Investigation of EEGNet and its variants for ADHD Detection: Joint Spatial, Temporal, and Frequency Domain Learning

Kuricheti Prerana
Amrita School of Artificial Intelligence
Coimbatore
Amrita Vishwa Vidyapeetham
India

Niharika Sharma
Amrita School of Artificial Intelligence
Coimbatore
Amrita Vishwa Vidyapeetham
India

Diya Dinesh
Amrita School of Artificial Intelligence
Coimbatore
Amrita Vishwa Vidyapeetham
India

Pavani Satwika
Amrita School of Artificial Intelligence
Coimbatore
Amrita Vishwa Vidyapeetham
India

Amrutha Veluppal
Amrita School of Artificial Intelligence
Coimbatore
Amrita Vishwa Vidyapeetham
India

***Abstract*—Early and reliable detection of Attention Deficit Hyperactivity Disorder (ADHD) through EEG analysis holds much promise for developing objective and accessible neurodiagnostics. However, the conventional EEG- based ADHD detection approaches usually process the spatial, temporal, and frequency information separately ,which leads to fragmented representations and limit the model's ability to learn representations from EEG data. In the study, we propose an end to end ADHD detection framework that uses EEGNet and modified EEGNet which are both compact convolutional networks used to classify different EEG-based brain computer interface applications. The models learns frequency, temporal and spatial domain representations together which enables the model to extract EEG patterns which reflect ADHD related neural activity. To avoid data leakage and evaluate the model's performance, the dataset was evaluated using subject-wise splitting through Group K-Fold cross validation. The model achieved an accuracy of 97% when it was evaluated using sample-wise splitting. However, this setup introduces data leakage due to overlap of subject specific EEG characteristics across training and testing sets. These results highlight EEGNet as an efficient approach for ADHD detection from EEG signals.**

***Index Terms*—ADHD, EEG, EEGNet, Spatiotemporal Analysis,Data Leakage**

I. Introduction

Attention Deficit Hyperactivity Disorder (ADHD) is a common neurodevelopmental condition in children, characterized by persistent patterns of inattention, hyperactivity, and impulsivity. Conventional diagnostic procedures primarily depend on clinical observation and subjective behavioral assessments, which can lead to diagnostic uncertainty and delayed intervention. Consequently, objective neuroscience-based diagnostic approaches are gaining increasing attention. In this context, electroencephalography (EEG) has emerged as a promising noninvasive modality, as it captures neural oscillatory activity that reflects functional brain abnormalities associated with ADHD. EEG-based signal analysis has been widely explored across cognitive, affective, and neurological disorder domains, demonstrating its effectiveness in modeling neural dynamics through interpretable feature representations and signal complexity analysis [1], [2].

According to neurophysiological studies, ADHD is associated with functional abnormalities in the frontal, central, and parietal brain areas. While central regions (C3, C4, Cz) linked to motor inhibition are linked to hyperactivity and impulsivity, frontal areas (F3, F4, Fz) responsible for executive control and attention exhibit poor regulation. Changes in activation patterns are also seen in the posterior parietal and occipital areas (P3, P4, O1, O2), which control attentional modulation and sensory integration. A neurophysiological foundation for objective EEG characterization of ADHD is provided by these regional dysfunctions, which show up in the EEG as decreased Alpha and Beta power in posterior regions and increased slow-wave activity (Delta and Theta) over frontal–central areas.

The integration of deep learning techniques into EEG analysis has significantly advanced automated neurodisorder detection. Convolutional Neural Networks (CNNs), in particular, are capable of learning hierarchical spatial–temporal representations directly from minimally processed EEG signals, enabling effective classification without reliance on handcrafted features. One of the foundational works in this area was introduced by Schirrmeister *et al.*, who demonstrated that CNN-based architectures can successfully decode EEG signals while offering improved interpretability and visualization [3]. This work laid the groundwork for subsequent clinical EEG classification studies.

Building upon these advances, several studies have applied deep learning models to pediatric EEG data for ADHD de-

tection. Bansal *et al.* achieved high classification accuracy by modeling complex non-linear EEG patterns using autoencoder-based feature extraction combined with deep residual networks [4]. Similarly, Ahire *et al.* proposed a machine learning framework for automated ADHD prediction by focusing on segment-level EEG categorization [5]. To enhance real-world applicability, practical systems such as ADHD-AID have also been developed, incorporating feature selection and multi-resolution signal analysis [6].

Despite these methodological advancements, data partitioning strategies play a critical role in determining the clinical validity of EEG-based ADHD classifiers. Recent investigations have shown that sample-wise splitting can lead to data leakage and artificially inflated performance estimates, as EEG segments from the same subject may appear in both training and testing sets [7]. Although subject-wise splitting typically results in lower reported accuracies, it provides a more realistic evaluation of generalization to unseen individuals. A recent comprehensive analysis by Del Pup *et al.* further emphasized the impact of partitioning strategies on cross-subject EEG deep learning performance [8].

Channel selection which is disorder-sensitive ,centered on important frontal, temporal, and central regions has been driven by the fact that full 19-channel EEG setups, although informative, are not feasible for portable or quick testing. Therefore, this work compares the lightweight EEGNet with the deeper Shallow ConvNet under clinically realistic conditions, assessing CNN-based models for young children ADHD detection using different data-splitting and channel configurations.

II. Methodology

A. Dataset

This particular study uses the publicly available "EEG Data for ADHD / Control Children" dataset from IEEE DataPort [3]. It consists of EEG recordings from 121 children (60 ADHD, 61 controls), aged 7-12 years which are obtained under eyes-open and eyes-closed resting-state conditions using the 10-20 system. 19 channels were recorded for six-minute trials at 128 Hz using the NEURON-SPECTRUM system, with ADHD diagnoses evaluated by psychiatrists following DSM-IV criteria. This dataset serves as a standard reference for evaluating EEG-based ADHD detection methods.

A comprehensive preprocessing pipeline was used to remove noise and artifacts from EEG signals. The following steps were applied:

B. Preprocessing

1) Band Pass Filtering: A 4th-order Butterworth bandpass filter with a passband of 0.5–40 Hz was applied to remove baseline drift and high-frequency noise:

$$H(s) = \frac{1}{\sqrt{1 + (s/\omega_c)^{2n}}},$$

where $H(s)$ is the magnitude response of the Butterworth filter, $s = j\omega$ is the complex frequency variable, $j = \sqrt{-1}$ is the imaginary unit, $n = 4$ denotes the filter order, and ω_c represents the cutoff angular frequency corresponding to the specified lower and upper cutoff frequencies of the bandpass filter.

2) Artifact Removal via FastICA:Independent Component Analysis (ICA) was used to remove artifacts caused by eye movements and muscle activity. ICA is a well-known and effective technique for identifying unwanted sources in multichannel electroencephalographic (EEG)

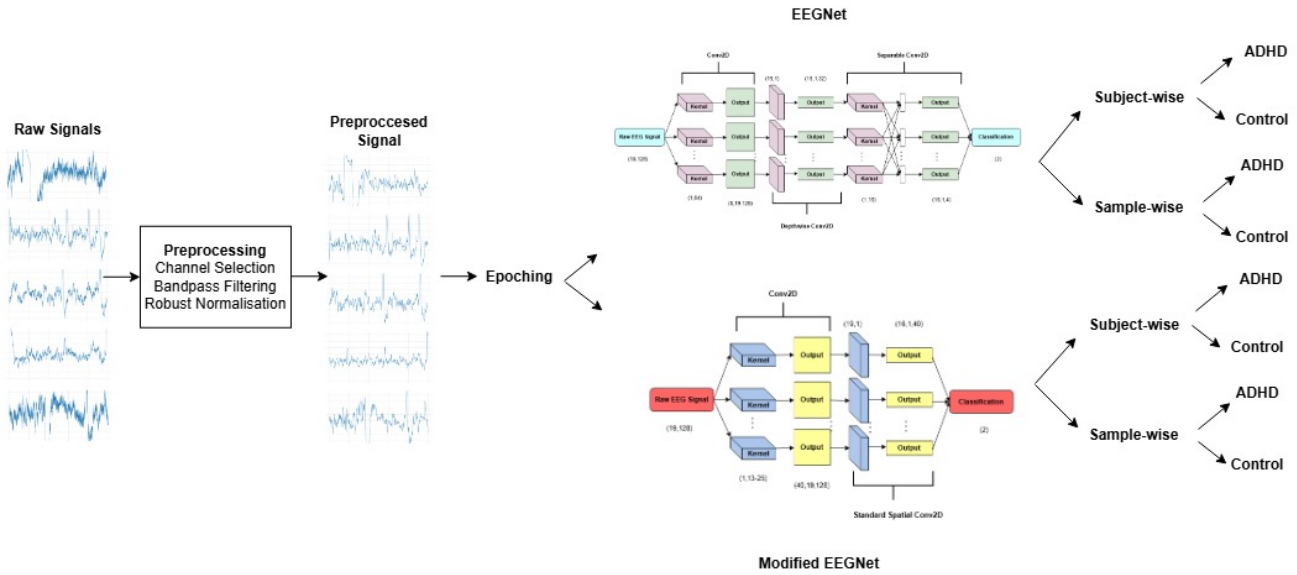

Fig. 1: Overall pipeline of ADHD Classification

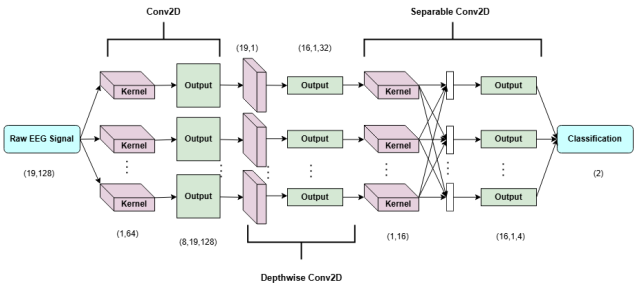

Fig. 2: Illustration of the EEGNet architecture.

data. In this case, we applied the FastICA algorithm, which uses the hyperbolic tangent (tanh) nonlinearity, to break down the EEG signals into statistically independent components.

C. Data Segmentation and Balancing

The EEG signals were divided into fixed-length epochs after preprocessing. We used a sliding window method with the following specifications:

- Epoch Duration: 10 seconds ($T = 10 \times 128 = 1280$ samples).
- Overlap: To maximize data availability and capture transitionary dynamics, there is a 50% overlap between successive epochs (stride of 5 seconds).

We standardized the data to strictly contain $N_{segs} = 40$ epochs per subject in order to guarantee a balanced representation of all subjects during training and avoid model bias toward subjects with longer recordings. A conditional sampling technique was used to accomplish this:

1) Subsampling: To bring the count of valid segments down to 40, we used random selection without replacement for subjects who produced more than 40 valid segments.
2) Oversampling: To achieve the desired count, we used random sampling with duplication for subjects with fewer than 40 segments.

D. Core EEGNet Architecture (19-Channel)

At the heart of our proposed framework is EEGNet, a streamlined convolutional neural network tailored for EEG-based classification. The model takes in an input tensor $X \in \mathbb{R}^{C \times T \times 1}$, where $C = 19$ denotes the number of EEG channels (following the standard 10-20 system) and $T = 1280$ corresponds to the temporal dimension (10 seconds at 128 Hz). It is illustrated in Fig 2.

The core feature extraction module consists of two major convolutional blocks:

1) Block 1: Temporal and Spatial Filtering:

- Temporal Convolution: A 2D convolution with 16 filters and kernel size $(1, 64)$ is applied to extract temporal–frequency features.
- Depthwise Convolution: A depthwise convolution with kernel size $(C, 1)$ aggregates information across EEG channels using a depth multiplier of 2 ($D = 2$).

Batch Normalization, ELU activation, Average Pooling $(1, 4)$, and SpatialDropout2D ($p = 0.3$) are applied to stabilize learning and reduce temporal resolution.

2) Block 2: Separable Convolution: A separable convolution with 32 filters and a kernel size of $(1, 16)$ helps to separate the learning of temporal and spatial features. This is followed by Average Pooling $(1, 8)$ to further reduce the temporal dimension. To help prevent overfitting, we apply SpatialDropout2D ($p = 0.3$) after the pooling layer.

E. Output Layer

The final set of feature maps is flattened and passed through a fully connected classification head with the following structure:

- Dense layer with 64 units and ELU activation.
- Dropout layer with $p = 0.5$.
- Final dense layer with $N_{\text{classes}} = 2$ and Softmax activation produces class probabilities (ADHD vs. Control).

F. Modified EEGNet Architecture (ShallowConvNet)

For our comparative analysis, we chose to work with the Shallow ConvNet (ShallowNet), which is a modified EEGNet inspired by the Filter Bank Common Spatial Pattern (FBCSP) pipeline [3]. This architecture is designed specifically to extract logarithmic band-power features.

The model takes an input tensor shaped like $(C, T, 1)$ and processes it through several stages:

1) Temporal Convolution: We start with a standard 2D convolution using 40 filters and a kernel size of $(1, 13)$.We limit the weights by applying a max-norm constraint ($||w||_2 \leq 2$).
2) Spatial Convolution: Following is a 2D convolution with kernel size of $(C, 1)$ and 40 filters, also applying max-norm regularization.
3) Normalization & Power pooling: Batch Normalization is followed by a non-linearity, specifically the squaring function (x^2). The signal is downsampled through Average Pooling with pool size $(1, 35)$ and stride $(1, 7)$ after that.
4) Log-Transformation: A logarithmic activation function is employed to obtain the approximation of signal log-variance after which Dropout ($p = 0.5$) and the classification head are applied.

G. Wilcoxon Rank-Sum Based Channel Selection

Wilcoxon rank-sum test was used to identify most prominent EEG channels for ADHD detection. The test was performed for each EEG channel separately to compare the feature distributions in the ADHD and control groups. The closest discriminative channels were those revealing statistically significant differences. This channel selection method enhances the computational efficiency by discarding irrelevant or noisy channels and thus effectively reducing input dimensionality, while the classification performance is barely affected.

- EEGNet (9-ch): The number of temporal convolution filters has been increased to 32 with a kernel size of $(1, 32)$ in order to improve temporal feature extraction. The separable convolution layer was increased to 64 filters and the depth multiplier was lowered to $D = 1$ in order to keep the representational capacity with a reduced channel input.
- ShallowConvNet (9-ch):To capture broader temporal changes appropriate for the smaller spatial configuration, the average pooling window was changed to $(1, 75)$ with a stride of $(1, 15)$, and the temporal convolution kernel size was also changed to $(1, 25)$.

H. Training Procedure

1) Cross-Validation: In order to ensure that our evaluation is subject-independent, we used a Group K-Fold strategy with $k = 5$ for EEGNet validation. This implies that all segments from a certain subject were allocated only to the training or validation folds, thus completely avoiding any data leakage.

2) Loss Function: Categorical Focal Loss: To address class imbalance, Categorical Focal Loss(FL) was employed:

$$FL(p_t) = -\alpha(1 - p_t)^{\gamma} \log(p_t),$$

with p_t denotes the predicted probability assigned by the model to the true class label, $\gamma = 2.0$ and $\alpha = 0.25$.

3) Optimization and Class Balancing: Training was done with the Adam optimizer. For the 19-channel models, an initial learning rate of 1×10^{-3} was used, while the 9-channel variants utilized a finer learning rate of 3×10^{-4}. A ReduceLROnPlateau scheduler lowered the learning rate by a factor of two after epochs with no improvement (patience=10). Class weights calculated using the "balanced" heuristic were used to account for the minority-class under-representation.

III. Results and Discussion

As shown in Fig. 3, the Wilcoxon rank-Sum analysis reveals that there are prominent spatial differences in EEG power distributions between ADHD and Control groups among multiple regions of the scalp.

Wilcoxon rank-sum analysis was performed to evaluate EEG-derived features which exhibit significant differences between ADHD and Control groups. As illustrated in Fig. 3, marked spatial variations in power distribution are observed across the scalp , which indicates that these are region-specific changes in neural activity linked with ADHD.

Considerable proportion of the extracted features demonstrated highly important group-level differences, with several features yielding $p - values < 0.05$. Such extremely low p-values suggests the strong discriminative capability between ADHD and Control subjects which reflects the consistent neurophysiological deviations across the dataset.The idea that ADHD is defined by unique oscillatory patterns rather than localized variations is supported by the fact that these variances are noticeable in clinically significant scenarios.

Another group of features showed moderate significance, suggesting that they still contribute to class differentiation but with lower reliability. A small subset of features did not show any meaningful statistical differences ($p > 0.05$), indicating limited diagnostic relevance.

Figure 3 represents the spatial distribution of EEG band-power and the differences between Control and ADHD groups using topomaps. The average band-power for both groups in the Delta, Theta, Alpha, Beta, and Gamma frequency ranges is shown in the first two rows. There are significant visual differences , and ADHD individuals have more slow-wave activity (Delta and Theta), most of them are over frontal and central regions, whereas people from control group have higher Alpha, Beta, and Gamma activity, especially in the posterior and parietal areas.

The ADHD–Control difference maps, which are showing the localized clusters of increased slow-frequency power in ADHD and decreased high-frequency power in posterior regions, were shown in the third row. These spatial patterns are consistent enough with identified neurophysiological indicators of ADHD, such as decreased fast-wave synchronization and cortical under-arousal. In line with the Wilcoxon results, the boxplots in the last row also summarize the worldwide band-power distributions and display significant differences over all five bands.

A. Band-Wise EEG Power Differences

1) Delta Band (0.5–4 Hz): The ADHD group exhibited increased Delta band power predominantly over frontal and central regions, whereas the Control group showed weaker and more diffuse Delta activity. Wilcoxon rank-sum analysis revealed a statistically significant group difference ($p = 0.0190$), indicating enhanced slow-wave synchronization associated with ADHD.

2) Theta Band (4–8 Hz): Theta activity was consistently higher in ADHD subjects, especially in midline frontal and temporal areas and in these areas theta activity was consistently higher in ADHD participants. The control group showed reduced Theta power and greater consistency. The common Theta-related attentional problems observed in ADHD has been confirmed by the significant group contrast (p=0.0414).

3) Alpha Band (8–12 Hz): Alpha power over posterior regions was significantly decreased in people who have ADHD, whereas Alpha rhythms were higher in people from control group. The reduction suggests that cortical inhibition and resting-state stability are affected in ADHD. There was still a significant difference between the groups (p = 0.0203).

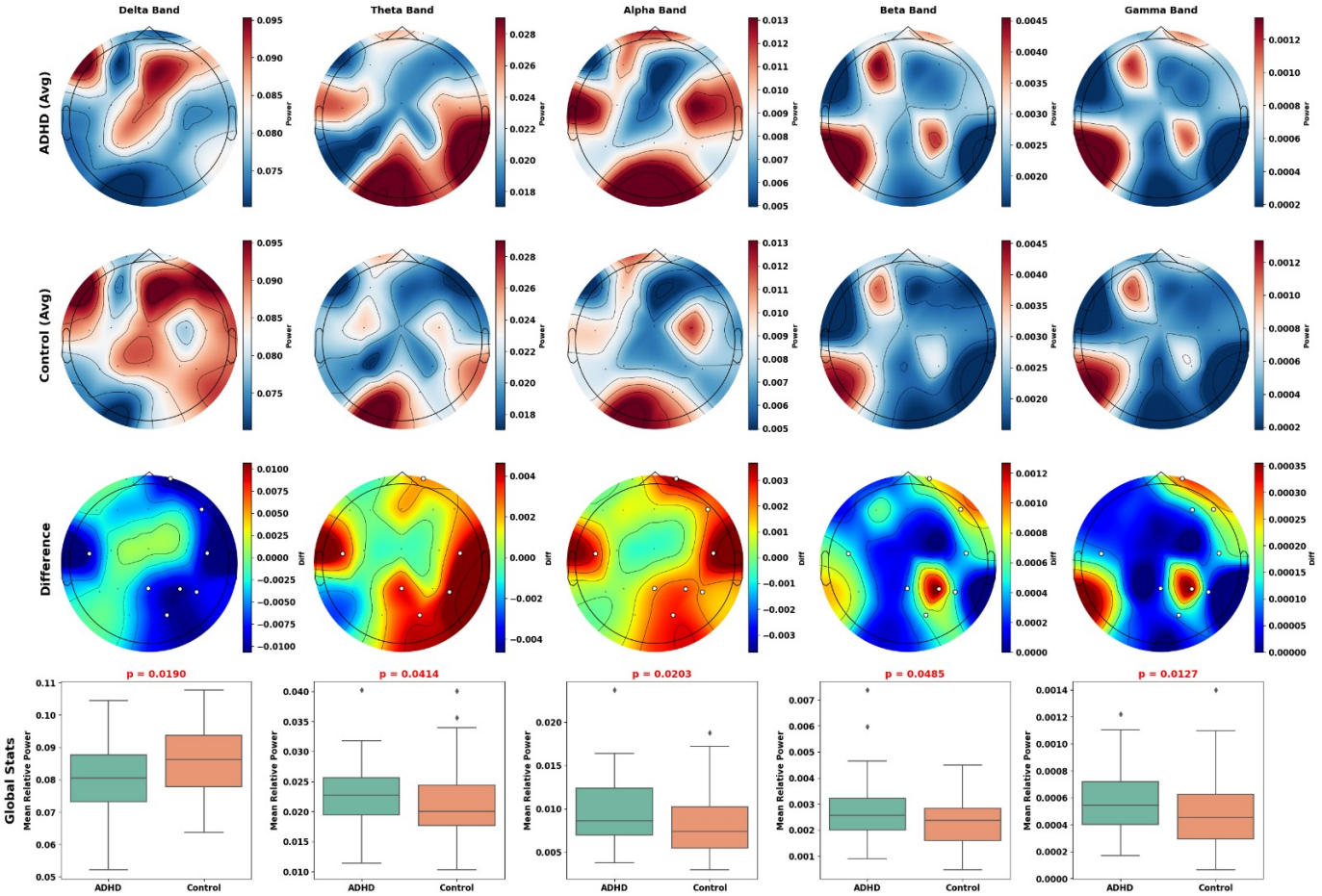

Fig. 3: Topographic power distribution across ADHD and Control groups.

4) Beta Band (12–30 Hz): In comparison to controls, the ADHD group showed lower beta activity in both frontal and sensorimotor areas. Reduced control from above and cognitive ability are linked to lower beta power. A significant difference was confirmed by the Wilcoxon test (p = 0.0485).

5) Gamma Band (30–45 Hz): ADHD patients had significantly decreased gamma power, especially in temporal and parieto-occipital locations. Memory for work, perceptual binding, and high-frequency cortical integration are all related to gamma oscillations. High-frequency processing in ADHD fails which is shown by the highly significant group differences (p = 0.0127).

B. Model Performance Comparison

The classification performance for different model architectures and channel configurations is summarized in Table I. EEGNet obtained an accuracy of **72.98%** using the entire 19-channel montage, whereas the modified EEGNet achieved a higher accuracy of **77.46%**, indicating its enhanced capacity to model diverse spatio-temporal dynamics relevant to ADHD. Both models maintained competitive performance with only a marginal reduction in accuracy when restricted to Wilcoxon-significant electrodes, with modified EEGNet reaching **73.95%** and EEGNet achieving **71.43%**. This suggests that the majority of discriminative information present in the full montage is preserved within a compact, statistically selected channel subset, supporting the feasibility of lightweight and wearable EEG systems.

In addition to accuracy-based evaluation, training and validation loss curves obtained under sample-wise data splitting revealed tightly coupled convergence behavior for both EEGNet and ShallowConvNet. The rapid reduction and close alignment of training and validation losses indicate artificially optimistic learning dynamics caused by data leakage, where subject-specific information is shared across splits. Notably, EEGNet exhibited smoother convergence patterns, while ShallowConvNet demonstrated faster loss minimization due to its higher model capacity. These observations further reinforce the necessity of strict subject-wise evaluation protocols to obtain clinically valid and generalizable performance estimates for EEG-based ADHD detection.

C. Impact of Sample-Wise Splitting (Data Leakage)

The high accuracies of 97.57% and 99.51% which were obtained with sample-wise splitting for EEGNet and the modified EEGNet indicate that there is clear data leakage, as EEG

TABLE I: Model Performance Comparison Across Channel Configurations

Model	Channels	Accuracy (%)	F1-Score	Precision
EEGNet	19	72.98	75.68	68.93
EEGNet with ADHD focused channels	9	71.43	75.57	72.22
Modified EEGNet	19	77.46	80.25	72.47
Modified EEGNet with ADHD focused channels	9	73.95	77.24	75.26
EEGNet (Sample Splitting)	19	97.57	98.00	98.00
Modified EEGNet (Sample Splitting)	19	99.51	99.00	99.00

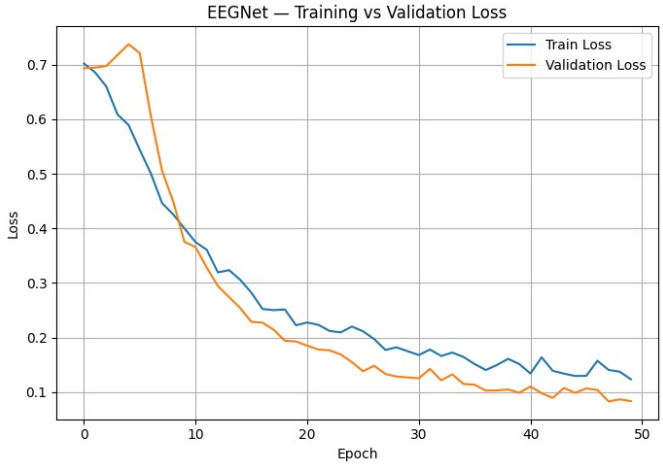

Fig. 4: Training and validation loss curves for EEGNet under sample-wise data splitting, illustrating tightly coupled convergence behavior indicative of data leakage.

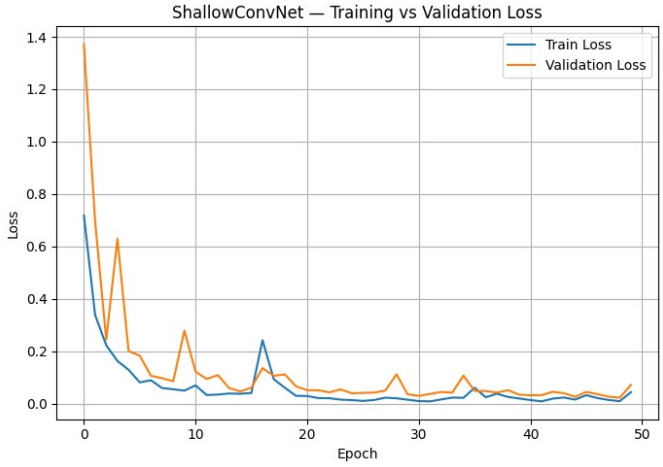

Fig. 5: Training and validation loss curves for ShallowConvNet under sample-wise data splitting, showing rapid loss minimization and close train–validation alignment caused by data leakage.

segments from the same subject appear in both training and test sets. As a result of this the model learns subject-specific signal features instead of brain patterns associated with ADHD which is leading to performance that is artificially inflated. Consequently, obtaining accurate and clinically relevant EEG-based diagnostic results requires subject-wise GroupKFold validation.

IV. Conclusion

This study evaluates EEGNet and its variants for pediatric ADHD detection using multi-channel EEG under strict, leak-free subject-wise validation.The results demonstrate that CNNs may learn discriminative neural patterns without the need for manually created features, and that channel selection based on Wilcoxon minimizes performance loss while reducing dimensionality. This paper highlights how sample-wise splitting inflates accuracy, demonstrating EEGNet's better stability and generalization and highlighting the need for an acceptable subject-level evaluation for EEG-based diagnosis.

References

[1] A. Vishal, S. Deepan, and V. Amrutha, "Single-Modality Emotion Detection: EEG-Based Feature Engineering and Interpretability," in *Proc. 11th Int. Conf. on Bio Signals, Images, and Instrumentation (ICBSII)*, Chennai, India, 2025, pp. 1–9, doi: 10.1109/ICBSII65145.2025.11013139.

[2] D. Sasidharan, V. Sowmya, and E. A. Gopalakrishnan, "Significance of gender, brain region and EEG band complexity analysis for Parkinson's disease classification using recurrence plots and machine learning algorithms," *Phys. Eng. Sci. Med.*, vol. 48, pp. 391–407, 2025, doi: 10.1007/s13246-025-01521-5.

[3] R. T. Schirrmeister *et al.*, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Human Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.

[4] 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 Sciences*, vol. 15, no. 1, p. 95, 2025. https://doi.org/10.3390/brainsci15010095

[5] Ahire, N., Awale, R. N., & Wagh, A. (2025). *Electroencephalogram (EEG)-based prediction of attention deficit hyperactivity disorder (ADHD) using machine learning*. Applied Neuropsychology: Adult, 32(4), 966–977. https://doi.org/10.1080/23279095.2023.2247702

[6] 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. 188, 2024.

[7] G. Brookshire, J. Kasper, N. M. Blauch, Y. C. Wu, R. Glatt, D. A. Merrill, S. Gerrol, K. J. Yoder, C. Quirk, and C. Lucero, "Data leakage in deep learning studies of translational EEG," *Frontiers in Neuroscience*, vol. 18, p. 1373515, 2024. https://doi.org/10.3389/fnins.2024.1373515

[8] F. Del Pup, A. Zanola, L. F. Tshimanga, A. Bertoldo, L. Finos, and M. Atzori, "The role of data partitioning on the performance of EEG-based deep learning models in supervised cross-subject analysis," *arXiv preprint arXiv:2505.13021*, 2025.