Journal of Clinical and Nursing Research, 2025, Volume 9, Issue 1
http://ojs.bbwpublisher.com/index.php/JCNR
Online ISSN: 2208-3693
Print ISSN: 2208-3685

Utilizing Machine Learning Techniques to Enhance Attention-Deficit Hyperactivity Disorder Diagnosis Using Resting-State EEG Data

Lina Han[1], Liyan Li[1], Yanyan Chen[1], Xiaohan Wu[1], Yang Yu[1], Xu Liu[2], Zihan Yang[2], Ling Li[2], Xinxian Peng[1]*

[1]Changchun Sixth Hospital, Changchun 130000, Jilin Province, China

[2]The School of Communication Engineering, Jilin University, Changchun 130000, Jilin Province, China

***Corresponding author:** Xinxian Peng, peng18946617300@163.com

Abstract: *Objective:* This study investigates the auxiliary role of resting-state electroencephalography (EEG) in the clinical diagnosis of attention-deficit hyperactivity disorder (ADHD) using machine learning techniques. *Methods:* Resting-state EEG recordings were obtained from 57 children, comprising 28 typically developing children and 29 children diagnosed with ADHD. The EEG signal data from both groups were analyzed. To ensure analytical accuracy, artifacts and noise in the EEG signals were removed using the EEGLAB toolbox within the MATLAB environment. Following preprocessing, a comparative analysis was conducted using various ensemble learning algorithms, including AdaBoost, GBM, LightGBM, RF, XGB, and CatBoost. Model performance was systematically evaluated and optimized, validating the superior efficacy of ensemble learning approaches in identifying ADHD. *Conclusion:* Applying machine learning techniques to extract features from resting-state EEG signals enabled the development of effective ensemble learning models. Differential entropy and energy features across multiple frequency bands proved particularly valuable for these models. This approach significantly enhances the detection rate of ADHD in children, demonstrating high diagnostic efficacy and sensitivity, and providing a promising tool for clinical application.

Keywords: Attention-deficit hyperactivity disorder; Machine learning; EEG signals; Feature extraction; Ensemble learning models; Diagnosis

Online publication: February 17, 2025

1. Introduction

Attention-Deficit Hyperactivity Disorder (ADHD), commonly referred to as hyperactivity disorder, is among the most prevalent neuropsychiatric disorders in childhood [1]. It is characterized by age-inappropriate inattention, reduced attention span, excessive activity irrespective of context, emotional impulsivity, cognitive impairments, and learning difficulties, while intellectual capacity typically remains normal or near normal [2,3]. In recent years, ADHD has been widely recognized as a neurodevelopmental disorder with a biological basis and significant impairments in executive function [4]. Given its high prevalence and severe impacts, ADHD substantially affects children's academic performance and overall well-being [5].

Advancing the diagnosis and treatment of ADHD is of paramount importance. Current diagnostic methods rely heavily on subjective approaches, such as interviews, observations, and rating scales, which are often time-consuming, labor-intensive, and susceptible to bias [6]. To address these challenges, researchers have explored neurophysiological methods, particularly the utility of electroencephalography (EEG) in identifying ADHD. EEG has demonstrated high sensitivity and specificity, gaining approval from the U.S. Food and Drug Administration (FDA) for its use in ADHD diagnostics [7]. A notable feature of ADHD is the high prevalence of EEG abnormalities observed in affected individuals.

Technological advancements have driven substantial progress in EEG research, particularly in its medical applications. The integration of EEG signals into ADHD diagnosis, research, and treatment has become a shared focal point in these fields. However, studies focusing on resting-state EEG in ADHD patients remain limited. Leveraging artificial intelligence algorithms to analyze resting-state EEG signals holds both theoretical and practical value for advancing ADHD diagnostics.

As machine learning continues to evolve, it has been increasingly adopted in healthcare-related fields. Within psychiatric research, machine learning has shown promising results in the classification and diagnosis of various mental disorders. EEG signals, which first garnered researchers' attention in the early 20th century, have since undergone extensive algorithmic and theoretical development. These advancements have positioned EEG as a valuable tool for diagnosing brain-related disorders. By processing and analyzing EEG signals, researchers can extract rich physiological information, enabling precise and actionable conclusions.

2. Materials and methods

2.1. Study design and participants

2.1.1. Subjects

2.1.1.1. ADHD group

The ADHD group comprised 29 patients (17 boys and 12 girls, aged 4–13 years) who visited the outpatient clinic between 2022 and 2024. Diagnoses were made by attending physicians or specialists based on the criteria outlined in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) [8]. The diagnostic criteria included:

(1) Developmentally inappropriate symptoms of hyperactivity/impulsivity and/or inattention persisting for at least six months;

(2) Symptoms manifesting across multiple settings (e.g., home and school);

(3) Significant impairment in daily functioning caused by the symptoms;

(4) Initial onset of symptoms and associated impairments during early to middle childhood;

(5) Symptoms not attributable to other medical or psychiatric conditions.

Patients who had been taking psychotropic medications or stimulants were excluded unless such medications had been discontinued for at least 48 hours prior to enrollment. The study was approved by the Ethics Committee of Changchun Sixth Hospital (ethics number: 202203).

2.1.1.2. Control group

The control group included 28 healthy children (17 boys and 11 girls, aged 4–13 years) who underwent routine health checkups during the same period. Participants in the control group were matched with the ADHD group in terms of age, IQ, and gender. Structured interviews confirmed the absence of significant psychological or behavioral symptoms, severe physical illnesses, or psychiatric disorders.

Exclusion criteria for all participants encompassed physical illnesses, neurological disorders, genetic conditions, or a history of psychiatric disorders. Written informed consent was obtained from both children and their parents. The study protocol received approval from the ethics committee. General clinical data and resting-state EEG data were collected from all participants.

2.2. Data collection

Resting-state EEG recordings were obtained in a quiet environment, with participants awake and their eyes closed, using a 32-channel EEG system (Nicolet V32, Natus Medical Incorporated). Electrodes were positioned following the international 10–20 system, with conductive paste ensuring electrode-scalp resistance below 5 kΩ. The EEG data were preprocessed using the EEGLAB toolbox in MATLAB to remove artifacts caused by powerline noise, eye movements, muscle activity, and perspiration, yielding artifact-free EEG signals suitable for analysis.

2.3. Methods

2.3.1. Signal processing

EEG signals are inherently weaker and more random than other bioelectric signals, such as electrocardiograms (ECG) and electrooculograms (EOG). As a result, collected EEG data often contain components from EOG, ECG, and electromyographic (EMG) signals, which can interfere with subsequent analyses. To ensure the accuracy of results, preprocessing was performed using the EEGLAB toolbox in MATLAB. The following steps were undertaken:

(1) Manual inspection and removal of noisy segments: Each channel's data were visually inspected, and segments exhibiting excessive signal fluctuations were manually excluded.

(2) Electrode localization: Electrode positions were mapped to ensure accurate spatial alignment of the EEG data.

(3) Exclusion of unnecessary channels: Channels not required for analysis, such as A1 and A2, were removed.

(4) Bandpass filtering: A finite impulse response (FIR) bandpass filter was applied to retain signal frequencies within the range of 0.5 Hz to 45 Hz, relevant for analysis.

(5) Resampling: Based on the Nyquist sampling theorem, a sampling rate exceeding 90 Hz was required to capture the target frequency range. For computational efficiency, the signals were resampled at 128 Hz.

(6) Faulty channel detection and interpolation: Faulty channels were identified, and their signals were corrected using interpolation based on data from neighboring channels, preserving the original features of the EEG signals.

(7) Independent Component Analysis (ICA): ICA was applied to decompose the EEG signals into independent components. The ICLable algorithm was used to identify and remove artifacts such as EOG and ECG. This algorithm classified components into seven types: EEG, EMG, EOG, ECG, line noise, channel noise,

and others. Components classified as EEG with a probability of at least 80% were retained for further analysis.

2.3.2. Machine learning model

Ensemble learning combines multiple learners to solve a single problem. The core principle involves constructing and integrating multiple models to enhance prediction accuracy and stability [9]. Unlike individual models, ensemble learning algorithms train multiple weak classifiers on different data subsets and aggregate their decisions. This approach effectively reduces bias and variance, improving generalization and providing robust resistance to noise, thus minimizing the risk of overfitting. Ensemble learning is particularly effective for high-dimensional, noisy real-world datasets [10].

In this study, ensemble learning algorithms, including AdaBoost, GBM, LightGBM, RF, XGB, and CatBoost, were applied to the classification task. Systematic evaluation and optimization of these models demonstrated the superior effectiveness of ensemble learning algorithms in this context.

2.3.3. Differential entropy

Differential entropy, a key concept in information theory, measures the information content of continuous random variables [11]. It generalizes the concept of discrete Shannon entropy, making it particularly suitable for analyzing continuous signals. Unlike Shannon entropy, which applies to discrete random variables, differential entropy accommodates variables with continuous value ranges. This property renders it highly applicable in fields such as signal processing, machine learning, and statistical analysis.

Differential entropy quantifies the average uncertainty of a random variable [12]. Higher differential entropy typically indicates greater complexity or information content within a signal, which is especially relevant for analyzing natural signals such as EEG data. By computing differential entropy, it is possible to measure the complexity of a signal, which provides valuable insights for applications such as feature extraction and pattern recognition [13]. The formula for calculating differential entropy is as follows:

$$H(X) = -\int_{-\infty}^{\infty} f(x)\, log\, f(x) dx \tag{1}$$

Here, *H(X)* represents the differential entropy, *X* is the continuous random variable, and *f(x)* denotes the probability density function of *X*.

2.3.4. Wavelet packet decomposition for energy calculation

Wavelet packet decomposition (WPD) is a sophisticated signal processing technique that enables comprehensive frequency analysis by recursively dividing a signal into its frequency components [14]. Unlike standard wavelet decomposition, which focuses on either low-frequency or high-frequency components, WPD subdivides both, allowing for a more detailed and adaptable representation of the signal's structure.

The process of wavelet packet decomposition involves the iterative application of low-pass and high-pass filters, which split the signal into two sub-bands at each decomposition level: a low-frequency sub-band and a high-frequency sub-band [15]. This recursive division generates a complete binary tree structure, offering enhanced resolution and flexibility for analyzing complex signals.

The mathematical framework of WPD can be expressed as follows:

$$\phi^{2k}(t) = \sqrt{2} \sum_n h(n) \phi^k (2t - n) \tag{2}$$

$$\phi^{2k+1}(t) = \sqrt{2}\sum_n g(n)\phi^k(2t-n) \quad (3)$$

Here, $h(n)$ and $g(n)$ are the low-pass and high-pass filters, $\varphi^0(t)=\varphi(t)$ is the scaling function, and $\varphi^1(t)=\varphi(t)$ is the wavelet function. The recursive decomposition of $x(t)$ the input signal into low-frequency and high-frequency components is described by:

$$x_{j+1,2k}(t) = \sum_m h(m-2n)\, x_{j,k}(t) \quad (4)$$

$$x_{j+1,2k+1}(t) = \sum_m g(m-2n)\, x_{j,k}(t) \quad (5)$$

where $x_{j,k}(t)$ represents the wavelet coefficients for the k sub-band at the j level. The signal can then be reconstructed as:

$$x(t) = \sum_{k=0}^{2j-1} x_{j,k}(t) \quad (6)$$

Here, j and k denote the decomposition level and sub-band, respectively. Once the signal is decomposed into specific frequency bands, the energy of the wavelet packet coefficients is calculated to quantify the signal's intensity within each frequency band. The energy calculation formula is:

$$E_{j,k} = \sum_m \left|x_{j,k}(t)\right|^2 \quad (7)$$

The energy distribution across different frequency bands provides critical insights into the signal's characteristics, supporting advanced tasks such as feature extraction and pattern recognition. The flexibility and precision of WPD make it particularly effective for analyzing complex signals like EEG, where detailed frequency resolution is essential.

2.3.5. Evaluation metrics

For this binary classification problem, precision and recall were prioritized as key evaluation metrics. Precision measures the proportion of true positive predictions among all positive predictions, while recall evaluates the classifier's ability to identify all actual positive instances.

The Receiver Operating Characteristic (ROC) curve is a graphical representation that illustrates the trade-off between the true positive rate and the false positive rate across various thresholds. The area under the ROC curve (AUC) is a widely used metric for binary classification tasks, particularly in medical diagnostics.

AUC values range from 0 to 1, with higher values indicating better classification performance. An AUC value closer to 1 signifies superior classifier effectiveness, making this metric crucial for evaluating and comparing models in scenarios requiring high predictive accuracy, such as EEG-based diagnostic tasks.

4. Results

To evaluate the impact of feature extraction on the performance of machine learning models, raw EEG data (without feature extraction) were initially input into the models. The resulting accuracy and AUC values are summarized in **Table 1**, with the corresponding ROC curve illustrated in **Figure 1**.

After applying feature extraction, notable changes in model performance were observed. The updated results, including accuracy and AUC values, are presented in **Table 2**, with the corresponding ROC curve depicted in

Figure 2.

This comparative analysis highlights the importance of feature extraction in enhancing the predictive capabilities of machine learning models when applied to EEG data.

Table 1. Results obtained using raw EEG data

	AdaBoost	**CatBoost**	**GBM**	**LightGBM**	**RF**	**XGB**
Accuracy	0.76	0.81	0.79	0.82	0.79	0.81
AUC	0.84	0.88	0.86	0.89	0.84	0.88

Table 2. Results obtained after feature extraction

	AdaBoost	**CatBoost**	**GBM**	**LightGBM**	**RF**	**XGB**
Accuracy	0.80	0.96	0.87	0.94	0.93	0.93
AUC	0.89	0.99	0.94	0.99	0.98	0.98

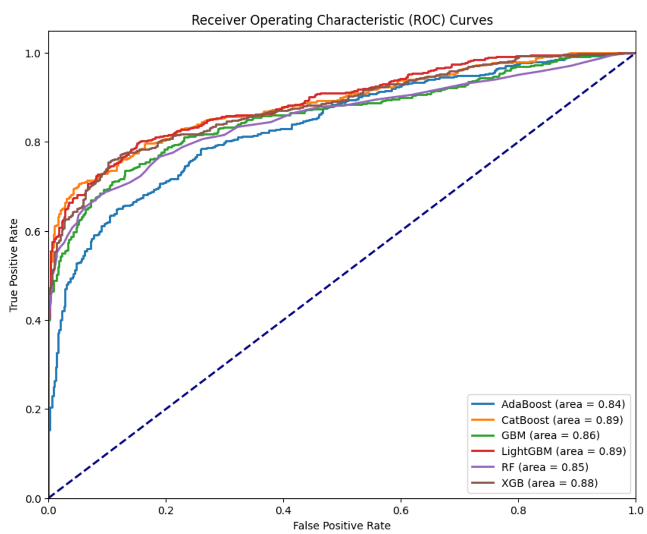

Figure 1. ROC curve obtained using raw EEG signals

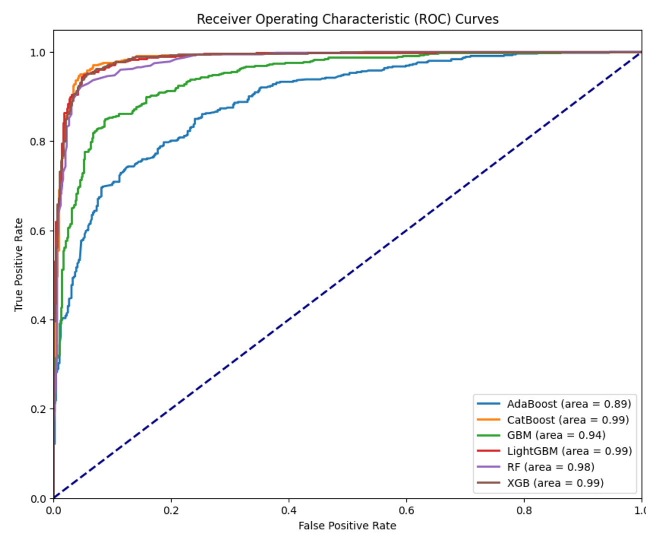

Figure 2. ROC curve obtained after feature extraction

5. Discussion

Efforts to identify objective diagnostic markers for ADHD have prompted numerous resting-state EEG studies. In this study, EEG signals were collected from 28 ADHD patients and 29 healthy controls. By extracting features from these EEG signals and applying ensemble learning models, the classification performance of features in a single modality was assessed. The results yielded promising data, facilitating the identification of objective differences linked to brain activity. The strong discriminatory ability of machine learning algorithms underscores their significant clinical potential and affirms the feasibility of using EEG signals for ADHD identification.

This study leveraged machine learning techniques and advancements in EEG signal analysis to develop more robust models. The results confirmed the effectiveness of the extracted features, such as differential entropy and energy characteristics across different frequency bands, for machine learning-based ADHD identification. However, several limitations should be addressed in future research:

(1) Sample size limitations: ADHD is a highly specific condition. The dataset in this study, which included augmented EEG signal clips from the same individuals, differs considerably from larger datasets comprising signals from multiple individuals. Expanding the sample size is essential for drawing scientifically valid and objective conclusions, thereby reducing errors linked to data insufficiency.

(2) Model generalizability: Future studies should prioritize adapting methods to ensure robust model performance across various datasets, thereby facilitating cross-dataset ADHD classification. Establishing a shared platform or database based on this study could help overcome challenges related to sample scarcity. This aligns with current trends toward scalable and generalizable technological solutions.

(3) Clinical phenotype differentiation: This study did not account for the clinical phenotypes of ADHD. However, an accurate clinical diagnosis requires a clear identification of these phenotypes.

(4) Differentiation from other psychiatric disorders: While this study excluded comorbid psychiatric disorders in the ADHD group, epidemiological studies and clinical experience suggest a high prevalence of comorbid conditions among ADHD patients.

These findings emphasize the importance of expanding sample sizes, developing a comprehensive ADHD database, integrating EEG data across different age groups, and incorporating comorbid psychiatric disorders into future research. Additionally, developing more stable models will further enhance the accuracy and applicability of machine-learning approaches in clinical practice.

6. Conclusion

This study analyzed EEG modality data from both patients and healthy control groups. Following the preprocessing of EEG signals from both groups, relevant features were extracted and classified using ensemble learning models in machine learning.

The results demonstrated that, even prior to feature extraction, the models achieved commendable performance, emphasizing the robust feature extraction capabilities of ensemble learning when applied to raw data. Importantly, feature extraction led to a significant improvement in model performance, validating the effectiveness of the extracted features, such as differential entropy and energy characteristics across various frequency bands, for machine learning tasks.

These findings provide a highly efficient approach for the diagnostic identification of ADHD, offering valuable support for its diagnosis, research, and treatment. Furthermore, this paves the way for more accurate and accessible diagnostic tools in clinical practice.

Funding

This study received financial support from the Jilin Province Health and Technology Capacity Enhancement Project (Project Number: 222Lc132).

Authors' contribution

Conceptualization: Xinxian Peng

Methodology: Xu Liu, Zihan Yang, Ling Li

Formal analysis: Liyan Li, Yanyan Chen, Xiaohan Wu, Yang Yu

Writing – original draft: Lina Han
Writing – review & editing: all authors

Disclosure statement

The authors declare no conflict of interest.

References

[1] Weigard A, McCurry KL, Shapiro Z, et al., 2023, Generalizable Prediction of Childhood ADHD Symptoms from Neurocognitive Testing and Youth Characteristics. Transl Psychiatry, 13(1): 225. https://doi.org/10.1038/s41398-023-02502-6

[2] Brandt V, Kerner Auch Koerner J, Palmer-Cooper E, 2019, The Association of Non-obscene Socially Inappropriate Behavior With Attention-Deficit/Hyperactivity Disorder Symptoms, Conduct Problems, and Risky Decision Making in a Large Sample of Adolescents. Front Psychiatry, 10: 660. https://doi.org/10.3389/fpsyt.2019.00660

[3] Rosenthal EA, Broos HC, Timpano KR, et al., 2024, Does Emotion-Related Impulsivity Relate to Specific ADHD Symptom Dimensions, and Do the Effects Generalize Across Comorbid Internalizing and Externalizing Syndromes? J Atten Disord, 28(2): 178–188. https://doi.org/10.1177/10870547231210283

[4] Mahone EM, Denckla MB, 2017, Attention-Deficit/Hyperactivity Disorder: A Historical Neuropsychological Perspective. J Int Neuropsychol Soc, 23(9–10): 916–929. https://doi.org/10.1017/S1355617717000807

[5] Loe IM, Feldman HM, 2007, Academic and Educational Outcomes of Children with ADHD. J Pediatr Psychol, 32(6): 643–654. https://doi.org/10.1093/jpepsy/jsl054

[6] Emser TS, Johnston BA, Steele JD, et al., 2018, Assessing ADHD Symptoms in Children and Adults: Evaluating the Role of Objective Measures. Behav Brain Funct, 14(1): 11. https://doi.org/10.1186/s12993-018-0143-x

[7] Chen H, Yang Y, Odisho D, et al., 2023, Can Biomarkers be Used to Diagnose Attention Deficit Hyperactivity Disorder? Front Psychiatry, 14: 1026616. https://doi.org/10.3389/fpsyt.2023.1026616

[8] McCabe GA, Smith MM, Widiger TA, 2023, Psychopathy and Antisocial Personality Disorder in the Fifth Edition of the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders: An Attempted Replication of Wygant et al. (2016). Personal Disord, 14(6): 636–648. https://doi.org/10.3389/fpsyt.2023.1026616

[9] Poirion OB, Jing Z, Chaudhary K, et al., 2021, DeepProg: An Ensemble of Deep-Learning and Machine-Learning Models for Prognosis Prediction Using Multi-Omics Data. Genome Med, 13(1): 112. https://doi.org/10.1186/s13073-021-00930-x

[10] Zhao M, Ye N, 2024, High-Dimensional Ensemble Learning Classification: An Ensemble Learning Classification Algorithm Based on High-Dimensional Feature Space Reconstruction. Applied Sciences, 14(5): 1956. https://doi.org/10.3390/app14051956

[11] Liu ZP, Gao R, 2018, Detecting Pathway Biomarkers of Diabetic Progression with Differential Entropy. J Biomed Inform, 82: 143–153. https://doi.org/10.1016/j.jbi.2018.05.006

[12] Shi LC, Jiao YY, Lu BL, 2013, Differential Entropy Feature for EEG-Based Vigilance Estimation. 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 6627–6630. https://doi.org/10.1109/EMBC.2013.6611075

[13] Garbaczewski P, 2006, Differential Entropy and Dynamics of Uncertainty. Journal of Statistical Physics, 123(2):

315–355. https://doi.org/10.1007/s10955-006-9058-2

[14] Daubechies I, 1990, The Wavelet Transform, Time-Frequency Localization and Signal Analysis. IEEE Transactions on Information Theory, 36(5): 961–1005. https://doi.org/10.1109/18.57199

[15] Pan Y, Zhang L, Wu X, et al., 2019, Structural Health Monitoring and Assessment Using Wavelet Packet Energy Spectrum. Safety Science, 120(3): 652–665. https://doi.org/10.1016/j.ssci.2019.08.015

[16] He H, Chen Y, Lan B, 2021, Damage Assessment for Structure Subjected to Earthquake Using Wavelet Packet Decomposition and Time-Varying Frequency. Structures, 34(12): 449–461. https://doi.org/10.1016/j.istruc.2021.07.087