

Classification of ADHD and Healthy Children Using Multi-Band and Spatial Features of EEG

Md Bayazid Hossain

Computer Science and Engineering
Pabna University of Science and
Technology, Pabna 6600, Bangladesh.
Email: bayazid204@gmail.com

Md Anwarul Islam Himel

Computer Science and Engineering
Pabna University of Science and
Technology, Pabna 6600, Bangladesh.
Email: himelhim31@gmail.com

Md Abdur Rahim

Computer Science and Engineering
Pabna University of Science and
Technology, Pabna 6600, Bangladesh.
Email: rahim@pust.ac.bd

Shabbir Mahmood

Computer Science and Engineering
Pabna University of Science and
Technology, Pabna 6600, Bangladesh.
Email: shabbir.cse@pust.ac.bd

Abu Saleh Musa Miah

Computer science and engineering
The University of Aizu
Aizuwakamatsu, Japan.
Email: abusalehcse.ru@gmail.com

Jungpil SHIN

Computer science and engineering
The University of Aizu
Aizuwakamatsu, Japan.
Email: jpshin@u-aizu.ac.jp

Abstract—Attention Deficit Hyperactivity Disorder (ADHD) is a common neurodevelopmental disorder in children, characterized by difficulties in attention, hyperactivity, and impulsivity. Early and accurate diagnosis of ADHD is critical for effective intervention and management. Electroencephalogram (EEG) signals have emerged as a non-invasive and efficient tool for ADHD detection due to their high temporal resolution and ability to capture neural dynamics. In this study, we propose a method for classifying ADHD and healthy children using EEG data from the benchmark dataset. There were 61 children with ADHD and 60 healthy children, both boys and girls, aged 7 to 12. The EEG signals, recorded from 19 channels, were processed to extract Power Spectral Density (PSD) and Spectral Entropy (SE) features across five frequency bands, resulting in a comprehensive 190-dimensional feature set. To evaluate the classification performance, a Support Vector Machine (SVM) with the RBF kernel demonstrated the best performance with a mean cross-validation accuracy of 98.2% and a standard deviation of 0.0089, indicating high robustness and precision. These results highlight the potential of spatial features in conjunction with machine learning for accurately classifying ADHD using EEG data. This work contributes to developing non-invasive, data-driven tools for early diagnosis and assessment of ADHD in children.

Index Terms—ADHD Classification, Electroencephalography (EEG), Machine Learning, Spatial Feature of EEG, Non-invasive Diagnosis, Cross-Validation Accuracy

I. INTRODUCTION

ADHD is a neurodevelopmental disorder marked by persistent patterns of inattention, hyperactivity, or impulsivity that interfere with daily functioning. These behaviors occur frequently across various settings, such as home, school, work, and go beyond typical behavior for most individuals [1]. Children with ADHD often struggle with sitting still, planning ahead, completing tasks, or maintaining full awareness of their surroundings. The prevalence of ADHD is estimated to be around 12.1% in boys and 3.9% in girls [2]. ADHD is a prevalent condition frequently identified in children. Its

symptoms typically start in childhood and may persist into adolescence and adulthood. ADHD often occurs alongside other conditions, including conduct issues, learning disabilities, sleep disturbances, anxiety, or depression, which can complicate its diagnosis and treatment [3]. For individuals with ADHD, symptoms often impact daily functioning. These symptoms can create challenges in completing tasks, disrupt school, work, and other activities, and put a strain on social relationships. Children with ADHD face a higher likelihood of injuries, social difficulties, family stress, and academic struggles. Adolescents and adults with ADHD are also at a greater risk of engaging in risky behaviors [4]. EEG is a valuable technique that offers insights into the brain's background activity and serves as an indicator of the foundation for cognition and behavior [5]. According to previous study, EEG plays an important role in the evaluation of neural function in children with ADHD [6]. Therefore, it can be a useful gadget for investigating and diagnosing the abnormal behavior of ADHD children. In this paper, the continuity of attention will be investigated for ADHD and control children and using this continuity a new approach will be introduced. In the following, firstly EEG recording will be described, then the methodology of analysis will be introduced and finally the results will be demonstrated.

II. LITERATURE REVIEW

There are many studies which employed EEG analysis for diagnosing ADHD. Recent studies have extensively employed EEG signal analysis to investigate ADHD. In Armin Allahverdy et al. (2016) the utility of nonlinear EEG features, achieving 96.7% accuracy in ADHD classification using visual attention tasks [7].

Ali Ekhlesi et al. (2020) explores disrupted information flow in ADHD children compared to healthy peers using directed phase transfer entropy (dPTE) on EEG data. Significant differences in connectivity in delta, beta, and theta bands were

also observed, particularly in frontal and occipital regions. These findings underline the disrupted effective connectivity in ADHD brains, providing insights into its neurological underpinnings [8]. Mohammad Reza Mohammadi et al., (2016) proposed a method using nonlinear features of EEG, such as fractal dimensions, approximate entropy, and Lyapunov exponent. Feature selection methods such as mRMR and DISR were employed to enhance classification accuracy. A Multi-Layer Perceptron (MLP) neural network achieved accuracies of 92.28% and 93.65% with mRMR and DISR methods, respectively [9]. Farhana et al. (2023) presented an emotion recognition method using EEG signals by extracting narrow-band spatial features from theta, alpha, beta, and gamma bands. The method involves preprocessing EEG data, sub-band decomposition using Butterworth bandpass filters, and calculating short-term entropy and energy features. A common Spatial Pattern (CSP) was employed to derive discriminative features, which were classified using a Support Vector Machine (SVM) with a polynomial kernel. The proposed approach outperforms several existing methods in binary emotion classification (valence and arousal), achieving accuracies of 96.15% for DEAP and 99.95% for SEED datasets [10]. In 2013, Nazhvani et al, used N2 and P2 peaks of ERP to diagnose ADHD and achieved 92.9% accuracy [11]. These findings collectively underscore the potential of advanced EEG analysis in developing objective and precise ADHD diagnostic tools.

III. DATASET DESCRIPTION AND DATA PREPROCESSING

A. Dataset Description

The study involved 61 children diagnosed with ADHD and 60 healthy controls, consisting of both boys and girls aged 7 to 12 years. ADHD diagnoses were confirmed by an experienced psychiatrist using DSM-IV criteria, and the children had been on Ritalin for up to six months. In contrast, none of the children in the control group had a history of psychiatric disorders, epilepsy, or high-risk behaviors [12]. EEG recordings were conducted following the 10-20 standard, utilizing 19 channels (Fz, Cz, Pz, C3, T3, C4, T4, Fp1, Fp2, F3, F4, F7, F8, P3, P4, T5, T6, O1, O2) at a sampling frequency of 128 Hz. The reference electrodes, A1 and A2, were positioned on the earlobes. Since visual attention is often impaired in children with ADHD, the EEG protocol was designed around a visual attention task. During this task, participants viewed images of cartoon characters and were asked to count the number of characters in each image. The number of characters varied randomly between 5 and 16, and the images were sufficiently large enough to they were clearly visible and easily countable. To have a continuous stimulus during the signal recording, each image was displayed immediately and uninterrupted after the child's response. As a result, the total duration of the recording depended on the child's response speed throughout the task.

B. Data Preprocessing

The EEG signals were digitized at the sampling rate of 128 Hz and recorded in the frequency range of 0.1-60 Hz. A bandpass Butterworth filter of 0.5 Hz to 50 Hz with order five was applied to continuous EEG data in order to eliminate artefacts. We have also tried using less than 50 Hz which decreases the classification accuracy. Then, an average number of 29 epochs of 1280 points (10 seconds) with 640 points (50%) overlap were selected for each subject. Trial lengths exceeding 10 seconds are usually employed to analyze slower dynamics in the signal or to achieve higher frequency resolution. The resulting data shape being (61 + 60) subject \times average 29 epochs \times 19 channel \times 1280 points.

IV. METHODOLOGY

The proposed methodology for ADHD classification using EEG signals involves several key stages given in fig 1, starting from preprocessing to final classification. Initially, preprocessed EEG trials are subjected to subband decomposition, where the EEG signals are divided into multiple frequency bands (delta, theta, alpha, beta, gamma) to isolate distinct rhythmic components of brain activity. For each subband, two essential features, Power Spectral Density (PSD) and Spectral Entropy, are extracted. PSD represents the distribution of power across various frequency components, while spectral entropy quantifies the randomness or irregularity of the signal, both of which are crucial in distinguishing ADHD characteristics. Once the features are extracted from all subbands,

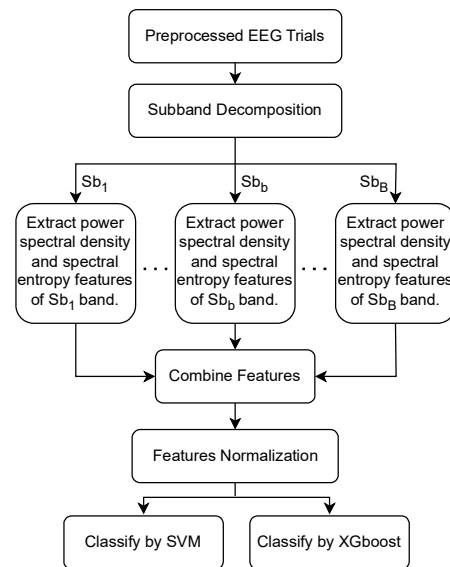


Fig. 1. The methodology figure.

they are combined into a single feature set to enhance the robustness of the model by incorporating information from multiple subbands. The combined feature set is then passed

through a normalization step to scale the features and reduce the effect of varying magnitudes across different features. This step is critical for ensuring fair contribution of all features during classification.

Finally, the normalized feature set is fed into two classifiers Support Vector Machine (SVM) and XGBoost to perform ADHD classification. SVM is a widely used classifier that finds an optimal hyperplane to separate different classes, while XGBoost is a scalable and efficient gradient-boosting framework known for its high performance in complex classification tasks. We applied different SVM kernels (linear, polynomial, sigmoid, and RBF) to determine the best fit for the EEG data. The RBF kernel outperformed others, achieving 98.2% accuracy, effectively capturing nonlinear patterns in brain activity. Using both classifiers, the methodology ensures that the best possible classification performance can be achieved for the ADHD data set.

A. Feature Extraction

In previous study they had used non linear features to classify ADHD. Here we have use two spatial features band power and spectral entropy. PSD helps analyze brain activity by measuring the power distribution across different frequency bands, which is known to differ in ADHD children [13]. And Spectral Entropy quantifies signal randomness, reflecting cognitive control and brain organization. Since ADHD is linked to altered neural oscillations and irregular brain activity, these features provide valuable insights for classification [14]. First of all we found power spectral density(PSD) using Eqn(1). From the PSD the calculate band power Eqn(2). And using Eqn(3) calculate the band energy feature. After extracting features of all bands then normalize the data using Eqn(4).

$$PSD(f) = \frac{1}{L} \sum_{k=1}^L \frac{1}{N_w \cdot F_s} |FFT_k(f)|^2 \quad (1)$$

Where:

- L : Number of segments the signal is divided into.
- $FFT_k(f)$: Discrete Fourier Transform (DFT) of the k -th segment at frequency f .
- N_w : Number of samples in each segment.
- F_s : Sampling frequency of the signal.
- $|\cdot|^2$: Squared magnitude (to compute power).
- The factor $\frac{1}{N_w \cdot F_s}$: Normalizes the power to account for the sampling frequency and segment length.

1) *Band Power*: The band power calculates the total power of the signal within a specific frequency band $[f_{low}, f_{high}]$. The equation is :

$$P_{band} = \sum_{f \in [f_{low}, f_{high}]} PSD(f) \quad (2)$$

Where:

- $PSD(f)$: Power Spectral Density at frequency f ,
- $[f_{low}, f_{high}]$: Frequency range of the band (e.g., alpha, beta, etc).

2) *Band Entropy*: The band entropy is calculated using the normalized band-limited PSD. The formula for entropy is derived from Shannon's entropy:

$$E_{band} = \sum_{f \in [f_{low}, f_{high}]} p(f) \cdot \log(p(f)) \quad (3)$$

Where:

- $p(f) = \frac{PSD(f)}{\sum_{f \in [f_{low}, f_{high}]} PSD(f)}$: Normalized PSD values within the band,
- $\sum p(f) = 1$: Ensures $p(f)$ is a probability distribution.
- A small value (e.g., 1×10^{-8} is added to $p(f)$ in the code to avoid $\log(0)$.

B. Feature Normalization

The extracted features normalized by following formula:

$$z = \frac{x - \mu}{\sigma} \quad (4)$$

Where:

- z : The normalized value.
- x : The original feature value.
- μ : Mean of the feature values.
- σ : The standard deviation of the feature values.

This feature normalization ensures that the normalized features have a mean of 0 and a standard deviation of 1.

C. Classification

1) *Classification By SVM*: The performs a comparative evaluation of different kernel functions (linear, rbf, poly, and sigmoid) for an SVM model. To evaluation use 10-fold cross-validation. For each kernel, we initialize an SVM model (SVC)

TABLE I
COMPARISON OF MODEL PERFORMANCE

Kernel	Mean Accuracy	Standard Deviation
Linear	93.36%	0.0100
Rbf	98.20%	0.0089
Poly	95.63%	0.0129
Sigmoid	72.50%	0.0489

with the specified kernel type, along with fixed regularization parameter $C = 1.0$, a scaling gamma value (gamma='scale').

The model's performance is assessed on the training data using cross-validation. For each kernel, here calculates the individual accuracy scores across the 10 folds, as well as the mean accuracy and standard deviation. The result of this experiment given on Table I. This analysis helps compare the effectiveness of different kernel functions in terms of their classification accuracy and variability.

2) *XGBoost Classifier*: The Classifier with specific hyperparameters tailored for training a robust model. The classifier is set to use 100 trees (n_estimators) and a learning rate of 0.1, ensuring a balanced step size for weight updates during boosting. The maximum depth of each tree is limited to 6 (max_depth), controlling the complexity of the model and preventing overfitting. Additionally, the training process uses

80% of the samples (subsample) and 80% of the features (colsample_bytree) for each tree, introducing randomness to improve generalization. This configuration is designed to provide a balance between accuracy and generalization, making the model robust and effective.

V. EXPERIMENTAL RESULTS

The experimental evaluation of the proposed methodology was conducted using two classifiers: SVM with an RBF kernel and XGBoost. The results indicate that both models achieved high performance, but SVM outperformed XGBoost in terms of accuracy and other metrics in Table II.

TABLE II
COMPARISON OF MODEL PERFORMANCE

Methodology	Accuracy	Precision	Recall	F1-Score
SVM(rbf kernel)	98.75%	0.98	0.99	0.99
XGBClassifier	96.92%	0.95	1.00	0.97

Figure 2 shows the accuracy curve for both training and validation phases across multiple epochs. It can be observed that the validation accuracy stabilizes after an initial rapid increase, reaching a maximum value close to 99%, indicating that the model generalizes well to unseen data without significant overfitting.

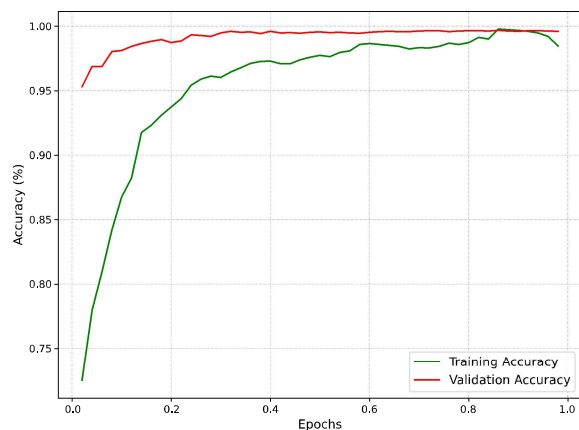


Fig. 2. Accuracy Curve

SVM with an RBF kernel outperformed XGBoost, achieving higher accuracy (98.2% vs. 96.92%). This can be attributed to SVM's ability to find an optimal decision boundary in high-dimensional feature spaces, which is crucial given the complexity of EEG signals. The RBF kernel effectively captures non-linear patterns in brain activity, making it well-suited for differentiating ADHD and healthy children. On the other hand, while XGBoost is powerful in handling structured data, it may not generalize as well to highly non-linear relationships in EEG data. Additionally, SVM's robustness against small dataset sizes and high-dimensional features likely contributed to its superior performance.

Overall, these results demonstrate the effectiveness of the proposed methodology, with SVM providing slightly superior performance in terms of accuracy and F1-score.

VI. CONCLUSIONS

This study demonstrated that multi-band EEG features, combined with machine learning, achieve high accuracy (98.2%) in classifying ADHD and healthy children. The use of Power Spectral Density and Spectral Entropy features proved to be highly effective in capturing ADHD-related neural patterns, reinforcing the potential of EEG as a non-invasive and objective diagnostic tool. The superior performance of SVM with an RBF kernel highlights its ability to handle the complex, non-linear nature of EEG data, making it a strong candidate for real-world applications.

Despite these promising results, some limitations should be considered. The size of the data set, though reliable for initial analysis, remains relatively small for a broader generalization. Future studies should validate the approach in larger and more diverse populations.

REFERENCES

- [1] Health Topics. National Institute of Mental Health. <https://www.nimh.nih.gov/health/topics/attention-deficit-hyperactivity-disorder-adhd#hts-intro>
- [2] R. Kessler, L. Adler, P. Berglund, J. G. Green, K. McLaughlin, J. Fayyad, et al., "The effects of temporally secondary co-morbid mental disorders on the associations of DSM-IV ADHD with adverse outcomes in the US National Comorbidity Survey Replication Adolescent Supplement (NCS-A)," *Psychological medicine*, vol. 44, pp. 1779-1792, 2014.
- [3] Timothy E Wilens, Thomas J Spencer "Understanding Attention-Deficit/Hyperactivity Disorder From Childhood to Adulthood" <https://pmc.ncbi.nlm.nih.gov/articles/PMC3724232/>
- [4] Smitha Bhandari, MD "ADHD in Teens" <https://www.webmd.com/add-adhd/childhood-adhd/adhd-teens> July , 2023
- [5] Gregory A Light, Lisa E Williams, Falk Minow "Electroencephalography (EEG) and Event-Related Potentials (ERP's) with Human Participants" *Curr Protoc Neurosci*. 2010 Jul, CHAPTER Unit-6.2524. doi 10.1002/0471142301.ns0625s52
- [6] A. Lenartowicz and S. K. Loo, "Use of EEG to diagnose ADHD," *Current psychiatry reports*, vol. 16, pp. 1-11, 2014.
- [7] Armin Allahverdy, Alireza Khorrami, Ali Motie "Detecting ADHD Children using the Attention Continuity as Nonlinear Feature of EEG" *Frontiers in Biomedical Technologies*, Volume 3, Issue 1-2, June 2016.
- [8] Ali Ekhlasi, Ali Motie Nasrabadi, Mohammad Reza Mohammadi "Direction of information flow between brain regions in ADHD and healthy children based on EEG by using directed phase transfer entropy" *Cognitive Neurodynamics*, Springer Nature B.V., 16 April 2021.
- [9] Mohammad Reza Mohammadi, Ali Khaleghi, Ali Moti Nasrabadi, Safa Rafeivand, Moslem Begol and Hadi Zarafshan "EEG Classification of ADHD and Normal Children Using Non-linear Features and Neural Network" *The Korean Society of Medical & Biological Engineering and Springer* 2016.
- [10] Ifat Farhana, Jungpil SHIN, Sabbir Mahmood, Md. Rabiul Islam and Md. Khademul Islam Molla "Emotion Recognition Using Narrowband Spatial Features of Electroencephalography" *IEEE Access*, vol. 11, pp. 44019-44033, 2023.
- [11] A. D. Nazhvani, R. Boostani, S. Afrasiabi, and K. Sadatnezhad, "Classification of ADHD and BMD patients using visual evoked potential," *Clinical neurology and neurosurgery*, vol. 115, pp. 2329-2335, 2013.
- [12] Ali Motie Nasrabadi, Armin Allahverdy, Mehdi Samavati, Mohammad Reza Mohammadi "EEG data for ADHD / Control children" <https://iee-dataport.org/open-access/eeg-data-adhd-control-children#files>
- [13] Elena I. Rodríguez-Martínez "Frequency coupling of low and high frequencies in the EEG of ADHD children and adolescents in closed and open eyes conditions" <https://doi.org/10.1016/j.ridd.2019.103520>
- [14] Anastasia G. Papaioannou, Eva Kalantzi "Complexity analysis of the brain activity in Autism Spectrum Disorder (ASD) and Attention Deficit Hyperactivity Disorder (ADHD) due to cognitive loads/demands induced by Aristotle's type of syllogism/reasoning. A Power Spectral Density and multiscale entropy (MSE) analysis" <https://doi.org/10.1016/j.heliyon.2021.e07984>