NTU Journal of Engineering and Technology (2026) 5 (1) : 39-60
DOI: https://doi.org/10.56286/ntujet.v5i1

IRAQI Academic Scientific Journals

P-ISSN: 2788-9971 E-ISSN: 2788-998X
NTU Journal of Engineering and Technology
Available online at: https://journals.ntu.edu.iq/index.php/NTU-JET/index

EEG-Based Diagnosis of ADHD in Children and Adolescents: A Comprehensive Review of Machine Learning and Deep Learning Approaches

Sarah Talal Mohammed Taher[1], Mohammed Sabah Jarjees[1], Muhammad Abul Hasan[2]
[1]Department of Medical Instrumentation Techniques Engineering, Technical Engineering College of Mosul, Northern Technical University, Mosul, Iraq,
[2]Department of Biomedical, NED University of Engineering & Technology.
sarah.tm@ntu.edu.iq, mohammed.s.jarjees@ntu.edu.iq, abulhasan@neduet.edu.pk

Article Informations

Received: 25-10- 2025,
Revised: 29-11- 2025,
Accepted: 01-12-2025,
Published online: 22-03-2026

Corresponding author:
Name: Sarah Talal Mohammed Taher
Affiliation: Northern Technical University
Email: sarah.tm@ntu.edu.iq

Key Words:
ADHD,
EEG,
Deep Learning,
Machine Learning,
Classification.

ABSTRACT

Attention Deficit Hyperactivity Disorder (ADHD) is a major challenge in pediatric neurodevelopment and one of the most common conditions in children and adolescents. Clinical diagnosis remains difficult due to symptom overlap with other disorders, increasing the need for objective tools. Electroencephalography (EEG) has emerged as a non-invasive technique that, when combined with artificial intelligence, enables more reliable diagnostic models. This study aims to review pediatric and Adolescent EEG-based ADHD studies published between 2015 and 2025, covering both traditional signal-processing and artificial intelligence approaches.Fifty studies were included in this review study.The originality of this review lies in providing a decade-long, child-focused synthesis that highlights accuracy trends, methodological progress, and challenges related to clinical translation. Reported performance varied widely: traditional spectral analyses with Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) achieved 62–88%, whereas deep learning models including Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and attention-based hybrids,often exceeded 99%. The limitations across the literature include small sample sizes, restricted multi-center datasets, and insufficient evaluation of ADHD subtypes.

1. Introduction

Attention Deficit Hyperactivity Disorder (ADHD) is one of the most common neurodevelopmental disorders affecting children and adolescents. According to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5), individuals with ADHD may exhibit symptoms related to inattention, hyperactivity, or impulsivity, and the disorder is categorized into three subtypes: inattentive (ADHD-I), hyperactive/impulsive (ADHD-PH), and combined (ADHD-C). These symptoms typically appear before the age of twelve and may persist for at least six months, leading to difficulties in academic, social, and behavioral functioning. The overlap of ADHD symptoms with other conditions makes clinical diagnosis challenging and increases the need for objective diagnostic approaches [1, 2].

Electroencephalography (EEG) has emerged as a valuable tool for investigating the neural activity patterns associated with ADHD. Unlike subjective behavioral assessments, EEG provides an objective biological measure of brain function. Event-Related Potentials (ERP) derived from EEG recordings enable precise temporal evaluation of cognitive and executive functions, including attention, response inhibition, and working memory [3, 4].

The purpose of this review is to analyze and summarize pediatric studies conducted between 2015 and 2025 that used EEG signals for ADHD diagnosis. This includes examining traditional signal-processing methods and machine-learning techniques, in addition to describing advances in deep-learning models that have contributed to improved diagnostic performance. The novelty of this review lies in providing a comprehensive decade-long synthesis focused exclusively on children and adolescents, identifying accuracy trends, methodological developments, research gaps, and factors affecting clinical translation.

Early EEG-based ADHD research reported abnormalities in cortical arousal, including increased slow-wave activity in the theta band and reduced neural excitability [5, 6]. These findings supported the development of quantitative biomarkers, leading to the approval of a medical device by the U.S. Food and Drug Administration (FDA) that measures the ratio of high- to low-frequency EEG activity as an auxiliary diagnostic indicator [7, 8].

Subsequent studies indicated that reliance on simple frequency-based measures, such as the theta/beta ratio (TBR), may lack the consistency and specificity required for broad clinical use. Consequently, research has shifted toward ERP-based analyses, which offer more precise insight into the brain’s processing speed and efficiency [5,9].

Despite progress, several limitations continue to appear across studies. Many investigations relied on small, single-center datasets, lacked standardized EEG preprocessing protocols, and did not sufficiently explore ADHD subtypes or gender differences. Although deep-learning approaches have recently demonstrated improved accuracy, their clinical applicability remains limited due to challenges in interpretability and generalization.

Advances in signal-processing techniques, such as Independent Component Analysis (ICA) and wavelet decomposition, have improved EEG data quality and supported more effective feature extraction. These methods have enabled the transformation of raw EEG signals into compact and informative representations that reflect neural patterns associated with ADHD.

This review provides a detailed overview of EEG-based approaches for diagnosing pediatric and adolescent ADHD, emphasizing the evolution of analytical techniques and the integration of machine learning and deep learning models. It incorporates nearly all relevant studies that used EEG, EEG-ERP, or EEG combined with Magnetic Resonance Imaging (MRI), providing a comprehensive and up-to-date perspective on the field. By synthesizing a decade of research, it establishes a clear framework for understanding existing methods and guiding future investigations toward reliable and explainable diagnostic solutions.

This review aims to summarize and evaluate pediatric and adolescent EEG-based ADHD studies, covering traditional signal-processing methods and artificial intelligence approaches, while highlighting methodological developments, analytical techniques, and diagnostic performance.

The remainder of this review is organized as follows: First, the previous studies on EEG-based ADHD diagnosis are reviewed, highlighting the evolution of methods from traditional signal processing and machine learning to advanced deep learning approaches. Next, the review methodology is presented, including the criteria for study selection and categorization of research phases. This is followed by a discussion of the main findings, emphasizing observed trends, strengths, limitations, and current research gaps. Finally, the review concludes by summarizing key insights and suggesting potential directions for future research.

2. Literature Review

In 2015, Sangal and Sangal conducted a study on the EEG in children who have ADHD and those who do not, during tasks that required auditory and visual attention. This study found that the TBR is higher for children with ADHD than for normal children, but the main difference was the decrease in

Beta-1 wave power, especially in Broca's area (electrode sites F7 and FC5). The power of Beta-1 appears to have good sensitivity in Broca's area, reaching 86%. This appears to be important as a biomarker disorder, as it is reliable at excluding those without the condition (low specificity), which could lead to some false positives. It's meant to support diagnosis when ADHD is already suspected, not for general screening. Furthermore, since the study only included children aged 6 to 14, the findings may not apply to younger kids or adults [10, 11].

In the same year, the study of Snyder et al. combined ADHD assessment with the TBR. This study included 275 individuals diagnosed with behavioral problems. This combination improved diagnostic accuracy, increasing from 61% to 88%. The study also detected misdiagnoses due to failure to meet criterion E (symptoms not better explained by another disorder), which was 91%. The results also supported the use of the EEG biomarker, as it improved the diagnosis of ADHD in particularly complex cases. However, there were some limitations, most notably that due to the head movement of young children, it was difficult to record EEG signals [12, 13].

In the same year, Helgadóttir et al. used EEG coherence measures with chronological age to develop a multivariate diagnostic classifier for diagnosing ADHD disorder. The study included 310 children with ADHD disorder and 351 children with normal behavior. Their ages ranged from 5.8 to 14 years, from Iceland. Researchers developed the model by using statistical pattern recognition. They tested it through cross-validation and on an independent group that was not used in training. The result showed that the age-adjusted classification method had higher accuracy (76% in the independent test set; 81% in the cross-validation) compared to the age-unassigned approach (76%; 73%). Chronological age emerged as an important feature in the classification process; also, the sample well reflected clinical reality. This made the study distinct and used the same devices and software in the recorded EEG, although there were several limitations, including the fact that the study used the same equipment, which may affect the possibility of applying the model in other places. All participants were from Iceland only, which reduced the possibility of generalizing the results. The study did not follow up on cases over the long term because it was a cross-sectional study [14, 15].

In 2016, Mohammadi et al. proposed a model to distinguish between patients with ADHD and normal children. The model was based on the analyzed EEG signal. The study was applied to 30 children with ADHD and 30 normal children. The EEG was recorded for all children while they performed the cognitive task. To extract the pattern from the EEG signals, the study used non-linear features such as Approximate Entropy (ApEn), Fractal Dimension (FD), and Lyapunov exponent; also employed feature selection methods Double Input Symmetrical Relevance and minimum Redundancy Maximum Relevance (DISR and mRMR) to select optimal features; and classified them using a Multi-Layer Perceptron (MLP) type of neural network. The model achieved a classification accuracy of 92.28% with mRMR, and the classification accuracy of the model with DISR reached 93.65%. However, one of the limitations of the study was the small sample size, which limited the possibility of generalizing the results, as well as the absence of a validation study [16, 17].

In the same year, Kamida et al. conducted a study that included 80 children with ADHD and 59 typically developing children, aged 4 to 15 years. EEG signals were recorded using the 10–20 system, and spectral analysis was applied. The results indicated an increase in beta-wave power in children with ADHD, particularly in the frontal lobe, while alpha power slightly decreased in the occipital region. Nevertheless, a notable shortcoming of the study was the small sample size and the absence of differentiation among ADHD subtypes [18, 19].

In 2017, Jahanshahloo et al. suggested an automatic system for diagnosing ADHD in children. ERPs were recorded from 60 participants (30 children with ADHD and 30 controls) under visual and auditory stimuli using three electrodes. After preprocessing, many features were extracted, including band power, fractal FD, autoregressive coefficients, and wavelet coefficients. The results showed that the combination of fractal and wavelet features, referred to as Fra-wave, achieved a higher accuracy of 99.43% using a variant-Support Vector Machine (v-SVM) classifier, compared to 97.65% using wavelet features alone. Additionally, the performance of auditory stimuli in classification was better than that of visual stimuli [20, 21].

In the same year, Rahadian et al. used a method that combined the Learning Vector Quantization 2 Neural Network (LVQ2NN) with the genetic algorithm to classify the type of ADHD. The study used data collected from 100 samples (45 suffered from symptoms of ADHD). The severity of symptoms was assessed in the study by a psychological expert. The results showed that the combination of the genetic algorithm and LVQ2NN increased and improved the classification accuracy, reaching 89.5%, while classification using LVQ2NN alone achieved an accuracy of 80%. The study supported the idea that the genetic algorithm helped in finding optimal weights for LVQ2NN, which improved performance, and the highest classification accuracy reached 95% using this method. However, there were some limitations, including that the number of samples was very small, only 100 samples, and that the data was limited to a single source, which may have limited the possibility of generalizing the results [22, 23].

In 2018, Farnia et al. examined the relationship between severe symptoms of ADHD in children and the TBR of the electrical brain activity EEG recorded from two regions, CZ and FZ, in two cases (resting state with open eyes without a mental task & during the mental task by neurofeedback). The study included 61 children with ADHD and 59 in the control group. The result showed that there was a moderate correlation between the TBR and the severity of the symptoms of the disease. It also showed that the ratio of theta/beta was higher in the FZ region without any task for children with this disease. Theta/beta without a test in Fz (sensitivity=62%; specificity=71%) and in Cz (sensitivity=51%; specificity=73%) differentiated the two groups only at a medium level. However, the samples did not include disorders other than ADHD, and the children's IQ was not measured accurately and directly; the signals were recorded only from two regions, CZ and FZ, without assessing executive functions. The researchers recommended the use of more accurate tools, such as quantitative EEG (qEEG) [24, 25].

In the same year, the study of Lopez Marcano et al. used a Gaussian Mixture Model with a Universal Background Model to diagnose ADHD in children by analyzing EEG data. The study included 61 children with ADHD and 60 normal children, all of whom were male and aged from 6 to 8 years. The EEG signals were recorded while children performed the Attention Network Task (ANT) in addition to the resting cases with closed and open eyes. The model achieved an accuracy of 92% when using the data of the attention task, while the accuracy decreased when using the rest state data. These results indicated that the model's detection performance was highly dependent on the task condition during signal recording. A notable limitation of the study was the small sample size, the inclusion of only male participants, and the reliance on mothers' reports for ADHD diagnosis, all of which limited the generalizability of the results to females and other age groups [26, 27].

In the same year, Shiva Khoshnoud et al. analyzed the EEG signals for 12 children with ADHD and 12 children with normal brains by using nonlinear dynamic analysis methods. All of the children were in a rest state with closed eyes. The EEG was recorded from 19 channels, and 3 features were extracted (Largest Lyapunov Exponent (LLE), ApEn, and Multifractal Singularity Spectrum). The result showed a clear difference between the two groups (ADHD & normal). LLE was higher in the left medial frontal lobe, but ApEn was lower in the prefrontal cortex of children with ADHD. The study used SVM as a classifier, and its classification accuracy was 83.3%. This was one of the first studies that used a nonlinear statistical method to analyze EEG signals. A notable limitation was the small sample size of only 12 children per group, which restricted the generalizability of the findings [28, 29].

In the same year, in the study introduced by Vuckovic et al., the EEG signals were recorded from 31 patients with spinal cord injuries and 10 healthy people. The study used EEG lab tools to remove segments that exceeded 100 microvolts, manually reviewed the signals, and re-referenced them to the mean. To remove non-neural components (eye movement), the researchers used Infomax ICA after signal cleaning. The researchers trained three classification models (Linear Discriminant Analysis (LDA), SVM, and Artificial Neural Network (ANN)) and achieved classification accuracy exceeding 88%. The study showed that the method used to process the EEG signals was significantly effective in improving classification performance and data quality, which indicated that this study applied to ADHD research, as EEG signals used to classify ADHD required accurate noise removal [30, 31, 32].

In a recent study in 2019, Vahid et al. developed the deep learning model known as EEG Net to distinguish between children with ADHD and typically developing children by analyzing the EEG during ERP. The study relied solely on neurophysiological measures without using traditional diagnostic tools. The model achieved an accuracy of 83%, and in some cases, up to 86%. The results showed that in the superior parietal cortex region, neural processes associated with attentional selection contributed more to classification. A notable limitation of the model was its inability to distinguish ADHD subtypes, and further development was needed to improve its reliability in clinical applications. Additionally, several participants were not accurately classified [33, 34].

In the same year, a study conducted by Chen et al. used CNN as a deep learning model to distinguish ADHD in children, which depended on the recorded EEG signals from 107 participants (50 children with ADHD and 57 normal). The study applied CNN to analyze the brain activity patterns, and Gradient Weighted Class Activation Mapping (Grad-CAM) was used as an interpretation tool to identify the brain areas and frequencies that contributed to the diagnosis individually for each case. The results showed that the accuracy of the model in classification reached 90.2%, and the highest ACU value was 0.96 ± 0.01. The Grad-CAM results demonstrated that the model could detect spatial and frequency variations in EEG that differed from one child to another, which reflected the individualized nature of ADHD. However, there were some limitations in this study, as the generalizability was limited because it relied on data from a single source without multi-site validation. Additionally, the flexibility of the model in processing the initial data may have been affected because the study used predefined transforms [35, 36].

In the same year, Zhang and Li used Convolutional Neural Networks (CNNs) with transfer learning techniques to classify the children into normal and those suffering from ADHD disorder, depending on the EEG signal recorded in the resting state from 32 channels for each participant. The study used the Phase Lag Index (PLI) indicator to build brain network maps that were used as inputs to the CNN model. The CNN model had been trained previously to classify the images. Some of its layers were transferred to the new specific model for classifying ADHD. The model achieved an accuracy of 94.39%, a sensitivity of 97.83%, and a quality of 91.80%. This result reflected the role of transfer learning in improving the performance compared to creating the model from zero, and this study showed that transfer learning was effective in diagnosing the ADHD disorder. However, there were some limitations, the most important of which was that it had not been tested on data from external sources and did not mention performance across different ages or clinical settings [37, 38].

In the same year, the study of T. Wang and S. aimed to diagnose ADHD. A 3D-CNN model was used to analyze the Fractal Dimension Complexity Map (FDCM), which the researchers extracted from structural Magnetic Resonance Imaging (sMRI) data. The study applied the model to ADHD-200 data, which included a group of children and adolescents aged 7 to 12 years. The model's classification accuracy reached 69%, and based on the results achieved by the model, it outperformed previous methods that relied solely on sMRI. However, there were some limitations: the model did not rely on fMRI and EEG data, so the study did not directly reflect brain functions and neural activity. Therefore, the study was considered to have not included functional data but rather focused on structural imaging data, which limited the comprehensive assessment of the disorder [39, 40].

In 2020, Khaleghi et al. extracted five groups of features from the recorded EEG signals: linear features, shape features, temporal features, frequency features, and time-frequency features. All participants were in a relaxed state during data collection. The dataset included 60 children, 30 diagnosed with ADHD and 30 typically developing, aged between 7 and 12 years. The five feature groups were compared to identify the most effective ones for ADHD diagnosis. Statistical analysis showed that non-linear features achieved the highest classification accuracy (86.40%) and an Area Under Curve (AUC) value of 0.899. Frequency features followed, with a classification accuracy of 80.44%. Based on these findings, the study concluded that nonlinear features are a promising tool for ADHD detection. However, the study had several limitations. It did not specify which individual non-linear features were most discriminative. Additionally, there was no standardized protocol for EEG recording, the sample size was relatively small, and the model lacked a diagnostic threshold for classifying subjects as ADHD or control [41, 42].

In the same year, Moghaddari et al .developed a CNN-based model to diagnose ADHD in children with this disorder. EEG data were recorded during a continuous mental task. There were 31 children with ADHD and 30 healthy children. The EEG signals were divided into small segments after preprocessing and converted into RGB color images based on frequency bands (delta, theta, alpha, beta, and gamma). These images were fed into a CNN. The network consists of 13 layers to extract features and classify cases. The model's accuracy reached 97.47% at the segment level and 98.48% at the individual level. The study had some limitations, including the small sample size and reliance on EEG recording during a single task. Additionally, the model's performance had not yet been validated in real clinical settings [43].

In the same year, the recent study of TaghiBeyglou et al. used recorded EEG signals from the dataset known as the NBML dataset. It contains 328 epochs, and the number of participants is 61 (children with ADHD and normal children). The study extracted the time and frequency features by using Common Spatial Pattern (CSP) and nonlinear features. To optimize the representation, the study used a filter bank and time windowing. The study showed that the K-Nearest Neighbors (KNN) with the proposed method achieved a classification accuracy of 83.33%. This result is higher than previous results on the same data set. Based on the results reached by the study, this method is considered relatively effective, but there were some limitations to this study, the most important of which was that it had not been tested on external data, and the number of participants was only 61, which was a limited number, making it difficult to generalize the results [44, 45].

In the same year, another study by Van Dijk et al. analyzed EEG data from two studies (iSPOT-A and ICAN) to calculate the TBR by using five methods for 328 children with ADHD and 151 normal children, with the recorded EEG signals from 2 channels (Fz and Cz). The result showed no important clinical difference between the two groups, and the range of effect size was from 0.102 to 0.215. The results also showed a moderate negative relationship between TBR and age, and that TBR alone is not sufficient to diagnose ADHD. This is what the study confirmed. Finally, the difference in brain signal processing methods explains part of the variance in the results. However, there were some limitations to this study, the most prominent of which was that it was limited to the TBR index only and did not take into account other factors, such as sleep quality and participant selection criteria, which may have affected the results [46, 47].

Another study in the same year, Ciodaro et al., used resting-state EEG signals to diagnose ADHD disorder. The study included 86 children; two models were used, one of which depended on XGBoost and another based on the Residual CNN network, to diagnose ADHD disorder. The model based on XGBoost that used alpha and beta energy achieved an accuracy of 86.3% and an F1-score of 73.3%, while the model based on the Residual CNN network that used spectral images for alpha and beta activity achieved an accuracy of 90% and an F1-score of 76%. Activation maps in the second model showed differences in activity in the frontal and temporal lobes. There were some limitations: the number of samples, only 86 children, was not distributed between the affected and healthy individuals. The generalizability of the results was limited because the model had not been tested on external data [48, 49].

In the same year, Amado-Caballero et al. The study included a group of 148 children, using daily activity logs that lasted approximately 24 hours. Among the participants were 73 children with mixed-type ADHD according to DSM-5 criteria and 75 healthy children. The study indicated that the children with ADHD did not receive any medication during the study period. The researchers used a CNN model to analyze the activity on the spectrograms, and the study demonstrated high diagnostic accuracy, in addition to a sensitivity value reaching 97.62%, and the specificity reached 99.52%. However, despite the results achieved by the study, there were some limitations, most notably the lack of neural coverage when compared to studies using EEG and functional Magnetic Resonance (fMRI) data. The study may also not have fully represented the neurological complexities of ADHD [50, 51].

In the same year, the study of Altınkaynak et al. To diagnose children with ADHD, EEG signals recorded from 23 children with this disorder and 23 healthy children were used. The signals were recorded from the channels Fz, Cz, Pz, and Oz during the auditory oddball task by using a Biopac device at a frequency of 2500 Hz. The extracted features were P300 latency/amplitude, non-linear FD, and wavelet coefficients. The classifiers that were used in this study were MLP, SVM, and k-NN. The MPL classifier achieved an accuracy of 91.3%. However, there were some limitations, such as the small number of participants (small sample size), the study did not address the differentiation between types of ADHD, the number of channels was small, and finally, hearing problems may have had a significant impact [52, 53].

In 2021, Ekhlasi et al. extracted effective connectivity from the EEG signal to distinguish between the normal children and children with ADHD disorder through the indicator that analyzes the connectivity and direction between the brain regions known as directed Phase Transfer Entropy (dPTE). To represent the properties of the connection between the electrodes, the matrix of dPTE was converted to the vectors called Effective Connectivity Vectors (ECVs). These vectors were used as the input to a model based on the ANN, and the genetic algorithm was used to improve the performance and reduce the number of features of the model. The study analyzed the five frequency bands (Delta, Theta, Beta, Alpha, and Gamma), and the accuracy of the model reached 96% when combining the features of all bands in a single vector known as the global Effective Connectivity Vector (gECV), but the study suffered from some limitations, such as the risk of overfitting in the result because the number of extracted features was relatively large and the training and test data was small (62 children: 31 with ADHD and 31 in the control group), which made it impossible to generalize the results to larger and more diverse samples [54, 55].

In the same year, the study of Peng et al. proposed a model consisting of two branches of the 3D CNN; each branch processed a different type of imaging data. This model was known as SSANN (Summation-based Synergetic Artificial Neural Network), dependent on the MRI AND sMRI images to diagnose the ADHD disorder. The outputs from 2 branches were combined through the summation process and passed through the neural network, where the classification was performed. The model can classify the children into normal and those with ADHD. The result showed that the accuracy of the model reached 72.89% on the ADHD-200 dataset. This study outperformed other studies that used the same dataset in accuracy and the value of AUC. However, there were some limitations, such as resizing the images due to the difference in the dimensions of MRI images and sMRI images. This scaling may have led to the loss of information. Also, the model was trained and tested on the ADHD-200 dataset, and because the samples were small, these results could not be generalized. Also, when merging data of different resolutions, it may lead to the loss of some important features [56, 57].

In the same year, Maya-Piedrahita et al. used EEG signals recorded from 67 children during the Reward Stop-Signal Task (RSST) with the Hidden Markov Model (HMM). This model was trained on the experiences of failure to curb the response only for each participant. The study used Probability Product Kernel (PPK) to calculate the similarity between the models. The SVM is used as a classification tool. The maximum accuracy was 90%, especially in the decreasing reward condition, indicating ADHD patients' sensitivity to rewards. A notable limitation of this study was that it relied only on failure trials and did not use spectral analysis [58, 59].

In the same year, Mustafa Tosun recorded EEG signals by using a Nihon Kohden device through 16 different channels according to the 10-20 system, with recording through a closed-eye rest state and other cases, to use the EEG data to distinguish between children with ADHD and normal children. The signals were divided into segments, 30 seconds for each segment, and 50 spectral properties were extracted from each segment. The study used three models (LSTM, SVM, and Feed-Forward Back Propagation Neural Network (FFBPNN)), and the LSTM model achieved an accuracy of 88.88% for the Fp1-F7 channels in the closed-eyes rest state. The spectral entropy property contributed to improved performance. The study showed that using specific recording conditions and channels can help reduce the number of electrodes without affecting performance. The study suggests using algorithms such as sign, cosine, and grasshopper optimisation to improve feature selection. However, there were some limitations, most notably that the number of participants was not detailed clearly [60, 61].

In the same year, Pham et al. diagnosed the children with ADHD disorder by using the Ensemble Learning method. The study used nonlinear features (FDs) extracted from recorded EEG signals. The dataset used in this study included 61 children with ADHD and 60 children with normal behavior. Their age ranges from 7 to 12 years. The EEG signals were recorded from 19 channels during the visual attention task. The study used ensemble learning and specific feature techniques instead of the neural network; the model achieved a classification accuracy of 98.33%. This result was reflected, but the study had some limitations, most notably that it was not tested on external data and relied on a single data source, which limited the generalizability of the results [62].

In the same year, the study by Bakhtyari et al. presented a new framework by combining ConvLSTM with an attention mechanism to diagnose ADHD from EEG signals. The study used a dynamic connectivity tensor as an alternative to extract features as input to the model, as the model enabled the temporal and spatial characteristics of the signal. The experiment was applied to 400 EEG samples. The model achieved an accuracy of 99.75% using fold cross-validation. This study was considered one of the first studies to use this combination on EEG data. Nevertheless, the study had some limitations, chiefly that it relied on a single dataset. The very high accuracy suggested possible overfitting, and the model could not be directly compared to traditional clinical methods [63, 64].

In the same year, the study by Cabarcas-Mena et al. conducted a study to diagnose ADHD. The study included a group of 44 children, 22 of whom were diagnosed with ADHD and 22 of whom were healthy. The researchers recorded EEG signals using a low-cost Emotiv EPOC device. The study used CNNs to analyze recurrence plots extracted from the ERPs. The classification accuracy for this study reached 77.78%. However, there were some limitations, most notably the small sample size, which limited the generalizability of the results. The study also had some limitations, including the use of a low-cost EEG recording device, which may have affected the quality of the recorded signals. Finally, there was difficulty in obtaining children's cooperation during the recording [65, 66].

In 2022, Cheng et al. proposed a model to study transitions of mental and cognitive states during the Continuous Performance Test (CPT) based on LSTM. The study was applied to 67 children (34 ADHD sufferers and another 33 who were normal). The EEG was recorded from 2 electrodes (O1 and O2), and the accuracy of the model reached 90.5%. This result reflected the importance of beta wave activity, especially in the occipital region, as a distinct neural indicator. However, the results cannot be generalized because the study was conducted on a single, small dataset [67, 68].

In the same year, the study by Ghasemi et al. analyzed the brain activity of children who suffer from ADHD and compared it to the brain activity of normal children by using the signals of the ERP through the auditory and visual stimuli. The ERP data was analyzed into four main frequency bands (Delta, Theta, Alpha, and Beta). The study used seven machine learning algorithms, such as Logistic Regression (LR), Generalized Linear Model (GLM), Deep Learning, SVM, and other models, after extracting features. The study used new energy-based features, including Absolute and Relative Band Powers (ABP, RBP), testing different combinations to find the most effective ones. Some models, particularly deep learning, achieved very high accuracy—up to 100% for delta, beta, and alpha bands, and 98.15% for theta. Nevertheless, using more features may increase overfitting and reduce generalizability [69, 70].

In the same year, Catherine Joy et al. introduced a study using EEG signals recorded from 16 channels for the children (5 ADHD and 5 normal), whose ages ranged from 7 to 12 years. The EEG signals were recorded for two cases, eyes-open and eyes-closed, at 256 Hz for 300 sec and divided into sections of 25 sec. The study used Tunable Q-Factor Wavelet Transform (TQWT) to analyze data and extract FD features (Katz and Higuchi). For classification, the study used an ANN with 10-fold cross-validation. The study achieved a classification accuracy of 100%. Although this result of the study had several limitations, the most notable was that the number of children from whom EEG signals were recorded was only ten, which was a very small number. Additionally, the study did not use data from cognitive tasks, and the data were private, meaning that they were not available to the public except upon request [71, 72].

In the same year, Majid Mafi et al. used EEG signals to analyze the connectivity between brain channels (EEG connectivity). The study was conducted on 121 children, 61 of whom were diagnosed with ADHD, while the remaining 60 children were normal, ranging in age from 7 to 12 years. In this study, two indicators were used:

wavelet coherence and synchronization likelihood to analyze the connectivity between brain channels. This study presented two models using high-dimensional CNNs, 6D and 4D, to classify children. The results showed high accuracy in classifying cases, as the classification accuracy reached 99.17% with 100% recall. The study also reported that the positive accuracy reached 98.39%. Despite this strong performance, a notable limitation was that the models were not tested on external data, which limited the generalizability of the results [73, 74].

In 2023, Esas and Latifoğlu used 2 decomposition techniques, local mode decomposition and variational mode decomposition, to divide the EEG signal that was recorded from 19 channels into subbands. The study used the dataset available on IEEE DataPort (EEG ADHD and Control Children Dataset). This dataset includes children who suffered from ADHD and normal children. The subbands that were decomposed from the EEG signals were used as input to the deep learning model, and the classification accuracy of the system reached 95%. The sensitivity reached 97%, and the specificity was 94%. To test the importance of each channel, the study examined the model using data from a single channel (Fp1). The results showed that this channel alone achieved an accuracy higher than 87% when combined with signal analysis techniques. The result showed the important role of the frontal lobe in diagnosing ADHD disorder. Nevertheless, a notable limitation was that the study was conducted on a small sample, which restricted the generalizability of the findings [75, 76].

In the same year, Alkahtani et al. utilized intelligent techniques (deep learning and machine learning) to develop an intelligent model to diagnose ADHD in children. The study was applied to a general dataset that included raw EEG records for 61 children with ADHD and 60 children with normal behavior. All children underwent a task requiring visual attention. The study extracted the feature after precise preprocessing of the raw data. The feature extracted depends on time, frequency, and entropy. After that, the study used selection feature techniques to determine the indicator that was most relevant to the case, such as Least Absolute Shrinkage and Selection Operator (LASSO) regularization and recursive feature elimination (RFE). The study used several classification models, such as CatBoost, Random Forest, CNNs, and LSTMs. The result showed that the CNN model achieved a high accuracy, reaching 97.75%, and the CatBoost model achieved an accuracy reaching 95.13%. This result reflects the importance of selecting suitable features to improve the performance. To further improve the classification of brain patterns associated with ADHD, the study combined REF and Principal Component Analysis (PCA). However, despite the excellent results, the study confirmed the need for further experimental validation using larger and more representative samples to ensure the generalizability of the results [77, 78].

In the same year, the recent study of Cura et al. built EEG Feature Maps (EEG-FMs) from EEG signals for 15 children with ADHD and 18 children with normal development, depending on time features such as Hjorth parameters, skewness, and kurtosis, and nonlinear features such as Lyapunov exponent, FD, and entropy extracted from recorded EEG signals. These features were converted to images, and the deep features were extracted from them. The SVM classifier was used to classify the deep features. The result showed accuracy above 80.2% for the images generated from time features. As for images based on non-linear features, they achieved a classification accuracy of more than 93.5%, and when combining the features, the accuracy reached 100%. Despite these results, there were some limitations, the most prominent of which were that the model was not tested on external data, the method was not tested on cognitive tasks, and the sample size was very small, with only 33 participants [79, 80].

In the same year, in a study by Yağmur Can et al., the researchers used EEG signals recorded from 121 children (61 children with ADHD and 60 healthy children). They extracted temporal, frequency, and nonlinear features and used the LASSO method to select the most relevant features. The study used four deep learning algorithms, but the SVM algorithm with LASSO achieved the highest classification accuracy, reaching 96.3%. The results indicated the importance of signals from temporal, parietal, and occipital regions as potential biological indicators. The major limitations of this study were that the data were not diverse and were relatively small, which may have limited the generalizability of the results [81, 82].

In 2024, Chugh et al. proposed a hybrid model based on deep learning to improve the diagnosis of ADHD disorder, the hybrid model combining CNNs and LSTM. The study used two generally available data groups (ADHD dataset, FOCUS dataset) for training and testing the model, whereas the CNN part extracted the spatial features, and the LSTM part analyzed the time relationship in the EEG data. The result showed that the performance accuracy of the model reached 98.86% on the ADHD dataset, and the performance accuracy of the model reached 98.28% on the FOCUS dataset. Also, the result showed that the children with ADHD have higher energy in the theta band by using topographic analysis for the brain signals. Nevertheless, despite the strong performance of the study, some limitations were noted, the most prominent of which were the difficulty of interpreting the model due to the nature of deep networks, the challenge of identifying the most influential features for the same reason, and the small sample size, especially in the FOCUS group, where gender distribution bias may have affected the generalizability of the results. To improve visual understanding of the model, the authors recommended using techniques such as saliency maps and t-Distributed Stochastic Neighbor Embedding (t-SNE) [83, 84].

In the same year, Alsharif et al. used EEG data that included 61 children with ADHD and 60 children with normal. The statistical features extracted from the data, such as the standard deviation and kurtosis, and the spectral features, such as entropy, were obtained after the raw data were preprocessed. To select optimal features and improve the performance of the model, the study used selective feature techniques and PCA and combined them with the chi-square test. The study used several deep and machine learning models. The accuracy of the SVM model with PCA reached 94.86%, while the Convolutional Neural Network–Bidirectional Long Short-Term Memory (CNN-BiLSTM) model achieved 94.50% accuracy, and the Gated Recurrent Unit–Transformer model (GRU-Transformer model) with PCA and Chi-square achieved the highest accuracy, reaching 95.59%. These results indicate that the improvement in diagnostic accuracy was due to the use of feature selection techniques. However, the study was applied to a small number of models, which limits the generalizability of the results and makes it difficult to interpret the deep models to know which features most affect the diagnosis [85, 86].

In the same year, Pedrollo et al. used the Random Forest (RF) algorithm with the Genetic Algorithm (GA) for optimization to analyze resting-state EEG signals for the diagnosis of ADHD. The study included 856 participants. Two techniques were used (Artefact Subspace Reconstruction (ASR) and ICA) to process the signal and remove noise. The analysis focused on the theta frequency band. The result showed that the classification accuracy of the proposed model reached 88.6%. These results reflect the possibility of using this method as an effective tool, but there were some limitations. The study did not rely on signals during cognitive tasks but only relied on them during resting states, which may have limited the generalization of the results [87, 88].

In the same year, Pappula and Anwar presented a study that used EEG signals and converted them into spectrograms to improve the diagnosis of ADHD. They extracted features and classified cases using the ResNet-18 model. The model achieved high accuracy with an F1 score of 0.9. The model was able to identify prominent brain regions related to the disorder. The study demonstrated the power of combining EEG signals with deep learning. The study proposed an effective digital system for diagnosis at a low cost that can be used in schools. However, one of the most prominent limitations of this study was the need for the model to be evaluated on external data to increase its reliability [89, 90].

In the same year, the recent study of Mondal et al. developed a deep learning model known as ADHD-Net to diagnose ADHD disorder by using recorded EEG signals from 61 children with ADHD and 60 children with a normal condition; their ages range from 7 to 12 years. The researchers processed and refined the signals through filtering and ICA. To reduce the data dimensions, they used PCA. To improve the diagnostic accuracy, they used ADHD-Net with an attention mechanism during model training. The results showed that the model's accuracy reached 97.96% in distinguishing between people with the disease and healthy people. However, there were some limitations: the number of participants was relatively limited, and the model was not tested on data from cognitive tasks. The model needed to be tested on a broader and more diverse dataset to ensure its effectiveness, as the study focused on a specific age group of 7–12 years [91, 92].

In the same year, the study introduced by Yousefimehr et al. improved ADHD detection using a cost-sensitive LightGBM-based model and focused on reducing classification errors, which are critical in clinical classification. The model used in this study achieved a high accuracy of 94%. The model demonstrated high efficiency and speed, making it suitable for scientific applications. However, this study did not rely on physiological data such as EEG, nor did it focus heavily on model interpretation, which reduces the understanding of classification decisions. Finally, the model was not compared to deep learning models [93, 94].

In the same year, Singh et al. proposed a hybrid model for diagnosing ADHD based on CNN and the Transformer layer. In this study, they used EEG and fMRI data. The study included EEG data that contained more than 100 records and used 20 electrical channels distributed across different regions of the brain. The classification accuracy reached 97.53%. With the use of the confusion matrix, the accuracy rose to 99%. Despite the impressive results of the study, there were some limitations, most notably that the details of the samples were not clear, and there was a weakness in interpreting the model's decisions [95].

In the same year, a study was presented by Latifi et al. This study aimed to diagnose ADHD in children using a type of deep neural network known as Siamese CNN. The researchers used this network to analyze brain maps derived from Power Spectral Density (PSD). In this study, the researchers highlighted the most effective features in diagnosing the disorder using the Grad-CAM model interpretation technique. The model achieved a high accuracy of 99.17%, reflecting the role of using deep learning models with interpretation techniques in improving classification accuracy. The results also showed that theta wave activity is a biomarker for distinguishing ADHD, especially in the occipital and frontal lobes. Nevertheless, the study had some limitations, including that it relied solely on EEG signals and did not use fMRI data, which may have restricted the understanding of the full functional aspects of ADHD [96].

In 2025, Bansal et al. introduced a model based on EEG for diagnosing ADHD. The study was applied to 121 children (61 with ADHD and 60 normal). The study used an autoencoder to analyze 50,000 EEG samples to extract the features. The Reptile Search Algorithm was used to select significant features. The ResNet and Double Augmented Attention were used to build the classification model. The proposed model showed distinct performance and accuracy of 99.42%, recall of 99.82%, precision of 99.03%, and an F1-score of 99.42%, while the AUC value was close to 1.0. But there were some limitations, mainly that some children used the drug Ritalin, which may have affected EEG patterns, and the study was limited to cognitive tasks without including other aspects such as memory and auditory attention. The study recommended conducting various EEG signal recordings without drug intervention in the future [97].

In the same year, Hossain et al. proposed a model using an SVM with the Radial Basis Function (RBF) kernel to diagnose children with ADHD and others as normal. The study applied to the recorded EEG signal from 121 children (61 with ADHD and 60 with normal). Their ages range from 7 to 12 years. The EEG signal was recorded from 19 channels. The features were extracted by using two techniques, PSD and Spectral Entropy (SE), through five frequency bands. In the same year, the model achieved an average accuracy of 99.2%, with a standard deviation of 0.0079. This result reflects the higher model accuracy, and it was robust. The results also indicated that spatial and frequency features are effective in diagnosing ADHD in a non-invasive manner. However, the limitations lay in the fact that the study was conducted on a small number of samples, which required verifying the model on larger and more diverse samples [98].

In the same year, Kim et al. used the machine learning algorithm XGBoost with EEG signals to diagnose ADHD disorder. The study was applied to 168 participants (107 with ADHD and 61 normal). The EEG was analyzed through five frequency bands (delta, theta, alpha, beta, and gamma) after being recorded from 19 channels. To ensure the reliability and accuracy of the model, the Leave-One-Subject-Out (LOSO) technique was used. The result showed that the accuracy of the proposed model reached 90.81% and the F1-score 0.9347. The SHAP analysis showed the most important features were in the middle beta frequency, especially in the O1 electrode. But there were some limitations: the sample size was relatively small, EEG signals were used without data from cognitive tasks, and 81 of the patients were taking medication, which could have negatively affected the EEG signals. Finally, there was a numerical imbalance between the two groups, but this was addressed in the study by dividing the records [99].

Another study in the same year, Hu et al., proposed an advanced model known as SCANet to diagnose ADHD disorder. By using EEG signals. The model accurately extracts important features because it uses attention mechanisms between signal channels and their time with a multi-branch, multi-scale design. The study used a public dataset containing three categories (Attention Deficit Disorder (ADD), ADHD, and normal). The model's accuracy on this dataset was 99.78%, and the F1-score was 89.14%. To interpret the results of the model and the brain channels that significantly influence diagnosis and the determination of time zones, the researchers used Grad-CAM. Despite the strong results, the model's effectiveness was not tested on external data from multiple centers or on cases accompanied by other disorders. This limits the possibility of generalizing the results [100].

In the same year, Balamurugan et al. presented a study that proposed classifying ADHD using deep learning and brain signals from data collected from 120 children, 60 of whom had ADHD and 60 of whom had normal brain signals. The study stated that 11 brain features were extracted using the SVM algorithm, and the top 10 EEG channels were selected. The study stated that the researchers tested six machine learning models, and the Fine KNN model achieved the best accuracy among them, reaching 98%, and also outperformed the other models in terms of sensitivity. The study confirmed that EEG signals are effective in non-invasive diagnosis and that incorporating machine learning provides higher speed and accuracy. However, there were some limitations because the study included a very small number of samples, which limited the generalizability of the results, and the model was not tested on external data [101]. The schematic in Fig. 1 summarizes the common workflow for EEG-based ADHD diagnosis across the reviewed studies, highlighting the key methodological steps from data collection to classification.

3. Methodology

This review focused on studies conducted between 2015 and 2025 that used EEG brain signals recorded from children and adolescents for ADHD diagnosis. Three studies involving adults and studies not using EEG signals were excluded, except for one study (Ref. 48), which, although not purely EEG-based, was included due to its methodological relevance and contribution to ADHD prediction using cost-sensitive LightGBM. The studies were taken from solid scientific databases (IEEE Xplore, PubMed, Scopus, and Web of Science),

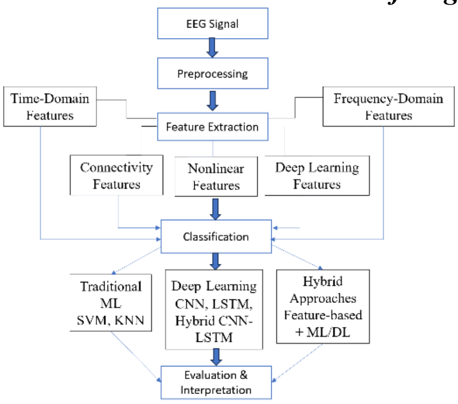

Fig. 1. Schematic overview of the EEG-based ADHD diagnosis workflow used in the reviewed studies. The diagram illustrates the typical steps from EEG data acquisition, preprocessing, feature extraction, and selection to classification using machine learning and deep learning models.

including all of them using EEG brain signals recorded from children and adolescents. These studies provided a sufficient and good methodology for feature extraction and classification using machine learning and deep learning. By searching the databases, 53 studies were initially identified, and 3 studies were excluded because they concerned adults. 50 studies were included in the final review. We were unable to access the full text of 3 studies (Refs 48, 46, and 31), but they were included in the table and chart to ensure transparency in the selection process. The study's selection process is illustrated in Fig. 2, showing the number of records identified, screened, excluded, and finally included in the review.

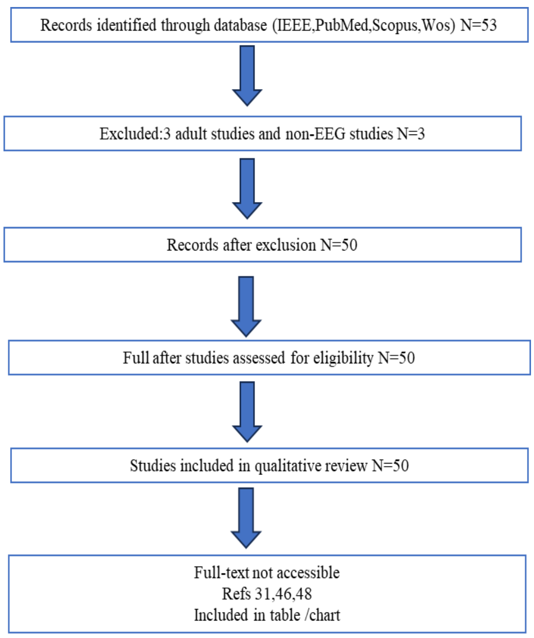

Fig. 2. PRISMA flow diagram

The main criteria for including studies were

1. Type of data used: Only studies that relied on EEG signals were included. Some studies combined EEG with MRI or fMRI. ERP signals were also used in some studies to assess neural responses to stimuli. Studies included both resting-state EEG and EEG recorded during specific cognitive tasks.
2. Processing and feature extraction methods: Studies used traditional methods such as frequency spectrum analysis (TBR, beta power, alpha power), correlation and concordance analysis (EEG coherence), nonlinear feature extraction (ApEn entropy, Lyapunov exponent, FD), and ICA of ERP.
3. Modern methods for deep learning and artificial intelligence: CNN, LSTM, hybrid CNN-LSTM, EEGNet, and attention mechanisms. Also, studies that used a method of converting EEG brain signals into images and spectral maps (PSD) to apply neural networks.
4. Classification models :traditional models, such as SVM, KNN, Random Forest, and MLP. Modern models of deep learning such as CNN, LSTM, hybrid CNN-LSTM, ResNet, SCANet, and Siamese CNN.
5. Feature Selection and Performance Optimization: Some studies used feature selection or performance optimization techniques to enhance classification accuracy, such as genetic algorithms, PCA, and feature selection.
6. Accuracy and evaluation criteria: Most studies focused on measuring the accuracy, sensitivity, F1-score, MCC, and quality of the model, as accuracy in some early studies reached 70%, and in some recent studies that used deep learning, it reached 99%.

4. Result and Discussion

The fifty studies mentioned in this review over the past ten years showed a marked variation in the accuracy of ADHD classification using EEG signals. The studies included children and adolescents. As shown in Table 1, in early studies such as the study by Sangal Sangal in 2015 and the 2018 study by Farnia et al., the accuracy ranged from 62% to 88% when using traditional theta/beta wave analysis and traditional classification techniques such as KNN and SVM. Accuracy has increased significantly and noticeably with the development of deep learning methods. Some recent studies, such as the Hu et al. (2025) study using the SCANet model and the Bansal et al. (2025) study using an

Autoencoders with ResNet and Attention, respectively, achieved classification accuracies of 99.78% and 99.42%. As shown in Table 1, some studies used small or limited datasets, while others used larger datasets or used performance optimization techniques such as feature selection, genetic algorithms, or PCA to improve models. This contributed to increased accuracy and improved other important metrics such as F1-score or MCC. The general trend from the table shows a shift from traditional wavelet analysis to hybrid deep learning models and modern techniques for discriminating between children with and without ADHD, and also shows improved performance over the years.

Table 1. Summary of EEG-Based ADHD Diagnostic Studies Included in the Literature Review (2015–2025).

Ref. no.	Year	Methodology	Data Used	Extracted Features	Accuracy	Limitations
Sangal and Sangal [10, 11]	2015	EEG analysis, TBR	EEG from children (6-14 years)	Theta/Beta Ratio, Beta-1 wave power (esp. Broca's area electrodes F7, FC5)	Sensitivity 86% (Beta-1 power)	Low specificity (false positives), limited age range (6-14), not for general screening
Snyder et al. [12, 13]	2015	Combination of ADHD assessment and TBR	EEG from 275 individuals with behavioral problems	Theta/Beta Ratio	Increased from 61% to 88% diagnostic accuracy	Difficulty recording EEG in young children due to head movement
Helgadóttir et al. [14, 15]	2015	EEG coherence + chronological age, multivariate classifier	EEG from 310 ADHD and 351 controls (Icelandic children 5.8-14 years)	EEG coherence, age	76% (independent test), 81% (cross-validation)	Same equipment only, single geographic area, cross-sectional (no longitudinal follow-up)
Mohammadi et al. [16, 17]	2016	Non-linear feature extraction + MLP neural network	EEG from 30 ADHD and 30 controls	ApEN, FD, Lyapunov exponent	92.28% (mRMR), 93.65% (DISR)	Small sample size, no validation study
Kamida et al. [18, 19]	2016	Spectral analysis	EEG from 80 ADHD and 59 controls (4-15 years)	Beta and alpha wave power	Not explicitly stated	Small sample size, no ADHD subtype distinction
Jahanshahloo et al. [20, 21]	2017	ERP recording + fractal & wavelet features + v-SVM classifier	ERP from 30 ADHD and 30 controls	Band power, FD, AR coefficients, wavelet coefficients	99.43% (fra-wave features), 97.65% (wavelet alone)	-
Rahadian et al. [22, 23]	2017	LVQ2NN + genetic algorithm	Data from 100 participants (45 ADHD)	Psychological symptom severity, LVQ2NN features	89.5% (LVQ2NN + GA), 80% (LVQ2NN alone), max 95%	Small sample, single data source
Farnia et al. [24, 25]	2018	EEG TBR during rest and neurofeedback	EEG from 61 ADHD and 59 controls	TBR (CZ and FZ regions)	Sensitivity up to 62%, specificity up to 73%	No other disorders were included, IQ was not directly measured, there

						were limited electrode sites, and there was no executive function data.
Lopez Marcano et al. [26, 27]	2018	Gaussian Mixture Model + Universal Background Model	EEG from 61 ADHD and 60 controls (male, 6-8 years)	EEG during the ANT and rest	92% (attention task), lower at rest	Small male-only sample, diagnosis based on mothers' reports only
Shiva Khoshnoud et al. [28, 29]	2018	Nonlinear dynamic analysis + SVM classifier	EEG from 12 ADHD and 12 controls	Largest Lyapunov Exponent, ApEn, Multifractal Spectrum	83.30%	Very small sample size
Vuckovic et al. [30, 31, 32]	2018	EEG preprocessing (ICA) + classification (ANN, SVM, LDA)	EEG from 31 spinal injury patients and 10 healthy controls	Cleaned EEG signals	>88% accuracy	Study on spinal injury patients, but methods relevant for ADHD EEG noise removal
Vahid et al. [33, 34]	2019	Deep learning (EEG Net) on ERP data	ERP EEG from children with ADHD and controls	Neurophysiological features during ERP	83% - 86%	Cannot distinguish ADHD subtypes, some misclassification, first deep learning study on EEG in ADHD
Chen et al. [35, 36]	2019	CNN + Grad-CAM interpretation	EEG from 50 ADHD and 57 controls	Brain activity patterns, spatial and frequency features	90.2% accuracy, AUC 0.96	Single-source data, no multi-site validation, limited flexibility due to predefined transforms
Zhang and Li [37, 38]	2019	CNN with transfer learning	Resting-state EEG (32 channels) from children	Phase Lag Index, brain network maps	94.39% accuracy, 97.83% sensitivity	No external dataset validation, no clinical or age diversity assessment
T. Wang and S. Kamata [39, 40]	2019	3D-CNN on FD complexity maps from sMRI	Structural MRI images (ADHD-200 dataset, children 7-12 years)	Fractal Dimension Complexity Map	69% accuracy	No functional data (fMRI, EEG), only structural imaging, limiting assessment of brain function
Khaleghi [41, 42]	2020	Extraction of 5 feature groups + statistical analysis	EEG from 30 ADHD and 30 controls (7-12 years)	Linear, shape, temporal, frequency, and nonlinear features	86.4% (nonlinear features best)	Small sample, no standardized EEG protocol, no diagnostic threshold

Moghaddari et al. [43]	2020	CNN on RGB images converted from EEG frequency bands	EEG from 31 ADHD and 30 controls	Frequency band images (delta to gamma)	97.47% (segment level), 98.48% (individual)	Small sample, single-task EEG recording, no clinical validation
TaghiBeyglou et al. [44, 45]	2020	Feature extraction (CSP, nonlinear) + KNN classifier	NBML EEG dataset (61 children)	Time and frequency features	83.33%	Small sample, no external data validation
Van Dijk et al. [46, 47]	2020	Analysis of TBR using 5 methods	EEG from 328 ADHD and 151 controls	Theta/Beta Ratio	No significant clinical difference	Only TBR was used, with no sleep quality or participant criteria, limiting interpretation.
Ciodaro et al. [48, 49]	2020	XGBoost and Residual CNN on resting EEG	EEG from 86 children	Alpha and beta energy	86.3% (XGBoost), 90% (CNN)	Small sample, no external validation
Amado-Caballero et al. [50, 51]	2020	CNN on daily activity spectrograms	148 children (73 ADHD mixed-type, 75 healthy)	Activity spectrograms	Sensitivity 97.62%, Specificity 99.52%	Lack of neural coverage (EEG, fMRI) may not fully represent ADHD neurological complexity.
Altınkaynak et al. [52, 53]	2020	EEG features (P300 latency/amplitude, FD, wavelet) + MLP, SVM, k-NN classifiers	EEG from 23 ADHD and 23 healthy children	P300 features, FD, wavelet coefficients	91.3% (MLP)	Small sample size, no ADHD subtype differentiation, few EEG channels, and possible hearing issues
Ekhlasi et al. [54.55]	2021	Effective connectivity (dPTE) + ANN + genetic algorithm	EEG from 62 children (31 ADHD, 31 controls)	ECVs (dPTE over 5 frequency bands)	96%	Small sample, risk of overfitting, limited generalizability
Peng et al. [56, 57]	2021	SSANN (3D CNN with two branches for MRI and sMRI)	ADHD-200 MRI/sMRI dataset	MRI and sMRI image features	72.89%	Image resizing may lose info, small dataset, limited generalizability
Maya-Piedrahita et al. [58, 59]	2021	EEG + Hidden Markov Model + SVM classifier	EEG from 67 children during the Reward Stop-Signal Task	EEG during failure trials only	90%	Only failure trials were used; no spectral analysis
Tosun [60.61]	2021	Spectral features + LSTM, SVM, and FFBPNN classifiers	EEG from an unspecified number of children	50 spectral features per 30-sec segment	88.88% (LSTM on Fp1-F7)	Number of participants unclear

Pham et al. [62]	2021	Ensemble learning on FD features	EEG from 61 ADHD and 60 controls	Nonlinear FD features	98.33%	No external data testing, single data source
Bakhtyari et al. [63, 64]	2021	ConvLSTM with attention mechanism on EEG connectivity tensors	400 EEG samples	Dynamic connectivity tensor features	99.75% (cross-validation)	Single dataset, possible overfitting, no clinical comparison
Cabarcas-Mena et al. [65, 66]	2021	CNN on recurrence plots from ERP data recorded with a low-cost EEG device	EEG from 44 children (22 ADHD, 22 healthy)	Recurrence plots from ERP	77.78%	Small sample, low-cost device affects quality, difficulty with child cooperation.
Cheng et al. [67, 68]	2022	LSTM model on EEG from two electrodes during the CPT task	EEG from 67 children (34 ADHD, 33 controls)	EEG from O1 and O2 electrodes	90.50%	Small single dataset, limited generalizability
Ghasemi et al. [69, 70]	2022	Machine learning (LR, GLM, DL, SVM) on ERP frequency band features	ERP data from children with auditory/visual stimuli	Band power (absolute and relative) features	Up to 100% (deep learning), 98.15% (theta band)	Possible overfitting with many features
Joy et al. [71, 72]	2022	TQWT + fractal features + ANN	EEG from 10 children (5 ADHD, 5 normal)	FD (Katz, Higuchi)	100%	Very small sample, no cognitive task data, private dataset
Mafi et al. [73, 74]	2022	Wavelet coherence + synchronization likelihood + high-dimensional CNN	EEG from 121 children (61 ADHD, 60 controls)	EEG connectivity features	99.17% accuracy, 100% recall	No external data testing
Esas and Latifoğlu [75, 76]	2023	Local mode and variational mode decomposition + deep learning	EEG ADHD and Control Children dataset from IEEE DataPort	Decomposed EEG subbands	95% accuracy, 97% sensitivity, 94% specificity	Small sample size
Alkahtani et al. [77, 78]	2023	Feature extraction + feature selection + multiple classifiers (CatBoost, RF, CNN, LSTM)	EEG from 61 ADHD and 60 controls during a visual task	Time, frequency, and entropy features	97.75% (CNN), 95.13% (CatBoost)	Single dataset, needs a larger sample validation
Cura et al. [79, 80]	2023	EEG-FMs from time & nonlinear features + SVM	EEG from 33 children (15 ADHD, 18 normal)	Hjorth parameters, skewness, kurtosis, Lyapunov exponent, FD, entropy	>80.2% (time features), >93.5% (nonlinear), 100% combined	Small sample, no external testing, no cognitive tasks
Yağmur Can et al. [81, 82]	2023	EEG temporal, frequency, and nonlinear features + LASSO + SVM	EEG from 121 children (61 ADHD, 60 normal)	Temporal, frequency, and nonlinear features	96.3% (SVM with LASSO)	Limited data diversity and size

Chugh et al. [83, 84]	2024	Hybrid CNN-LSTM deep learning on EEG data	ADHD dataset + FOCUS dataset	Spatial features (CNN), temporal features (LSTM)	98.86% (ADHD dataset), 98.28% (FOCUS dataset)	Difficult model interpretation, small biased sample in the FOCUS dataset
Alsharif et al. [85, 86]	2024	Statistical and spectral EEG features + PCA + Chi-square + Deep Learning models	EEG from 121 children (61 ADHD, 60 normal)	Standard deviation, kurtosis, entropy	95.59% (GRU-Transformer with PCA & Chi-square)	Small model variety, limited interpretability
Pedrollo et al. [87, 88]	2024	RF model + Genetic Algorithm on resting-state EEG	EEG from 856 participants	Theta band features	88.60%	Resting state only, no cognitive task data
Pappula & Anwar [89, 90]	2024	EEG spectrograms + ResNet-18 deep learning	EEG dataset (size not specified)	Spectrogram features	High accuracy, F1=0.9	No external data evaluation
Mondal et al. [91, 92]	2024	ADHD-Net (CNN + attention) + PCA + ICA preprocessing	EEG from 121 children (61 ADHD, 60 normal)	EEG features after filtering and ICA	97.96%	Small sample, no cognitive task data
Yousefimehr et al. [93, 94]	2024	Cost-sensitive LightGBM model	Not EEG-based	N/A	94% accuracy	No physiological data used, low interpretability, and no comparison with deep learning.
Singh et al. [95]	2024	Hybrid CNN + Transformer on EEG & fMRI data	EEG (100+ records) + fMRI	EEG and fMRI combined features	97.53%, 99% with a confusion matrix	Unclear sample details, poor model interpretability
Latifi et al. [96]	2024	Siamese CNN on EEG PSD brain maps + Grad-CAM interpretation	EEG data	PSD	99.17%	EEG only, no fMRI data
Bansal et al. [97]	2025	Autoencoder + Reptile Search Algorithm + ResNet + Attention	EEG from 121 children (61 ADHD, 60 normal)	Features extracted from 50,000 EEG samples	99.42% accuracy, 99.82% recall, 99.03% precision	Some children medicated with Ritalin, only for cognitive tasks, were recommended more diverse EEG recordings.
Hossain et al. [98]	2025	SVM with RBF kernel on PSD and SE features	EEG from 121 children (61 ADHD, 60 normal)	PSD, SE	99.2% average accuracy	Small sample size
Kim et al. [99]	2025	XGBoost on EEG five frequency bands + LOSO cross-validation	EEG from 168 participants	Delta, theta, alpha, beta, and gamma bands	90.81% accuracy, F1=0.9347	Small sample, medicated patients, no cognitive task data, class

			(107 ADHD, 61 normal)			imbalance handled
Hu et al. [100]	2025	SCANet deep model with multi-branch attention mechanisms	Public EEG dataset (ADD, ADHD, normal)	Attention on channels & time scales	99.78% accuracy, F1=89.14%	No external/multicenter validation, no comorbid disorder testing
Balamurugan et al. [101]	2025	EEG feature extraction + six machine learning models, including Fine KNN	EEG from 120 children (60 ADHD, 60 normal)	Top 11 EEG features selected via SVM	98% accuracy	Small sample, no external validation

The fifty studies included in this review reflect the rapid evolution of ADHD classification methods using EEG signals over the past decade. Early studies (2015–2017) primarily relied on traditional TBR analysis and simple classifiers such as KNN and SVM (e.g., Sangal & Sangal, Snyder et al., Helgadóttir et al.), achieving moderate accuracy ranging from 61% to 92%. These studies were generally limited by small, geographically restricted samples, low age and gender diversity, and a lack of ADHD subtype differentiation.

From 2018 onwards, the focus shifted to advanced feature extraction and deep learning approaches. ERP-based features and CNN architectures (e.g., Farnia et al., Vahid et al., and Chen et al.) improved classification accuracy to 83–90%, while hybrid models combining CNN, LSTM, or attention mechanisms (e.g., Chugh et al., Bansal et al., and Hu et al.) achieved high accuracies exceeding 98%, along with improved F1-scores and MCC values. These models leveraged the combination of temporal, spatial, and spectral EEG features, highlighting the added value of multi-dimensional analysis. Fig. 3 summarizes the evolution of classification accuracy over time, showing the performance achieved by each of the fifty studies included in this review.

Despite the high accuracy achieved by recent deep learning models, several challenges remain regarding their interpretability and clinical integration. Many architectures, such as hybrid CNN-LSTM and SCANet, operate as black-box models, making it difficult for clinicians to understand the rationale behind predictions. This lack of transparency limits trust and may hinder adoption in real-world clinical settings. Additionally, most high-performing models have been trained on single-site or small datasets, raising concerns about overfitting and limited generalizability across different populations. To facilitate clinical translation, future studies should focus on using multi-center datasets, incorporating diverse age groups and ADHD subtypes, and applying standardized EEG recording and preprocessing protocols. Techniques for model explanation, such as attention maps or feature importance metrics, should also be integrated to provide interpretable insights, enabling clinicians to make informed decisions based on EEG-based predictions.

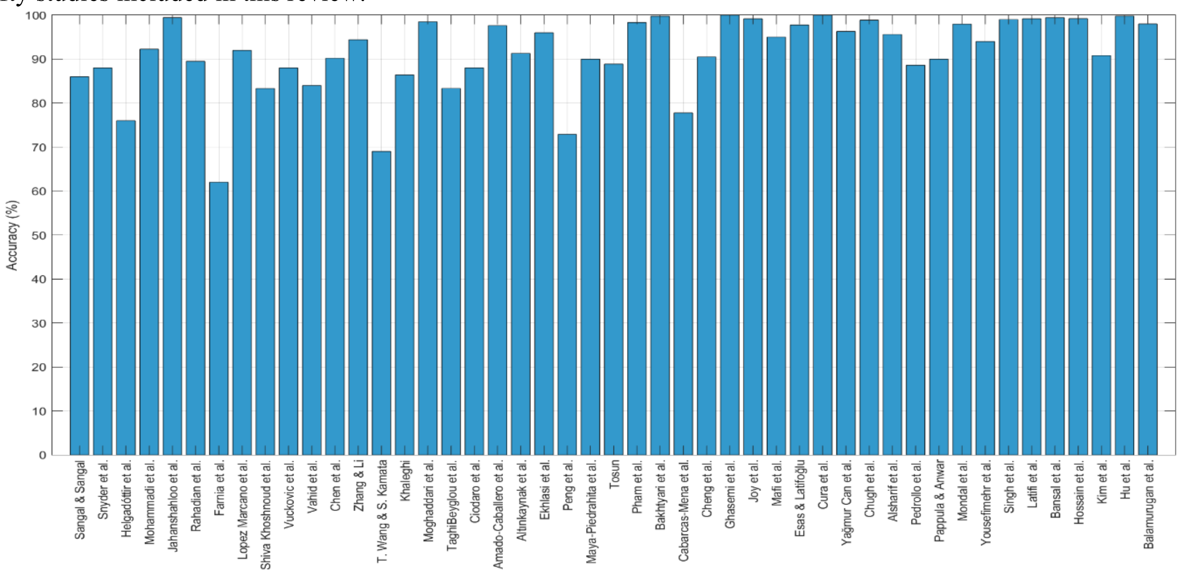

Fig. 3. EEG-based ADHD classification accuracy over time. Each bar corresponds to the accuracy reported by an individual study from 2015 to 2025.

- **Limitations and Future Directions**

Nevertheless, some limitations remain, including small or geographically limited sample sizes, limited gender or age diversity, and the failure to address disease subtypes in some studies. Additionally, several high-performing models relied on single-site or small datasets, raising concerns about overfitting and limited generalizability. Future research should prioritize multi-center datasets, inclusion of diverse age groups and ADHD subtypes, and incorporation of standardized EEG recording protocols. Furthermore, enhancing model interpretability, especially for complex architectures like SCANet and hybrid CNN-LSTM, is crucial for clinical applicability. This suggests the need to use standardized protocols and expand the dataset to enhance the generalizability and practical application of these findings in ADHD diagnosis.

5. Conclusion

This review demonstrates that ADHD classification using EEG signals from children and adolescents has become more accurate and efficient with advances in deep learning techniques. Traditional methods, relying on TBR analysis and simple classifiers such as KNN and SVM, achieved limited accuracy, whereas hybrid deep learning models—including CNN-LSTM, SCANet, and attention-based architectures—consistently attained the highest performance. These models improve understanding of neural structures but often function as “black boxes,” making it difficult to identify which features or neural patterns drive classification. In contrast, some traditional machine learning models offer better interpretability of the features used but provide limited anatomical insight, even when achieving high accuracy.

Furthermore, most studies did not consider ADHD subtypes, and datasets were often geographically restricted or included only male participants, limiting generalizability. These findings highlight the need for larger, more diverse datasets that include all ADHD subtypes and both sexes, as well as standardized EEG recording and preprocessing protocols. Incorporating interpretability techniques into deep learning models can support clinicians in understanding model predictions, linking methodological advances directly to practical clinical applications, and showcasing the significant progress made in this field over the past decade.

What distinguishes this review from previous studies is its comprehensive focus on pediatric populations, including both children and adolescents, while considering a wide range of datasets, sample sizes, and EEG-based modalities. Unlike other reviews that mainly emphasize feature extraction, machine learning techniques, or ADHD subtypes individually, our work synthesizes methodological advancements alongside sample diversity and clinical considerations, providing a more holistic perspective on EEG-based ADHD diagnosis.

References

[1] F. Svenaeus, “Diagnosing mental disorders and saving the normal,” *Med. Heal. Care Philos.*, vol. 17, no. 2, pp. 241–244, 2014, doi: 10.1007/s11019-013-9529-6.

[2] K. Sayal, V. Prasad, D. Daley, T. Ford, and D. Coghill, “ADHD in children and young people: prevalence, care pathways, and service provision,” *The Lancet Psychiatry*, vol. 5, no. 2, pp. 175–186, Feb. 2018, doi: 10.1016/S2215-0366(17)30167-0.

[3] P. Arpaia *et al.*, “A Systematic Review on Feature Extraction in Electroencephalography-Based Diagnostics and Therapy in Attention Deficit Hyperactivity Disorder,” *Sensors*, vol. 22, no. 13, 2022, doi: 10.3390/s22134934.

[4] G. Polanczyk, M. S. de Lima, B. L. Horta, J. Biederman, and L. A. Rohde, “The Worldwide Prevalence of ADHD: A Systematic Review and Metaregression Analysis,” *Am. J. Psychiatry*, vol. 164, no. 6, pp. 942–948, 2007, doi: 10.1176/ajp.2007.164.6.942.

[5] A. Lenartowicz and S. K. Loo, “Use of EEG to diagnose ADHD,” *Curr. Psychiatry Rep.*, vol. 16, no. 11, p. 498, Nov. 2014, doi: 10.1007/s11920-014-0498-0.

[6] J. F. Lubar, “Discourse on the development of EEG diagnostics and biofeedback for attention-deficit/hyperactivity disorders,” *Biofeedback Self. Regul.*, vol. 16, no. 3, pp. 201–225, 1991, doi: 10.1007/BF01000016.

[7] T. A. Thomas and A. Gaffney, “Neuropsychiatric EEG-based assessment aid system,” *US Pharm*, vol. 42, pp. 25–27, 2017.

[8] B. Albrecht, H. Uebel-von Sandersleben, K. Wiedmann, and A. Rothenberger, “ADHD History of the Concept: the Case of the Continuous Performance Test,” *Curr. Dev. Disord. Reports*, vol. 2, no. 1, pp. 10–22, 2015, doi: 10.1007/s40474-014-0035-1.

[9] K. W. Lange, S. Reichl, K. M. Lange, L. Tucha, and O. Tucha, “The history of attention deficit hyperactivity disorder,” *Atten. Defic. Hyperact. Disord.*, vol. 2, no. 4, pp. 241–255, Dec. 2010, doi: 10.1007/s12402-010-0045-8.

[10] R. B. Sangal and J. M. Sangal, “Use of EEG beta-1 power and theta/beta ratio over Broca’s area to confirm diagnosis of attention deficit/hyperactivity disorder in children,” *Clin. EEG Neurosci.*, vol. 46, no. 3, pp. 177–182, 2015, doi: 10.1177/1550059414527284.

[11] P. Marshall, J. Hoelzle, and M. Nikolas, “Diagnosing Attention-Deficit/Hyperactivity Disorder (ADHD) in young adults: A qualitative review of the utility of assessment measures and recommendations for improving the diagnostic process,” *Clin. Neuropsychol.*, vol. 35, no. 1, pp. 165–198, 2021, doi: 10.1080/13854046.2019.1696409.

[12] S. M. Snyder, T. A. Rugino, M. Hornig, and M. A. Stein, "Integration of an EEG biomarker with a clinician's ADHD evaluation," *Brain Behav.*, vol. 5, no. 4, pp. 1–17, 2015, doi: 10.1002/brb3.330.

[13] A. R. Clarke, R. J. Barry, R. McCarthy, and M. Selikowitz, "EEG differences between good and poor responders to methylphenidate and dexamphetamine in children with attention-deficit/hyperactivity disorder," *Clin. Neurophysiol.*, vol. 113, no. 2, pp. 194–205, 2002, doi: https://doi.org/10.1016/S1388-2457(01)00736-2.

[14] H. Helgadóttir *et al.*, "Electroencephalography as a clinical tool for diagnosing and monitoring attention deficit hyperactivity disorder: A cross-sectional study," *BMJ Open*, vol. 5, no. 1, 2015, doi: 10.1136/bmjopen-2014-005500.

[15] N. Patel, M. Patel, and H. Patel, "Patel N, Patel M, Patel H. ADHD and Comorbid Conditions, Current Directions in ADHD and Its Treatment. Norvilitis JM (ed), 2012. doi:10.5772/30279. http://www.intechopen.com/books/current-directions-in-adhd-and-its-treatment/adhd-and-comorbidity," *Curr. Dir. ADHD Its Treat.*, 2012, doi: 10.5772/30279.

[16] M. R. Mohammadi, A. Khaleghi, A. M. Nasrabadi, S. Rafieivand, M. Begol, and H. Zarafshan, "EEG classification of ADHD and normal children using non-linear features and neural network," *Biomed. Eng. Lett.*, vol. 6, no. 2, pp. 66–73, 2016, doi: 10.1007/s13534-016-0218-2.

[17] F. Alba-Sanchez, O. Yanez-Suarez, and H. Brust-Carmona, "Assisted diagnosis of Attention-Deficit Hyperactivity Disorder through EEG bandpower clustering with self-organizing maps," in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, 2010, pp. 2447–2450. doi: 10.1109/IEMBS.2010.5626360.

[18] Y. Hirooka *et al.*, "EEG Power Spectrum Analysis in Children with ADHD," *Yonago Acta Med.*, vol. 59, pp. 169–173, 2016.

[19] A. R. Clarke, R. J. Barry, R. McCarthy, and M. Selikowitz, "Age and sex effects in the EEG: development of the normal child," *Clin. Neurophysiol.*, vol. 112, no. 5, pp. 806–814, 2001, doi: https://doi.org/10.1016/S1388-2457(01)00488-6.

[20] H. R. Hanshahloo, M. Shamsi, E. Ghasemi, and A. Kouhi, "Automated and ERP-Based Diagnosis of Attention-Deficit Hyperactivity Disorder in Children," *J. Med. Signals Sens.*, vol. 7, no. 1, pp. 26–32, 2017, doi: 10.4103/2228-7477.199152.

[21] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 2, pp. 129–150, 2011, doi: https://doi.org/10.1016/j.acha.2010.04.005.

[22] B. A. Rahadian, C. Dewi, and B. Rahayudi, "The performance of genetic algorithm learning vector quantisation 2 neural network on identification of the types of attention deficit hyperactivity disorder," *Proc. - 2017 Int. Conf. Sustain. Inf. Eng. Technol. SIET 2017*, vol. 2018-Janua, pp. 337–341, 2017, doi: 10.1109/SIET.2017.8304160.

[23] Y. Pangaribuan and M. Sagala, "Menerapkan Jaringan Saraf Tiruan untuk Mengenali Pola Huruf Menggunakan Metode Perceptron," *J. Tek. Inform. Unika St. Thomas*, vol. 2, no. 2, pp. 53–59, 2017.

[24] S. Farina, A. Alipour, A. B. Shafat, S. Fattahi, J. A. Rad, and M. Ghahremanloo, "Investigating the Relationship between Theta/Beta Ratio and Intensity of Disease in Children with ADHD," vol. 5, no. 1, pp. 1–2, 2018.

[25] J. F. Saad, M. R. Kohn, S. Clarke, J. Lagopoulos, and D. F. Hermens, "Is the Theta/Beta EEG Marker for ADHD Inherently Flawed?," *J. Atten. Disord.*, vol. 22, no. 9, pp. 815–826, 2018, doi: 10.1177/1087054715578270.

[26] J. L. Marcano, M. A. Bell, and A. A. (Louis) Beex, "Classification of ADHD and non-ADHD subjects using a universal background model," *Biomed. Signal Process. Control*, vol. 39, pp. 204–212, 2018, doi: https://doi.org/10.1016/j.bspc.2017.07.023.

[27] M. R. Rueda *et al.*, "Development of attentional networks in childhood," *Neuropsychologia*, vol. 42, no. 8, pp. 1029–1040, 2004, doi: https://doi.org/10.1016/j.neuropsychologia.2003.12.012.

[28] S. Khoshnoud, M. A. Nazari, and M. Shamsi, "Functional brain dynamic analysis of ADHD and control children using nonlinear dynamical features of EEG signals," *J. Integr. Neurosci.*, vol. 17, no. 1, pp. 17–30, 2018, doi: 10.3233/JIN-170033.

[29] M. Ikawa, M. Nakanishi, T. Furukawa, S. Nakaaki, S. Hori, and S. Yoshida, "Relationship between EEG dimensional complexity and neuropsychological findings in Alzheimer's disease," *Psychiatry Clin. Neurosci.*, vol. 54, no. 5, pp. 537–541, 2000, doi: 10.1046/j.1440-1819.2000.00749.x.

[30] A. Vuckovic, V. J. F. Gallardo, M. Jarjees, M. Fraser, and M. Purcell, "Prediction of central neuropathic pain in spinal cord injury based on EEG classifier," *Clin. Neurophysiol.*, vol. 129, no. 8, pp. 1605–1617, 2018, doi: https://doi.org/10.1016/j.clinph.2018.04.750.

[31] M. S. Jarjees, "PhD: The causality between Electroencephalogram (EEG) and Central Neuropathic Pain (CNP), and the effectiveness of neuromodulation strategies on cortical excitability and CNP in patients with spinal cord injury," pp. 27–287, 2017, [Online]. Available: http://theses.gla.ac.uk/7985/

[32] A. Vuckovic, M. S. Jarjees, M. A. Hasan, M. Purcell, and M. Fraser, "Chapter 10 - EEG biomarkers of pain and applications of machine learning," in *Spinal Cord Injury Pain*, C. N. Sang and C. E. Hulsebosch, Eds., Academic Press, 2022, pp. 199–225. doi: https://doi.org/10.1016/B978-0-12-818662-6.00019-4.

[33] A. Vahid, A. Bluschke, V. Roessner, S. Stober, and C. Beste, "Deep learning based on event-related EEG differentiates children with ADHD from healthy controls," *J. Clin. Med.*, vol. 8, no. 7, 2019, doi: 10.3390/jcm8071055.

[34] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya, "Deep learning for healthcare

applications based on physiological signals: A review," *Comput. Methods Programs Biomed.*, vol. 161, pp. 1–13, Jul. 2018, doi: 10.1016/j.cmpb.2018.04.005.

[35] H. Chen, Y. Song, and X. Li, "Use of deep learning to detect personalized spatial-frequency abnormalities in EEGs of children with ADHD," *J. Neural Eng.*, vol. 16, no. 6, p. 66046, 2019.

[36] M. Arns, "EEG-Based Personalized Medicine in ADHD: Individual Alpha Peak Frequency as an Endophenotype Associated with Nonresponse," *J. Neurother.*, vol. 16, no. 2, pp. 123–141, 2012, doi: 10.1080/10874208.2012.677664.

[37] Z. Zhang and X. Li, "Use transfer learning to promote identification ADHD children with EEG recordings," in *2019 Chinese Automation Congress (CAC)*, 2019, pp. 2809–2813. doi: 10.1109/CAC48633.2019.8997426.

[38] K. Weiss, T. M. Khoshgoftaar, and D. D. Wang, *A survey of transfer learning*, vol. 3, no. 1. Springer International Publishing, 2016. doi: 10.1186/s40537-016-0043-6.

[39] T. Wang and S. Kamata, "Classification of Structural MRI Images in ADHD Using 3D Fractal Dimension Complexity Map," in *2019 IEEE International Conference on Image Processing (ICIP)*, 2019, pp. 215–219. doi: 10.1109/ICIP.2019.8802930.

[40] D. Dai, J. Wang, J. Hua, and H. He, "Classification of ADHD children through multimodal magnetic resonance imaging," *Front. Syst. Neurosci.*, vol. 6, no. SEPTEMBER, pp. 1–8, 2012, doi: 10.3389/fnsys.2012.00063.

[41] A. Khaleghi, P. M. Birgani, M. F. Fooladi, and M. R. Mohammadi, "Applicable features of electroencephalogram for ADHD diagnosis," *Res. Biomed. Eng.*, vol. 36, no. 1, pp. 1–11, 2020, doi: 10.1007/s42600-019-00036-9.

[42] H. Zarafshan, A. Khaleghi, M. R. Mohammadi, M. Moeini, and N. Malmir, "Electroencephalogram complexity analysis in children with attention-deficit/hyperactivity disorder during a visual cognitive task," *J. Clin. Exp. Neuropsychol.*, vol. 38, no. 3, pp. 361–369, 2016, doi: 10.1080/13803395.2015.1119252.

[43] M. Moghaddari, M. Z. Lighvan, and S. Danishvar, "Diagnose ADHD disorder in children using convolutional neural network based on continuous mental task EEG," *Comput. Methods Programs Biomed.*, vol. 197, p. 105738, Dec. 2020, doi: 10.1016/j.cmpb.2020.105738.

[44] B. TaghiBeyglou, N. Hasanzadeh, F. Bagheri, and M. Jahed, "+," in *2020 28th Iranian Conference on Electrical Engineering (ICEE)*, 2020, pp. 1–5. doi: 10.1109/ICEE50131.2020.9260711.

[45] S. Khoshnoud, M. Shamsi, and M. A. Nazari, "Non-linear EEG analysis in children with attention-deficit/hyperactivity disorder during the rest condition," in *2015 22nd Iranian Conference on Biomedical Engineering (ICBME)*, 2015, pp. 87–92. doi: 10.1109/ICBME.2015.7404122.

[46] H. van Dijk *et al.*, "Different Spectral Analysis Methods for the Theta/Beta Ratio Calculate Different Ratios But Do Not Distinguish ADHD from Controls," *Appl. Psychophysiol. Biofeedback*, vol. 45, no. 3, pp. 165–173, 2020, doi: 10.1007/s10484-020-09471-2.

[47] M. Le Van Quyen and A. Bragin, "Analysis of dynamic brain oscillations: methodological advances," *Trends Neurosci.*, vol. 30, no. 7, pp. 365–373, Jul. 2007, doi: 10.1016/j.tins.2007.05.006.

[48] G. Ciodaro, A. J. Najafabadi, and B. Godde, "Resting State EEG Classification of Children With ADHD," in *2020 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, 2020, pp. 1–6. doi: 10.1109/SPMB50085.2020.9353628.

[49] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, in KDD '16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 785–794. doi: 10.1145/2939672.2939785.

[50] P. Amado-Caballero *et al.*, "Objective ADHD Diagnosis Using Convolutional Neural Networks over Daily-Life Activity Records," *IEEE J. Biomed. Heal. Informatics*, vol. 24, no. 9, pp. 2690–2700, 2020, doi: 10.1109/JBHI.2020.2964072.

[51] D. Wagner, K. Kalischewski, J. Velten, and A. Kummert, "Activity recognition using inertial sensors and a 2-D convolutional neural network," in *2017 10th International Workshop on Multidimensional (nD) Systems (nDS)*, 2017, pp. 1–6. doi: 10.1109/NDS.2017.8070615.

[52] M. Altınkaynak *et al.*, "Diagnosis of Attention Deficit Hyperactivity Disorder with combined time and frequency features," *Biocybern. Biomed. Eng.*, vol. 40, no. 3, pp. 927–937, 2020, doi: 10.1016/j.bbe.2020.04.006.

[53] R. Quian Quiroga, O. W. Sakowitz, E. Basar, and M. Schürmann, "Wavelet Transform in the analysis of the frequency composition of evoked potentials," *Brain Res. Protoc.*, vol. 8, no. 1, pp. 16–24, 2001, doi: https://doi.org/10.1016/S1385-299X(01)00077-0.

[54] A. Ekhlasi, A. M. Nasrabadi, and M. Mohammadi, "Classification of the children with ADHD and healthy children based on the directed phase transfer entropy of EEG signals," *Front. Biomed. Technol.*, vol. 8, no. 2, pp. 115–122, 2021, doi: 10.18502/fbt.v8i2.6515.

[55] T. T. Erguzel, S. Ozekes, O. Tan, and S. Gultekin, "Feature Selection and Classification of Electroencephalographic Signals: An Artificial Neural Network and Genetic Algorithm Based Approach," *Clin. EEG Neurosci.*, vol. 46, no. 4, pp. 321–326, 2015, doi: 10.1177/1550059414523764.

[56] J. Peng, M. Debnath, and A. K. Biswas, "Efficacy of novel Summation-based Synergetic Artificial Neural Network in ADHD diagnosis," *Mach. Learn. with Appl.*, vol. 6, no. January, p. 100120, 2021, doi: 10.1016/j.mlwa.2021.100120.

[57] B. B. Lahey *et al.*, "DSM-IV field trials for attention deficit hyperactivity disorder in children and adolescents," *Am. J. Psychiatry*, vol. 151, no. 11, p. 1673—1685, Nov. 1994, doi: 10.1176/ajp.151.11.1673.

[58] M. C. Maya-Piedrahita, D. Cárdenas-Peña, and A. A. Orozco-Gutierrez, "Diagnosis of attention

deficit and hyperactivity disorder (ADHD) using Hidden Markov Models," *Eur. Signal Process. Conf.*, vol. 2021-Janua, pp. 1205–1209, 2021, doi: 10.23919/Eusipco47968.2020.9287676.

[59] N. A. Ann, A. Juan, and T. Lucas, "Facultat de Matemàtiques i Informàtica Universitat de Barcelona Machine Learning: Modelos Ocultos de Markov (HMM) y Redes Director," 2017.

[60] M. Tosun, "Effects of spectral features of EEG signals recorded with different channels and recording statuses on ADHD classification with deep learning," *Phys. Eng. Sci. Med.*, vol. 44, no. 3, pp. 693–702, 2021, doi: 10.1007/s13246-021-01018-x.

[61] H. Chen, W. Chen, Y. Song, L. Sun, and X. Li, "EEG characteristics of children with attention-deficit/hyperactivity disorder," *Neuroscience*, vol. 406, pp. 444–456, 2019, doi: https://doi.org/10.1016/j.neuroscience.2019.03.048.

[62] P. T. Viet Huong, N. A. Tu, and T. A. Vu, "Ensemble learning in detecting ADHD children by utilizing the non-linear features of EEG signal," *CEUR Workshop Proc.*, vol. 3026, no. Computing4Human 2021, pp. 129–141, 2021.

[63] M. Bakhtyari, S. Mirzaei, and H. Amiri, "Combination of ConvLSTM and Attention mechanism to diagnose ADHD based on EEG signals," in *2021 7th International Conference on Signal Processing and Intelligent Systems (ICSPIS)*, 2021, pp. 1–5. doi: 10.1109/ICSPIS54653.2021.9729359.

[64] X. SHI, Z. Chen, H. Wang, D.-Y. Yeung, W. Wong, and W. WOO, "Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting," in *Advances in Neural Information Processing Systems*, C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, Eds., Curran Associates, Inc., 2015. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2015/file/07563a3fe3bbe7e3ba84431ad9d055af-Paper.pdf

[65] Y. P. Cabarcas-Mena, A. G. Marrugo, and S. H. Contreras-Ortiz, "Classification of Cognitive Evoked Potentials for ADHD Detection in Children using Recurrence Plots and CNNs," in *2021 XXIII Symposium on Image, Signal Processing and Artificial Vision (STSIVA)*, 2021, pp. 1–6. doi: 10.1109/STSIVA53688.2021.9592021.

[66] I. M. Mercado-Aguirre, K. Gutiérrez-Ruiz, and S. H. Contreras-Ortiz, "Acquisition and Analysis of Cognitive Evoked Potentials using an Emotiv Headset for ADHD Evaluation in Children," in *2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA)*, 2019, pp. 1–5. doi: 10.1109/STSIVA.2019.8730225.

[67] C. Belmonte *et al.*, "Towards high-accuracy classifying attention-deficit/hyperactivity disorders using CNN-LSTM model Towards high-accuracy classifying attention-deficit / hyperactivity disorders using CNN-LSTM model."

[68] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A Search Space Odyssey," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, 2017, doi: 10.1109/TNNLS.2016.2582924.

[69] E. Ghasemi, M. Ebrahimi, and E. Ebrahimie, "Machine learning models effectively distinguish attention-deficit/hyperactivity disorder using event-related potentials," *Cogn. Neurodyn.*, vol. 16, no. 6, pp. 1335–1349, 2022, doi: 10.1007/s11571-021-09746-2.

[70] M. Sabeti, R. Boostani, and E. Moradi, "Event related potential (ERP) as a reliable biometric indicator: A comparative approach," *Array*, vol. 6, p. 100026, 2020, doi: https://doi.org/10.1016/j.array.2020.100026.

[71] R. C. Joy *et al.*, "Detection and Classification of ADHD from EEG Signals Using Tunable Q-Factor Wavelet Transform," *J. Sensors*, vol. 2022, pp. 1–17, 2022, doi: 10.1155/2022/3590973.

[72] I. W. Selesnick, "Wavelet Transform With Tunable Q-Factor," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3560–3575, 2011, doi: 10.1109/TSP.2011.2143711.

[73] M. Mafi and S. Radfar, "High Dimensional Convolutional Neural Network for EEG Connectivity-Based Diagnosis of ADHD," *J. Biomed. Phys. Eng.*, vol. 12, no. 6, pp. 645–654, 2022, doi: 10.31661/jbpe.v0i0.2108-1380.

[74] G. Michelini *et al.*, "Atypical functional connectivity in adolescents and adults with persistent and remitted ADHD during a cognitive control task," *Transl. Psychiatry*, vol. 9, no. 1, 2019, doi: 10.1038/s41398-019-0469-7.

[75] M. Y. Esas and F. Latifoğlu, "Detection of ADHD from EEG signals using new hybrid decomposition and deep learning techniques," *J. Neural Eng.*, vol. 20, no. 3, 2023, doi: 10.1088/1741-2552/acc902.

[76] A. Saxena, "An Introduction to Convolutional Neural Networks," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 10, no. 12, pp. 943–947, 2022, doi: 10.22214/ijraset.2022.47789.

[77] H. Alkahtani, T. H. H. Aldhyani, Z. A. T. Ahmed, and A. A. Alqarni, "Developing System-Based Artificial Intelligence Models for Detecting the Deficit Hyperactivity Disorder," *Mathematics*, vol. 11, no. 22, 2023, doi: 10.3390/math11224698.

[78] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.-A. Muller, "Deep learning for time series classification: a review," *Data Min. Knowl. Discov.*, vol. 33, no. 4, pp. 917–963, 2019, doi: 10.1007/s10618-019-00619-1.

[79] O. K. Cura, S. K. Atli, S. Y. Sen, and A. Akan, "Classification of ADHD by Using Multiple Feature Maps of EEG Signals and Deep Feature Extraction," *Eur. Signal Process. Conf.*, pp. 1065–1069, 2023, doi: 10.23919/EUSIPCO58844.2023.10289929.

[80] Y. K. Boroujeni, A. A. Rastegari, and H. Khodadadi, "Diagnosis of attention deficit hyperactivity disorder using non-linear analysis of the EEG signal," *IET Syst. Biol.*, vol. 13, no. 5, pp. 260–266, 2019, doi: 10.1049/iet-syb.2018.5130.

[81] Y. Can, İ. Bigat, F. Nassehi, A. Eken, and O. Eroğul, "Determination of the Optimal EEG-based Features to Detect ADHD by Machine Learning Algorithms," in *2023 Medical*

Technologies Congress (TIPTEKNO), 2023, pp. 1–4. doi: 10.1109/TIPTEKNO59875.2023.10359207.

[82] M. McVoy, S. Lytle, E. Fulchiero, M. E. Aebi, O. Adeleye, and M. Sajatovic, "A systematic review of quantitative EEG as a possible biomarker in child psychiatric disorders," *Psychiatry Res.*, vol. 279, pp. 331–344, 2019, doi: https://doi.org/10.1016/j.psychres.2019.07.004.

[83] N. Chugh, S. Aggarwal, and A. Balyan, "The Hybrid Deep Learning Model for Identification of Attention-Deficit/Hyperactivity Disorder Using EEG," *Clin. EEG Neurosci.*, vol. 55, no. 1, pp. 22–33, 2024, doi: 10.1177/15500594231193511.

[84] A. D. Nazhvani, R. Boostani, S. Afrasiabi, and K. Sadatnezhad, "Classification of ADHD and BMD patients using visual evoked potential," *Clin. Neurol. Neurosurg.*, vol. 115, no. 11, pp. 2329–2335, 2013, doi: 10.1016/j.clineuro.2013.08.009.

[85] N. Alsharif, M. H. Al-Adhaileh, and M. Al-Yaari, "Diagnosis of attention deficit hyperactivity disorder: A deep learning approach," 2024. doi: 10.3934/math.2024517.

[86] A. Alim and M. H. Imtiaz, "Automatic Identification of Children with ADHD from EEG Brain Waves," *Signals*, vol. 4, no. 1, pp. 193–205, 2023, doi: 10.3390/signals4010010.

[87] G. R. Pedrollo, L. B. Bagesteiro, A. R. Franco, and A. Balbinot, "ADHD Diagnosis through Resting-State EEG Frequency Analysis with Random Forest." *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.*, vol. 2024, pp. 1–4, Jul. 2024, doi: 10.1109/EMBC53108.2024.10782906.

[88] G. Biau and E. Scornet, "A random forest guided tour," *TEST*, vol. 25, no. 2, pp. 197–227, 2016, doi: 10.1007/s11749-016-0481-7.

[89] M. Pappula and S. M. Anwar, "An ADHD Diagnostic Interface Based on EEG Spectrograms and Deep Learning Techniques," *Proc. 20th Int. Symp. Med. Inf. Process. Anal. SIPAIM 2024*, pp. 3–6, 2024, doi: 10.1109/SIPAIM62974.2024.10783602.

[90] R. Baboli, M. Cao, E. Martin, J. M. Halperin, K. Wu, and X. Li, "Distinct structural brain network properties in children with familial versus non-familial attention-deficit/hyperactivity disorder (ADHD)," *Cortex*, vol. 179, pp. 1–13, 2024, doi: https://doi.org/10.1016/j.cortex.2024.06.019.

[91] S. Mondal, D. Biswas, P. Mukherjee, and A. H. Roy, "ADHD-Net: A Robust Deep Learning-based Framework for Detecting ADHD Patients using EEG," in *2024 Global Conference on Communications and Information Technologies (GCCIT)*, 2024, pp. 1–6. doi: 10.1109/GCCIT63234.2024.10862881.

[92] Ö. Kasim, "Identification of attention deficit hyperactivity disorder with deep learning model," *Phys. Eng. Sci. Med.*, vol. 46, no. 3, pp. 1081–1090, 2023, doi: 10.1007/s13246-023-01275-y.

[93] B. Yousefimehr, M. Ghatee, and A. Heydari, "Improving ADHD Detection with Cost-Sensitive LightGBM," in *2024 14th International Conference on Computer and Knowledge Engineering (ICCKE)*, 2024, pp. 109–113. doi: 10.1109/ICCKE65377.2024.10874782.

[94] D. Hardy, G. Malléus, and J.-N. Méreur, "The foundations of NGN," *Networks*, pp. 611–620, 2002, doi: 10.1007/978-3-642-55498-8_24.

[95] A. R. Singh, G. Singh, and N. Saluja, "Unlocking the Potential of EEG: Convolutional Neural Networks for ADHD Classification," *2024 IEEE 3rd World Conf. Appl. Intell. Comput. AIC 2024*, no. November, pp. 269–273, 2024, doi: 10.1109/AIC61668.2024.10731030.

[96] B. Latifi, A. Amini, and A. Motie Nasrabadi, "Siamese based deep neural network for ADHD detection using EEG signal," *Comput. Biol. Med.*, vol. 182, p. 109092, Nov. 2024, doi: 10.1016/j.compbiomed.2024.109092.

[97] J. Bansal, G. Gangwar, M. Aljaidi, A. Alkoradees, and G. Singh, "EEG-Based ADHD Classification Using Autoencoder Feature Extraction and ResNet with Double Augmented Attention Mechanism," *Brain Sci.*, vol. 15, no. 1, 2025, doi: 10.3390/brainsci15010095.

[98] M. B. Hossain, M. A. I. Himel, M. A. Rahim, S. Mahmood, A. S. M. Miah, and J. Shin, "Classification of ADHD and Healthy Children Using EEG Based Multi-Band Spatial Features Enhancement," 2025, [Online]. Available: http://arxiv.org/abs/2504.04664

[99] J. W. Kim, B. N. Kim, J. I. Kim, C. M. Yang, and J. Kwon, "Electroencephalogram (EEG) Based Prediction of Attention Deficit Hyperactivity Disorder (ADHD) Using Machine Learning," *Neuropsychiatr. Dis. Treat.*, vol. 21, no. February, pp. 271–279, 2025, doi: 10.2147/NDT.S509094.

[100] H. Hu *et al.*, "SCANet: An Innovative Multiscale Selective Channel Attention Network for EEG-Based ADHD Recognition," *IEEE Sens. J.*, vol. 25, no. 11, pp. 20920–20932, 2025, doi: 10.1109/JSEN.2025.3560349.

[101] A. M. Balamurugan, V. S., and V. N. M., "Machine Learning Approach for ADHD Diagnosis in Children using EEG," in *2025 International Conference on Intelligent Systems and Computational Networks (ICISCN)*, 2025, pp. 1–7. doi: 10.1109/ICISCN64258.2025.10934201.