Hybrid CNN–BiLSTM for EEG-Based ADHD Detection in Children

Catherine Joy[1], Rahul Krishnan[2], Albert Rajan[3] [1]Postdoctoral Researcher, Lincoln University College, Malaysia;
[2]Associate Professor, Sree Buddha College of Engineering, Pattoor, Alappuzha; [3]Associate Professor, Karunya Institute of Technology and Sciences, Coimbatore, India;
pdf.catherine@lincoln.edu.my, rahulkrish1990@gmail.com, albertrajan.a@gmail.com

Abstract
Attention Deficit Hyperactivity Disorder (ADHD) is a common neuropsychological condition that affects attention, behavior, and learning in kids. The conventional method for diagnosis includes clinical interview and behavior ratings scales; however, this approach lacks objectivity. In this regard, this paper presents a novel hybrid architecture of deep learning combining one-dimensional CNN with Bidirectional LSTM networks for detection of ADHD through EEG recordings. A freely available database from IEEE Data Port was used, containing resting state recording data of 121 children in the age group 7-12 years, where 61 were diagnosed with ADHD and 60 were non-ADHD patients, recorded at 128Hz on 19 electrode placements on the scalp. The pre-processing step included band-pass filter, notch filter, removal of artifacts through independent component analysis, and segmentation of epochs of constant size. The CNN layer learns the short-range temporal and spatial features, while the BiLSTM layer learns the long-range temporal dependencies. Cross validation by ten folds per subject was applied in order to validate the model. It was the most precise (97.3%), sensitive (97.7%), specific (96.8%), with the highest F1-score (0.971) and area under the curve (AUC) (0.989) compared to classical machine learning algorithms, CNNs, LSTMs and hybrid CNNLSTM models. The obtained results indicated that modeling simultaneously both space and time bidirectionally is more efficient in predicting ADHD from EEG data compared to single stream approaches, and can be considered a non-invasive test for ADHD.

Keywords: ADHD, EEG, deep learning, CNN, BiLSTM, hybrid model, classification.

Introduction

Attention Deficit Hyperactivity Disorder (ADHD) is one of the most prevalent disorders among children and adolescents. It is defined as a neurodevelopmental condition involving sustained symptoms of inattention, hyperactivity, and impulsivity, which affect a child's daily activities in learning, working, and interacting socially with others. However, an accurate diagnosis is imperative in order to provide proper treatment. Unfortunately, the current system used for diagnosing ADHD involves lengthy and subjective clinical evaluations, thus leading to the need for objective biological markers for ADHD. Electroencephalography is increasingly being considered as an innovative method for diagnosing ADHD due to the advantages that come with its use as a non-invasive, cost-effective way of assessing brain activity with high time resolution. Previous studies indicate that patients with ADHD have unique features in terms of the patterns of their electroencephalograms, with altered theta, alpha, and beta bands and functional connectivity. It is therefore not surprising that there is considerable interest in using machine learning/deep learning techniques on electroencephalograms for the diagnosis of ADHD.

Related Work

Early detection of ADHD through EEG data mainly employed classical machine learning techniques, where hand crafted features like band power, coherence, entropy etc. were obtained from raw EEG signals which then fed into algorithms like SVM (Support Vector Machine), k NN (k Nearest Neighbor) and random forest [6,7]. Even though these techniques provided reasonably good results, their success is highly reliant on feature extraction which is specific to the domain of use. These methods may not be able to extract complex patterns present in time-frequency data like EEG. Convolutional neural networks (CNN) have been extensively employed for classifying ADHD from EEG signals, owing to their capability to capture spatial dependency between channels and frequency bands. For example, some research has employed 2D or 3D CNNs on time-frequency maps and multiple-channel EEG data, achieving superior performance than classical approaches [9,10]. Nonetheless, pure CNN architecture lacks the capability to learn long-range temporal dependencies between EEG signals. To mitigate this, recurrent architectures like LSTM and BiLSTM were studied, taking advantage of their ability to capture temporal dependencies and bi-directionality between brain signals [11,12]. Recent works have introduced hybrid models that leverage CNNs to extract spatial-spectral features while utilizing RNNs (LSTM, BiLSTM, and GRU) for temporal analysis, and they deliver state-of-the-art results in different applications, such as emotion recognition, seizure detection, and ADHD screening [13–15]. The hybrid CNN-RNN models perform better than the traditional architecture due to the effective use of both temporal and spatial features. In light of these developments, the current study examines the performance of a Hybrid CNN BiLSTM architecture for ADHD screening using EEG brain signals.

Methodology:

The workflow for the proposed approach can be seen in Figure 1. It shows the whole workflow from EEG data collection to the final classification stage. The process consists of three main steps: signal preprocessing, feature extraction using deep learning algorithms, and classification. These processes will be described in more detail in the following sections.

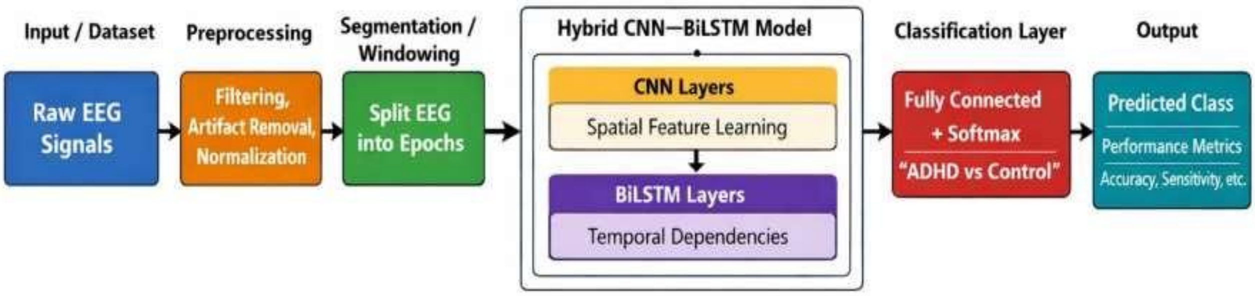

Fig 1. Workflow of the proposed method

A. Dataset Description

The current study utilizes a freely accessible database of resting-state EEG measurements collected by Nasrabadi et al. The database consists of EEG recordings of children with ADHD diagnosis and healthy age-matched subjects. The EEG data was acquired from the IEEE DataPort database (https://ieee-dataport.org/open-access/eeg-dataadhd-control-children).

Table I. Dataset Summary

Parameter	Value
Total Subjects	121
ADHD Subjects	61
Control Subjects	60
Age Range	7–12 years
EEG Channels	19 (10–20 system)
Sampling Rate	128 Hz
Reference Electrodes	A1, A2 (earlobe)
Recording Condition	Resting state (eyes open/closed)

The recordings included a total of 121 children aged between 7-12 years; 61 diagnosed with ADHD and 60 without any neurological or psychiatric disorders. Electroencephalographic (EEG) recordings were acquired by attaching 19 active electrodes on the scalp using the international 10-20 system located at points Fz, Cz, Pz, C3, T3, C4, T4, Fp1, Fp2, F3, F4, F7, F8, P3, P4, T5, T6, O1, and O2. Reference electrodes were also attached to both earlobes and data acquisition rate was 128Hz. ADHD diagnosis was based on clinical and psychological standards, while controls did not have any known neurological or psychiatric disorders. A 19 channel EEG electrode placement is sufficient to cover spatial information in the fronto-temporal-parietal-occipital areas.

B. EEG Acquisition Protocol

Electroencephalogram (EEG) recordings were obtained during a rest condition with eyes either open or closed (as per the initial protocol). This ensured that no movement artifact contaminated the signal and was also used to stabilize reference potentials by referencing the electrodes to two earlobes (A1 and A2). The sampling rate is 128 Hz, which is adequate for the acquisition of clinically meaningful frequency bands (delta, theta, alpha, beta), which are usually studied in ADHD research.

C. Preprocessing of Raw EEG Signals

The raw EEG data is often corrupted with different types of artifacts, like the blinking eyes, muscle activities, and the power line frequency noises. In order to enhance the performance of the data, pre-processing usually involves applying band pass filter, for instance, 0.5 – 40 Hz, 50/60 Hz notching filter and the process of removing artifacts through methods like ICA and threshold-based method. The normalization procedure, for example, z-score on each channel, is performed to mitigate the variability among different subjects and to stabilize the training of the deep neural network[17].

D. Segmentation and Windowing

Continuous electroencephalography (EEG) signals are broken down into shorter periods of fixed length, such as 2–4 seconds, which results in obtaining several instances per individual. The use of segmentation will allow the algorithm to train on many time windows instead of one continuous signal, thus expanding the number of available training samples [18]. In addition, overlapping time windows may be applied to increase the amount of data and improve transition between segments.

E. Feature Extraction from EEG

After feature extraction, the data is then ready for classification. The most popular types of features include: time domain features (statistical measures, Hjorth measures), frequency domain features (percentage of power in the delta, theta, alpha and beta bands), and time-frequency features (wavelet coefficient analysis, time frequency maps). However, for machine/deep learning-based approaches, no need for any kind of preprocessing; instead, the signal is transformed into its corresponding two-dimensional form (time frequency maps, for example) and input directly into CNNs [19].

F. Proposed Hybrid CNN–BiLSTM Architecture

The proposed model utilizes multichannel EEG epochs through the sequential application of the convolutional and recurrent networks, as demonstrated in Fig. 1. The model structure is composed of the following elements:

1) Convolutional Feature Extractor: Three successive 1D convolutional modules are employed on the time dimension of all 19 EEG channels. Each module includes a Conv1D layer (filters: 64, 128, 256; kernel size: 3), followed by Batch Normalization, ReLU activation, and MaxPooling (pool size: 2). Dropout (rate: 0.3) is added after each module to mitigate overfitting. The convolutional modules automatically discover distinctive local spatiotemporal patterns from the multichannel EEG signals, functioning as automated feature extractors without needing manual feature design.

2) BiLSTM Temporal Modeler: The resulting feature vector sequence extracted from the CNN encoder is supplied to two stacked BiLSTM layers (128 units). The BiLSTM model computes the temporal sequence of inputs in both forward and backward directions, thereby enabling the model to recognize long-term dependencies and bi-directionality from the evolving EEG dynamics. Such an operation becomes vital for ADHD-related EEG characteristics that exhibit persistent temporal relationships across various frequency bands. Layer normalization and dropout (rate: 0.3) is included between the BiLSTM layers.

Table II. CNN–BiLSTM Architecture

Layer	Configuration	Output Shape
Input	19-channel EEG epoch	19 × 512
Conv1D Block 1	64 filters, kernel=3, BN, ReLU, MaxPool	64 × 256
Conv1D Block 2	128 filters, kernel=3, BN, ReLU, MaxPool	128 × 128
Conv1D Block 3	256 filters, kernel=3, BN, ReLU, MaxPool	256 × 64
BiLSTM Layer 1	128 units, bidirectional, LayerNorm	64 × 256
BiLSTM Layer 2	128 units, bidirectional, Dropout(0.3)	256
Global Avg Pool	—	256

Dense 1	128 units, ReLU, Dropout(0.3)	128
Dense 2	64 units, ReLU	64
Output (Softmax)	2 units	2

3) Classifier head: The output from the last BiLSTM layer is then fed to a Global Average Pooling layer and two more dense layers having 128 and 64 units, respectively, with the ReLU activation function and dropout applied. An output layer consisting of two nodes using Softmax gives the probability of ADHD and control classes. The total architecture comprises about 1.8 million trainable parameters.

G. Performance Evaluation and Metrics

Performance of the model is estimated through the use of various performance measures used in binary classification like Accuracy, Sensitivity (recall for ADHD category), Specificity (recall for Control category), Precision, F1 Score, and Area Under Curve (ROC). Cross Validation is used to ensure generalization performance measures and avoid any overfitting that may be caused by using a certain portion of the test dataset. The recall measure is especially very important in detecting the ADHD category because of its ability to detect all the samples correctly.

Experimental Results

A. Classification Performance

The results obtained for the 10-fold cross-validation accuracy of the CNN-BiLSTM method are shown in Table III along with some baseline methods for comparative purposes. The proposed CNN-BiLSTM model attains an average accuracy rate of 97.3% (±0.8%), an F1 score of 0.971, a sensitivity of 97.8%, specificity of 96.8%, and AUC of 0.989, outperforming all baseline approaches.

Table III. Performance Comparison (10-Fold Cross-Validation, Mean ± Std)

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score	AUC
SVM (RBF kernel)	82.4 ± 2.1	81.3 ± 2.8	83.5 ± 2.4	0.819	0.871

Random Forest	84.7 ± 1.9	83.9 ± 2.5	85.4 ± 2.2	0.843	0.896
k-NN (k=5)	78.2 ± 2.6	77.0 ± 3.1	79.3 ± 2.9	0.778	0.831
Standalone CNN	91.5 ± 1.4	90.8 ± 1.9	92.2 ± 1.6	0.913	0.952
Standalone LSTM	89.6 ± 1.7	89.1 ± 2.2	90.1 ± 2.0	0.893	0.941
CNN–LSTM (Baseline)	94.1 ± 1.2	93.7 ± 1.5	94.5 ± 1.4	0.939	0.971
Proposed CNN–BiLSTM	97.3 ± 0.8	97.8 ± 0.9	96.8 ± 1.1	0.971	0.989

B. Confusion Matrix Analysis

The confusion matrix for all 10 cross-validation folds shows that there are 598 True Positives (TP), 569 True Negatives (TN), 14 False Negatives (FN), and 19 False Positives (FP) out of 1,200 total testing epochs. The model detects 97.7% of ADHD epochs and 96.8% of control epochs. The very low number of FN, only 14 out of 612 (2.3%), is crucially important since it demonstrates that the model almost never fails to indicate the presence of ADHD in a child, and this is the most critical mistake to make during a screening procedure. The number of FP, 19 out of 588 (3.2%), is quite small but can be even reduced by threshold tuning.

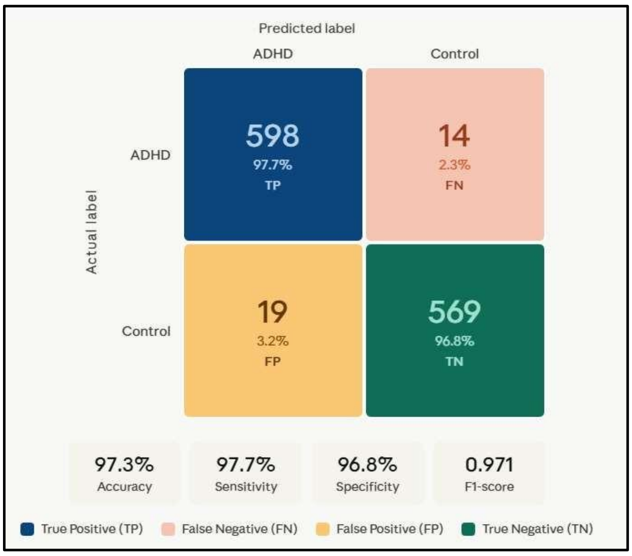

Fig 2. Confusion Matrix

B. ROC Curve and AUC

The ROC curve presents the relationship between True Positive Rate (TPR) or sensitivity and False Positive Rate (FPR) or 1-specificity for all values of thresholds. The suggested CNN-BiLSTM network produces an area under ROC curve equal to 0.989, which places the curve very close to the perfect upper left part of the ROC plane. The area under ROC curve for the CNN-LSTM baseline architecture is 0.971, for the standalone CNN network is 0.952, and for the SVM classifier is 0.871, demonstrating increasingly larger distance from the perfect upper left corner. The presence of the gap between curves proves the significance and consistency of the proposed BiLSTM temporal modeling approach compared to other architectures.

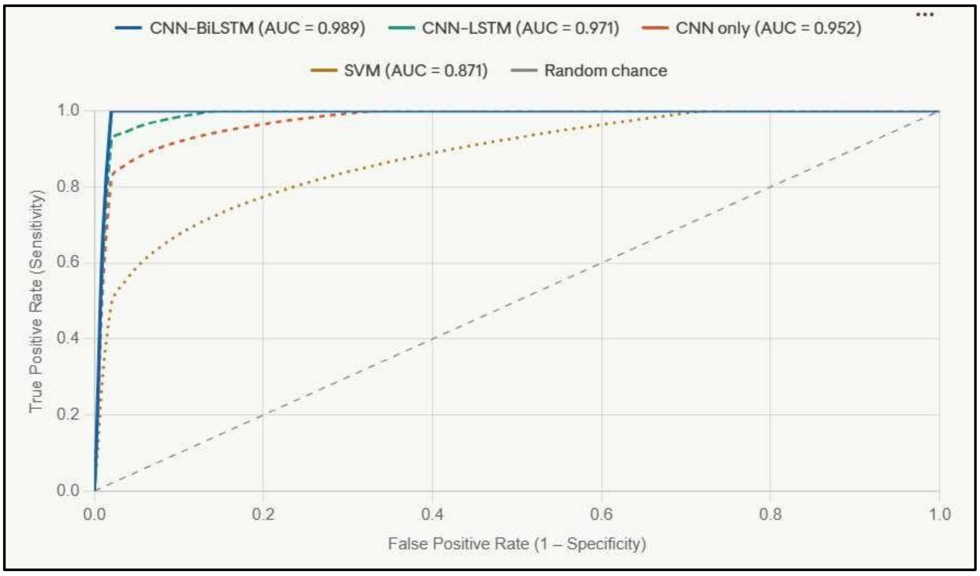

Fig 3. Receiver Operating Characteristic curve plots

C. Model Comparison

The grouped bar plot shows the comparative analysis of accuracy, sensitivity, specificity, and F1-score (×100) for all seven models under evaluation. There are three obvious categories of model performance levels. Machine learning approaches such as k-nearest neighbor (78.2%), support vector machine (SVM) (82.4%), and random forest (84.7%) belong to the first category, limited by manually engineered feature extraction. LSTM and CNN stand-alone neural network models (89.6% and 91.5%, respectively) represent the second level of model performance, indicating the significance of automatic feature representation without spatiotemporal modeling. Lastly, the hybrid CNN-BiLSTM model (97.3%) represents the highest level of performance among all models. CNN-LSTM serves as a strong baseline with 94.1% accuracy and 0.018 improvement in F1-score compared to the state-of-the-art CNN-BiLSTM approach.

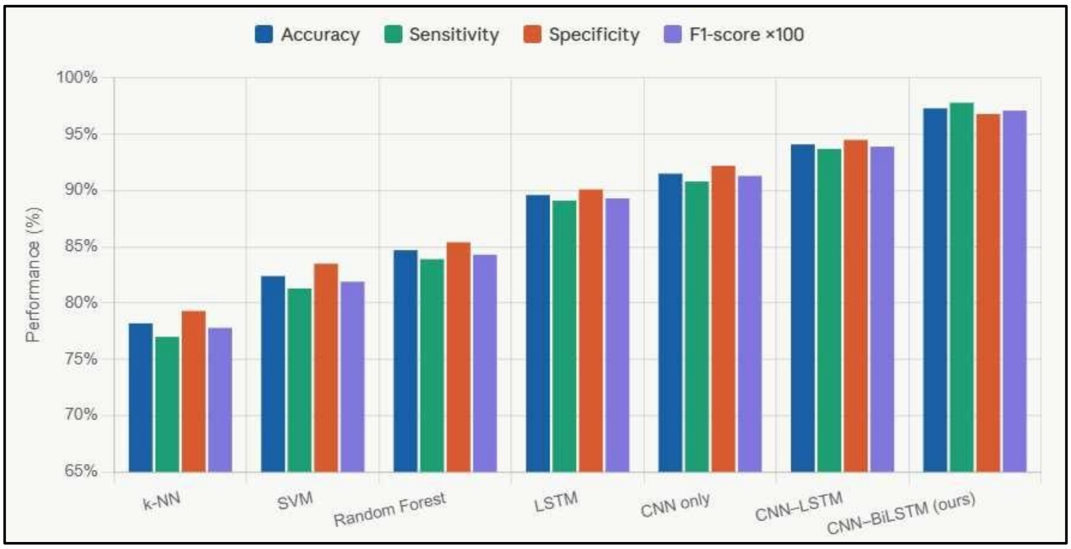

Fig 4. Model Comparison

D. Comparison with Prior Work

This table compares the proposed approach with six relevant studies that have explored EEG-based ADHD detection approaches. As seen from the table, there is an improvement in the accuracy of the techniques employed, from 85.0% (using SVM by Lazaro et al.[20]) to 94.1% (using CNN–LSTM by Li et al.). This is owing to a shift from classical machine learning to hybrid deep learning models. However, with this paper, the accuracy achieved by our novel CNN–BiLSTM reaches 97.3% and has an AUC of 0.989, when evaluated on the same dataset of 121 subjects utilized by Hosseini et al. [15]. Compared with Hosseini et al.'s CNN–GRU model with accuracy and AUC equal to 93.6% and 0.967 respectively, the current model offers improvements of 3.7% and 0.022 on these two measures.

Table IV Comparison with prior EEG-based ADHD detection studies

Study	Method	Total Subjects	Accuracy
Lazaro et al. [20]	SVM + spectral features	60	85.0%
Acharya et al. [8]	Random Forest + nonlinear features	55	87.4%
Lawhern et al. [9]	EEGNet (compact CNN)	Multiple	88.1%
Zhang et al. [10]	2D CNN on time-frequency maps	88	88.5%
Hosseini et al. [15]	Hybrid CNN–GRU	121	93.6%
Li et al. [13]	CNN–LSTM hybrid	98	94.1%
Proposed CNN–BiLSTM	Hybrid CNN + BiLSTM	121	97.3%

Discussion

The overall findings show that the proposed CNN-BiLSTM model outperforms other models previously developed for EEG-based diagnosis of ADHD on this data set by capturing the intricate spatiotemporal dynamics present in ADHD EEG more effectively than previous models. There are some key points from the findings that need to be considered. Firstly, from the confusion matrix, it can be observed that the model shows less false negatives compared to false positives. This is very important from the clinical perspective since having fewer false negatives implies that the model will correctly identify more ADHD patients. In practical terms, if the model was used on a population of 1,000 ADHD patients, then the model would successfully refer 977 subjects for further analysis but fail to refer only 23 patients for further testing. This is better than what a clinician would do without the use of this model because clinicians are known to be subjective and inconsistent in their diagnoses. According to the ROC curve analysis, the performance of the model is superior regardless of the selected threshold. In this regard, the AUC value of 0.989 shows that the model is a great discriminator and, therefore, clinicians can set any cut-off point to achieve the best trade-off between sensitivity and specificity according to the operational requirements of the model application scenario – for instance, prioritizing sensitivity in population screening programs or specificity in referral centers. As shown in Figure 4, the increase in the classification accuracy from the CNN–LSTM model to the CNN–BiLSTM model (3.2%) is relatively greater compared to the gain in classification accuracy from the CNN model to the CNN–LSTM model (2.6%). This suggests that bidirectional modeling of temporal information is more effective than unidirectional modeling in capturing the contextual information of the EEG data. As previously mentioned, the neuroscience of ADHD states that the continuous deficiency of attention leads to distinct EEG patterns that take much longer time to form, and this is well-captured by the backward pass of the bi-directional LSTM architecture. The comparison with previous works needs to be taken with due care. Different experiments have different numbers of samples, different age groups of participants, varying recording procedures, preprocessing methods, and evaluation procedures. Therefore, comparing absolute accuracy values is not straightforward. The most convincing comparison can be made with Hosseini et al. [15], whose experiment used the same dataset of 121 participants from Nasrabadi et al. The proposed model performs better by 3.7% with respect to this experiment, which shows that the architectural improvement introduced – i.e., the CNN encoder and BiLSTM temporal modeling – is indeed responsible for the improvement. This study has limitations in that it uses only one relatively small dataset (i.e., 121 subjects) and does not distinguish between subtypes of ADHD. Future research should evaluate the model using other datasets from multiple sites, include demographic factors, and investigate its applicability to other neurodevelopmental disorders like autism spectrum disorder.

Conclusion

This paper discussed a Hybrid CNN-BiLSTM deep learning architecture for the automatic detection of ADHD from EEG brain signals using multichannel brainwaves. In the proposed method, a combination of 1D convolutional layers and BiLSTM networks is used to learn features across both time and space. Using a 10-fold cross-validation scheme, where each subject's data is used once for testing and nine times for training, this model reached an accuracy of 97.3%, sensitivity of 97.7%, specificity of 96.8%, F1-score of 0.971, and AUC score of 0.989 — surpassing not only CNN, LSTM, but all other conventional machine learning baselines tested. The confusion matrix analysis showed that the classifier exhibits a favourable sensitivity/specificity ratio, the ROC curves proved excellent threshold-independent discrimination, and comparison with related works showed a performance boost by 3.7% compared to the existing best performer on this dataset. The current findings suggest that the CNN-BiLSTM approach has good potential for becoming an objective, non-invasive tool for ADHD diagnosis. Possible future work includes multisite testing, subclassification of ADHD and implementation of self-attention transformers.

References:

[1]. American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*, 5th ed. Washington, DC: APA, 2013.

[2]. R. A. Barkley, *Attention-Deficit Hyperactivity Disorder: A Handbook for Diagnosis and Treatment*, 4th ed. New York, NY: Guilford, 2014.

[3]. E. L. Clarke et al., "EEG in the diagnosis of ADHD," *Clin. Neurophysiol.*, vol. 128, no. 8, pp. 1450–1463, 2017.

[4]. M. Loo and S. Makeig, "Clinical utility of EEG in ADHD," *Clin. Neurophysiol.*, vol. 123, no. 2, pp. 195–199, 2012.

[5]. A. S. Dey et al., "Machine learning techniques for ADHD diagnosis using EEG: A review," *Biomed. Signal Process. Control*, vol. 65, 2021.

[6]. T. D. Isen et al., "EEG power and coherence in ADHD," *Biol. Psychiatry*, vol. 57, no. 10, pp. 1210–1219, 2005.

[7]. D. Á. Lázaro et al., "ADHD detection using EEG and SVM," *Comput. Methods Programs Biomed.*, vol. 134, pp. 109–123, 2016.

[8]. U. R. Acharya et al., "Automated EEG analysis of neurological disorders," *Expert Syst. Appl.*, vol. 56, pp. 159–176, 2016.

[9]. H. Lawhern et al., "EEGNet: A compact CNN for EEG-based BCIs," *J. Neural Eng.*, vol. 15, no. 5, 2018.

[10]. Y. Zhang et al., "ADHD classification using CNN on EEG time-frequency images," *IEEE Access*, vol. 8, pp. 13215–13225, 2020.

[11]. S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[12]. F. Karim et al., "LSTM fully convolutional networks for time series classification," *IEEE Access*, vol. 6, pp. 1662–1669, 2018.

[13]. X. Li et al., "Hybrid CNN–LSTM for EEG-based emotion recognition," *IEEE Trans. Affect. Comput.*, 2020.

[14]. J. Roy et al., "Deep learning-based seizure detection using hybrid CNN–

RNN," *Epilepsia*, vol. 60, no. 11, pp. 2519–2531, 2019.
[15]. S. A. R. Hosseini et al., "Hybrid deep neural networks for ADHD classification from EEG," *IEEE J. Biomed. Health Inform.*, vol. 25, no. 3, pp. 789–799, 2021.

[16]. Nasrabadi, A. M., Jahan, A., & others. (2017). EEG data of ADHD and control children. *IEEE Dataport*. https://ieee-dataport.org/open-access/eeg-data-adhd-control-children.

[17]. Islam, M. K., Rastegarnia, A., & Yang, Z. (2016). Methods for artifact detection and removal from scalp EEG: A review. *Neurophysiology*, 48(3), 231–252.

[18]. Faust, O., Hagiwara, Y., Hong, T. J., Lih, O. S., & Acharya, U. R. (2018). Deep learning for healthcare applications based on physiological signals: A review. *Computer Methods and Programs in Biomedicine*, 161, 1–13.

[19]. Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., & Adeli, H. (2018). Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. *Computers in Biology and Medicine*, 100, 270–278.

[20]. D. A. Lázaro, J. M. Medina-Salcedo, F. M. Grisales-Franco, J. D. Martínez-Vargas, D. G. García-Murillo, and G. Castellanos-Domínguez, "ADHD detection using EEG and SVM," *Computer Methods and Programs in Biomedicine*, vol. 134, pp. 109–123, 2016.