<https://data.mendeley.com/datasets/sfwkvmmd5/1>), consisting of EEG signal recordings during neurofeedback training involving 60 students with hyperactivity, regardless of sex [54]. These students were divided into two groups: treatable (30 individuals) and non-treatable (30 individuals). The criteria for treatability or intractability were made by observing the modifications in the brain signal patterns during a course of 30–45 neurofeedback training sessions. Each patient underwent 10 to 20 treatment sessions, with each session comprising 20 tests. Consequently, a total of 1200 training instances were recorded, with a sampling frequency of 500 Hz and a duration of 1.1 seconds during stimulation. This recording encompassed 1100 ms of data, including 100 ms before stimulus onset and 1000 ms after stimulus onset. The data were collected using 32 EEG channels.

The data were stored in text file format and imported into MATLAB software (MATLAB R2020b, MathWorks, Natick, MA, USA) for analysis. During the neurofeedback sessions, patients were presented with two different images and instructed to close their eyes while recalling the details of these images. Neurofeedback sessions were conducted three times a week, alternating between even and odd days, with each session lasting 1 hour. Comprehensive evaluations of the treatment process were conducted during the 5th, 15th, and 30th sessions, while psychiatric evaluations occurred at the 10th, 20th, and 40th sessions. The implementation process followed a sequence of pre-tests, tests, and post-tests.





**Fig. 1. Block diagram of the proposed method.**

### 3.2 Data Pre-Processing

Brain signals typically fall within the significant frequency range of 0.5 to 45 Hz and undergo a comprehensive six-step data preprocessing sequence:

- **Spectral Analysis:** The first step involves spectral analysis. It begins by applying the Fourier transform to transfer the signal from the time domain to the frequency domain. This analysis facilitates a comprehensive examination of the signal's spectral characteristics, including the frequency range and power spectrum density function. If the observed frequency range is deemed unsuitable for brain signal analysis, data recording is repeated.

- **Noise Removal with Notch Filter:** In the second step, a notch filter is employed to eliminate specific noise sources, such as power line interference. This frequency-selective filter has a cutoff frequency of 50 Hz and a bandwidth equal to 0.1 times of the sampling frequency.

- **Butterworth Intermediate Filter:** The third step introduces the Butterworth filter, chosen for its quasi-linear phase properties. This filter is apt for preprocessing and minimizing noise, as it leads to linear alterations in signal patterns upon its use. The ideal order of the filter, established via McClain's algorithm, is determined to be 10. Its design includes cutoff frequencies that span from 0.1 to 45 Hz, with a ripple factor set at 0.1 Hz for both passband and stopband, expressed in linear terms. Additionally, it features a frequency transition rate of 1 Hz.

- **Wavelet Transformation for Low-Frequency Coefficients:** The fourth stage employs a wavelet bank filter to mitigate noise with unidentified origins at low frequencies. This entails matching the mother signal with the raw signal to selectively remove undesired low-frequency generalizations. Signal reconstruction is also considered. Notably, these low-frequency noises typically stem from motion artifacts, and their reduction enhances the signal-to-noise ratio (SNR) through the application of the wavelet bank filter.

- **Wavelet Transformation for High-Frequency Coefficients:** Moving to the fifth step, a wavelet transform bank

filter is utilized once more to diminish noise from unidentified sources at high frequencies. As in prior stages, this involves matching the mother signal with the raw signal to eliminate unwanted high-frequency details. These details typically arise from high-frequency noise sources outside the desired signal's operational range. In this study, a 5-level decomposition approach is employed, featuring the Daubechies 10 wavelet mother. The first-level details are classified as high-frequency noise, while the fifth-level generalities are categorized as low-frequency noise. After noise removal, signal reconstruction is performed.

- **Moving Average Filter:** The final step integrates a moving average filter to smoothen the signal and attenuate high-frequency noise and noise attributable to motion artifacts. Considering a sampling frequency of 500 Hz, equivalent to 500 samples per second, a window length of approximately 20 ms is employed. This choice allows effective adjustment of high-frequency noise while retaining the integrity of the signal pattern based on a window of 10 samples.

In conclusion, these six steps meticulously preprocess brain signal data, rendering it amenable to subsequent analysis and interpretation.

### 3.3 Data Processing

After reducing the signal noise by performing the filtering operation in the preprocessing stage using a Butterworth filter, the alpha and beta frequency bands are separated and extracted with the Phase Locking Value (PLV) index of the desired feature, which shows the connectivity between the brain lobes. To reduce the feature space using a genetic algorithm and *t*-test, the optimal electrodes are selected by the *p*-value index and the data are classified into two classes, treatable and non-treatable.

The *p*-value index helps to decide whether to accept or reject the null hypothesis without referring to the statistical distribution table. An electrode corresponding to a brain channel is selected if the *p*-value for that channel in both al-

pha and beta bands is less than 0.05; if the  $p$ -value exceeds 0.05, the electrode is not selected. These numbers are obtained with the help of a  $t$ -test.

The  $t$ -test is a statistical method in which samples are randomly selected and there is no perfect normal distribution. The accuracy of the test depends on various factors such as the distribution patterns used and the types of influences on the collected samples. After performing the test, a value is obtained as a statistical inference of probability. In the proposed method, a two independent samples  $t$ -test is used. This test is done when samples from two different groups, species, or populations are studied and compared. In this study, data were extracted from two treatable and non-treatable groups. This test is also referred to as the independent samples  $t$ -test, and its formula is as follows:

$$t = \frac{m_A - m_B}{\frac{\sqrt{S^2}}{\sqrt{n_A}} + \frac{\sqrt{S^2}}{\sqrt{n_B}}} \quad (1)$$

where  $m_A$  and  $m_B$  are the average of samples from two different groups or populations,  $n_A$  and  $n_B$  are the respective sample sizes, and  $S^2$  is the standard deviation or common variance of the two samples.

The processing of EEG signals into alpha and beta frequency bands employs a 10th order band-pass Butterworth filter. Given the Butterworth filter's nearly linear phase response, it is deemed appropriate for preprocessing and noise reduction, as it ensures that the signal patterns undergo linear shifts following the filter's application. The filter's order is determined using McClain's algorithm, with the optimal order identified as 10. In terms of filter design, cutoff frequencies range from 0.1 to 45 Hz, and a ripple coefficient of 0.1 Hz is maintained for both the passband and stopband.

### 3.4 Feature Extraction

The primary focus of this research is to explore connectivity among different brain lobes while utilizing neurofeedback training. This investigation relies on feature extraction, specifically the utilization of the PLV. When two brain regions are functionally connected, the variation in the instantaneous phases of signals emanating from these regions should remain relatively consistent. It's important to note that the instantaneous phase possesses a physical interpretation primarily for narrowband signals. Consequently, the initial step in computing the PLV involves filtering the signals [55].

To calculate this index, the phase of each signal is first obtained using the Hilbert transform and then the phase difference is calculated. If the phase difference is slightly different in the experiments, PLV is close to 1 and otherwise close to 0. Given multiple trials or epochs of narrowband filtered brain signals from two EEG channels, the amount of phase locking can be defined as follows:

$$PLV_t = \frac{1}{N} \left| \sum_{n=1}^N e^{j\theta(t,n)} \right| \quad (2)$$

where  $N$  is the number of tests and  $\theta(t,n)$  is the difference between the instantaneous phase of the two signals at time  $t$  and test  $n$ . Therefore, calculating the PLV requires filtering the data in the desired frequency band and then extracting the instantaneous phase through the Hilbert transform, which transforms real signals into a complex representation. PLV takes values in the range  $[0, 1]$ , where PLV equal to 0 indicates that there is no phase synchronization and PLV equal to 1 indicates that the relative phase between the two signals is the same during the experiments.

### 3.5 Reducing the Feature Space

The genetic algorithm employed for electrode selection in the proposed method can be summarized as follows:

(1) Data Input: Begin by inputting a data matrix with dimensions of  $32 \times 550 \times 1200$  for each frequency band. These dimensions represent the number of test repetitions, signal samples, and device channels, respectively.

(2) Initial Population Generation: Create the initial population by generating random binary chromosomes, where the number of bits equals the number of features (32 in this case, corresponding to the 32 electrodes used). Each chromosome represents a combination of electrodes. Evaluate the correctness of each chromosome by passing it to a classifier, comparing its output label with the real label matrix, and calculating the accuracy (accuracy = 1 - error). Select two parent chromosomes based on their accuracy and repeat this process through multiple generations.

(3) Cost Function: Define a cost function based on classification error, which should be minimized. The error is calculated as the square of the difference between the actual label and the label obtained from the classifier.

(4) Population Evolution: Create a new population in each generation based on the fitness criteria of each chromosome, using both crossover and mutation operations.

(5) Selection: Select the desired population based on the fitness function, retaining the best-performing chromosomes.

(6) Termination Conditions: Continuously check for termination conditions. If they are not satisfied, return to step 4 for further iterations. Termination conditions include a slight change in error for two consecutive algorithm stages. If no significant error reduction is observed in consecutive executions of the algorithm, the chromosome with the highest selection accuracy, containing binary genes with a value of 1, is considered as the optimal electrode set. The threshold for detecting a significant change in error reduction is set at  $10^{-6}$ .

In this study, the crossover operation is performed on parent chromosomes using the exclusive or exclusive-

OR (XOR) operator, while mutation employs the flip-bit method with a mutation rate of 0.01. The percentage of crossover and mutation is set at 0.7 and 0.3, respectively. The objective function to minimize is the classification error. Termination criteria includes a small change in error for two consecutive iterations and no substantial error reduction in consecutive executions of the algorithm. Accuracy and error are as following:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (3)$$

$$\text{Error} = 1 - \text{Accuracy} \quad (4)$$

The value of the fitness function (FF) is determined from the following equation, in which the number of repetition steps of the algorithm are  $m_1, m_2, \dots, m_{n-1}, m_n$ .

$$\text{FF} = \text{Error}(m_a) - \text{Error}(m_{n-1}) = 10^{-6} \quad (5)$$

Fig. 2 illustrates the cost function of the genetic algorithm for selecting the optimal electrode, demonstrating its convergence towards minimization. This cost function is derived from the classification error. It's important to note that the algorithm continues to iterate until the difference in error between consecutive iterations falls below a threshold of  $10^{-6}$ . This threshold value is determined through repetition and serves as a criterion for identifying the algorithm's convergence to its absolute minimum. Therefore, when the difference in error attains the specified threshold, it signifies that the algorithm has achieved its optimal performance. This triggers the cessation of the iterative process.

## 4. Analysis of the Results

This section delves into the outcomes of predicting the treatability of ADHD patients through neurofeedback training. It encompasses the results generated from the six fundamental steps of the proposed method, which includes data collection, preprocessing, processing, feature extraction, feature space reduction, and classification. For this purpose, we present the simulation results and accuracy values to validate the efficacy of the proposed method.

### 4.1 Data

The data used in this study were recorded in a multi-channel format, which allowed the selection of the optimal channel. It should be noted that the algorithm presented in this article is coded in the MATLAB R2020b environment. First, the primary dataset is entered into the preprocessing structure to reduce the noise. Fig. 3 shows the output from the MATLAB software as an example of an electroencephalogram signal during neurofeedback training before the noise removal. This signal is drawn for a period of

1.1 seconds from the recording of the electroencephalogram signal during neurofeedback, the horizontal axis is time in seconds and the vertical axis is voltage in microvolts, which shows the range of the signal. The dataset is saved as a compressed file that contains three matrices. The first matrix is called data, which is a  $32 \times 55 \times 1200$  three-dimensional matrix and shows the number of test repetitions, the number of signal samples, and the number of neurofeedback recording channels, respectively. The sampling rate is set at 500 Hz, and the duration of signal recording is 1.1 seconds, resulting in a total of 550 signal samples (calculated as  $500 \times 1.1$ ). Accompanying this dataset is a label matrix of dimensions  $1 \times 1200$ . This matrix classifies the data into two categories: label 1 for treatable patients and label 0 for incurable patients. In the scope of this study, we have focused exclusively on data from the C3, FZ, F4, CZ, C4, and F3 channels. This selection is based on the involvement of the frontal lobe and the central cerebral cortex in neurofeedback training. These regions are targeted for three primary reasons:

A—Due to the fact that the sensory-motor cortex and the frontal cortex are layers of the brain that are responsible for making decisions at the highest advanced level and due to the neurofeedback training (i.e., remembering moving images) as an advanced brain operation, it seems that from a physiological point of view, the frontal and central cerebral cortex are involved lobes that have a higher decision-making level than other brain lobes.

B—The manifestation of brain imagination process occurs in the central cerebral cortex.

C—Usually the alpha band can be measured mostly in the occipital lobe, but here, because remembering is an advanced process, the alpha and beta bands are dominant in the frontal lobe and central cerebral cortex.

### 4.2 Evaluation of the Results of the Proposed Method

To evaluate the performance of the proposed algorithm, we utilize the accuracy metric as defined by Eqn. 3. This equation takes into account true positive (TP) rate, which is the proportion of treatable cases correctly identified; true negative (TN) rate, the proportion of untreatable cases accurately recognized; false positive (FP) rate, the rate at which treatable cases are incorrectly labeled; and false negative (FN) rate, the rate of misclassification for untreatable cases. A value of this accuracy index closer to 1 signifies the optimal performance of the proposed method. Our method demonstrated high efficacy in discerning the treatability and non-treatability of hyperactivity in patients using neurofeedback, achieving an accuracy rate of 90.6%.

### 4.3 Preprocessing

According to previous studies, some signal patterns are lost due to filtering, but the spectral analysis of the signal shows that the removed patterns are not in the frequency range of the alpha and beta bands [56–59]. In this method,

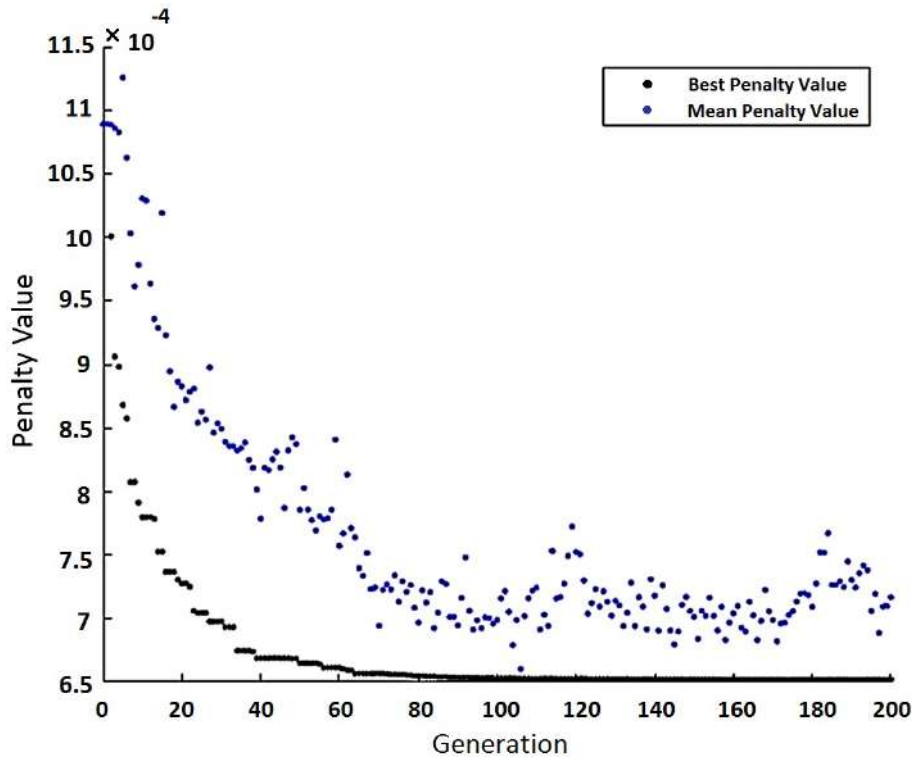


Fig. 2. Cost function in the genetic algorithm for optimal electrode selection.

after filtering, the SNR is calculated to check the effect of the filter on the signal. In this study, the target brain signal range is considered between 0.1 and 45 Hz. After filtering in the desired range (i.e., 0.1 to 45 Hz), a noise-free signal is obtained. To calculate the SNR, a random white noise must be created and passed through a 45 Hz high-pass filter. To get a high-frequency pink noise (because only 45 Hz and above values are considered), and to test the efficiency of the filters, the created pink noise is added to the signal. Therefore, if the filters effectively eliminate the targeted pink noise, it indicates that all noise within this specific frequency band has been removed. To compute the SNR, the powers of both the signal and the noise are measured separately. The SNR is then determined by comparing these two values, providing a measure of the signal's quality against the backdrop of noise.

Table 1 displays the SNR for various filtering stages before and after filter application. The rows of this table illustrate the different stages of filtering, while the columns detail the SNR changes before and after applying the filter. According to the table:

- The first filtering stage, using a notch filter, saw an increase in SNR by  $2 \pm 2.1$ .
- In the second stage, employing a Butterworth band-pass filter, the SNR improvement was  $2 \pm 1$ .
- For the third stage, which utilized wavelet denoising, the increase in SNR was  $2 \pm 1.2$ .
- The fourth stage involved a high-frequency wavelet transformation, leading to a notable SNR boost of  $3.4 \pm 0.8$ .

- Finally, in the fifth stage, with a moving average filter, the SNR further improved to  $3 \pm 0.8$ .

These results highlight that the fourth-stage application of high-frequency wavelet transformation yielded the most significant increase in SNR. Across the five filtering stages, the total enhancement in SNR amounted to  $7 \pm 1.2$ . The calculation of SNR is based on the following equation:

$$\text{SNR} = \frac{P_{\text{signal}}}{P_{\text{noise}}} \quad (6)$$

Fig. 3 shows the brain signal of treatable and non-treatable people before filtering, Figs. 4,5,6,7 show the steps before and after filtering the brain signal, and Fig. 8 shows the power spectrum density in the time-frequency domain using Welch's method to estimate the power spectrum.

#### 4.4 Processing

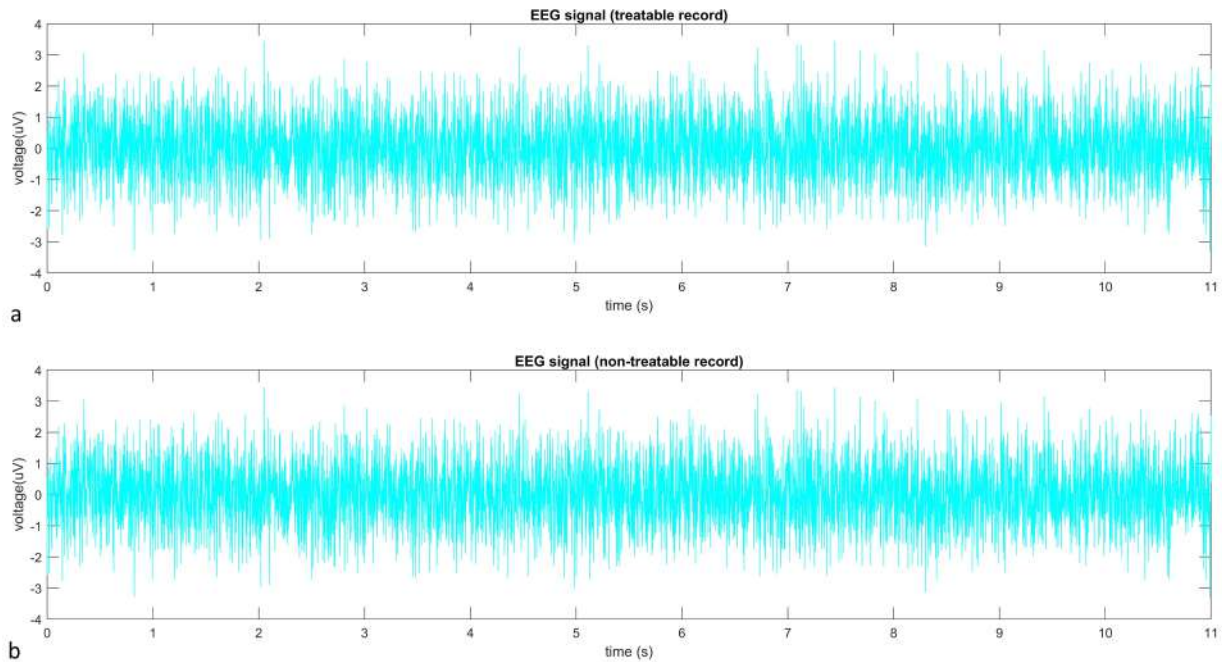
This study focuses on the analysis of functional connectivity to differentiate between treatable and non-treatable groups through the phase lock index. It utilizes the  $p$ -value as a statistical measure to quantify the strength of these connections. Specifically, brain channels are selected based on their  $p$ -values within the alpha and beta frequency bands; a  $p$ -value below 0.05 indicates selection, while a value above 0.05 leads to exclusion. The preprocessing stage involves filtering operations to diminish signal noise, followed by the extraction of alpha and beta bands using a



**Table 1. Signal-to-noise ratio (dB) of various filtering stages, before and after filter application.**

Filtering steps	SNR before application	SNR after application
First stage filter (Notch filter)	$55.1 \pm 2.5$	$66.5 \pm 1.3$
Second stage filter (Butterworth band pass)	$62.3 \pm 2.3$	$74.3 \pm 1.3$
Third stage filter: low frequency wavelet transform	$71.6 \pm 2.5$	$82.1 \pm 1.3$
Fourth stage filter: high frequency wavelet transform	$84.1 \pm 2.2$	$92.5 \pm 1.1$
Fifth stage filter: moving average trap	$90.5 \pm 2.1$	$98.3 \pm 1.3$

SNR, signal-to-noise ratio.



**Fig. 3. EEG signal of people before applying filtering on (a) treatable (b) non-treatable participants.** EEG, electroencephalogram.

10th order band-pass Butterworth filter with an unlimited impulse response. The selection of optimal electrodes was facilitated by a genetic algorithm and *t*-tests, employing the *p*-value as a criterion. Table 2 shows the EEG brain channels employed for neurofeedback training across the alpha and beta bands, with a separate calculation of the *p*-value for each frequency band. In this table, the rows contain the name of the EEG channels, and the columns contain the *p*-value for the alpha band and the beta band. As seen in Fig. 9, according to the desired statistical index, the selected channels are in the frontal lobe and central cerebral cortex.

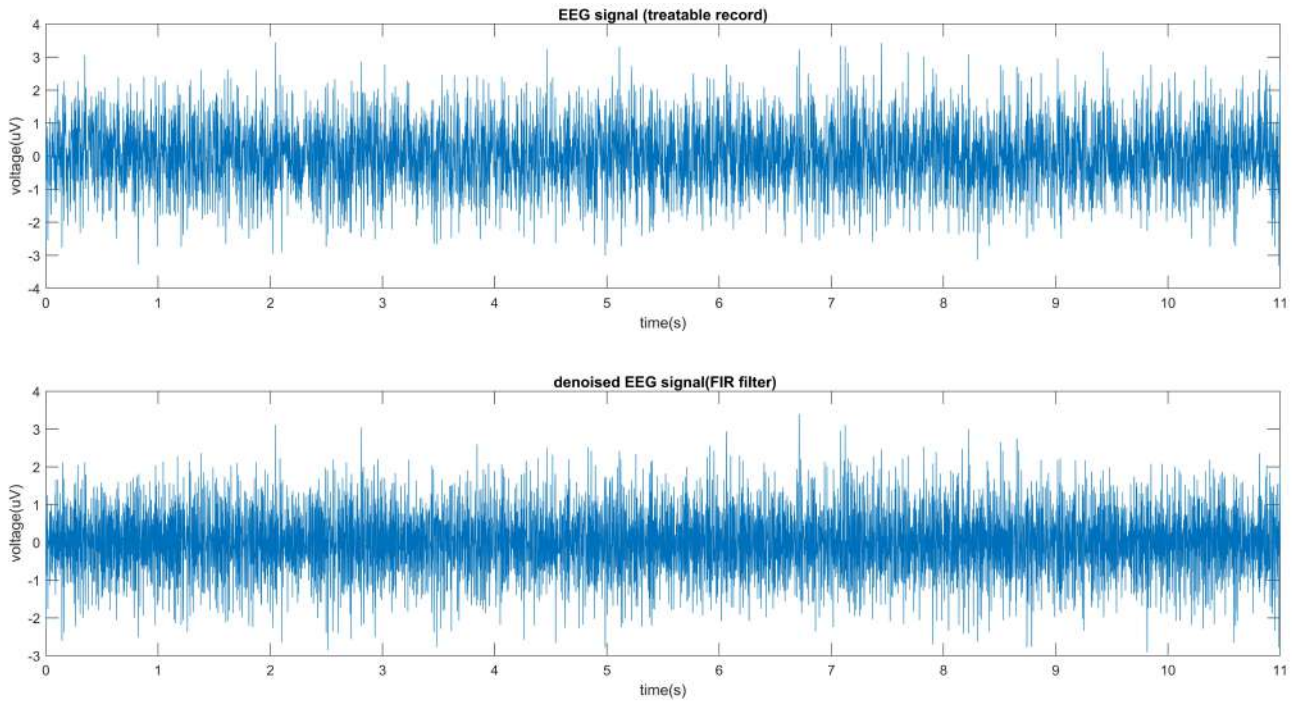
Figs. 10,11 illustrate the separation of alpha and beta frequency bands for both treatable and non-treatable patients, with each trial lasting 1.1 seconds. Specifically, a segment comprising trials 1 through 10, amounting to 11 seconds, corresponds to 550 samples at the 5th electrode. Table 3 details the EEG channels chosen via a *p*-value statistical test across both alpha and beta bands and compares these with channels pinpointed by a genetic algorithm as having the most significant brain connectivity. The comparison reveals a strong alignment between the electrodes identified through the genetic algorithm and those selected

based on the statistical test, underscoring the effectiveness of both methodologies in harmony. Importantly, the analysis suggests that the beta frequency band yields the most advantageous results.

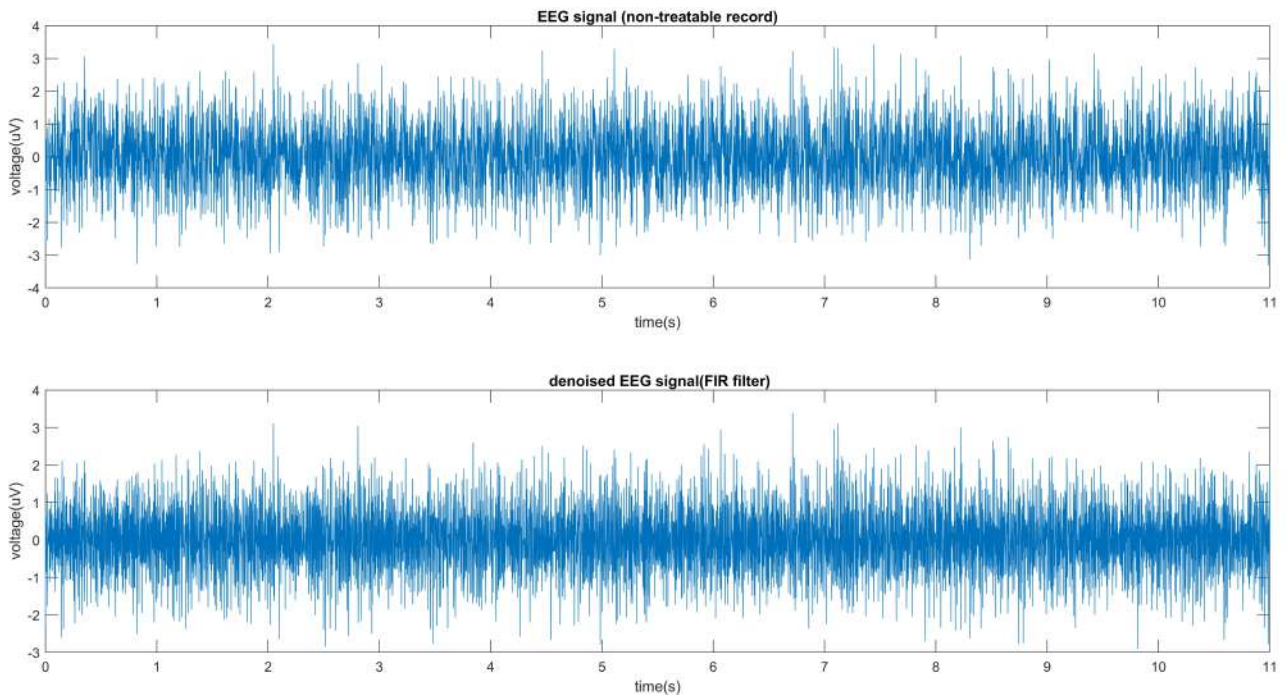
#### 4.5 Classification

In this research, a variety of classifiers, including support vector machine, nearest neighbor, and decision tree, were employed, utilizing both bagging and boosting methods for their integration. The input for these classifiers is a feature matrix that represents brain connectivity metrics derived from the optimal electrodes, identified by intersecting the *p*-value (PV) index with a genetic algorithm. This matrix has dimensions of  $1200 \times 36$ , where each row corresponds to a patient and each column to a feature. The output matrix, indicating treatability (label 1) or non-treatability (label 0), has dimensions of  $1200 \times 1$ , encompassing 600 treatable and 600 untreatable cases.

For assessing the performance of these classifiers, a five-fold cross-validation strategy was adopted. This involves dividing the dataset into five equal parts, ensuring that each part represents 20% of the data for testing and the



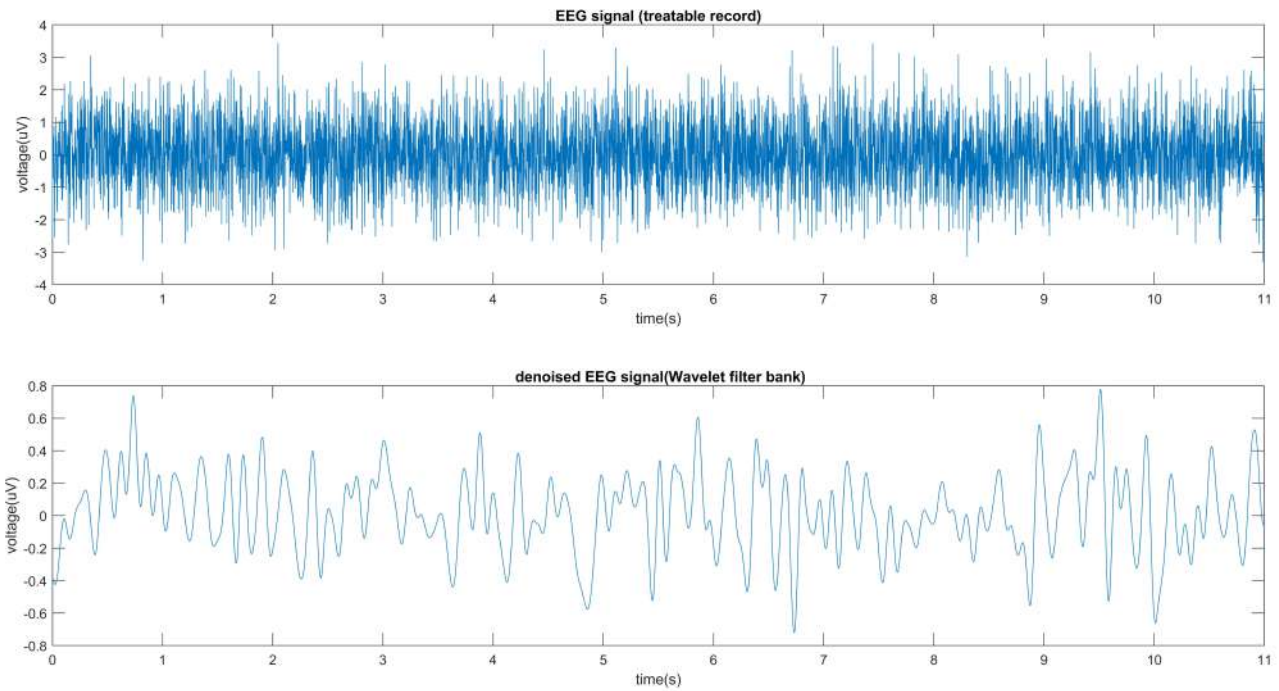
**Fig. 4. FIR filter for the brain signal of treatable participants.** FIR, finite impulse response.



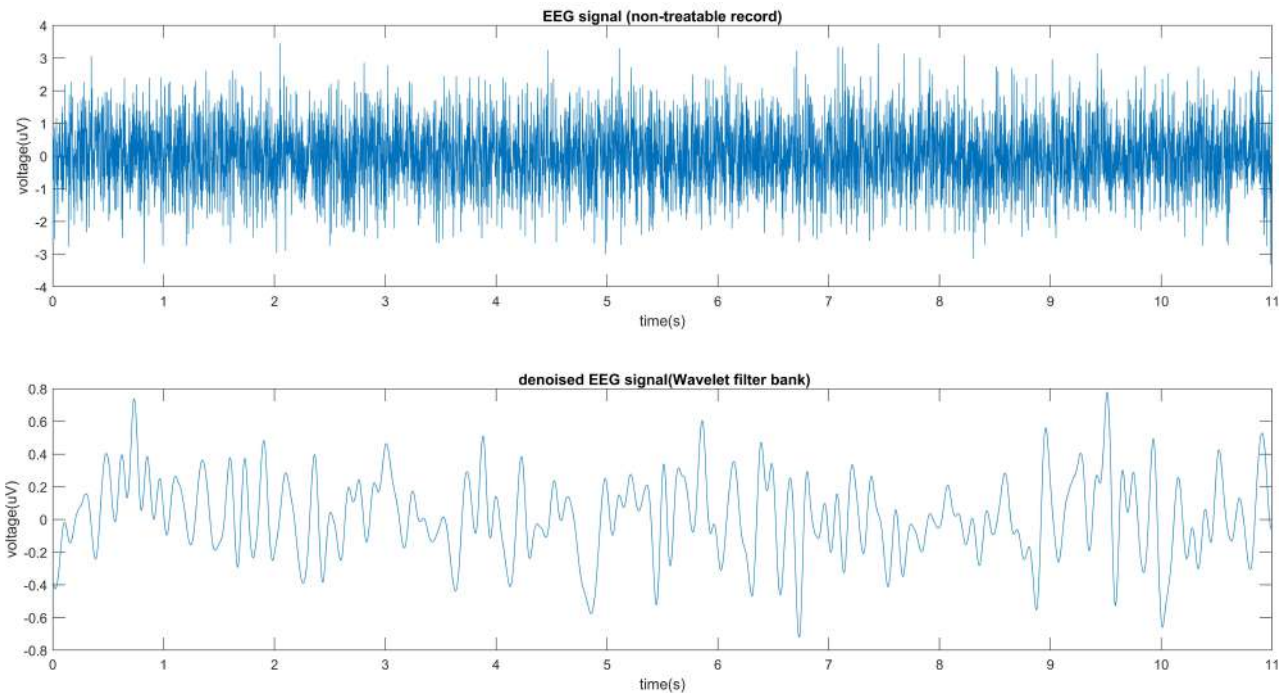
**Fig. 5. FIR filter for the brain signal of non-treatable participants.**

remaining 80% for training. This division is predicated on the rationale that a five-fold division strikes a balance between bias and variance, optimizing the learning process. In each iteration, one of the five subsets is used as the test set while the other four serve as the training set. This process is rotated until each subset has been utilized as a test set once, allowing for a comprehensive evaluation across the entire

dataset. The results are then aggregated, with classification metrics such as accuracy and standard deviation reported over 20 runs of the model. These metrics are presented in classification tables, where rows correspond to each fold ( $k = 1$  to 5) and columns to the classifier type. The tables also delineate the mean accuracy and its standard deviation and provide a clear understanding of the models' performance



**Fig. 6. Wavelet bank filter for treatable participants.**



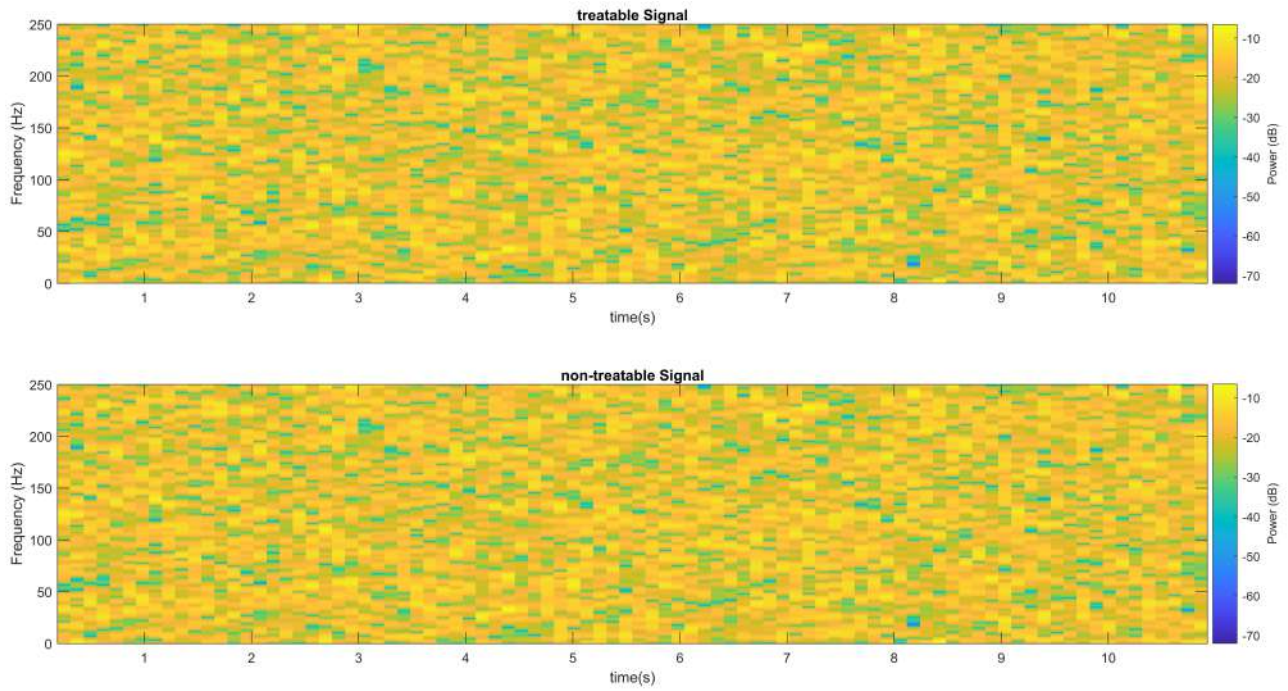
**Fig. 7. Wavelet bank filter for non-treatable participants.**

variability [47]. In this approach, the data folds are distinctly segregated for training and testing purposes, ensuring that each phase of the evaluation process is conducted independently. However, the individual accuracies derived from each of the five folds are collectively averaged to ascertain the model's overall accuracy across all stages.

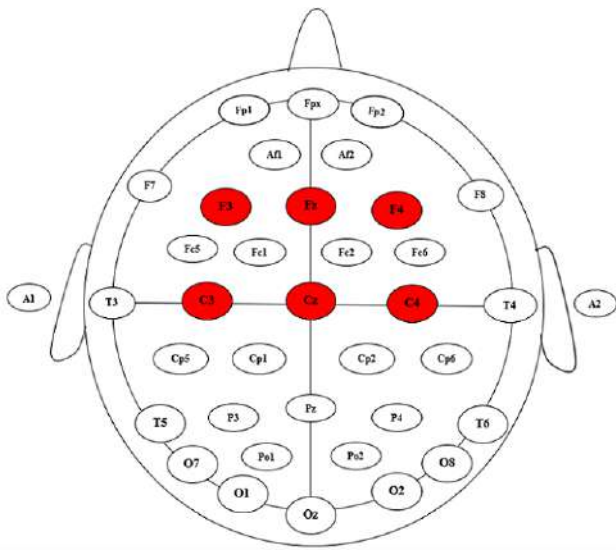
#### 4.5.1 Classification of Support Vector Machine

A support vector machine (SVM) is a supervised learning algorithm designed to classify data samples by finding the optimal hyperplane which acts as a decision boundary in a high-dimensional space. The essence of SVM is to segregate data points into distinct groups, such that data points on either side of this boundary share similar-





**Fig. 8. Power spectrum density in the time-frequency domain for (a) treatable (b) non-treatable participants.**



**Fig. 9. Optimal electrodes effective in the treatment of hyperactivity disease with neurofeedback.**

ities and are assigned to the same category. The algorithm aims to achieve this separation with the highest margin, thereby enhancing the model's predictive accuracy. When new data samples are introduced, they are positioned within this multi-dimensional space and classified based on which side of the hyperplane they fall on. The primary objective of SVM is to discover a hyperplane in an  $N$ -dimensional space (where  $N$  represents the number of features) that effectively differentiates the data points, ensuring precise classification [60].

Fig. 12 illustrates the SVM classifier's performance for the alpha band, displaying a linear graph where the horizontal axis represents the  $k$ -fold classification iterations and the vertical axis denotes the classifier's kernel types. These kernel types encompass a range of options: linear kernel, second-order polynomial (2D), third-order polynomial (3D), micro-scale Gaussian (micro-G), medium-scale Gaussian (mean-G), and large-scale Gaussian (Large-G), with each point on the graph connecting the number of kernels utilized. Among these, the small-scale Gaussian kernel stands out for achieving the highest classification accuracy, recording an 82.22% accuracy rate for the alpha band and 83.06% for the beta band. Similarly, Fig. 13 focuses on the SVM classifier's performance for the beta band, adopting the same graphical representation to detail the relationship between  $k$ -fold classification iterations and kernel types, thereby underscoring the consistency in methodology and results across different frequency bands.

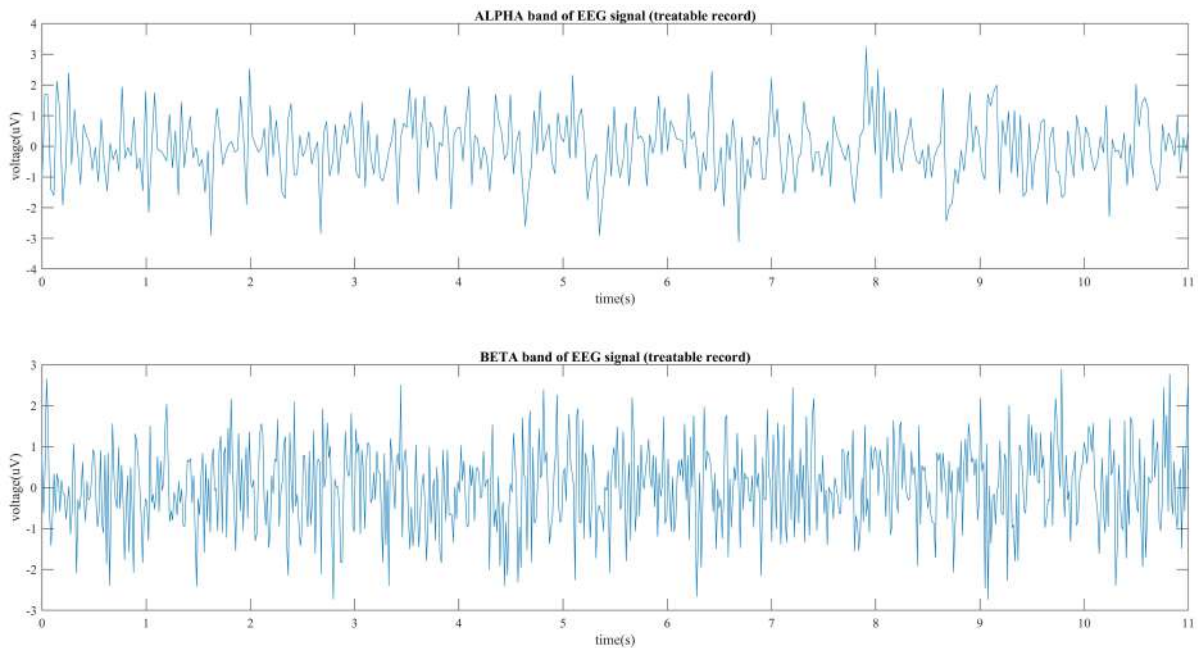
#### 4.5.2 Nearest Neighbour Classification

The nearest neighbor algorithm employs feature similarity to forecast the values of new data points, meaning it assigns values to new data based on their resemblance to points in the training set. Similar to the SVM classifier, this method organizes its output with rows indicating  $k$  values ranging from 1 to 5 and columns representing the distance metrics used to identify the nearest neighbor. Notably, the algorithm achieves its highest accuracy when utilizing a weighted Murkowski distance with  $k$  equal to 5, recording accuracy rates of 76.14% for the alpha band and 76.94% for the beta band, respectively.



**Table 2. EEG channels and corresponding  $p$ -values in the alpha and beta frequency bands.**

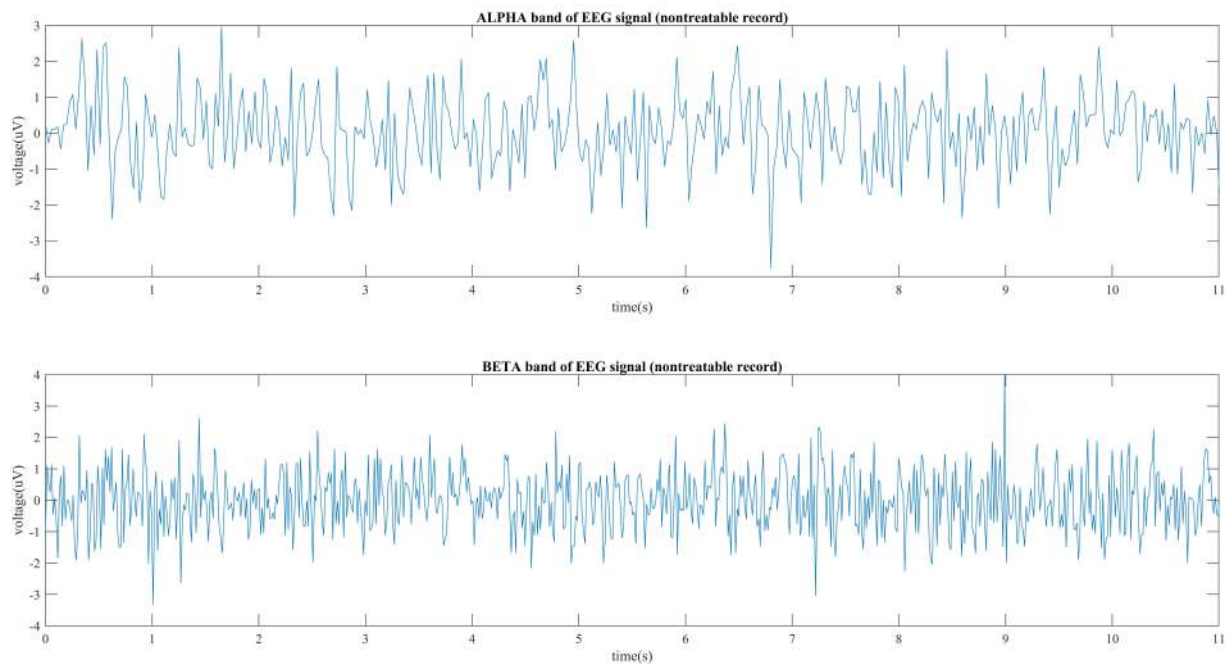
Brain channel according to $p$ -value	Alpha band	Beta band	To choose or not to choose ( $p < 0.05$ )	Brain channel according to $p$ -value	Alpha band	Beta band	To choose or not to choose ( $p < 0.05$ )
Cz	0.004	0.03	Choose	Fp1	0.7	0.3	Not to choose
C4	0.003	0.0001		T7	0.5	0.2	
F4	0.04	0.006		F7	0.7	0.4	
Fz	0.03	0.003		Oz	0.2	0.2	
C3	0.003	0.001		PO3	0.1	0.09	
F3	0.004	0.0002		AF3	0.2	0.08	
P7	0.1	0.07		FC5	0.6	0.06	
P4	0.08	0.06	FC1	0.9	0.5		
Pz	0.2	0.09	CP5	0.08	0.07		
P3	0.09	0.07	CP1	0.06	0.08		
P8	0.3	0.1	CP2	0.09	0.06		
O1	0.3	0.2	CP6	0.1	0.2		
O2	0.2	0.5	AF4	0.5	0.2		
T8	0.08	0.1	FC2	0.07	0.06		
F8	0.4	0.2	FC6	0.8	0.4		
Fp2	0.1	0.07	PO4	0.6	0.3		



**Fig. 10. Alpha and beta frequency bands for treatable participants.**

Fig. 14 depicts the nearest neighbor classifier's performance for the alpha band, employing a similar layout to previous classifiers with the horizontal axis indicating the  $k$ -fold classification iterations and the vertical axis detailing the number of points for different distance metrics used in classification. These metrics are represented on a line graph and include Euclidean distance (SQ), Manhattan distance (COS), Murkowski distance (MG), and weighted Murkowski distance for  $k$  values of 5 (W5), 7 (W7), and 9 (W9), connecting the points corresponding to each distance type.

Similarly, Fig. 15 shows the nearest neighbor classifier for the beta band, maintaining the same graphical structure as for the alpha band, with the horizontal axis representing  $k$ -fold classification iterations and the vertical axis showing the different classification distance metrics. As in the alpha band, the data points are organized into six categories, denoted as SQ, COS, MG, W5, W7, and W9, with each category's points connected in a linear fashion to visualize the performance of each distance type in the classification process.



**Fig. 11. Alpha and beta frequency bands for non-treatable participants.**

**Table 3. EEG channels selected by the genetic algorithm and those identified through statistical testing.**

Channels selected by $p$ -value statistical test	Selected channels by genetic algorithm with maximum alpha band connectivity	Selected channels by genetic algorithm with maximum beta band connectivity
CZ	CZ	CZ
C4	CZ	C4
F4	F4	FZ
FZ	F4	FZ
C3	C3	C3
F3	FZ	F3

#### 4.5.3 Decision Tree Classification

A decision tree represents the decision-making process using a branching, tree-like structure. It is widely used in machine learning as a method for both classification and regression tasks [61]. The decision tree algorithm segments a dataset's features using a cost function, initially including features that may be irrelevant to the problem at hand. To refine the model, unnecessary branches are pruned to optimize the tree's structure. This process, known as pruning, helps adjust the tree's depth to prevent overfitting and maintain a manageable complexity. The algorithm employs predictive modeling to explore various decisions or solutions, aiming to achieve the desired output. This study examines three scales of decision tree models: micro-scale, medium-scale, and large-scale. Each scale has a distinct level of decomposition, with four levels for micro-scale, three for medium-scale, and two for large-scale trees. The ID3 algorithm, which focuses on maximizing information gain, serves as the primary learning function in these models.

Figs. 16,17 depict the decision tree classification for the alpha and beta bands, respectively. In both diagrams,

the horizontal axis denotes k-fold classification, and the vertical axis displays the various classification tree types, including large-scale tree (S-Tree), medium-scale tree (M-Tree), and micro-scale tree (C-Tree). Among these, the micro-scale tree (C-Tree) demonstrates the highest accuracy levels in the alpha and beta frequency bands, achieving accuracy rates of 67.94% and 68.78% in the alpha and beta bands, respectively.

#### 4.5.4 Classification Composition (Bagging and Boosting Method)

The integration of the two classification methods, bagging and boosting, yields superior outcomes. Bagging, a collective learning strategy, seeks to reduce the error associated with learning by employing an ensemble of homogeneous machine learning models [62–64]. Initially, the selection of basic model types and quantities takes place. Following this, a random sampling method is employed to choose a subset of data from the training dataset, which is then used to retrain each base model by replacement. In classification scenarios, a simple majority vote among

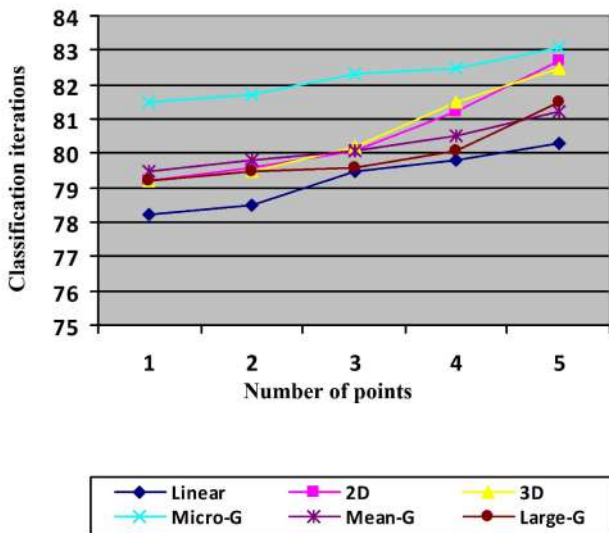


Fig. 12. Support vector machine classifier’s performance for the alpha frequency band. 2D, second-order polynomial; 3D, third-order polynomial; micro-G, micro-scale Gaussian; mean-G, medium-scale Gaussian; Large-G, large-scale Gaussian.

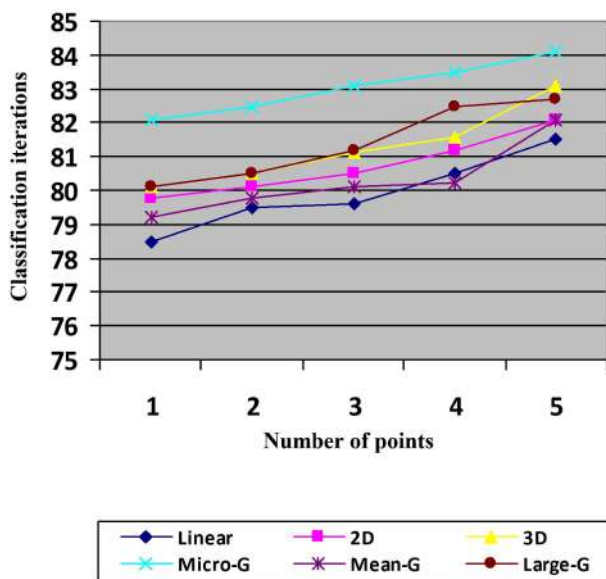


Fig. 13. Support vector machine classifier’s performance for the beta frequency band.

the models determines the assignment of new data to the class receiving the highest number of votes. For regression tasks, the average of the base models’ outputs is calculated. Boosting, whether applied in parallel or sequentially, significantly reduces error and optimizes classification outcomes. It transforms a weak learning system, which barely outperforms random guessing, into a strong classifier capable of accurately predicting sample labels. The adaptive (incremental) approach fine-tunes the classifier at each step to focus on previously misclassified samples, enhancing its

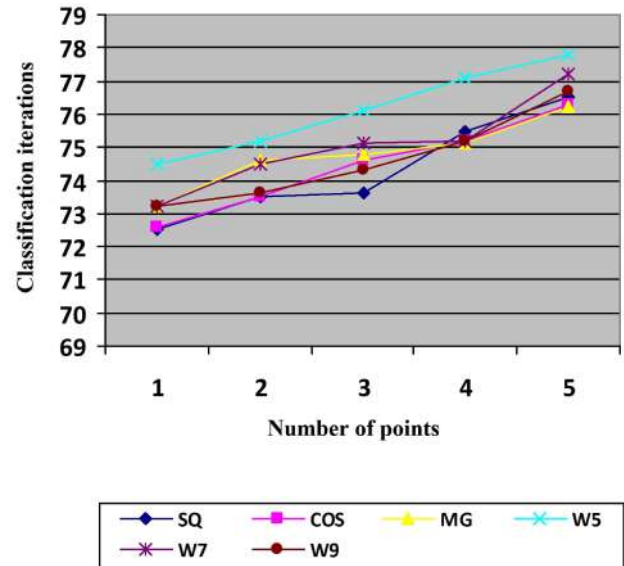


Fig. 14. K-Nearest neighbor classification for the alpha frequency band. SQ, Euclidean distance; COS, Manhattan distance; MG, Murkowski distance; W5/W7/W9, weighted Murkowski distance for k values of 5/7/9.

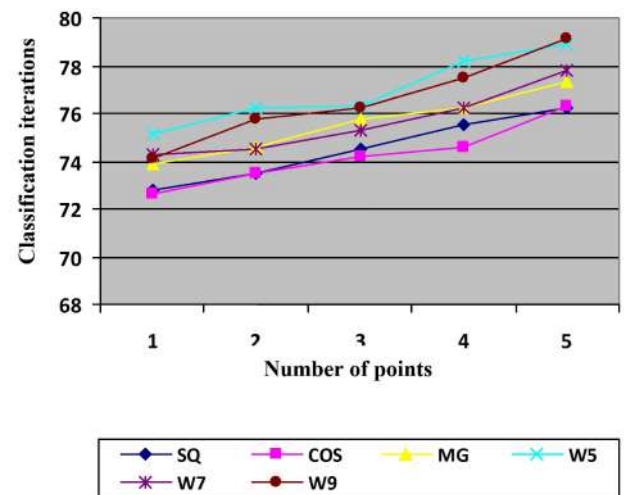


Fig. 15. K-Nearest neighbor classification for the beta frequency band.

accuracy over time. Despite AdaBoost’s vulnerability to noise and outliers, it excels at preventing overfitting, outperforming many other learning algorithms. The core requirement for the base classifier is to achieve just above 50% accuracy, a performance slightly better than random, which allows the algorithm to improve iteratively. Even classifiers with marginal accuracy improvements can contribute to enhanced overall performance by adopting a negative coefficient in the boosting process.

The AdaBoost algorithm enhances its classification accuracy by iteratively adding a weak classifier in each round. During these rounds, it adjusts the weights of the

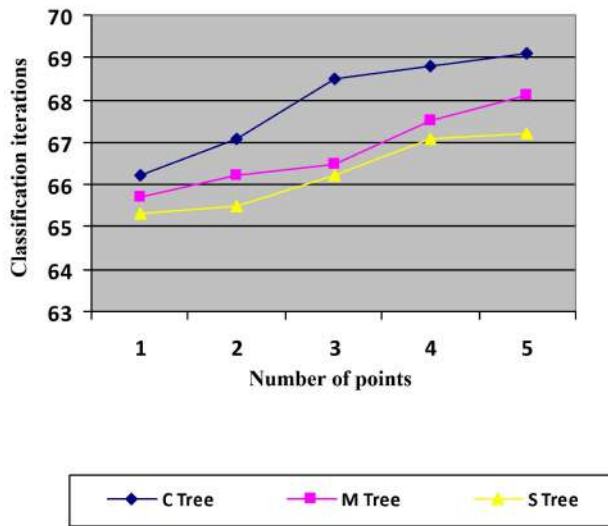


Fig. 16. Decision tree classification for the alpha frequency band.

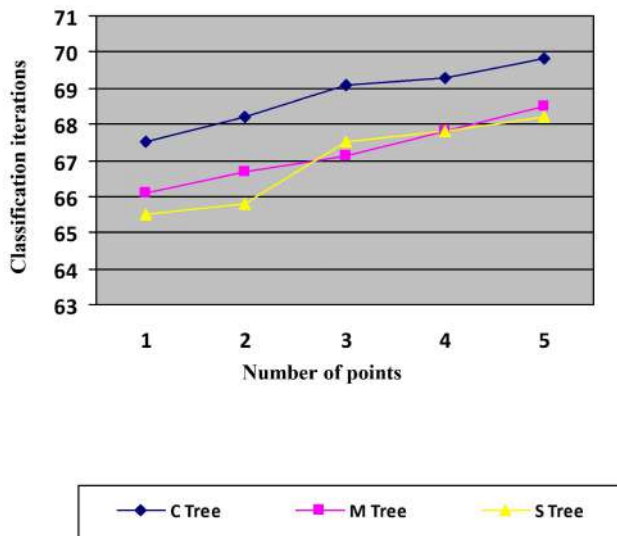


Fig. 17. Decision tree classification for the beta frequency band.

samples to reflect their significance, specifically increasing the weights of misclassified samples and decreasing the weights of those correctly classified. This adaptive weighting ensures that subsequent classifiers pay more attention to the samples that previous classifiers found challenging, thereby improving the model’s ability to learn from difficult examples [65].

Figs. 18,19 depict the performance of combined classification techniques for the alpha and beta frequency bands, respectively. These figures highlight the horizontal axis as representing k-fold classification, while the vertical axis indicates the performance of combined classification methods, including bagging, boosting, and AdaBoost. Notably, the adaptive boosting method stands out for its su-

perior accuracy, achieving 89.62% for the alpha band and 90.6% for the beta band, making it the most effective classification combination.

The examination of individual and combined classification methods reveals that certain classifiers yield the best results. These include the SVM classifier with a small-scale Gaussian kernel, the nearest neighbor classifier employing a weighted Minkowski distance with  $k = 5$ , and the tree classifier using a micro-scale decision tree. It is observed that the beta band consistently exhibits higher accuracy compared with the alpha band. Specifically, the SVM classifier with a small-scale Gaussian kernel for the beta band registers the highest accuracy at 83.06%, indicating its superiority for the analyzed features. Furthermore, among the combined classification approaches, the adaptive incremental method emerges as the most potent, especially for the beta band, with an impressive accuracy of 90.6%. Table 4 shows a comparative summary of the best-performing classifications for both the alpha and beta frequency bands, illustrating the effectiveness of these methods.

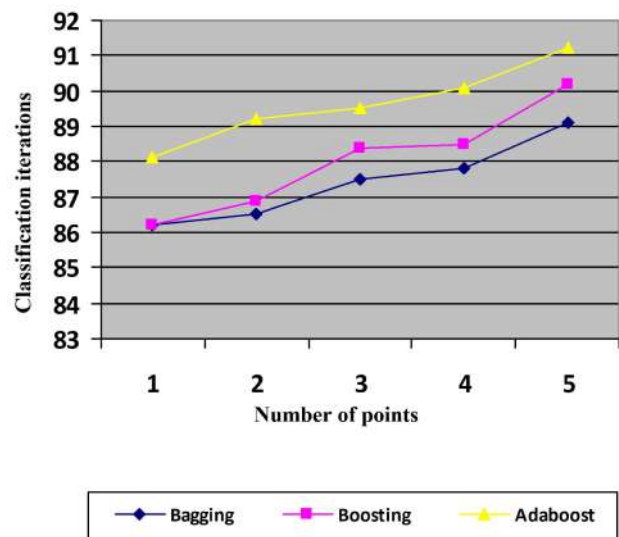


Fig. 18. Combination of classifications for the alpha frequency band.

#### 4.6 Comparing the Accuracy of the Proposed Method to Previous Studies

Table 5 (Ref. [4,53,66]) compares the accuracy of the proposed method and various previous studies. In a study focused on the electromagnetic tomography of the brain involving 18 patients with hyperactivity, the capability to predict treatment outcomes before starting neurofeedback treatment was demonstrated with a 70% accuracy rate.

Another study investigated the effectiveness of slow cortical potential (SCP) and the impact of combined theta/beta training across 18 sessions on 46 patients. This study found no significant predictors for the outcomes of



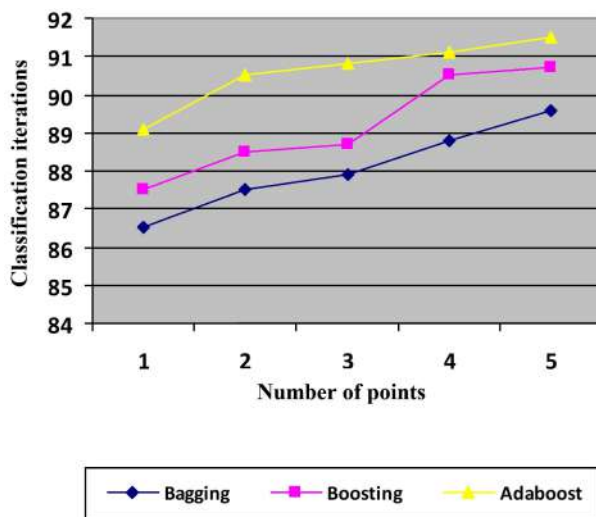
**Table 4. Comparison of the best output in different categories in the alpha and beta frequency bands.**

Classification	Best estimate	Accuracy of the alpha frequency band (%)	Accuracy of the beta frequency band (%)
Decision tree	Micro-scale tree	67.94 ± 1.42	68.78 ± 1.38
Support vector machine	Micro-scale Gossin	82.22 ± 1.4	83.06 ± 1.32
Nearest neighbor	Weighted Minkowski	76.14 ± 1.46	76.94 ± 1.58
Combination of categories	Adaptive incremental method	89.62 ± 1.32	90.6 ± 1.36

**Table 5. Comparison of accuracy percentages between the proposed method and various neurofeedback studies.**

Method	Number of patients	Accuracy (%)
Combined theta/beta training (18 sessions) and SCP [53]	46	45.5
SCP, alongside evaluating the patient's age and sex [66]	23	79.6
SCP [4]	20	89.9
Proposed method	60	90.6

SCP, slow cortical potential.

**Fig. 19. Combination of classifications for the beta frequency band.**

the comprehensive theta/beta+SCP training program. However, a significant decrease in theta activity during the theta/beta training portion was observed in responsive individuals, and an increase in alpha activity was noted in the SCP training part for those considered treatable, achieving an accuracy of 45.5% in predicting neurofeedback training outcomes.

In a third case, the evaluation of 23 children aged 9 to 12 years with hyperactivity, considering SCP criteria, age, and sex, showed that age significantly influences the initial conditions for neurofeedback exercises. EEG data analysis before training indicated a decline in slower frequency band activity and a decrease in the theta/beta ratio with age. This study also found no significant sex differences in neurofeedback learning performance among ADHD patients, with a predictive accuracy for successful treatment of 79.6%.

A fourth study examined 20 patients over the first six

sessions, highlighting a noticeable increase in positive SCP in the EEG of treatable individuals. This approach resulted in classifying treatable versus untreatable individuals with an accuracy of 89.9%.

Lastly, our proposed method was tested on 60 patients prior to neurofeedback treatment, assessing functional brain communication through the phase lock index. This method predicted treatability with an impressive accuracy of 90.6%. The results in Table 5 underscore the superiority of the proposed method in comparison with previous studies.

## 5. Conclusions

In this study, we introduced and validated a novel algorithm for predicting the responsiveness of hyperactive children to neurofeedback therapy, an essential tool in managing ADHD. Our study leveraged EEG data from the Mendelian database, involving 60 students aged 7–14 years undergoing neurofeedback sessions. Through a five-step process that included preprocessing to remove noise, extracting relevant frequency bands, evaluating functional relationships using the phase lock value, minimizing the feature space to identify optimal electrodes, and finally applying machine learning classifiers, we aimed to streamline the neurofeedback therapy process.

Our findings underscore the significance of selecting appropriate electrodes for neurofeedback, with the optimal ones identified in the frontal lobe and central cerebral cortex, notably at channels C3, FZ, F4, CZ, C4, and F3. This selection process, facilitated using the genetic algorithm and the *p*-value index, effectively reduced the number of necessary recording channels from 32 to 6, simplifying the treatment setup. Moreover, the application of SVM and boosting methods as classifiers enabled the accurate prediction of treatment effectiveness, achieving an impressive accuracy rate of 90.6%.

The implications of our study are manifold. Firstly, it offers a promising approach to predict the treatability of ADHD through neurofeedback therapy, potentially saving considerable time and resources by identifying non-responsive patients early in the treatment process. Secondly, by reducing the number of required electrodes, our algorithm simplifies the neurofeedback setup, making it more accessible and less burdensome for both practitioners and patients. Lastly, the high accuracy rate of our predictive model holds significant promise for enhancing the efficacy of neurofeedback therapy in treating hyperactivity disorders, potentially improving the quality of life for many children affected by ADHD.

In conclusion, our research presents a breakthrough in the optimization of neurofeedback therapy for hyperactive children, demonstrating that through the strategic analysis of EEG data and the application of machine learning, we could significantly enhance treatment predictability and efficiency. This advancement not only supports the clinical application of neurofeedback in psychological clinics but also marks a step forward in personalized medicine, tailoring interventions to the specific neural characteristics of individuals.

### Availability of Data and Materials

The data sets analyzed during the current study are available in the “Mendeley Data” repository (<https://data.mendeley.com/datasets/sfwkvmmd5/1>), EEG data (from real-time neurofeedback experiment).

### Author Contributions

MRY and AD designed the research study. MRY and NK analyzed the data. MRY, NK, and AD writing-original draft preparation. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

### Ethics Approval and Consent to Participate

Not applicable.

### Acknowledgment

Thanks to all the peer reviewers for their opinions and suggestions.

### Funding

This research received no external funding.

### Conflict of Interest

The authors declare no conflict of interest.

### References

[1] Molinero K, Hinckley JD. Adolescent Cannabis Use, Comorbid Attention-Deficit/Hyperactivity Disorder, and Other Internaliz-

ing and Externalizing Disorders. *Child and Adolescent Psychiatric Clinics of North America*. 2023; 32: 57–68.

[2] Deshpande G, Wang P, Rangaprakash D, Wilamowski B. Fully Connected Cascade Artificial Neural Network Architecture for Attention Deficit Hyperactivity Disorder Classification From Functional Magnetic Resonance Imaging Data. *IEEE Transactions on Cybernetics*. 2015; 45: 2668–2679.

[3] Dorvashi M, Behzadfar N, Shahgholian G. Classification of alcoholic and non-alcoholic individuals based on frequency and non-frequency features of electroencephalogram signal. *Iranian Journal of Biomedical Engineering*. 2020; 14: 121–130.

[4] Gevensleben H, Holl B, Albrecht B, Schlamp D, Kratz O, Studer P, *et al*. Distinct EEG effects related to neurofeedback training in children with ADHD: a randomized controlled trial. *International Journal of Psychophysiology*. 2009; 74: 149–157.

[5] Jalali P, Sho’ouri N. Neurofeedback Training Protocol Based on Selecting Distinctive Features to Treat or Reduce ADHD Symptoms. *Clinical EEG and Neuroscience*. 2020; 52: 414–421.

[6] Zuberer A, Brandeis D, Drechsler R. Are treatment effects of neurofeedback training in children with ADHD related to the successful regulation of brain activity? A review on the learning of regulation of brain activity and a contribution to the discussion on specificity. *Frontiers in Human Neuroscience*. 2015; 9: 135.

[7] Weber E, Köberl A, Frank S, Doppelmayr M. Predicting successful learning of SMR neurofeedback in healthy participants: methodological considerations. *Applied Psychophysiology and Biofeedback*. 2011; 36: 37–45.

[8] Zhao Y, Cui D, Lu W, Li H, Zhang H, Qiu J. Aberrant gray matter volumes and functional connectivity in adolescent patients with ADHD. *Journal of Magnetic Resonance Imaging*. 2020; 51: 719–726.

[9] McAlonan GM, Cheung V, Cheung C, Chua SE, Murphy DGM, Suckling J, *et al*. Mapping brain structure in attention deficit-hyperactivity disorder: a voxel-based MRI study of regional grey and white matter volume. *Psychiatry Research*. 2007; 154: 171–180.

[10] Cubillo A, Halari R, Smith A, Taylor E, Rubia K. A review of fronto-striatal and fronto-cortical brain abnormalities in children and adults with Attention Deficit Hyperactivity Disorder (ADHD) and new evidence for dysfunction in adults with ADHD during motivation and attention. *Cortex*. 2012; 48: 194–215.

[11] Kamida A, Shimabayashi K, Oguri M, Takamori T, Ueda N, Koyanagi Y, *et al*. EEG Power Spectrum Analysis in Children with ADHD. *Yonago Acta Medica*. 2016; 59: 169–173.

[12] Almeida LG, Ricardo-Garcell J, Prado H, Barajas L, Fernández-Bouzas A, Avila D, *et al*. Reduced right frontal cortical thickness in children, adolescents and adults with ADHD and its correlation to clinical variables: a cross-sectional study. *Journal of Psychiatric Research*. 2010; 44: 1214–1223.

[13] Fernández-Jaén A, López-Martín S, Albert J, Fernández-Mayoralas DM, Fernández-Perrone AL, de La Peña MJ, *et al*. Cortical thickness differences in the prefrontal cortex in children and adolescents with ADHD in relation to dopamine transporter (DAT1) genotype. *Psychiatry Research*. 2015; 233: 409–417.

[14] Almeida Montes LG, Prado Alcántara H, Martínez García RB, De La Torre LB, Avila Acosta D, Duarte MG. Brain cortical thickness in ADHD: age, sex, and clinical correlations. *Journal of Attention Disorders*. 2013; 17: 641–654.

[15] Boedhoe PSW, van Rooij D, Hoogman M, Twisk JWR, Schmaal L, Abe Y, *et al*. Subcortical Brain Volume, Regional Cortical Thickness, and Cortical Surface Area Across Disorders: Findings From the ENIGMA ADHD, ASD, and OCD Working Groups. *The American Journal of Psychiatry*. 2020; 177: 834–843.

- [16] Song S, Qiu J, Lu W. Predicting disease severity in children with combined attention deficit hyperactivity disorder using quantitative features from structural MRI of amygdaloid and hippocampal subfields. *Journal of Neural Engineering*. 2021; 18: 046013.
- [17] Mu S, Wu H, Zhang J, Chang C. Subcortical structural covariance predicts symptoms in children with different subtypes of ADHD. *Cerebral Cortex*. 2023; 33: 8849–8857.
- [18] Lin Q, Bu X, Chen H, Liang Y, Wang W, Yi Y, *et al.* Sex differences in microstructural alterations in the corpus callosum tracts in drug-naïve children with ADHD. *Brain Imaging and Behavior*. 2022; 16: 1592–1604.
- [19] Dubey AK, Saxena A, Kumar A, Rai A. Attention Deficit Hyperactivity Disorder (ADHD) Prediction System. In: 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS) (pp. 1321–1325). IEEE. 2022.
- [20] Behzadfar N. A brief overview on analysis and feature extraction of electroencephalogram signals. *Signal Processing and Renewable Energy*. 2022; 6: 39–64.
- [21] Karimi-Shahraki S, Khezri M. Identification of attention deficit Hyperactivity disorder patients using wavelet-based features of EEG signals. *Journal of Intelligent Procedures in Electrical Technology*. 2021; 12: 1–11.
- [22] Dehghani A, Soltanian-Zadeh H, Hossein-Zadeh GA. Neural modulation enhancement using connectivity-based EEG neurofeedback with simultaneous fMRI for emotion regulation. *NeuroImage*. 2023; 279: 120320.
- [23] Dehghani A, Soltanian-Zadeh H, Hossein-Zadeh GA. Probing fMRI brain connectivity and activity changes during emotion regulation by EEG neurofeedback. *Frontiers in Human Neuroscience*. 2023; 16: 988890.
- [24] Mosayebi R, Dehghani A, Hossein-Zadeh GA. Dynamic functional connectivity estimation for neurofeedback emotion regulation paradigm with simultaneous EEG-fMRI analysis. *Frontiers in Human Neuroscience*. 2022; 16: 933538.
- [25] Dehghani A, Soltanian-Zadeh H, Hossein-Zadeh GA. EEG coherence pattern through recalling positive autobiographical memories and neurofeedback. In: 2021 28th National and 6th International Iranian Conference on Biomedical Engineering (ICBME) (pp. 131–134). IEEE. 2021.
- [26] Dehghani A, Soltanian-Zadeh H, Hossein-Zadeh GA. Increased neuromodulation ability through EEG connectivity neurofeedback with simultaneous fMRI for emotion regulation. *arXiv [q-bio.NC]*. 2022. (preprint)
- [27] Dehghani A, Moghaddam HA, Moin MS. Retinal identification based on rotation invariant moments. In: 2011 5th International Conference on Bioinformatics and Biomedical Engineering (pp. 1–4). IEEE. 2011.
- [28] Dehghani A, Moradi A, Dehghani M, Ahani A. Nonlinear solution for radiation boundary condition of heat transfer process in human eye. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2011; 2011: 166–169.
- [29] Dehghani A, Moradi A, Jafari R. Modelling of transient heat transfer in human eye with a tumour using 3D finite element method. In: 2011 5th International Conference on Bioinformatics and Biomedical Engineering. IEEE. 2011.
- [30] Dehghani A, Moradi A, Jafari R. 3D Simulation of transient heat transfer in human eye using finite element method. In: 2010 International Symposium on Optomechatronic Technologies. IEEE. 2010.
- [31] Yousefi M, Dehghani A, Amini AA. Construction of Multi-Resolution Wavelet Based Mesh Free Method in Solving Poisson and Imaginary Helmholtz Problem. *International Journal of Smart Electrical Engineering*. 2022; 5: 215.
- [32] Yousefi MR. Solving the Inverse Problem of Magnetic Induction Tomography Using Gauss-Newton Iterative Method and Zoning Technique to Reduce Unknown Coefficients. *Journal of Computational Intelligence in Electrical Engineering*. 2022; 14: 67–78.
- [33] Dehghani A. Human recognition based on retinal bifurcations and modified correlation function [Internet]. *arXiv [eess.IV]*. 2021. (preprint)
- [34] Dehghani A, Moghaddam HA, Moin MS. Optic disc localization in retinal images using histogram matching. *EURASIP Journal on Image and Video Processing*. 2012; 2012: 1–11.
- [35] Dehghani A, Ghassabi Z, Moghaddam HA, Moin MS. Human recognition based on retinal images and using new similarity function. *EURASIP Journal on Image and Video Processing*. 2013; 2013: 1–10.
- [36] Dehghani A, Moin MS, Saghafi M. Localization of the optic disc center in retinal images based on the Harris corner detector. *Biomedical Engineering Letters*. 2012; 2: 198–206.
- [37] Dorvashi M, Behzadfar N, Shahgholian G. Electroencephalogram signal during neurofeedback training. *Signal and Data Processing*. 2022; 19: 163–174.
- [38] Dudek E, Dodell-Feder D. The efficacy of real-time functional magnetic resonance imaging neurofeedback for psychiatric illness: A meta-analysis of brain and behavioral outcomes. *Neuroscience and Biobehavioral Reviews*. 2021; 121: 291–306.
- [39] Dehghani A, Soltanian-Zadeh H, Hossein-Zadeh GA. Global Data-Driven Analysis of Brain Connectivity During Emotion Regulation by Electroencephalography Neurofeedback. *Brain Connectivity*. 2020; 10: 302–315.
- [40] Seif M, Yousefi MR, Behzadfar N. EEG Spectral Power Analysis: A Comparison Between Heroin Dependent and Control Groups. *Clinical EEG And Neuroscience*. 2022; 15500594221089366.
- [41] Nan W, Wan F, Vai MI, Da Rosa AC. Resting and Initial Beta Amplitudes Predict Learning Ability in Beta/Theta Ratio Neurofeedback Training in Healthy Young Adults. *Frontiers in Human Neuroscience*. 2015; 9: 677.
- [42] Janssen TWP, Bink M, Weeda WD, Geladé K, van Mourik R, Maras A, *et al.* Learning curves of theta/beta neurofeedback in children with ADHD. *European Child & Adolescent Psychiatry*. 2017; 26: 573–582.
- [43] Marzbani H, Marateb HR, Mansourian M. Neurofeedback: A Comprehensive Review on System Design, Methodology and Clinical Applications. *Basic and Clinical Neuroscience*. 2016; 7: 143–158.
- [44] Friston KJ. Functional and effective connectivity: a review. *Brain Connectivity*. 2011; 1: 13–36.
- [45] Mottola F, Blanchfield A, Hardy J, Cooke A. EEG neurofeedback improves cycling time to exhaustion. *Psychology of Sport and Exercise*. 2021; 55: 101944.
- [46] Deiber MP, Hasler R, Colin J, Dayer A, Aubry JM, Baggio S, *et al.* Linking alpha oscillations, attention and inhibitory control in adult ADHD with EEG neurofeedback. *NeuroImage. Clinical*. 2020; 25: 102145.
- [47] Gruzelier J. A theory of alpha/theta neurofeedback, creative performance enhancement, long distance functional connectivity and psychological integration. *Cognitive Processing*. 2009; 10: S101–S109.
- [48] Hynd GW, Lorys AR, Semrud-Clikeman M, Nieves N, Huettner MI, Lahey BB. Attention deficit disorder without hyperactivity: a distinct behavioral and neurocognitive syndrome. *Journal of Child Neurology*. 1991; 6: S37–S43.
- [49] Edgar JC, Fisk CL, IV, Chen YH, Stone-Howell B, Hunter MA, Huang M, *et al.* By our bootstraps: Comparing methods for measuring auditory 40 Hz steady-state neural activity. *Psychophysiology*. 2017; 54: 1110–1127.
- [50] Nan W, Wan F, Tang Q, Wong CM, Wang B, Rosa A. Eyes-Closed Resting EEG Predicts the Learning of Alpha Down-Regulation in Neurofeedback Training. *Frontiers in Psychology*. 2018; 9: 1607.

- [51] Collura TF, Guan J, Tarrant J, Bailey J, Starr F. EEG biofeedback case studies using live Z-score training and a normative database. *Journal of Neurotherapy*. 2010; 14: 22–46.
- [52] Arns M, Gunkelman J, Breteler M, Spronk D. EEG phenotypes predict treatment outcome to stimulants in children with ADHD. *Journal of Integrative Neuroscience*. 2008; 7: 421–438.
- [53] Okumura Y, Kita Y, Omori M, Suzuki K, Yasumura A, Fukuda A, *et al.* Predictive factors of success in neurofeedback training for children with ADHD. *Developmental Neurorehabilitation*. 2019; 22: 3–12.
- [54] Tuckute G, Hansen ST, Pedersen N, Steenstrup D, Hansen LK. Single-Trial Decoding of Scalp EEG under Natural Conditions. *Computational Intelligence and Neuroscience*. 2019; 2019: 9210785.
- [55] Cristianini N, Shawe-Taylor J. Background Mathematics. In: *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods* (pp. 165–172). Cambridge University Press: Cambridge. 2013.
- [56] Sideroff S. Case studies in applied psychophysiology: Neurofeedback and biofeedback treatments for advances in human performance. *The Sport Psychologist*. 2013; 27: 103–104.
- [57] Ebrahimzadeh E, Dehghani A, Asgarinejad M, Soltanian-Zadeh H. Non-linear processing and reinforcement learning to predict rTMS treatment response in depression. *Psychiatry Research: Neuroimaging*. 2024; 337: 111764.
- [58] Ebrahimzadeh E, Fayaz F, Rajabion L, Seraji M, Aflaki F, Hamoud A, *et al.* Machine learning approaches and non-linear processing of extracted components in frontal region to predict rTMS treatment response in major depressive disorder. *Frontiers in Systems Neuroscience*. 2023; 17: 919977.
- [59] Asgarinejad M, Saviz M, Sadjadi SM, Saliminia S, Kakaei A, Esmacili P, *et al.* Repetitive transcranial magnetic stimulation (rTMS) as a tool for cognitive enhancement in healthy adults: a review study. *Medical & Biological Engineering & Computing*. 2024; 62: 653–673.
- [60] Ter Harmsel JF, Noordzij ML, Goudriaan AE, Dekker JJM, Swinkels LTA, van der Pol TM, *et al.* Biofeedback and ambulatory biofeedback to enhance emotion regulation: A review of studies investigating non-psychiatric and psychiatric populations. *International Journal of Psychophysiology*. 2021; 159: 94–106.
- [61] Hurt E, Arnold LE, Lofthouse N. Quantitative EEG neurofeedback for the treatment of pediatric attention-deficit/hyperactivity disorder, autism spectrum disorders, learning disorders, and epilepsy. *Child and Adolescent Psychiatric Clinics of North America*. 2014; 23: 465–486.
- [62] Bartholdy S, Musiat P, Campbell IC, Schmidt U. The potential of neurofeedback in the treatment of eating disorders: a review of the literature. *European Eating Disorders Review*. 2013; 21: 456–463.
- [63] Yousefi MR, Dehghani A, Golnejad S, Hosseini MM. Comparing EEG-based epilepsy diagnosis using neural networks and wavelet transform. *Applied Sciences*. 2023; 13: 10412.
- [64] Yousefi MR, Dehghani A, Taghaavifar H. Enhancing the accuracy of electroencephalogram-based emotion recognition through Long Short-Term Memory recurrent deep neural networks. *Frontiers in Human Neuroscience*. 2023; 17: 1174104.
- [65] Budzynski TH, Budzynski HK, Evans JR, Abarbanel A. *Introduction to quantitative EEG and neurofeedback: Advanced theory and applications*. Academic Press: Cambridge, MA, USA. 2009.
- [66] Drechsler R, Straub M, Doehnert M, Heinrich H, Steinhausen HC, Brandeis D. Controlled evaluation of a neurofeedback training of slow cortical potentials in children with Attention Deficit/Hyperactivity Disorder (ADHD). *Behavioral and Brain Functions*. 2007; 3: 35.