Machine Learning-Based Polysomnography Data Analysis for ADHD Diagnosis: A Focus on Sleep Stage-Based Biomarkers

1st Amirhossein Eskorouchi
Department of Industrial and Systems Engineering
Mississippi State University, Mississippi State, MS, USA

2nd Haifeng Wang*
Department of Industrial and Systems Engineering
Mississippi State University, Mississippi State, MS, USA

3ed Jonathan W. Lee
SOM-Advanced Biomedical Education
University of Mississippi Medical Center, MS, USA

4th Vignesh H. Nayak
SOM-Advanced Biomedical Education
University of Mississippi Medical Center, MS, USA

5th Norma B. Ojeda
SOM-Advanced Biomedical Education
University of Mississippi Medical Center, MS, USA

6th Lir-Wan Fan
SOM-Peds-Neonatology
University of Mississippi Medical Center, MS, USA

***Abstract*—Attention-Deficit/Hyperactivity Disorder (ADHD) significantly impacts children and adults worldwide, posing challenges in neurodevelopmental diagnostics. This study aimed to enhance diagnostic accuracy by identifying sleep stage-based biomarkers for ADHD through the analysis of polysomnography (PSG) recordings, including EEG, EOG, EMG, and ECG data across five sleep stages. Our analysis included 20,294 epochs for ADHD and 19,968 epochs for non-ADHD from 48 patients. We applied graph metrics, particularly the average shortest path length, to quantify the efficiency of information transfer within brain networks and other physiological systems. We employed nested cross-validation to ensure robust model training and evaluation, with an outer 5-fold cross-validation and an inner GridSearchCV for hyperparameter tuning and model selection. The Random Forest model achieved an average accuracy of 72%, precision of 71%, recall of 85%, and F1-score of 76%. Feature Importance Analysis and Correlation Analysis highlighted that Sleep Stage 1 and Sleep Stage 3-4 were the most influential features for distinguishing ADHD from non-ADHD individuals. Permutation analysis further validated these findings. These results underscore the potential of sleep stage correlations and graph-based metrics as biomarkers for ADHD, demonstrating the model's ability to accurately classify ADHD using comprehensive PSG data.**

***Index Terms*—ADHD Diagnosis, Graph Metrics, Polysomnography, Machine Learning, Sleep Stages**

I. Introduction

Attention-Deficit/Hyperactivity Disorder (ADHD) is a widespread neurodevelopmental condition affecting many children, characterized by lack of attention, impulsive actions, and heightened activity levels. Additionally, many individuals with ADHD experience sensory processing issues [1]. Early detection and management can significantly improve the quality of life. ADHD affects approximately 7.6% of children, 5.6% of adolescents, and 6.76% of adults worldwide [2]. Current diagnostic methods, primarily based on subjective reports, can lead to misdiagnosis due to clinician variability. This highlights the need for objective biological markers.

This paper presents a novel approach to ADHD diagnosis by analyzing polysomnography (PSG) data using graph theory and machine learning. PSG captures various physiological signals, including electroencephalogram (EEG), electrooculography (EOG), electromyography (EMG), and electrocardiogram (ECG) during sleep. Graph theory is well-suited for examining complex physiological networks [3] and has been effectively applied to model functional networks like EEG, EOG, EMG, and ECG. [4].

Our research leverages PSG data to quantify the efficiency of information transfer within brain networks and other physiological systems [5], [6]. This comprehensive analysis across different sleep stages offers new insights into ADHD's neurobiological underpinnings. In summary, the contributions of this research are highlighted as follows: We proposed a novel approach to ADHD diagnosis by leveraging PSG data, including EEG, ECG, EOG, and EMG signals. We applied graph metrics, particularly the average shortest path length, to quantify the efficiency of information transfer within brain networks and other physiological systems.

The remaining part of the paper is organized as follows. Related works are reviewed in Section II. Section III describes the methodology. Results are demonstrated in Section IV. The conclusion and future work are discussed in Section V.

II. Related Works

Graph theory has been used to understand brain connectivity in ADHD. Shakur et al. (2023) combined eye movement data and EEG signals to identify multimodal biomarkers for ADHD

*Corresponding author, email: wang@ise.msstate.edu

[7]. Sen et al. (2018) developed a model for diagnosing ADHD and autism using MRI and resting-state fMRI data, achieving high classification accuracies [8].

Significant progress has been made in leveraging graph theory and machine learning to analyze ADHD using various physiological data. Altınkaynak et al. (2020) found that multilayer perceptrons outperformed other models in analyzing auditory evoked potentials for ADHD diagnosis [9]. Jia et al. (2022) introduced a Channel-Relationships-Based Graph Convolutional Network for EEG-based emotion recognition, adaptable for ADHD diagnosis [10]. Lev et al. (2022) combined eye-tracking with continuous performance tests to assess ADHD, showing how psychophysiological measures with machine learning could enhance diagnostic precision [11]. These studies highlight the need for multimodal approaches that leverage machine learning to improve ADHD diagnostics. Table I summarizes related works on ADHD diagnosis using graph theory and machine learning.

III. Methodology

A. Data Acquisition

Our study analyzed PSG data from 48 children, collected in 2023 from the University of Mississippi Medical Center (UMMC) pediatric sleep center. The participants included 25 children with ADHD and 23 controls. Each sample consists of approximately 8 hours of PSG recordings from 17 channels: 2 ECG, 5 EMG (2 leg and 3 chin movements), 2 EOG, and 8 EEG. Each 30-second epoch, sampled at 512 Hz, results in 15,360 samples per epoch per channel. Data were recorded across five sleep stages: Wake, Stage 1, Stage 2, Stage 3-4, and REM. The dataset includes 20,294 epochs for ADHD patients and 19,968 epochs for controls, providing a rich set of data points for analysis (see table II). Figure 1 illustrates PSG recordings across different channels for two epochs of sleep stage 3-4, serving as a visual reference rather than a detailed comparison across stages.

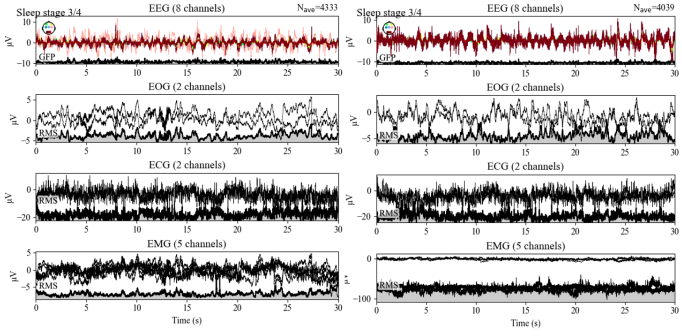

Fig. 1: PSG recordings across different channels for sleep stage 3-4 for a Non-ADHD (Left) and an ADHD (Right) patient.

B. Data Preprocessing

The EEG channels were processed using the reference electrode standardization technique (REST). A band-pass filter with a frequency range of 1 Hz to 100 Hz was applied for signal preprocessing, effectively attenuating unwanted frequencies outside this range. Additionally, a notch filter with a stop-band at 60 Hz was applied to reduce power line interference [12].

C. Feature Extraction

Feature extraction was essential for analyzing connectivity patterns in the PSG data, crucial for understanding neural interactions during sleep stages. For each 30-second epoch of each patient, we computed the correlation matrix across the 17 channels (C3, C4, Cz, F3, F4, O1, O2, Oz, EKG1, EKG2, E1, E2, CHIN1, CHIN2, CHINZ, RLEG, LLEG). The correlation matrix R was calculated using the pairwise Pearson correlation coefficient between signals from different sensors:

$$R_{ij}^{(e)} = \frac{\text{cov}(X_i^{(e)}, X_j^{(e)})}{\sigma_{X_i}^{(e)} \cdot \sigma_{X_j}^{(e)}} \tag{1}$$

where $\text{cov}(X_i^{(e)}, X_j^{(e)})$ is the covariance between sensors i and j during epoch e, and $\sigma_{X_i}^{(e)}$ and $\sigma_{X_j}^{(e)}$ are the standard deviations of sensors i and j during epoch e. These matrices quantify pairwise relationships between channels, with values ranging from -1 to +1.

Using these correlation matrices, we transformed the data into graph formats with NetworkX. Nodes represent the various physiological sensors, and edges are weighted by the absolute value of the correlation coefficients, indicating the strength of interaction between signals, forming the adjacency matrix A:

$$A_{ij}^{(e)} = |R_{ij}^{(e)}| \tag{2}$$

From these graphs, we extracted the average shortest path length, measuring the average minimum number of steps needed to connect any two nodes in the network, indicating network efficiency. For each epoch e, we constructed a graph $G^{(e)}$ using the adjacency matrix $A^{(e)}$ and calculated the shortest path lengths between all pairs of nodes. Longer path lengths may indicate impairments or less efficient interactions, which could be associated with ADHD. The average shortest path length $L^{(e)}$ for epoch e is computed as:

$$L^{(e)} = \frac{1}{N(N-1)} \sum_{i \neq j} d(i,j) \tag{3}$$

where N is the number of nodes (sensors), and $d(i,j)$ is the shortest path length between nodes i and j.

We created a comprehensive feature set by aggregating the extracted metrics across all epochs for each sleep stage and patient, resulting in a feature vector that captures temporal and spatial dynamics. The features included the average shortest path length across five sleep stages (Wake, Stage 1, Stage 2, Stage 3-4, and REM), forming a five-dimensional feature vector per patient. Each feature set was labeled with the subject's ADHD status, providing the target variable for the machine learning models. Aggregating features and calculating

TABLE I: Summary of related works on ADHD diagnosis using graph theory and machine learning.

Year	Explainability	Multi-channel	Validation Method	Data Type	Study Type	Ref
2018	✓	✓	Holdout Set	MRI, FMRI	Numerical	[8]
2020	×	×	Cross-Validation	Auditory Evoked	Numerical	[9]
2022	✓	✓	Cross-Validation	EEG	Numerical	[10]
2022	✓	✓	Cross-Validation	Eye-Tracking, CPT	Case Study	[11]
2023	✓	✓	Cross-Validation	EEG, EOG	Numerical	[7]
2024	✓	✓	Nested Cross-Validation	EEG, EOG, ECG, EMG	Case Study	This Paper

TABLE II: Number of patients in ADHD and Non-ADHD groups

ADHD				Non-ADHD			
Age $\leq$ 12		Age >12		Age $\leq$ 12		Age >12	
Female	Male	Female	Male	Female	Male	Female	Male
6	10	6	3	4	7	8	4

the average shortest path lengths for each sleep stage captured dynamic functional connectivity between physiological signals, offering a robust dataset for machine learning analysis and improving the classification of ADHD instances.

D. Machine Learning Models

We used a Random Forest classifier to predict ADHD status based on features comprising average shortest path lengths across five sleep stages. To ensure a rigorous evaluation and prevent overfitting, we applied nested cross-validation, a robust technique involving two layers of cross-validation: an outer loop and an inner loop. In the outer loop, the dataset was split into train-validation and test sets using a KFold object with 5 splits, ensuring that different portions of the data served as the test set in each iteration. The train-validation set from each outer fold was further divided into training and validation sets using another KFold object with 5 splits in the inner loop. This inner loop was responsible for hyperparameter tuning and model selection. We employed GridSearchCV to explore various hyperparameter combinations for the Random Forest model, using cross-validation on the training and validation sets to identify the best-performing model. After identifying the optimal model in the inner loop, it was evaluated on the independent test set from the outer loop, ensuring an unbiased performance estimate. Data augmentation, by inverting feature values, was used to balance the dataset and improve generalization.

IV. Experimental Results

A. Model Performance

The detailed results in Table III summarize the evaluation metrics (accuracy, precision, recall, F1 score), which are calculated based on the predictions and true labels for the test set. The Random Forest model achieved an average accuracy of 72%, precision of 71%, recall of 85%, and F1-score of 76%. These metrics highlight the model's strong ability to identify true positives, with a balanced trade-off between precision and recall, underscoring its reliability for ADHD prediction.

TABLE III: Evaluation Metrics of the Classifier for ADHD Prediction

Fold	Accuracy	Precision	Recall	F1-Score
1	0.70	0.75	0.75	0.75
2	0.68	0.62	0.89	0.73
3	0.63	0.67	0.83	0.74
4	0.68	0.54	1.00	0.70
5	0.89	1.00	0.80	0.89
Average	**0.72**	**0.71**	**0.85**	**0.76**

TABLE IV: Correlation Coefficient Matrices Between Different Sleep Stages for Non-ADHD (underlined) and ADHD Individuals

Stage	Wake	Sleep Stage 1	Sleep Stage 2	Sleep Stage 3-4	REM
Wake	1.00	0.61	0.63	0.48	0.59
Sleep Stage 1	0.59	1.00	0.21	0.09	0.58
Sleep Stage 2	0.71	0.82	1.00	0.75	0.35
Sleep Stage 3-4	0.47	0.63	0.81	1.00	0.10
REM	0.57	0.61	0.73	0.57	1.00

B. Feature Importance Analysis and Correlation Analysis in ADHD Diagnosis

We conducted a feature importance analysis using the Random Forest classifier to determine the significance of each feature. As illustrated in Figure 2, the importance scores reveal that Sleep Stage 1 and Sleep Stage 3-4 had the highest scores, indicating their significant contribution to the model's predictive performance. The Sleep Stage 2 and Rapid Eye Movement features showed lower importance scores, indicating a lesser impact on the model's predictions. Furthermore, a permutation feature importance analysis, which measures the decrease in model performance when each feature is randomly shuffled, was performed to confirm these findings. By shuffling these features—randomly mixing up the values for these sleep stages—their meaningful relationships with the ADHD label were disrupted, allowing us to assess their true importance to the model, as shown in Figure 3. This analysis supports the conclusion that Sleep Stage 1 and Sleep Stage 3-4 are the most influential features in predicting ADHD status.

Correlation analysis compares the sleep stage transitions between ADHD and non-ADHD individuals, as shown in table IV. A high correlation means that the transitions are smooth and consistent. In contrast, a lower correlation means the transitions are disrupted, indicating fragmented or irregular sleep cycles. For ADHD individuals, the transition between Sleep Stage 1 and Sleep Stage 3-4 is disrupted, as shown by

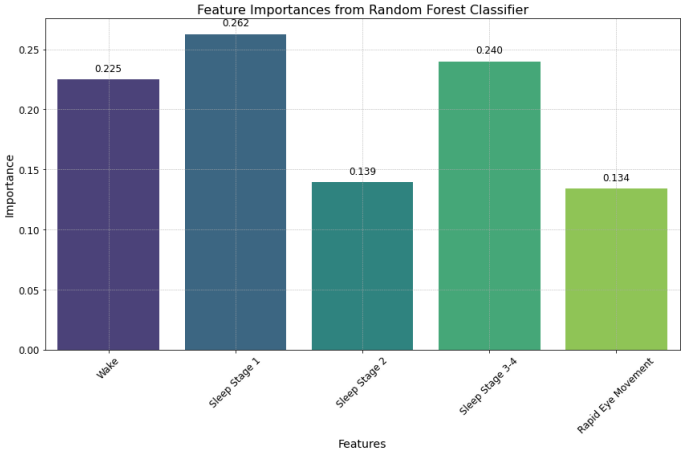

Fig. 2: Importance scores of the average shortest path lengths across different sleep stages for ADHD prediction.

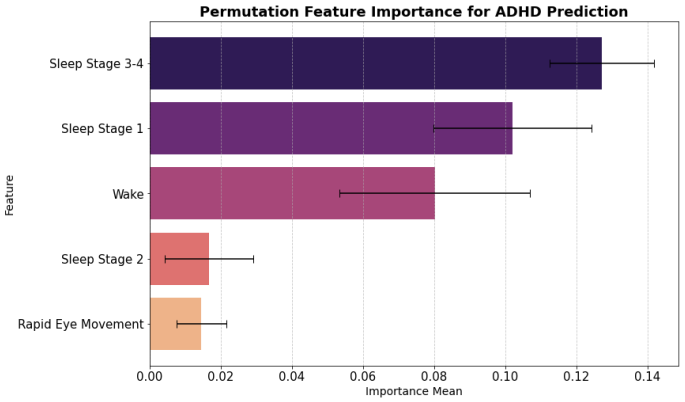

Fig. 3: Permutation Feature Importance Test.

a weak/none correlation (0.09). This negative value indicates that these two stages are not aligned, which may reflect difficulty in progressing from light to deep sleep. In contrast, non-ADHD individuals show a positive correlation (0.63) between these stages, suggesting smoother transitions into deep sleep, which is important for restorative sleep.

V. Conclusion

This study leveraged PSG data to enhance ADHD diagnostic accuracy using graph theoretical metrics and machine learning. Analyzing recordings from 48 patients across five sleep stages, we identified the average shortest path length as a robust feature for ADHD classification. The optimized Random Forest classifier achieved an average accuracy of 72%, precision of 71%, recall of 85%, and an F1-score of 76%. The model effectively distinguished between ADHD and non-ADHD individuals, with Sleep Stage 1 and Sleep Stage 3-4 identified as the most significant predictors. The permutation importance analysis further validated these findings, demonstrating the critical role of these features in the model's predictive performance.

Future work should expand the dataset and incorporate detailed signal correlation analysis to further understand differences in EEG, EOG, EMG, and ECG signals between ADHD and non-ADHD individuals. By comparing correlation matrices for each signal type, it may be possible to identify which signals exhibit the most significant connectivity differences. Integrating these features into machine learning models may uncover new biomarkers, enhancing diagnostic accuracy and clinical utility, and deepening our understanding of ADHD's neurophysiological underpinnings.

References

[1] R. C. Kessler, et al., "The World Health Organization adult ADHD self-report scale (ASRS): A short screening scale for use in the general population," Psychological Medicine, vol. 35, no. 2, pp. 245–256, 2005. https://doi.org/10.1017/S0033291704002892.

[2] D. Demontis, et al., "Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder," Nature Genetics, vol. 51, no. 1, pp. 63–75, 2019. https://doi.org/10.1038/s41588-018-0269-7.

[3] M. Rubinov, and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," NeuroImage, vol. 52, no. 3, pp. 1059–1069, 2010. https://doi.org/10.1016/j.neuroimage.2009.10.003.

[4] O. Sporns, "Structure and function of complex brain networks," Dialogues in Clinical Neuroscience, vol. 15, no. 3, pp. 247–262, 2013. https://doi.org/10.31887/dcns.2013.15.3/osporns.

[5] R. Ferri, et al., "Small-world network organization of functional connectivity of EEG slow-wave activity during sleep," Clinical Neurophysiology, vol. 118, no. 2, pp. 449–456, 2007. https://doi.org/10.1016/j.clinph.2006.10.021.

[6] S. Micheloyannis, et al., "Small-world networks and disturbed functional connectivity in schizophrenia," Schizophrenia Research, vol. 87, no. 1–3, pp. 60–66, 2006. https://doi.org/10.1016/j.schres.2006.06.028.

[7] A. H. Shakur, et al., "A rule-based exploratory analysis for discovery of multimodal biomarkers of ADHD using eye movement and EEG data," IISE Transactions on Healthcare Systems Engineering, vol. 13, no. 1, pp. 74-88, 2023. https://doi.org/10.1080/24725579.2022.2126036.

[8] B. Sen, et al., "A general prediction model for the detection of ADHD and Autism using structural and functional MRI," PloS one, vol. 13, no. 4, e0194856, 2018. https://doi.org/10.1371/journal.pone.0194856.

[9] M. Altınkaynak, N. Dolu, A. Güven, F. Pektaş, S. Özmen, E. Demirci, and M. İzzetoğlu, "Diagnosis of Attention Deficit Hyperactivity Disorder with combined time and frequency features," Biocybernetics and Biomedical Engineering, vol. 40, no. 3, pp. 927–937, 2020. https://doi.org/10.1016/j.bbe.2020.04.006.

[10] J. Jia, B. Zhang, H. Lv, Z. Xu, S. Hu, and H. Li, "CR-GCN: Channel-Relationships-Based Graph Convolutional Network for EEG Emotion Recognition," Brain Sciences, vol. 12, no. 8, 2022. https://doi.org/10.3390/brainsci12080987.

[11] A. Lev, Y. Braw, T. Elbaum, M. Wagner, and Y. Rassovsky, "Eye Tracking During a Continuous Performance Test: Utility for Assessing ADHD Patients," Journal of Attention Disorders, vol. 26, no. 2, pp. 245–255, 2022. https://doi.org/10.1177/1087054720972786.

[12] X. Lei, and K. Liao, "Understanding the influences of EEG reference: a large-scale brain network perspective," Frontiers in Neuroscience, vol. 11, p. 248735, 2017. https://doi.org/10.3389/fnins.2017.00205.