Review of Computer Engineering Research
2025 Vol. 12, No. 2, pp. 80-93
ISSN(e): 2410-9142
ISSN(p): 2412-4281
DOI: 10.18488/76.v12i2.4217

Artificial intelligence methods for identification of ADHD in children based on EEG signals

Noor Abdulmuttaleb Jaafar[1]
Rana Jassim Mohammed[2]
Shaymaa Taha Ahmed[3]
Qusay Kanaan Kadhim[4+]
Rasha Mahdi Abdulkader[5]

[1]*Department of Administration & Finance, University of Diyala, Baqubah 32001, Iraq.*
[1]*Email: noor84abd84@uodiyala.edu.iq*
[2,3,4,5]*Department of Computer Science, University of Diyala, Baqubah 32001, Iraq.*
[2]*Email: ad.ranamohammed@uodiyala.edu.iq*
[3]*Email: shaimaaAhmed@uodiyala.edu.iq*
[4]*Email: dr.qusay.kanaan@uodiyala.edu.iq*
[5]*Email: rashamhdi@uodiyala.edu.iq*

(+ Corresponding author)

Article History
Received: 13 March 2025
Revised: 28 April 2025
Accepted: 8 May 2025
Published: 22 May 2025

Keywords
Artificial intelligence
Attention deficit hyperactivity disorder
Bidirectional long short-term memory
Convolutional neural networks
Deep learning
Electroencephalograms.

ABSTRACT

Children with Attention Deficit Hyperactivity Disorder (ADHD) are among the most common neurodevelopmental disorders. The incidence of this disorder in society shows an increasing trend from the past to the present. Recent developments suggest that Artificial Intelligence and Electroencephalogram (EEG) analysis can accurately diagnose cases of ADHD in children. By combining a new type of Continuous Wavelet Transform (CWT) with Variational Mode Decomposition (VMD), a novel algorithm for self-adaptive signal processing that is more resilient to sampling and noise, the suggested method decomposed EEG signals using detection and removal of noise and extraction of relevant features. The classifier that uses Convolutional Neural Networks (CNNs) and Bidirectional Long Short-Term Memory (BiLSTM) networks, the study's Deep Learning (DL) algorithm used the EEG waves as input data. Results: An algorithm has been proposed that distinguishes between approximately 94% of individuals using a 17-channel EEG signal to compare healthy individuals with those who have ADHD. The proposed method, using the CNN-BiLSTM method to analyze EEG signals and process the data in a DL algorithm produced a classification accuracy of 98.69%. The combination of precise EEG with AI methods holds promise for improving our understanding of ADHD in children and developing more accurate diagnostic tools.

Contribution/Originality: This work aims to provide an improved methodology for integrating artificial intelligence methods to identify ADHD in children based on EEG signals and represents a major advance, contributing to improved diagnostic accuracy and decision-making.

1. INTRODUCTION

One of the most prevalent neuropsychiatric conditions affecting children nowadays is attention deficit hyperactivity disorder (ADHD) [1]. Although 5% of people in society suffer from ADHD on average, symptoms usually start to show before the age of twelve. Additionally, a child with this condition struggles to control their impulses and concentrate on one task at a time. For access to early intervention and treatment to enhance long-term outcomes, an accurate and timely diagnosis of ADHD is required [2].

Depending on this illness, children may grow up to exhibit dangerous behaviors like violence [3]. Clinicians diagnose ADHD based on symptoms including hyperactivity and inattention. The DSM-V's criteria for ADHD are evaluated; medical examinations, school records, family interviews, ADHD rating scales like Conners Rating, and neuropsychological testing are all part of the diagnostic process [4]. Electrodes are placed on a person's scalp to monitor and record neuron activation in the brain, generating electrical impulses and voltage variations.

This procedure is known as electroencephalography. An extended period of brain activity is recorded, and signals are measured using the non-invasive EEG technique. Therefore, an Electroencephalogram (EEG) study is utilized to address health issues concerning brain activity. Research fields include neurology, biomedical engineering, and brain computer interface studies that use EEG data, which contains information about a person's brain neuronal activity. Because of its low cost and non-invasiveness, the EEG is a critical source of data for many neurological studies [5].

One of the most significant subfields in Machine Learning (ML) is closely related to machine learning, Deep Learning (DL). Most people agree that machine-learning research started in the 1950s. Machine learning has emerged as a new kind of multidisciplinary field after years of expansion [6]. Field that encompasses biology, neurophysiology, statistics, probability theory, and approximation theory. Computer science is the study of how to replicate or recreate learning activities in the human brain using computers. In the realm of artificial intelligence, it is among the most advanced studies [7]. The deep learning concept in machine learning originates from an artificial neural network [8]. A multilayer artificial neural network serves as its foundation. In order to effectively and automatically integrate and extract the data's hidden characteristics, it integrates intrinsic legal knowledge from extensive data, it merges the data's low-level characteristics to produce a higher-level, more abstract feature that can symbolize the category of data attributes [9].

To find children with ADHD, a deep learning study based on EEG was conducted. According to the study's techniques, Convolutional Neural Networks (CNNs) and Bidirectional Long Short-Term Memory (BiLSTM) are two DL-based techniques that we used in this study. We then suggested a hybrid approach that combines these two techniques to distinguish ADHD. The findings of the categorization using various techniques and datasets in a different study involving 20 healthy and 20 ADHD persons showed that CNN performs noticeably better than Recurrent Neural Networks (RNNs) and achieves high accuracy [10]. According to a different study, utilizing CNN techniques to identify ADHD can achieve an accuracy of 90%. In a different study, 30 healthy children and 31 children with ADHD had their EEG data prepared and divided into four frequency ranges: beta, gamma, alpha, and theta. After that, the matrix was created [11].

A non-recursive technique for analyzing non-linear and non-stationary signals, the VMD was initially presented by Nouri and Tabanfar [9]. It has the ability to adaptively break down and convert a complex signal into a succession of quasi-orthogonal IMFs. It is possible to estimate the center frequency of each IMF online. VMD has a strong theoretical basis and good noise resilience when in contrast to the algorithms for Ensemble Empirical Mode Decomposition and Empirical Mode Decomposition [12]. It has been employed in mechanical diagnostics, underwater acoustic signal processing, and biological sciences. A class of denoising methods, which include component reconstruction, screening, and signal decomposition, is used to remove noise. The fundamental idea behind these methods is to identify and eliminate noise components based on the screening principle, as well as to extract signal components that were acquired using a signal decomposition approach. After that, denoising is achieved by reconstructing the signal's relevant components. Choosing the right screening criteria and breakdown algorithm is crucial. For instance, the wavelet analysis-based denoising technique has been applied extensively and produced positive outcomes in a variety of domains [13].

Using Continuous Wavelet Transform (CWT), Variational Mode Decomposition (VMD), and deep learning algorithms, we conducted a classification analysis in this work by identifying potential differences for the first time; the literature compares healthy people's EEG signals to individuals with ADHD [14]. CWT and VMD methods were used to separate EEG signals into subbands. The classification investigation was carried out using the subband data, which provides input information using the signal itself for the deep learning method. As a result, a novel model with high classification accuracy has been presented [15].

ADHD is a complex illness with a wide range of symptoms, as was previously established. Neural connections can be altered and symptomatology improved with prompt intervention and precise detection. However, because ADHD is complex, co-occurring disorders are common, and there is a global lack of qualified diagnosticians, it is sometimes difficult to diagnose ADHD. Thus, it is essential to investigate other strategies to improve early detection efficacy, including utilizing DL methods. These techniques could enhance existing diagnostic methods and help identify ADHD more quickly and effectively [16].

The following summarizes how our study went: Section 1 provides information on EEG and ADHD. The literature provides an overview of research on ADHD that uses deep neural networks and EEG. The titles of the second section cover EEG data, VMD, and CWT processes, as well as a deep learning approach. The route we took for our research is shown. The findings are examined, and the outcomes are presented in the third section. The four sections include the results and discussion. The conclusion is divided into five sections.

2. RELATED WORKS

Standardized surveys and self-assessments are frequently used in traditional diagnostic techniques, which might result in problems with discrepancies and inaccurate diagnoses [17]. Given these challenges, the integration of AI methods for identifying ADHD represents a major advance, helping to enhance diagnostic accuracy and decision-making [18]. Current diagnostic strategies struggle to adequately capture EEG signals of ADHD symptoms in children, highlighting the necessity of more precise and thorough approaches [19]. In the case of ADHD, researchers can use Convolutional Neural Networks (CNNs) to analyze brain imaging data [20]. CNNs in ADHD Diagnosis: CNNs, a form of deep learning. CNN finds and extracts local characteristics from EEG data using convolution techniques, then combines these features to create higher-level features [21]. This ability makes CNN particularly effective for EEG data classification by extracting pixel values and their feature vectors, which enhances the network's understanding and leads to accurate classification. The CNN is therefore being used more and more by researchers to study neurological, mental, and cerebral diseases [22]. The robustness of CNNs to variations, such as shifts and distortions in the input data, increases their suitability for neuroimaging studies. This robustness is critical when dealing with data with inherent variability [23]. CNNs have a record of accomplishment of superior performance in a wide range of EEE-based tasks, which translates well to neuroimaging tasks [24]. These advanced analytical tools enhance the accuracy of diagnosis, which can lead to a more effective prognosis. By leveraging CNN, Researchers and healthcare professionals may be able to overcome enduring challenges in the treatment of ADHD. This better performance demonstrates CNN's ability to diagnose ADHD accurately, indicating that DL approaches can be used to detect ADHD early and accurately, facilitating prompt intervention and treatment.

Research on ADHD patients' brain anatomy and function indicates that some brain regions are not functioning as well. The aforementioned functional impairment will also be exacerbated by other symptoms of ADHD, for instance, agitation, violent conduct, and ADHD comorbidities, such as conduct disorder, anxiety, depression, and oppositional defiant disorder. A study examined the effects of therapy on kids and teenagers with ADHD by Ma [25]. Four distinct machine learning classifiers were employed in the investigation to find pertinent traits from kids who were seeking ADHD therapy. The classifiers that were used were Deep Net, Endpoint Detection and Response (EDR), Logistic Regression (LR), and Classification and Regression Trees (CART). It is noteworthy that the Deep Net-based classifier outperformed the CART, EDR, and LR classifiers, having a strong AUC (area under the curve) of 0.72.

The absence of established standards for the use and interpretation of EEG data is the main reason why its use as a diagnostic tool for ADHD is still up for dispute. In order to solve this problem, Mafi and Radfar [26] used a

CNN model to identify ADHD in children using EEG data. The study included 61 subjects in total, 30 of whom were deemed to be typically developing and 31 of whom had an ADHD diagnosis.

A deep learning model was developed by Wang et al. [27] to precisely identify children with ADHD. In this study, raw EEG data were converted into visual representations and then fed into a CNN model. EEG recordings from a representative sample of children and teenagers were included in the dataset, which they utilized for their study. In particular, the dataset comprised 102 individuals, 51 of whom were children with ADHD and 51 of whom were typically developing. Interestingly, the deep learning model achieved a remarkable accuracy rate of 94.67%.

In a study by Dubreuil-Vall et al. [28], 40 volunteers 20 with an ADHD diagnosis and 20 healthy controls were asked to provide their EEG data. The goal of the study was to create a method for utilizing EEG data to identify ADHD. As inputs for a CNN model, the researchers used spectrogram pictures produced from the EEG data. The study's conclusions showed that the data classification accuracy was 88%.

For the purpose of predicting ADHD by Hegvik et al. [29] combined nonlinear characteristics, Discrete Wavelet Transform decomposition methods and Empirical Model Decomposition. 123 children and teenagers' EEG data was evaluated. ADHD symptoms were present in a subgroup that made up 45%. Utilizing the previously mentioned analytical techniques, the researchers sought to distinguish and categorize people according to their particular diagnosis.

This research presents an improved methodology for future exploration into integrating AI methods to identify ADHD, which represents a major advance in enhancing diagnostic accuracy and decision-making. A thorough search for articles was conducted on techniques of artificial intelligence methods for machine learning systems and model-based deep learning to diagnose and detect attention-deficit hyperactivity disorder based on EEG signals. Until February 2025, by computing the pooled sensitivity, specificity, positive and negative likelihood ratios, and area under the curve (AUC), the diagnostic value was assessed, as shown in Table 1.

Table 1. Some modern AI methods in the field of Identification of ADHD.

References	Methods	Accuracy	Recall	Precision
Khullar, et al. [30]	2DCNN-LSTM	97.00 %	92.00 %	91.00 %
TaghiBeyglou, et al. [31]	ERSP + CNN	91.46%	97.22 %	92.00 %
Loh, et al. [32]	DL (CNN+Grad CAM)	96.04%	96.26%	95.99%
Nouri and Tabanfar [9]	First-National-EEG CNN	94.52%	95.00 %	75.00 %
Ramalakshmi, et al. [33]	GSN - CNN	96.00 %	75.74 %	80.22 %
Dişli, et al. [34]	CWT-DCNN	95.99%	94.28%	95.30%
Abooelzahab, et al. [35]	DL (CNN–LSTM)	82.68%	78.5%	87.10%
Cao, et al. [36]	CNN- Bi-LSTM-AM	96.87 %	91.00 %	87.00 %

3. MATERIALS AND METHODS

The EEG data from the database in our investigation (https://ieee-dataport.org/open-access/eeg data-adhd-control-children) serves as a component of the CWT and VMD algorithms. The four subband indications were produced experimentally using the VMD algorithm. It has been observed that when fewer than four or more subband signals are used, a deep learning algorithm's classification success drops. Each of the 17 channels of the EEG signal was used in a classification analysis; the deep learning method was applied to each channel independently and in various combinations. After subband signals were acquired using CWT and VMD, success rates for classification at the deep learning algorithm's output were compared with those attained before the CWT and VMD processes. The effectiveness of CWT and VMD in detecting ADHD independently on the EEG data has been established. The best approach is examined, and how using CWT and VMD jointly affects the categorization success is discussed.

According to the literature review, EEG signals are utilized in deep learning and machine learning methods to diagnose ADHD. It is evident from the literature review that these methods have produced classification successes of up to an accuracy of 98.69%. High success rates have been achieved with the new procedure, which will add to the knowledge of ADHD diagnosis. Furthermore, our study especially looked at how the CWT and VMD techniques contribute to the diagnosis of ADHD. The results from very effective classification research with the fewest EEG channels are given using contributions from CWT and VMD. Next, we integrated the CNN and Bi-LSTM models to propose a new DL-based architecture for predicting ADHD in children. Through this merging, the model is able to utilize the various representations that each path has learned in a single design. Nonlinear combinations of the concatenated features are learned by a 1026-unit Dense layer. Overfitting is avoided by dropout regularization. The last Softmax output layer predicts class probabilities. Concatenating the CNN's complementary outputs. Figure 1 depicts the study's overall flow, and the integrating models' parameters are displayed in Table 2.

Table 2. CNN model parameters.

Model	Layer	Parameter
CNN	Conv	1026 filters, kernel size 3
	Batch norm	-
	Max pooling	Pool size 2
	Flatten	-
	Dense	128 units
	Batch norm	-
	Dropout	Rate 0.4
	Dense	66 units
	Conv	128 filters, kernel size 3
	Batch norm	-
	Max pooling	Pool size 2
	Bidirectional LSTM	66 units
	Dense	1026 units
	Dropout	Rate 0.4
	Conv	128 filters, kernel size 3
	Batch norm	-
	Max pooling	Pool size 2
CNN- BiLSTM	Bidirectional LSTM	66 units
	Dense	1026 units
	Dropout	Rate 0.4
Output layers	Dense	2 units, softmax
	Dense	1026 units

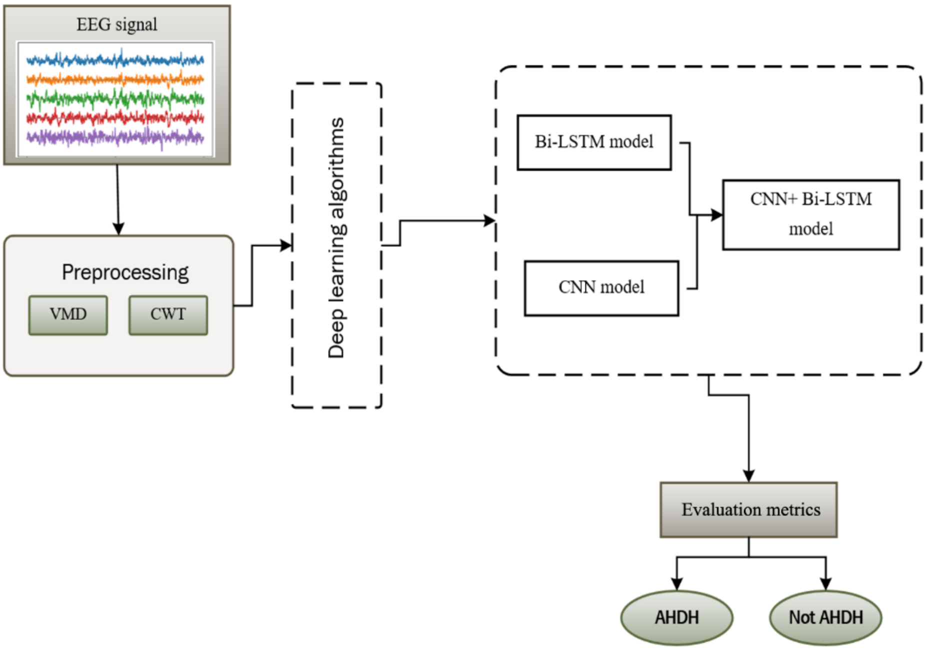

Figure 1. A Block diagram for the suggested model.

3.1. Data Collection

Every piece of experimental study data came from children aged 5 to 16 who had never been assessed or treated for ADHD before being given a diagnosis at a children's hospital. Every patient in the US uses the video EEG monitoring device CADWELL. Electrodes were affixed to their correct positions using conductive paste, following the global guidelines for placing electrodes in the 20/10 system, and the monitoring period lasted for at least twenty-four hours. The patient's EEG is compared, the different values of the other waves are recorded, and the long-range EEG is processed [16].

3.2. Preprocessing

This subsection describes how the Continuous Wavelet Transform (CWT) and the Variational Mode Decomposition (VMD) between each channel scalogram are used to convert EEG signals into scalograms.

3.2.1. Continuous Wavelet Transform (CWT).

A time-frequency diagram can be created by analyzing the time-varying data using the Continuous Wavelet Transform. When employing the pattern recognition approach, choosing the conversion method in the time-frequency domain is essential. For this kind of transformation, the wavelet transform is ideal. Since this conversion technique works incredibly well for non-stationary signals like electrocardiograms, electromyograms, and EEGs [37]. Signals are represented in terms of time and frequency [38]. We refer to these representations as scalograms. The absolute value of a signal's CWT coefficients is called a scalogram. First, computing the CWT filter bank and then performing these steps in MathWorks [39] employed in this investigation obtained the scalograms of the signals. When obtaining the CWT of several signals with identical characteristics, the ideal way is to precalculate Matlab's CWT filter bank. A filter bank's default wavelet is the analytical Morse wavelet. To customize Morse wavelets to your requirements, you can alter their time-bandwidth and symmetry properties. Equation 1 includes the generalized Morse wavelet function, which is provided by default in the MATLAB CWT filter bank and is utilized in the study [40]. The variable U (α) represents the Heaviside function. It is given in Equation 2 and is the normalizing constant. The function defined by the equation is clearly dependent on two parameters, "γ" and "β".

$$\beta,\gamma[\alpha\alpha] = U[\alpha]N\beta,\gamma[\alpha\alpha]\beta \;\; exp[-(\alpha\alpha)\gamma] \qquad (1)$$

$$N\beta,\gamma = 2(2\gamma/\beta)\beta/\gamma \qquad (2)$$

Impact wavelets or analytical Morlet wavelets are further options. Equation 3 provides the Morlet wavelet function. Here, 5 or 6 are taken, and c is the coefficient [41].

$$m(x) = \pi 1/4\ exp(icx)\, exp(-x2/2) \qquad (3)$$

For greater computational efficiency while evaluating several filters, once the filters have been precalculated, you can provide the CWT with the filter bank as input. The filter group allows for the visualization of wavelets in both time and frequency. Wavelets in the filter bank have a quality factor that you can set [41]. Figure 2 displays the scalogram EEG signal created by the CWT transformation together with an example signal from the data set.

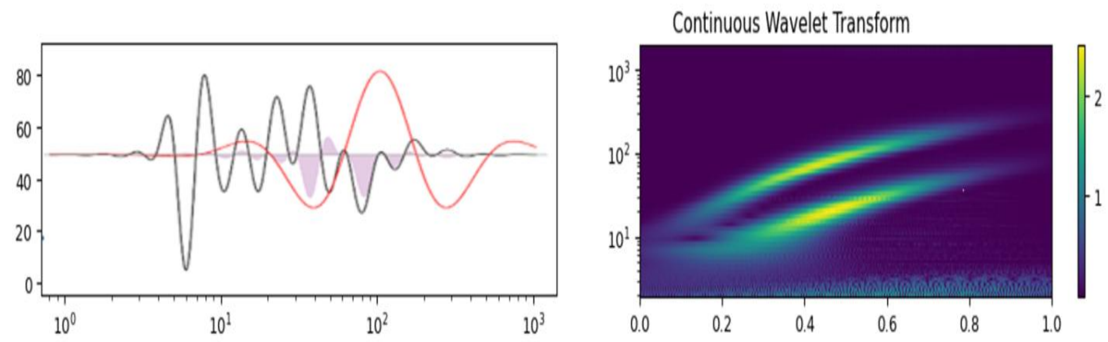

Figure 2. Signal and CWT in ADHD.

3.2.2. VMD Algorithm

Adaptive decomposition of a multi-component signal can yield many quasi-orthogonal Intrinsic Mode Functions (IMFs) at the same time using VMD, a novel signal processing approach. The IMF is characterized by VMD as an AM-FM signal, which is amplitude-modulated and frequency-modulated, and it has a solid mathematical foundation [42].

A non-recursive parsing technique called VMD was created to handle non-stationary and time-varying signals. Time series can be separated into replicable submodes with restricted bandwidth using VMD. A real-valued signal 'f' can be broken down by the VMD into a specific number of 'uk' sub-signals, each of which has specific bandwidth sparsity characteristics in the signal's spectrum. The center frequency (wk) is the compact center of each mode. The shifted signal VMD's 'H1' Gaussian smoothness is used to determine the bandwidth [43].

$$\{uk\}min\,, \{wk\}\, \{\textstyle\sum k = 1\, [\partial t\, [\delta\ (t)\ + j/\pi t]\, xvk\ (t)] \qquad (4)$$

Where the letters {uk}:={u1,…..,uK} and {wk}:={w1,…..,wK} stand for the collection of all modes and their associated center frequencies. Similarly, the sum of all modes is ∑k:=∑k = 1.

The set of all the modes and their center frequencies are denoted here by the abbreviations and, respectively. The sum of all modes is also. The VMD was applied to the EEG signal in this investigation in order to acquire subband signals. The deep learning algorithm's output shows the maximum learning success when three subband signals are chosen experimentally. As a result, VMD was used to split three subbands of the EEG signal [44].

3.3. Deep Learning

To automatically extract hierarchical features from data and learn representations, creating multi-layered artificial neural networks is the main goal of the machine learning subfield known as "deep learning." For complicated tasks, such as medical diagnosis, this makes it easier to create extremely powerful predictive models [45].

The CNN architecture in this study was trained for the diagnostic classification end-to-end on the labeled data challenge using convolution to create a representation of hierarchical features directly from the unprocessed EEG signals. The model consists of fully connected, max pooling, and dropout layers interspersed with convolutional layers that automatically learn spatial feature representations. Convolutions are fed the input EEG signals. A kernel size of three and two consecutive Conv layers with 66 filters were used to find temporal correlations. Nonlinear changes were activated using Rectified Linear Units (ReLUs). A 0.6 dropout rate was used between convolutional blocks to enhance generalization. Significant features were retained while Conv layers were reduced via max pooling layers. In order to facilitate high-level thinking, convolutional feature maps were connected to a series of dense layers after being flattened into a vector with progressively fewer units involved. For nonlinear combinations, ReLU activation was once more used [46]. ADHD and control groups were able to be binary classified by the final Softmax output layer's two nodes, as shown in Figure 3.

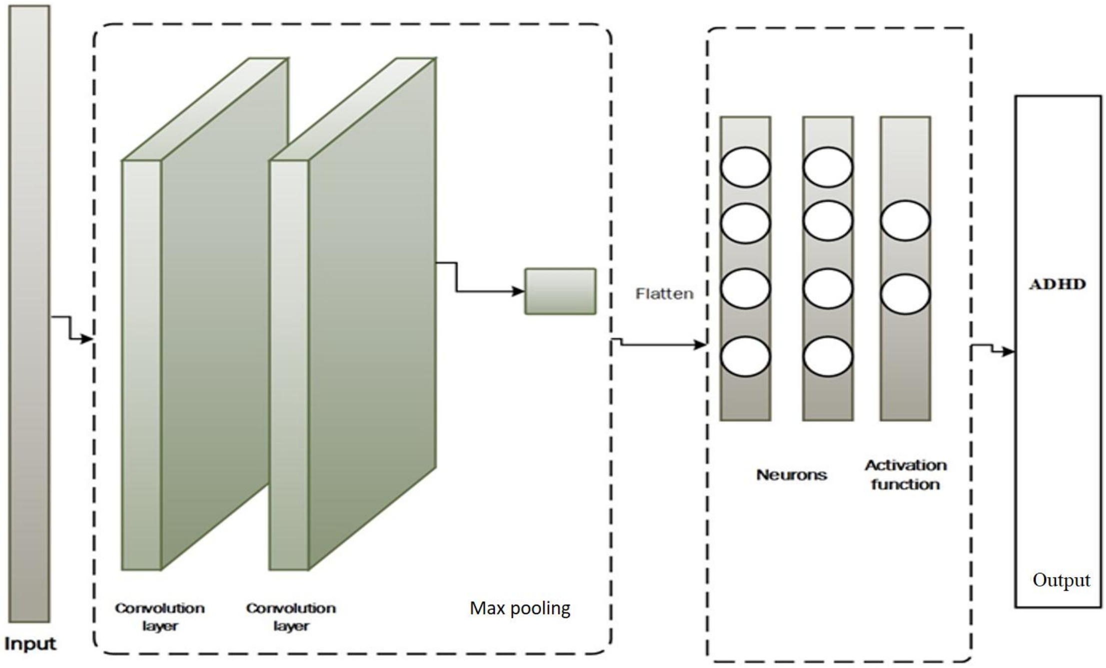

Figure 3. The design of the CNN model.

To obtain local features, the convolutional-BiLSTM approach also employs convolution. In the Conv layer, 189 filters were utilized. To learn temporal dependencies in both directions, the next steps are max pooling and a 66-unit bidirectional LSTM. The output of the BiLSTM is concatenated with the outputs of the other routes after being processed by a dropout and a dense layer of 1024 units.

$$C = \sum_{1}^{i} \sum_{j=1}^{x} I_{ij}\, F_{ij} \qquad (5)$$

Regarding CNNs, a convolution kernel or filter is represented by the letter F, and the rows and columns of EEG data I are represented by the letters I and J, respectively. First, EEG data is broken down into individual neurons, and then these neurons are flattened along the y and z directions. To detect and identify characteristics, a collection of x filters is installed in each layer of the network. Layer L creates feature maps of size X, which are suitably annotated [46].

$$c_i^L = B_i^L + \sum_{j=1}^{x(L-1)} F_{ij}^L * C_j^{L-1} \qquad (6)$$

The bias matrix is represented by the termB_i^L, In contrast, The filter that joins the jth feature map of the layer is denoted by the termF_{ij}^L.

$$ft = \sigma\,(WefXt + Wefht - 1 + WcfXt - 1 + Uf) \qquad (7)$$

$$it = \sigma\,(WxiXt + Whiht - 1 + WciXt - 1 + Ui) \qquad (8)$$

$$Ct = \sigma\,(ftCt - 1 + tanh(WxcXt + Whcht - 1 + \; U) \qquad (9)$$

$$ot = \sigma\,(WxoXt \; + \; Whoht - 1 + Wco\; Ct - 1 + \; Uo) \qquad (10)$$

$$ht = Ot\; X\; tanh(Ct) \qquad (11)$$

The previously described equations are commonly utilized in the sequential backward and forward processes. They represent the equations that make up the BiLSTM model. A gated cell that evaluates input data and decides whether to retain it based on its weight or significance can be thought of as the BiLSTM network. Three essential parts make up the BiLSTM model: the output, forget, and input gates. Determining which states should be stored in memory or deleted is the responsibility of the forget gate, represented by the symbol ft. The input gate considers the incoming signals when adjusting the value. The output gate, shown by the symbol, facilitates the transmission of the cell state to neighboring neurons. When CNN is used, the weight matrix W is employed by the hidden layer to process the input, yielding the final output ot. A memory cell, represented by the symbol Xt, is a crucial element of the BiLSTM model that is controlled by three different gates [46]. A memory cell, represented by the symbol ht, is a crucial part of the BiLSTM model and is controlled by three different gates, as shown in Figure 4.

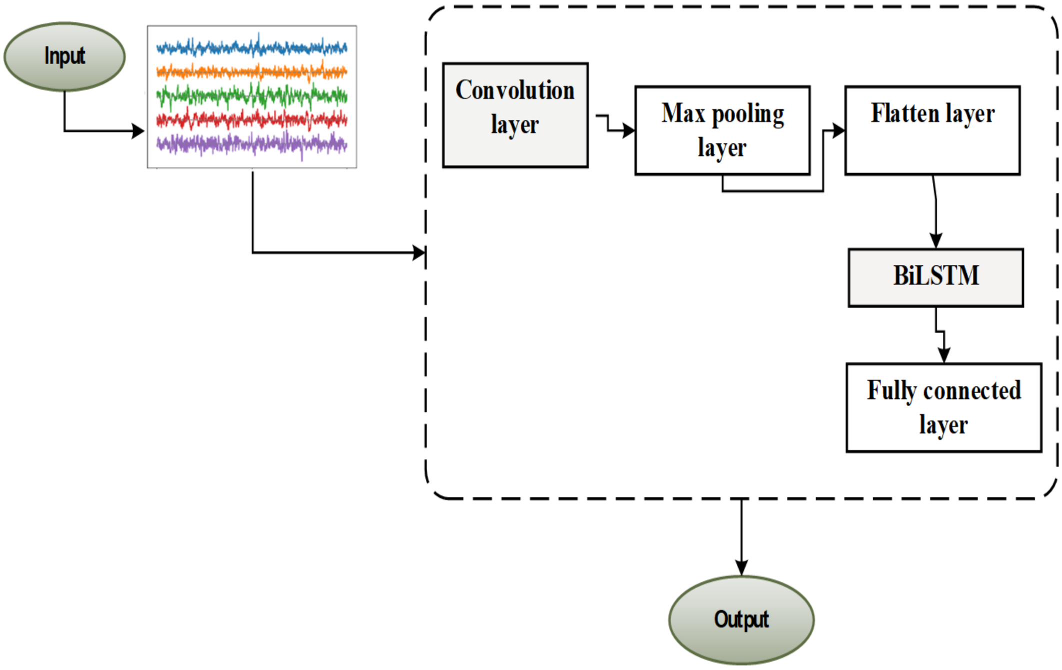

Figure 4. The architecture of CNN-BiLSTM.

4. PERFORMANCE MEASURES

Evaluating models' performance is essential to understanding their effectiveness. There are several indicators for evaluation, including accuracy. Among its fundamental components are False Negatives (FN) are ADHD cases that are mistakenly placed as controls (negative); True Negatives (TN) are control cases that the model correctly classifies as negative; and False Positives (FP) are control cases that are mistakenly projected to be ADHD (positive) [47, 48] as shown in Figure 5.

The equations of these measures are given in Equations 12, 13, and 14.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (12)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (13)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (14)$$

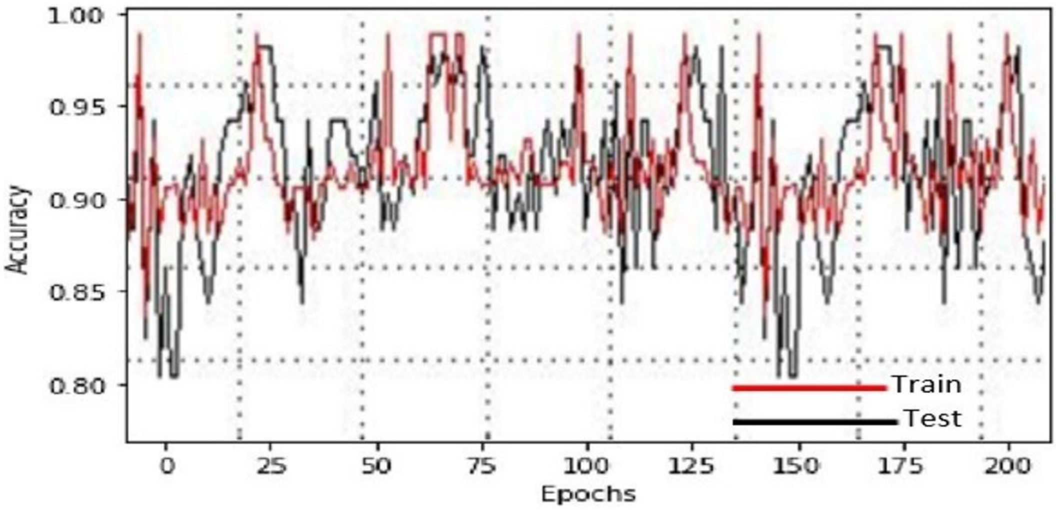

Figure 5. Transformer accuracy plot model.

5. RESULTS AND DISCUSSION

This study creates a comprehensive dataset by combining and randomly shuffling all brain wave difference samples and normal samples of ADHD children. Based on deep learning principles, the entire dataset is separated into training, test, and validation datasets. The data is divided into three sections, each representing 60%, 30%, and 10% of the total. Among these, the training dataset is used to optimize the neural network's parameters, such as its weights and biases, by fitting the data characteristics. Real-time monitoring of the effects of model training, real-time detection of model training problems, and parameter modification are conducted, and the validation dataset is used to confirm the model's effectiveness following several training cycles. The test dataset is used to measure the model's capacity for classification and generalization, as well as to assess the best model produced during training.

The EEG signal data of the either participants in this study, who were in the healthy or ADHD group, were used. There were 17 channels in the EEG recordings, and each channel was fed into a deep learning algorithm independently. The most efficient EEG channel was identified by examining each channel's categorization success independently.

Afterwards, the deep learning system was fed EEG readings from the frontal, parietal, occipital, temporal, and central lobes of the brain. The best brain region for identifying ADHD was therefore identified. Lastly, the classification success achieved when all 17 EEG channels were used simultaneously was computed when the channels were input into the deep learning system.

This study uses an 8-layer convolutional neural network to classify the retrieved features (3 convolutional layers, 2 max-pooling layers, 2 fully connected layers, and 1 output layer). The performance of the neural network was evaluated using K-fold cross-validation (K = 5).

This study's goal was to predict and report adult ADHD symptoms using deep learning approaches. CNN with BiLSTM was one of four deep learning algorithms used to differentiate between ADHD sufferers and control subjects. With scores ranging from 89% to 98%, the results showed noteworthy precision. Despite considerable heterogeneity in the various methods employed, the accuracy of predicting symptoms of ADHD in adults was quite high. CNN-BiLSTM, a widely used screening tool, makes it possible to identify risk variables linked to attention span shortening, a hallmark of adult ADHD. Applying deep learning algorithms is how this is accomplished. The CNN algorithm can be used to complete the task. The primary aim of this study is to identify ADHD by utilizing a dataset collected from people who were specially selected from the Peking University Institute of Mental Health's outpatient group. We believe that our research adds significantly to the growing body of knowledge about the accurate diagnosis of ADHD through the application of deep learning techniques.

A comparison between the built-in integrated deep learning model and other researchers' existing ADHD AI models is shown in Figure 6.

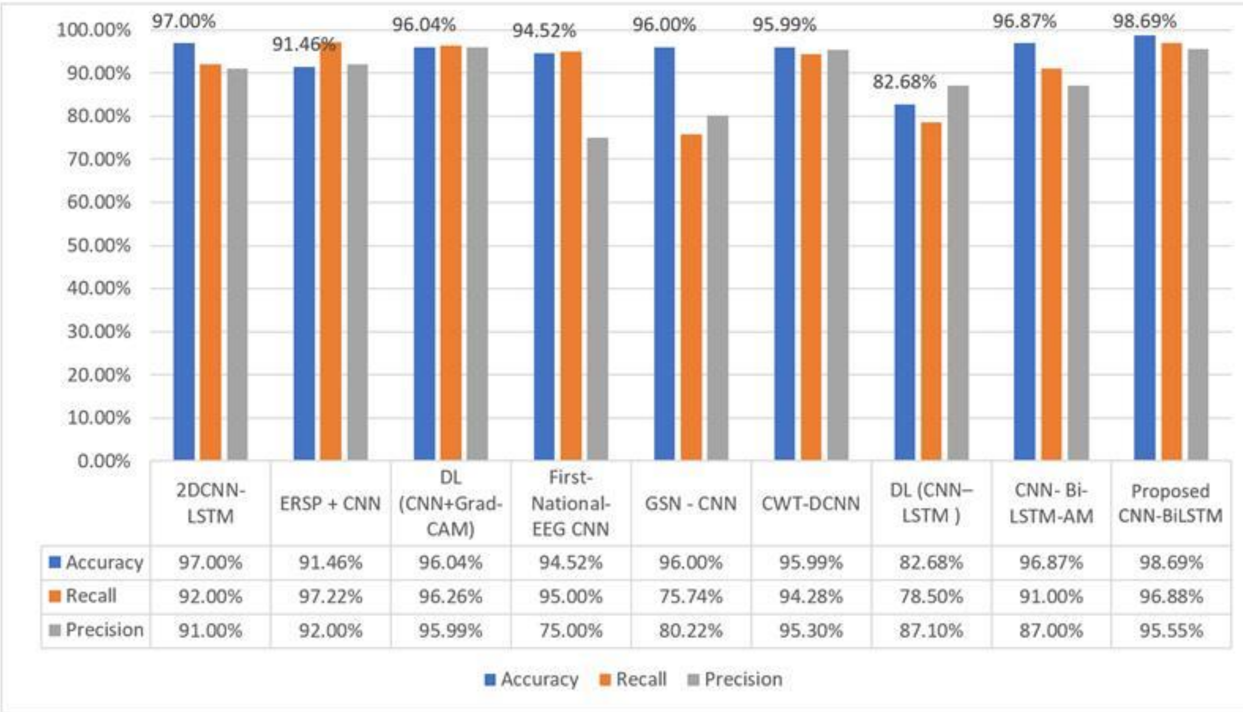

Figure 6. The created deep learning model is compared to existing models of ADHD.

6. CONCLUSION

Children with ADHD, their families, and society as a whole have suffered greatly; it is imperative that children with ADHD be diagnosed as quickly and simply as possible. The most accurate method for diagnosing and classifying ADHD is using brain electrical activity data. In this study, we provide a new computational framework for those AI methods to distinguish between ADHD profiles and typical child development as identified by careful EEG data analysis. Data preparation, intelligent feature extraction, strategic feature selection, and sophisticated classification techniques are just a few of the complex procedures of this extensive infrastructure. Preprocessing is done on the EEG signals to minimize noise and artifacts. The suggested system uses a unique preprocessing technique to convert 17 EEG channel segments into channel-wise CWT and VMD connection matrices. According to our data, the CNN with BiLSTM models outperformed other classification techniques and achieved high accuracy. This led to a high classification accuracy of 98.69% for ADHD, which is challenging for physicians and mental health experts to accomplish because of the identical clinical symptoms. Furthermore, we employ CWT and VMD, which can help us identify the crucial EEG channels for diagnosis. In addition, we will use AI techniques and work to create new algorithms to identify ADHD in patients with other comorbidities.

Funding: This study received no specific financial support.
Institutional Review Board Statement: Not applicable.
Transparency: The authors state that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] M. Y. Esas and F. Latifoğlu, "Detection of ADHD from EEG signals using new hybrid decomposition and deep learning techniques," *Journal of Neural Engineering*, vol. 20, no. 3, p. 036028, 2023.

[2] S. M. Naser, Y. H. Ali, and D. A.-J. OBE, "Deep learning model for cyber-attacks detection method in wireless sensor networks," *Periodicals of Engineering and Natural Sciences (PEN)*, vol. 10, no. 2, pp. 251-259, 2022.

[3] O. Attallah, "ADHD-AID: Aiding tool for detecting children's attention deficit hyperactivity disorder via EEG-based multi-resolution analysis and feature selection," *Biomimetics*, vol. 9, no. 3, p. 1-30, 2024.

[4] R. C. Joy *et al.*, "Detection and classification of ADHD from EEG signals using Tunable Q-Factor wavelet transform," *Journal of Sensors*, vol. 2022, no. 1, p. 1-17, 2022. https://doi.org/10.1155/2022/3590973

[5] S. T. Ahmed and S. M. Kadhem, "Early Alzheimer's disease detection using different techniques based on Microarray data: A review," *International Journal of Online & Biomedical Engineering*, vol. 18, no. 4, pp. 1–21, 2022. https://doi.org/10.3991/ijoe.v18i04.27133

[6] H. A. Mohammed, I. Nazeeh, W. C. Alisawi, Q. K. Kadhim, and S. T. Ahmed, "Anomaly detection in human disease: A hybrid approach using GWO-SVM for gene selection," *Revue d'Intelligence Artificielle*, vol. 37, no. 4, pp. 913–919, 2023.

[7] Z. K. Maseer, Q. K. Kadhim, B. Al-Bander, R. Yusof, and A. Saif, "Meta-analysis and systematic review for anomaly network intrusion detection systems: Detection methods, dataset, validation methodology, and challenges," *IET Networks*, vol. 13, no. 5-6, pp. 339-376, 2024. https://doi.org/10.1049/ntw2.12128

[8] S. T. Ahmed and S. M. Kadhem, "Using machine learning via deep learning algorithms to diagnose the lung disease based on chest imaging: A survey," *International Journal of Interactive Mobile Technologies*, vol. 15, no. 16, p. 95–112, 2021. https://doi.org/10.3991/ijim.v15i16.24191

[9] A. Nouri and Z. Tabanfar, "Detection of ADHD disorder in children using layer-wise relevance propagation and convolutional neural network: An EEG analysis," *Frontiers in Biomedical Technologies*, vol. 11, no. 1, pp. 14-21, 2024. https://doi.org/10.18502/fbt.v11i1.14507

[10] O. F. Alwan, Q. K. Kadhim, R. B. Issa, and S. T. Ahmed, "Early detection and segmentation of ovarian tumor using convolutional neural network with ultrasound imaging," *Revue d'Intelligence Artificielle*, vol. 37, no. 6, pp. 1503–1509, 2023. https://doi.org/10.18280/ria.370614

[11] J. Z.-L. Ji and J. T. Yanqing, "Machine learning-based automated detection of ADHD using heart rate variability data," *Epic Series in Computational Sciences*, vol. 101, pp. 49–57, 2024.

[12] B.-C. Yang *et al.*, "Complete ensemble empirical mode decomposition and wavelet algorithm denoising method for bridge monitoring signals," *Buildings*, vol. 14, no. 7, p. 2056, 2024. https://doi.org/10.3390/buildings14072056

[13] M. Pappula and S. M. Anwar, "An ADHD diagnostic interface based on EEG spectrograms and deep learning techniques," in *2024 20th International Symposium on Medical Information Processing and Analysis (SIPAIM)*, 2024: IEEE, pp. 1-4.

[14] D. Vimalajeewa, E. McDonald, S. A. Bruce, and B. Vidakovic, "Wavelet-based approach for diagnosing attention deficit hyperactivity disorder (ADHD)," *Scientific Reports*, vol. 12, no. 1, p. 21928, 2022. https://doi.org/10.1038/s41598-022-26077-2

[15] A. I. Altameemi, S. J. Mohammed, Z. Q. Mohammed, Q. K. Kadhim, and S. T. Ahmed, "Enhanced SVM and RNN classifier for cyberattacks detection in underwater wireless sensor networks," *International Journal of Safety & Security Engineering*, vol. 14, no. 5, pp. 1409–1417, 2024. https://doi.org/10.18280/ijsse.140508

[16] D. Yu and J. h. Fang, "Using artificial intelligence methods to study the effectiveness of exercise in patients with ADHD," *Frontiers in Neuroscience*, vol. 18, p. 1380886, 2024.

[17] B. S. Peterson *et al.*, "Tools for the diagnosis of ADHD in children and adolescents: A systematic review," *Pediatrics*, vol. 153, no. 4, pp. e2024065854, 2024. https://doi.org/10.1542/peds.2024-065854

[18] E. Sathiya, T. Rao, and T. S. Kumar, "Gabor filter-based statistical features for ADHD detection," *Frontiers in Human Neuroscience*, vol. 18, p. 1369862, 2024. https://doi.org/10.3389/fnhum.2024.1369862

[19] Y. Mao, X. Qi, L. He, S. Wang, Z. Wang, and F. Wang, "Advanced machine learning techniques reveal multidimensional EEG abnormalities in children with ADHD: A framework for automatic diagnosis," *Frontiers in Psychiatry*, vol. 16, p. 1-11, 2025. https://doi.org/10.3389/fpsyt.2025.1475936

[20] S. Luo, X. Meng, X. Niu, and H. Kong, "Revolutionising ADHD diagnosis: Deep learning in 3D medical imaging," in *Fourth International Conference on Signal Processing and Machine Learning (CONF-SPML 2024)*, 2024, vol. 13077: SPIE, pp. 110-118.

[21] H. S. M. Alsultani, S. T. Ahmed, B. J. Khadhim, and Q. K. Kadhim, "The use of spatial relationships and object identification in image understanding," *Technology*, vol. 9, no. 5, pp. 487-496, 2018.

[22] C. Surianarayanan, J. J. Lawrence, P. R. Chelliah, E. Prakash, and C. Hewage, "Convergence of artificial intelligence and neuroscience towards the diagnosis of neurological disorders—a scoping review," *Sensors*, vol. 23, no. 6, p. 3062, 2023. https://doi.org/10.3390/s23063062

[23] H. A. A. Mohammed, A. A. K. Jizany, I. M. Mahmood, and Q. K. Kadhim, "Predicting Alzheimer's disease using a modified Grey Wolf optimizer and support vector machine," *Ingenierie des Systemes d'Information*, vol. 29, no. 2, p. 669, 2024. https://doi.org/10.18280/isi.290228

[24] A. W. Salehi *et al.*, "A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope," *Sustainability*, vol. 15, no. 7, p. 5930, 2023. https://doi.org/10.3390/su15075930

[25] R. L. M. Ma, "Influence of relative age on diagnosis and treatment of attention-deficit/hyperactivity disorder in children," *Canadian Medical Association Journal*, vol. 184, no. 7, pp. 755–762, 2012.

[26] M. Mafi and S. Radfar, "High dimensional convolutional neural network for EEG connectivity-based diagnosis of ADHD," *Journal of Biomedical Physics & Engineering*, vol. 12, no. 6, p. 645, 2022.

[27] T. Wang *et al.*, "Prevalence of attention deficit/hyperactivity disorder among children and adolescents in China: A systematic review and meta-analysis," *BMC Psychiatry*, vol. 17, pp. 1-11, 2017. https://doi.org/10.1186/s12888-016-1187-9

[28] L. Dubreuil-Vall, G. Ruffini, and J. A. Camprodon, "Deep learning convolutional neural networks discriminate adult ADHD from healthy individuals on the basis of event-related spectral EEG," *Frontiers in Neuroscience*, vol. 14, p. 1-12, 2020. https://doi.org/10.3389/fnins.2020.00251

[29] T.-A. Hegvik, J. T. Instanes, J. Haavik, K. Klungsøyr, and A. Engeland, "Associations between attention-deficit/hyperactivity disorder and autoimmune diseases are modified by sex: A population-based cross-sectional study," *European Child & Adolescent Psychiatry*, vol. 27, pp. 663-675, 2018. https://doi.org/10.1007/s00787-017-1056-1

[30] V. Khullar, K. Salgotra, H. P. Singh, and D. P. Sharma, "Deep learning-based binary classification of ADHD using resting state MR images," *Augmented Human Research*, vol. 6, no. 1, p. 5, 2021. https://doi.org/10.1007/s41133-020-00042-y

[31] B. TaghiBeyglou, A. Shahbazi, F. Bagheri, S. Akbarian, and M. Jahed, "Detection of ADHD cases using CNN and classical classifiers of raw EEG," *Computer Methods and Programs in Biomedicine Update*, vol. 2, p. 100080, 2022. https://doi.org/10.1016/j.cmpbup.2022.100080

[32] H. W. Loh *et al.*, "Deep neural network technique for automated detection of ADHD and CD using ECG signal," *Computer Methods and Programs in Biomedicine*, vol. 241, p. 107775, 2023. https://doi.org/10.1016/j.cmpb.2023.107775

[33] K. Ramalakshmi, S. Rajagopal, M. B. Kulkarni, and H. Poddar, "A hyperdimensional framework: Unveiling the interplay of RBP and GSN within CNNs for ultra-precise brain tumor classification," *Biomedical Signal Processing and Control*, vol. 96, p. 106565, 2024. https://doi.org/10.1016/j.bspc.2024.106565

[34] F. Dişli, M. Gedikpınar, H. Fırat, A. Şengür, H. Güldemir, and D. Koundal, "Epilepsy diagnosis from EEG signals using continuous wavelet transform-based depthwise convolutional neural network model," *Diagnostics*, vol. 15, no. 1, p. 84, 2025. https://doi.org/10.3390/diagnostics15010084

[35] D. Abooelzahab, N. Zaher, A. H. Soliman, and C. Chibelushi, "A combined windowing and deep learning model for the classification of brain disorders based on electroencephalogram signals," *AI*, vol. 6, no. 3, p. 42, 2025. https://doi.org/10.3390/ai6030042

[36] X. Cao, S. Zheng, J. Zhang, W. Chen, and G. Du, "A hybrid CNN-Bi-LSTM model with feature fusion for accurate epilepsy seizure detection," *BMC Medical Informatics and Decision Making*, vol. 25, no. 1, p. 6, 2025. https://doi.org/10.1186/s12911-024-02845-0

[37] N. Alsharif, M. H. Al-Adhaileh, and M. Al-Yaari, "Accurate identification of attention-deficit/Hyperactivity disorder using machine learning approaches," *Journal of Disability Research*, vol. 3, no. 1, p. 20230053, 2024. https://doi.org/10.57197/JDR-2023-0053

[38] S. H. Dhahi, E. H. Dhahi, S. T. Ahmed, and Q. K. Kadhim, "Predicting Parkinson's disease using filter feature selection method," in *AIP Conference Proceedings*, 2024, vol. 3051, no. 1: AIP Publishing.

[39] MathWorks, "Wavelet toolbox™ for MATLAB® - Continuous Wavelet Transform (CWT)," 2019. Retrieved: https://www.mathworks.com/help/wavelet/continuous-wavelet-transform-cwt.html. 2019.

[40] W. Lee, D. Lee, S. Lee, K. Jun, and M. S. Kim, "Deep-learning-based ADHD classification using children's skeleton data acquired through the ADHD screening game," *Sensors*, vol. 23, no. 1, p. 246, 2022.

[41] H. W. Loh, C. P. Ooi, P. D. Barua, E. E. Palmer, F. Molinari, and U. R. Acharya, "Automated detection of ADHD: Current trends and future perspective," *Computers in Biology and Medicine*, vol. 146, p. 105525, 2022. https://doi.org/10.1016/j.compbiomed.2022.105525

[42] Y. Li, Y. Li, X. Chen, and J. Yu, "Denoising and feature extraction algorithms using NPE combined with VMD and their applications in ship-radiated noise," *Symmetry*, vol. 9, no. 11, p. 256, 2017. https://doi.org/10.3390/sym9110256

[43] A. T. Swadi and F. S. Miften, "ADHD detection using machine learning Algorithms and EEG brain signals," *Journal of Education for Pure Science*, vol. 13, no. 1, pp. 43-55, 2023.

[44] V. Welch *et al.*, "Use of mobile and wearable artificial intelligence in child and adolescent psychiatry: Scoping review," *Journal of medical Internet Research*, vol. 24, no. 3, p. e33560, 2022. https://doi.org/10.2196/33560

[45] I. Ashraf, S. Jung, S. Hur, and Y. Park, "A systematic literature review of neuroimaging coupled with machine learning approaches for diagnosis of attention deficit hyperactivity disorder," *Journal of Big Data*, vol. 11, no. 1, p. 140, 2024. https://doi.org/10.1186/s40537-024-00998-3

[46] G. Salazar de Pablo *et al.*, "Individualized prediction models in ADHD: A systematic review and meta-regression," *Molecular Psychiatry*, 29, pp. 3865–3873, 2024. https://doi.org/10.1038/s41380-024-02606-5

[47] N. Alsharif, M. H. Al-Adhaileh, S. N. Alsubari, and M. Al-Yaari, "ADHD diagnosis using text features and predictive machine learning and deep learning algorithms," *Journal of Disability Research*, vol. 3, no. 7, p. 20240082, 2024. https://doi.org/10.57197/JDR-2024-0082

[48] E. M. Hameed, H. Joshi, and Q. K. Kadhim, "Advancements in artificial intelligence techniques for diabetes prediction: A comprehensive literature review," *Journal of Robotics and Control*, vol. 6, no. 1, pp. 345-365, 2025. https://doi.org/10.18196/jrc.v6i1.22258