

An EEG Signals-based Deep Spatio-Temporal Feature Extraction with Dual Attention for Detecting Parkinson's Disease

A. Dhavapandiammal

dhavam.ap@gmail.com

Kalavathi Palanisamy

Method Article

Keywords: Common Spatial Pattern, Dual-Attention EEGNet, Parkinson's Disease, EEG Signals

Posted Date: July 9th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9318311/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: Competing interest reported. nil

An EEG Signals-based Deep Spatio-Temporal Feature Extraction with Dual Attention for Detecting Parkinson's Disease

A. Dhavapandiammal¹, Kalavathi Palanisamy^{2*}

^{1,2*} Department of Computer Science and Applications

The Gandhigram Rural Institute (Deemed to be University), Gandhigram, Dindigul 624 302

Abstract

Parkinson's disease (PD) is a neurological disorder, causing irregular brain oscillations that can vary rapidly, impairing motor and cognitive abilities. It is also very common, progressive, and complex. High-temporal-resolution electrophysiological monitoring, such as EEG, can provide valuable insights into brain changes associated with PD. Early and effective detection is crucial for adequate management and treatment planning. This paper proposes a Deep Spatio-Temporal Feature Extraction-based Dual-Attention EEGNet (CSFE-DAE) framework for PD detection using EEG signals. Initially, Common Spatial Pattern (CSP) is employed to extract discriminative spatial features by maximizing the variance differences between PD and healthy control (HC) signals. These features are then fed into a Dual-Attention (DA) EEGNet, which integrates both temporal attention to emphasize informative time segments and spatial attention to focus on the most relevant EEG channels. Data augmentation techniques, such as the addition of Gaussian noise, were integrated to enhance robustness against sensor variability and improve generalization. The empirical analysis demonstrates that the proposed CSFE-DAE achieves a mean accuracy of 97.59% (± 0.47) using 5-fold cross-validation on the publicly available UC San Diego resting-state EEG dataset. Additionally, it is implemented on the independent OpenNeuro dataset, showing improvement in external validation and generalization. These results show that CSFE-DAE is a competent and lightweight diagnostic tool that outperforms state-of-the-art EEG-based PD diagnostic techniques.

Keywords: Common Spatial Pattern, Dual-Attention EEGNet, Parkinson's Disease, EEG Signals.

1. Introduction

PD is the second most prevalent neurodegenerative disease, damages the dopaminergic neurons that control movement and results in neuronal death [1]. Primary motor and secondary non-motor symptoms are its defining feature [2]. Nonetheless the existence of motor symptoms is the most important indicator for PD diagnosis [3]. PD can be detected using a variety of

techniques including voice data, neuroimaging, and physiological measurements [4]. EEG has recently gained attention among these techniques as a diagnostic tool [5] – [7]. Rapid information about brain activity and connections is captured by EEG which facilitates the analysis of many neurological conditions including PD [8], [9]. Techniques for neuroimaging can be costly and not always available [10]. Nonetheless the sensors utilized in EEG signal processing methods are straightforward reasonably priced and more user-friendly for capturing brain activity. As a result, it is favored over other imaging techniques such as computed tomography (CT) magnetic resonance imaging (MRI) and functional MRI (fMRI) [11], [12]. In the early stages of PD abnormal brain oscillations will occur. Since EEG is a simple technique for identifying changes in brain activity with excellent temporal resolution it can be used to monitor such abnormalities [13],[14]. EEG signals have emerged as a promising sensor modality for PD diagnosis because of their non-invasiveness and capacity to detect abnormal variations linked to PD [15].

In recent years, healthcare has had a significant impact on the emergence of Artificial Intelligence, encompassing Deep Learning (DL) and Machine Learning (ML). These approaches were developed for the intuitive diagnosis and categorisation of PD using EEG signals [16]. EEG recording provides a rich source of information on the complex, nonlinear brain activities. A practical feature extraction approach is necessary to capture the features observed in the EEG. Diagnosing PD by extracting relevant features in EEGs has shown potential effectiveness with the aid of DL, and is also being actively researched. Recently, several studies have explored a variety of changes in EEG signals (linear/nonlinear) associated with PD [17] - [19], particularly in the frequency range (10-35 Hz) [20]. According to these studies, quantitative EEG analysis has the potential to improve PD diagnosis, prognosis, monitoring, and therapy planning. In recent years, numerous studies have been conducted on the automatic identification of PD from EEG signals. Several techniques have been implemented in conventional research works for PD diagnosis, utilizing various feature extraction and modelling approaches [21] – [24].

An ML-based classification algorithm utilized measurable EEG features retrieved from the domains (time/time-frequency) to distinguish between PD and HC [17], [25], [26]. These analyses have provided desirable insights by identifying specific EEG features related to the aberrant alterations associated with PD. However, the use of uniquely developed characteristics in classification may lead to limited efficiency when tested on previously unreported records from unknown instances [27]. The emerging application of DL to PD diagnosis leads to the building of relevant literature. Shah et al. [28] used DL models to classify PD and HC using

the signals of EEG to perform medication on vs off. Oh et al. [29] proposed a 13-layer convolutional neural network (CNN) model for raw resting EEG data. A pooling-based recurrent neural network (RNN) composed of multiple long short-term memory (LSTM) layers was developed by Xu et al. [30] to extract both long-term and short-term dependencies in EEG data. The model was applied to raw EEG data, achieving good performance. An automated Q-Wavelet transform was developed using ML-based methods in Khare et al. [21] to extract representative sub-bands from EEG data. Five features from each sub-band are retrieved and used to identify PD. Khare et al. [22] proposed the use of CNNs and the EEG time-frequency representations technique. The smoothed pseudo-Wigner-Ville distribution (SPWVD) technique generates time-frequency representations of EEGs, which are subsequently fed into CNNs for detecting PD. Lee et al. [23] presented a CNN-RNN (CRNN) model to extract significant EEG information, utilizing CNNs for spatial feature extraction and gated recurrent units (GRUs) for temporal feature discovery, which yielded superior classification results. Aljalal et al. [31] utilized a CSP to extract numerous features from EEG recordings using ML approaches to identify PD in both off- and on-medication settings. Several studies have attempted to diagnose PD using EEG signals. Khare and Bajaj [32] developed a computer-aided clinical decision-support system (CACDSS) that utilises EEG signals to detect PD automatically. The CACDSS analyzes and classifies signals using AOVMD and AOELM algorithms. AOVMD optimizes decomposition parameters to obtain representative modes.

In Soojin et al. [33], a different CRNN approach was used to classify PD, utilizing CSP as features. Delfan et al. [34] introduced a Hybrid Deep Spatiotemporal Attention-Based Model for PD detection. A novel method for identifying PD was presented by Fatma et al. [35], which makes use of ML approaches. An Artificial Neural Network (ANN) was developed in Kamalakannan et al. [36] to extract features, evaluate the EEG signal, and detect PD at an early stage. ICA-derived features, using a standard ML technique, are combined with the PD Oddball dataset in Smrdel [37] to predict PDs. Attention-based sparse graph convolutional networks (ASGCNN), represented in Chang et al. [38] and GNN were adapted by Neves et al. [39], demonstrating the power of graph learning techniques. CNNs with forged channels were also investigated in [40]. CNN, combined with an SVM classifier [41], has yielded promising results in the early-stage diagnosis of PD using EEG, a conventional approach [42].

Despite their effectiveness, the majority of these approaches exclusively concentrate on temporal or channel dynamics. These drawbacks serve as the foundation for our suggested approach, which uses CSP-based spatiotemporal filtering to extract discriminative EEG sensor signals and a DA mechanism to improve the signal-level representation to increase the accuracy

and generality of PD detection. However, most existing models struggle with two significant challenges: (i) direct reliance on raw EEG recordings without sufficiently separating discriminative spatial signal components, (ii) the lack of specialized dual attention mechanisms for selectively focusing on valuable spatial variability across electrodes and temporal fluctuations across time windows.

This study presents a CSFE-DAE model for detecting PD using EEG to bridge these gaps. The key highlights of this paper are:

- To the best of our knowledge, this is the first work to combine CSP-assisted EEGNet model equipped with dual-attention modules for identifying PD using EEG signals. As far as available literature indicates, this specific combination has not been explored previously for PD detection using EEG signals.
- A channel-attention mechanism is included to allow the network to adjust the importance of each EEG sensor, helping it focus on spatial regions that carry stronger PD-related information.
- Sensor-level Gaussian noise augmentation is used to improve the proposed model's tolerance to variability and minor disturbances that naturally occur during EEG acquisition.
- The generalisation of the proposed model is evaluated by applying it to an independent external dataset that was not used in the training process.

The remaining sections of the paper are arranged as follows: Section 2 outlines the methodology for the proposed model. The detailed architecture of the proposed framework is represented in Section 3. Section 4 discusses the experimental results of the suggested model and concludes in Section 5.

2. Proposed methodology for detecting PD using EEG signals

The CSFE-DAE model is a suitable deep learning framework developed to enhance the accuracy of PD detection using EEG signals. In contrast to conventional research works, the CSFE-DAE model is implemented by integrating CSP and a dual-attention EEGNet. By combining CSP features with DA-EEGNet, the proposed model achieves high accuracy, robustness to subject variability, and improved interpretability. Fig.1 illustrates the overall processes of the proposed CSFE-DAE model.

As presented in Fig.1, Raw EEG signals are collected from the scalp using a sensor for each patient. EEG captures brain activity. After that, preprocessing is performed to make the signals

suitable for further analysis. Then, Gaussian noise injection is first employed as a data augmentation technique to enhance robustness and generalization by simulating sensor variability. Thereafter, CSP is applied to extract discriminative spatial features from the augmented signals fed into a DA-EEGNet, where temporal attention emphasizes informative time segments and channel-level spatial attention focuses on the most relevant EEG signals to predict the PD. The Focal Loss function is used to address class imbalance, assigning more weight to difficult-to-classify samples and thereby improving overall accuracy.

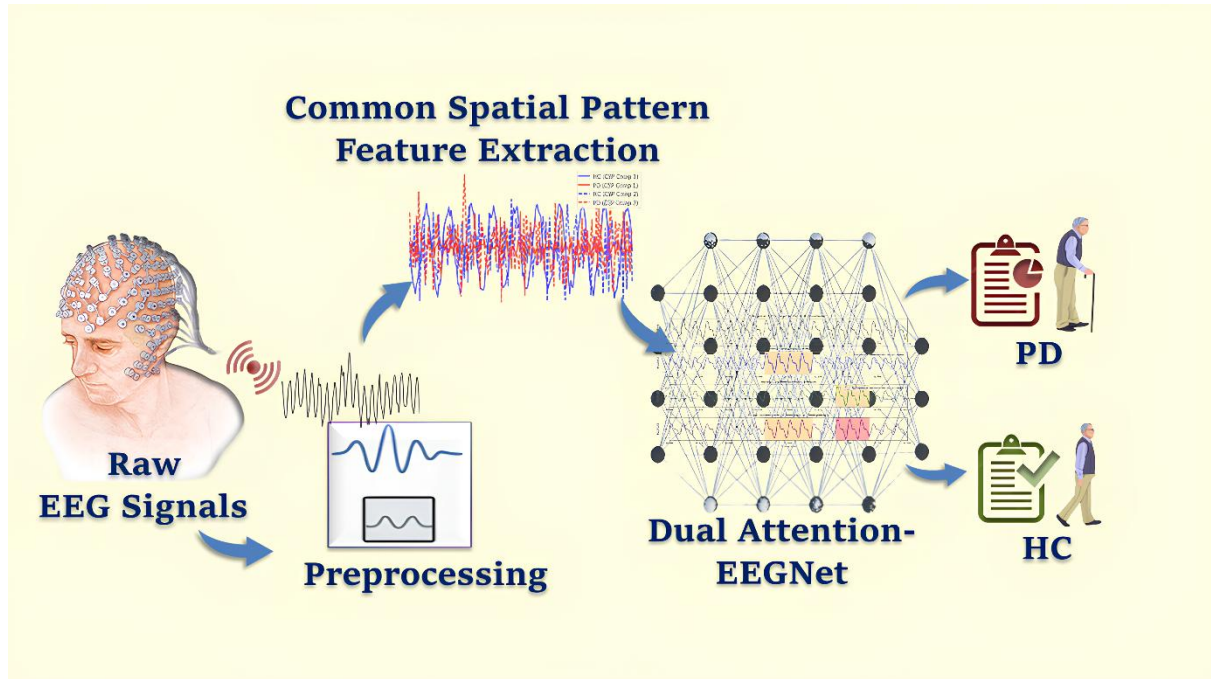


Fig.1. Overview of the proposed CSFE-DAE model

2.1 Dataset Description

Two publicly available EEG datasets were used in this study: ds002778 (UC San Diego resting-state EEG) [43] and ds003490 (OpenNeuro Parkinson's EEG) [44]. The ds002778 dataset included 15 PD and 16 HCs, each with one resting-state session recorded using 40 scalp electrodes at 512 Hz (~196 seconds per subject). The ds003490 dataset comprised 25 PD (two sessions per patient: ON and OFF medication) and 25 HC (one session each), recorded with 64 electrodes at 500 Hz. Only the resting-state segments (eyes-open / eyes-closed) were used for analysis, excluding the auditory oddball task.

2.2 Preprocessing

All EEG recordings from the two OpenNeuro datasets (ds002778 and ds003490) were processed using a unified pipeline to ensure consistency across sources. The raw signals were first downsampled to 256 Hz to reduce computational load while retaining essential spectral

information. Band-pass filtering between 0.5 and 45 Hz was applied to suppress slow drifts and high-frequency noise, followed by a notch filter to eliminate power-line interference. Independent Component Analysis (ICA) was then performed to identify and remove ocular and muscular artefacts. After artefact cleaning, all channels were re-referenced using the common average reference (CAR) approach. As reported in previous EEG studies across emotion analysis [45], segmenting continuous EEG into short, stationary epochs improves both spatio-temporal stability and classification reliability, the same strategy was adopted here.

The continuous EEG was segmented into non-overlapping 2-second epochs, and each channel within an epoch was normalised using z-score scaling. The initial segmentation resulted (ds002778) in an unequal number of epochs between the two groups, with the HC (642) and PD (1186) epochs. Based on the method described in [46], Gaussian noise augmentation was implemented only on the HC epochs. This increased the HC count while leaving the PD epochs remain unchanged, thereby creating a balanced training set. The final epoch distribution is shown in Table 1. The dataset (ds003490), used primarily for external evaluation, was processed with the same steps but without augmentation to preserve its original distribution. This unified procedure ensured that both datasets entered the model with comparable signal characteristics and compatible preprocessing structure.

Table 1. Class distribution before and after Augmentation (Training only)

Class	Before Augmentation	After Augmentation
HC	642	1186
PD	1186	1186
Total	1828	2372

The main preprocessing operations applied to the EEG signals are mathematically described below:

Signal Pre-processing and Augmentation

The mathematical representation of the raw EEG signal is given by,

$$X \in R^{C \times T} \quad (1)$$

where, C and T are the number of channels and time samples, respectively.

Then, band-pass filtering is presented as,

$$X_f = \mathcal{F}^{-1}[H(\omega) \cdot \mathcal{F}(\omega)] \quad (2)$$

F refers to the Fourier transform, and $H(\omega)$ defines a band-pass filter (1 inside the passband, 0 outside). Here, it retains only the frequency range relevant for PD analysis. Then, notch filtering removes line noise using,

$$X_{fn} = \mathcal{F}^{-1}[1 - N(\omega) \cdot \mathcal{F}(X_f)] \quad (3)$$

In equation (3), $N(\omega)$ represents a narrowband stop filter at the line frequency. Consequently, the Common Average Reference (CAR) is performed as,

$$X_{ref}(c,t) = X_{fn}(c,t) - \frac{1}{C} \sum_{i=1}^C X_{fn}(i,t) \quad (4)$$

Next, Independent Component Analysis (ICA) removes artifacts using,

$$X_{ref} = AS \quad (5)$$

Here, A refers to the mixing matrix, and S represents the independent components.

Subsequently, segmentation is accomplished, dividing the continuous EEG into discrete trials of a specified length. T_e Samples as mentioned below:

$$X_{clean}^{(k)} = X_{clean}[:, k:T_e + k] \quad (6)$$

Per-channel z-score normalization is mathematically done as,

$$X_{norm}(c,t) = \frac{X_{epoch}(c,t) - \mu_c}{\sigma_c} \quad (7)$$

In equation (7), μ_c and σ_c Represents the mean and standard deviation. Now, Gaussian noise is added to these normalized EEG signals in the training folds, which is given by:

$$X_{aug} = X_{norm} + N(0, \sigma^2) \quad (8)$$

where, $N(0, \sigma^2)$ Gaussian noise with mean 0 and variance σ^2 .

Common Spatial Feature Extraction

CSP is a supervised feature extraction method that identifies spatial filters maximizing variance differences between two classes of signals. In EEG analysis, CSP is employed to derive projections that enhance discrimination between PD and HC trials by emphasizing frequency band-specific alterations, such as changes in β -band (13–30 Hz) or θ -band (4–7 Hz) power over motor-related cortical areas.

To obtain discriminative features, CSP is explored on the preprocessed EEG data using the steps below:

Step 1: Compute the covariance matrices of signals for each class by

$$Cov^M = \frac{XX^T}{\text{trace}(XX^T)} \quad (9)$$

This normalization ensures scale invariance across each trial (Class 1) for PD and (Class 2) for HC.

Step 2: Compute mean composite covariance

CSP measures mean covariance matrices across trials for each class using,

$$\Sigma_1 = \frac{1}{N_1} \sum_{i=1}^{N_1} \Sigma_i^{(1)}, \Sigma_2 = \frac{1}{N_2} \sum_{i=1}^{N_2} \Sigma_i^{(2)} \quad (10)$$

Step 3: Compute composite covariance matrix

After Step 2, CSP integrates the covariance matrices into a single composite matrix by,

$$\Sigma_c = \Sigma_1 + \Sigma_2, \Sigma_c = U \Lambda U^T \quad (11)$$

This integrates information from both classes for whitening. Here, U and Λ Contains eigenvectors and diagonal eigenvalues.

Step 4: Compute the Whitening matrix and Whitened covariance

$$W = \Lambda^{-1/2} U^T, S_1 = W \Sigma_1 W^T, S_2 = W \Sigma_2 W^T \quad (12)$$

After whitening, the sum of whitened class covariances equals the identity matrix.

Step 5: Eigen-Decomposition of Whitened Covariance

$$S_1 = B D B^T, K = B^T W \quad (13)$$

where, B describes spatial filters in whitened space, and D represents a diagonal matrix of eigenvalues. Subsequently, CSP identifies eigenvectors corresponding to the largest and smallest eigenvalues, which are used to form spatial filters. The first few and the Last few filters maximize variance for PD (Class 1), HC (Class 2).

Step 6: Project Original EEG Signals through CSP Filters

$$Z = KX \quad (14)$$

Each row of Z is a spatially filtered EEG component.

Step 7: Compute Log-Variance Features

Here, CSP determines log-variance features for the selected m filters and the feature vector. These are given by,

$$f_k = \log \left(\frac{\text{var}(Z^k)}{\sum_{j=1}^m \text{var}(Z_j)} \right), f = [f_1, f_2, \dots, f_m] \quad (15)$$

These features are fed into an EEGNet classifier for the detection of PD.

3. Proposed CSFE-DAE method

The core idea is to improve the network's ability to concentrate on the brain signals' most relevant temporal (time-based) and spatial (channel-based) features. Selectively focusing on the most critical spatio-temporal features helps the model to achieve higher classification accuracy. By explicitly modelling dependencies in both the time and space domains, the network becomes more resilient to the low signal-to-noise ratio and inherent variability of EEG data.

Feature Extraction:

Fig.2 displays the proposed framework. The process begins with temporal convolution, which extracts frequency-specific patterns from the input CSP features. This stage transforms the normalized input into temporal representations. Subsequently, depthwise spatial convolution is applied to model inter-channel dependencies, followed by batch normalization. Next, a separable convolution is introduced to capture localized joint temporal–spatial interactions, enhancing the network’s ability to extract relevant EEG features. Further, a dual-attention mechanism is incorporated to strengthen feature discrimination.

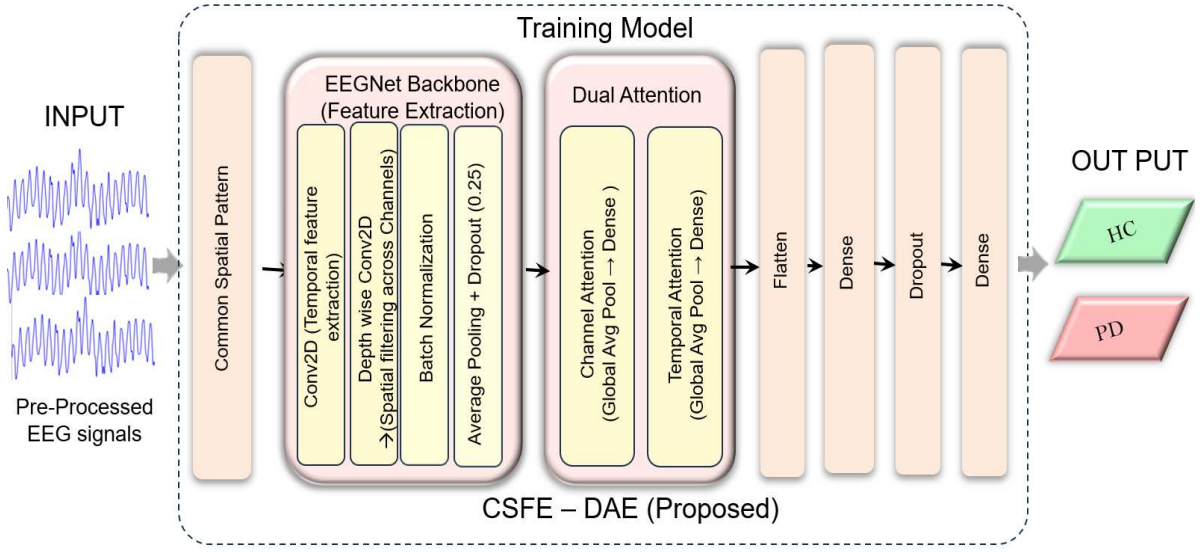


Fig.2. Architecture of the Proposed CSFE-DAE model

Dual Attention Mechanism: After CSP and EEGNet feature extraction, we obtain the feature representation, which is represented as:

$$F \in R^{CxT} \quad (16)$$

where, T and C denote the number of temporal dimensions and the EEG channels. Then, a DA mechanism is applied to enhance discriminative learning, which combines channel and temporal-level dependencies, as shown in Fig. 3.

Channel Attention: Channel attention adaptively reweights electrode allowance by learning importance scores for each channel, thereby suppressing noisy or less informative sensors, while emphasizing electrodes over motor-related cortical areas known to be affected in PD [46]. Here, we compute channel-wise importance by global average pooling across time as:

$$\alpha = \sigma(W_c \cdot AvgPool_t(F)), F' = \alpha \odot F \quad (17)$$

where, W_c is a learnable weight matrix, σ is the sigmoid activation and $\odot$ elementwise multiplication. By this operation, each channel is scaled according to its importance.

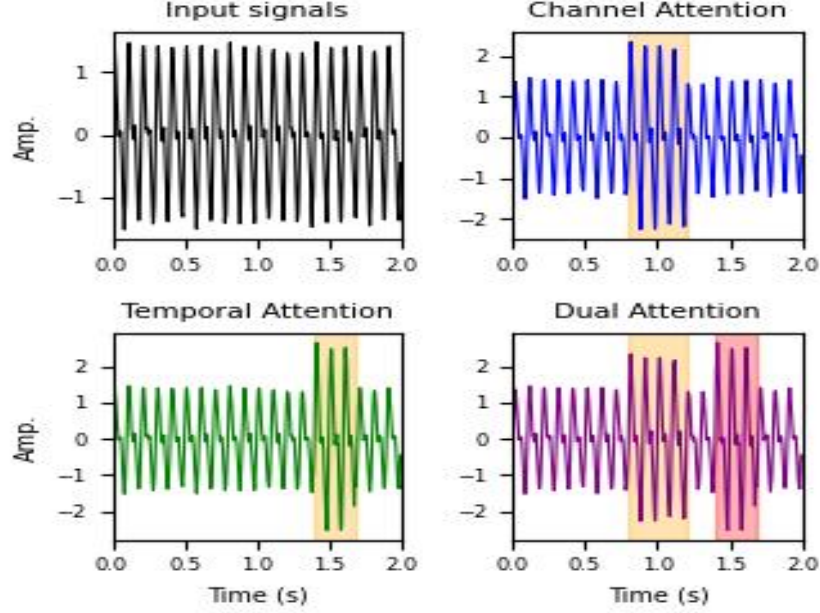


Fig.3. Dual Attention Mechanism

Temporal Attention: Temporal attention enhances this by selectively concentrating on informative signal intervals, capturing abnormal $\alpha - \beta$ dynamics, strongly linked to PD [47]. It highlights the most discriminative time steps. A channel-wise pooling is first performed to derive a temporal descriptor, followed by:

$$\beta = \text{softmax}(W_t \cdot \text{Avgpool}_t(F)), F'' = \beta \odot F' \quad (18)$$

where, SoftMax normalizes attention across time. Here, each segment captures dynamic oscillatory patterns relevant to PD.

Finally, the DA feature map is obtained as

$$\hat{F} = F'' \in R^{C \times T} \quad (19)$$

combines sensor-level and pattern-level importance and feeds it into the fully connected layer with SoftMax activation for final prediction combines sensor-level and pattern-level importance and feeds it into the fully connected layer with SoftMax activation for final prediction.

3.1 Focal Loss for Class Imbalance

The Focal Loss function is implemented to mitigate the effects of class imbalance, which dynamically scales the loss assigned to well-classified examples, focusing more on hard-to-

classify cases. We used $\gamma = 2.0$ and $\alpha = 0.25$, effectively handling imbalance in EEG classification tasks. Focal loss of class imbalance is described as,

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log p_t \quad (20)$$

where, $p_t = \hat{y}_t$ if correct class else $p_t = 1 - \hat{y}_t$ and α_t refers to the class weighting factor (e.g., inversely proportional to class frequency) and γ defines the focusing parameter $\gamma = 2$.

4. Performance metrics and experimental results

To validate the performance of PD recognition, the CSFE-DAE model is evaluated on the dataset [43]. Its efficiency is determined in terms of accuracy, loss, and the confusion matrix. Here, accuracy and loss measure the overall proportion of correctly and incorrectly classified classes, whereas the confusion matrix represents the distribution of predictions across actual classes.

The proposed CSFE-DAE approach is trained over 50 epochs using 5-fold cross-validation. The training and validation accuracy start at around ~ 0.93 and 0.91 . It gradually increases, stabilizing at nearly 0.97 after ~ 20 epochs, indicating that it effectively learns the discriminative features from EEG signals over time with some fluctuations due to cross-validation across different subject subsets, which is illustrated in Fig. 4. The close alignment of accuracy curves (training and validation) suggests good model generalization without significant over-fitting. The slight variance (shaded regions) reflects variability across folds, which indicates robust performance.

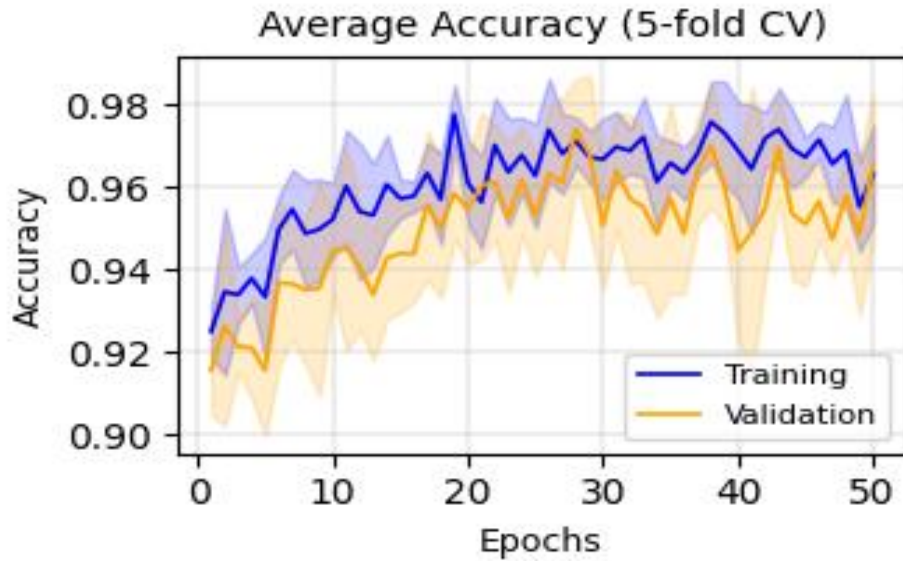


Fig.4. Accuracy of the proposed CSFE-DAE

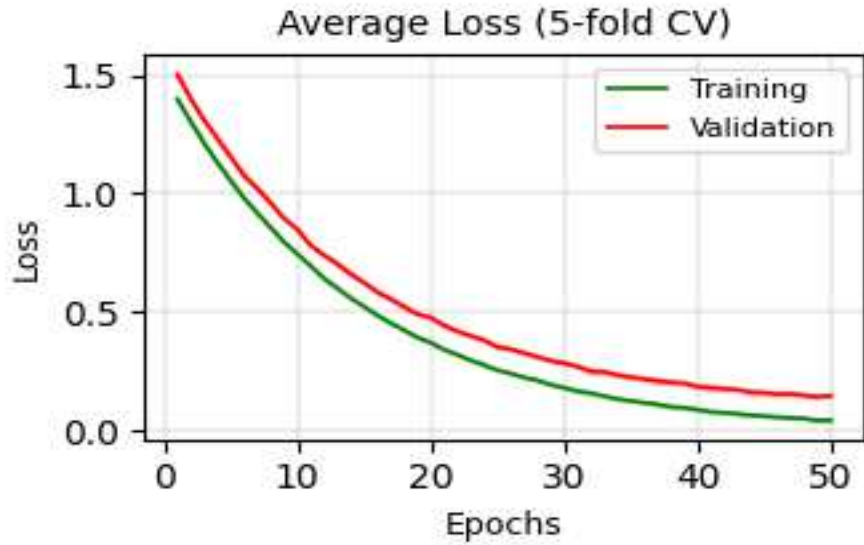


Fig.5. Loss of the CSFE-DAE (proposed)

In Fig.5, the training loss decreases sharply from ~ 1.4 to < 0.1 by epoch 50, showing that the network learns to minimize classification error effectively. The validation loss follows a similar descending pattern, reducing from ~ 1.5 to ~ 0.15 , confirming improved predictive capability on unseen samples. The decreasing gap between training and validation losses over time indicates stable convergence and absence of severe over-fitting.

Table 2. Comparison of the proposed CSFE-DAE Framework with Existing and State-of-the-Art Methods

Category	Method	Approach	Dataset	Accuracy (%)
Existing Works (UC San Diego Dataset)	Ullah et al. (2022) [48]	TFR + AlexNet CNN	US	94
	Bhurane et al. (2022) [49]	Multi-Pattern + Multi-Scale CNN	US	92
	Zhang et al. (2020) [50]	CSP + SVM	US	83
	Bera et al. (2025) [41]	CNN+SVM	US	81.72
	Zang et al. (2025) [42]	GEPD	US	84.3

	Neves et al. (2024) [39]	GNN	US	69.04
	Hamidi et al. (2023) [40]	Forged Channel CNN	US	90.3
	CSP+SVM (Baseline)	CSP + SVM	US	85.42
	EEGNet (Raw EEG)	Lightweight CNN	US	88.69
	CSP+EEGNet (without augmentation)	CSP Features+EEGNet	US	90.18
	CSP+EEGNet (with augmentation)	CSP Features+EEGNet	US	93.0
Our Study (Baselines and Proposed Model)	CSFE-DAE (Proposed)	CSP+Dual Attention	US	97.59
	CSFE-DAE (Validated)	CSP+Dual Attention	OpenNeuro	96.0

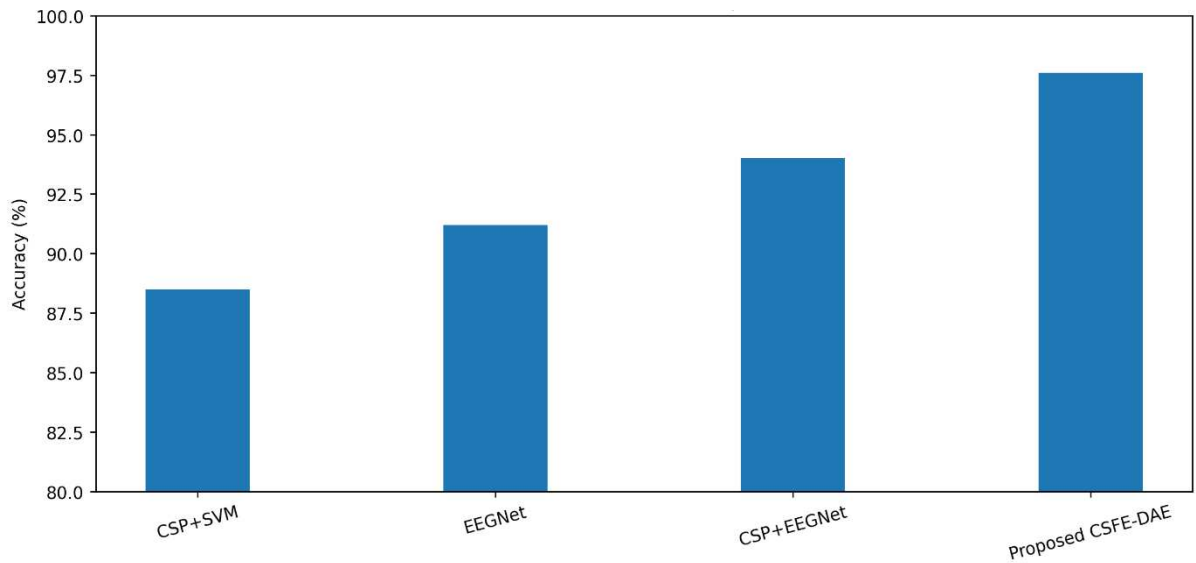


Fig. 6. Comparison of Model Performance

The proposed CSFE-DAE was systematically compared to recent methods using the same dataset [43]. Table 1 summarizes existing techniques, which report 81% to 40% accuracy, also with the internal baselines CSP+SVM, EEGNet, CSP+EEGNet from 85.42% to 93.28% accuracy. The comparison chart in Fig. 6 shows the performance of the proposed model against three baseline methods. our proposed CSFE-DAE outperforms, achieving $97.59 \pm 0.47\%$ accuracy, correctly classifying 1156/1186 PD and 617/642 HC samples across 5-fold cross-

validation (sensitivity 97.5%, specificity 96.1%). This improvement reflects the benefit of combining CSP filtering with a dual-attention EEGNet that captures both spatial and temporal discriminative patterns in PD EEG activity. Further, the proposed model maintained 96% accuracy on the external dataset [44] with sensitivity and specificity at 96%. These findings show that the model does not overfit and can be generalized to independent instances.

5. Conclusions and future work

This research work introduces an advanced EEG-sensor signal-based classification framework, CSFE-DAE, which leverages Gaussian noise augmentation to increase the dataset sizes to enhance PD detection. It integrates CSP with a DA EEGNet architecture, focusing on the signal in EEG data, to capture highly discriminative spatial features. Meanwhile, the DA mechanism focuses on the most relevant temporal and channel regions within EEG signals. The key contribution of this work lies in integrating CSP filtering with a DA variant of EEGNet for PD recognition from resting-state EEG. While CSP and EEGNet have been used separately, earlier studies have not combined them in a unified framework with coordinated channel- and feature-attention modules. This framework enhances sensitivity to PD-related patterns while suppressing irrelevant background fluctuations. In addition, Gaussian noise augmentation is applied only to the training data to improve robustness against sensor-level variations. The model's performance is further supported by evaluation on an external dataset, demonstrating that the approach generalizes beyond the training cohort. Together, these elements establish a methodological contribution that extends beyond existing CSP- or CNN-based approaches. The proposed model achieves a high accuracy using 5-fold cross-validation, with a reduced loss compared to the previous works. Additionally, it shows strong generalizability through external validation. These results demonstrate the effectiveness of CSFE-DAE as a promising method for the early detection of PD. Future research will extend the proposed approach to multi-class, cross-subject adaptation through transfer learning, as well as real-time sensor-adaptive clinical applications for the reliable classification of PD.

Conflict of interest: The authors declare no conflicts of interest.

References

- [1] S. Bhat, U. R. Acharya, Y. Hagiwara, N. Dadmehr, and H. Adeli, "Parkinson's Disease: Cause Factors, Measurable Indicators, and Early Diagnosis," *Computers in Biology and Medicine* 102 (2018):234–241.

- [2] Marinus J, Zhu K, Marras C, Aarsland D, van Hilten JJ (2018) Risk factors for non-motor symptoms in Parkinson's disease. *Lancet Neurol* 17(6):559–568. [https://doi.org/10.1016/S1474-4422\(18\)30127-3](https://doi.org/10.1016/S1474-4422(18)30127-3).
- [3] R. He, X. Yan, J. Guo, Q. Xu, B. Tang, and Q. Sun, “Recent advances in biomarkers for Parkinson's disease,” *Frontiers Aging Neurosci.*, vol. 10, p. 305, Oct. 2018.
- [4] Maitín AM, García-Tejedor AJ, Muñoz JPR (2020) Machine learning approaches for detecting Parkinson's disease from EEG analysis: a systematic review. *Appl Sci* 10(23):8662. <https://doi.org/10.3390/app10238662>.
- [5] Shi X, Wang T, Wang L, Liu H, Yan N (2019) Hybrid Convolutional Recurrent Neural Networks Outperform CNN and RNN in Task-state EEG Detection for Parkinson's Disease. In: 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference. pp. 939–944. <https://doi.org/10.1109/APSIP AASC47483.2019.9023190>.
- [6] Acharyya A, Jadhav PN, Bono V, Maharatna K, Naik GR (2018) Low-complexity hardware design methodology for reliable and automated removal of ocular and muscular artifact from EEG. *Comput Methods Progr Biomed* 158:123–133. <https://doi.org/10.1016/j.cmpb.2018.02.009>.
- [7] Al-Ani A, Koprinska I, Naik G (2017) Dynamically identifying relevant EEG channels by utilizing channels classification behaviour. *Expert Syst Appl* 83:273–282. <https://doi.org/10.1016/j.eswa.2017.04.042>.
- [8] Jindal K, Upadhyay R, Singh HS, Vijay M, Sharma A, Gupta K, Gupta J, Dube A (2018) Migraine disease diagnosis from EEG signals using Non-linear Feature Extraction Technique. In: 2018 IEEE International Conference on Computational Intelligence and Computing Research. pp. 1–4. <https://doi.org/10.1109/ICCIC.2018.8782341>.
- [9] Alyasseri ZAA, Alomari OA, Papa JP, Al-Betar MA, Abdulkareem KH, Mohammed MA, Kadry S, Thinnukool O, Khuwuthyakorn P (2022) EEG channel selection based user identification via improved flower pollination algorithm. *Sensors* 22(6):2092. <https://doi.org/10.3390/s22062092>.
- [10] Khare SK, Bajaj V, Acharya UR (2021) PDCNNNet: an automatic framework for the detection of Parkinson's disease using EEG signals. *IEEE Sens J* 21(15):17017–17024. <https://doi.org/10.1109/JSEN.2021.3080135>.
- [11] Tosun M (2021) Effects of spectral features of EEG signals recorded with different channels and recording statuses on ADHD classification with deep learning. *Phys Eng Sci Med* 44(3):693–702. <https://doi.org/10.1007/s13246-021-01018-x>.

- [12] Guttmann-Flury E, Sheng X, Zhu X (2022) Channel selection from source localization: a review of four EEG-based brain-computer interfaces paradigms. *Behav Res Methods*. <https://doi.org/10.3758/s13428-022-01897-2>.
- [13] Arasteh E, Mahdizadeh A, Mirian MS, Lee S, McKeown MJ (2021) Deep transfer learning for Parkinson's disease monitoring by image-based representation of resting-state EEG using directional connectivity. *Algorithms* 15(1):1–19. <https://doi.org/10.3390/a15010005>.
- [14] Bhardwaj S, Jadhav P, Adapa B, Acharyya A, Naik GR (2015) Online and automated reliable system design to remove blink and muscle artefact in EEG. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 6784–6787). IEEE. <https://doi.org/10.1109/EMBC.2015.7319951>.
- [15] C. Babiloni, M. F. De Pandis, F. M. Vecchio, P. Buffo, F. Sorpresi, G. B. Frisoni, and P. M. Rossini, "Cortical sources of resting-state electroencephalographic rhythms in Parkinson's disease related dementia and Alzheimer's disease," *Clinical Neurophysiology*, vol. 122, no. 12, pp. 2355–2364, Dec. 2011, doi: 10.1016/j.clinph.2011.03.029.
- [16] F. Ström and R. Koker, "A parallel neural network approach to prediction of Parkinson's Disease," *Expert Syst Appl*, vol. 38, no. 10, pp. 12470–12474, Sep. 2011, doi: 10.1016/j.eswa.2011.04.028.
- [17] C.-X. Han, J. Wang, G. S. Yi, and Y. Q. Che, "Investigation of EEG Abnormalities in the Early Stage of Parkinson's Disease," *Cognitive Neurodynamics* 7 (2013): 351–359.
- [18] Yuvaraj R, Rajendra Acharya U, Hagiwara Y, 2018. A novel Parkinson's Disease Diagnosis Index using higher-order spectra features in EEG signals. *Neural Comput. Appl.* 30, 1225–1235.
- [19] Carmona J, Suarez J, Ochoa J, 2017. Brain Functional Connectivity in Parkinson's disease – EEG resting analysis, in: VII Latin American Congress on Biomedical Engineering CLAIB 2016. Springer, Singapore, pp. 185–188.
- [20] Silberstein P, Pogosyan A, Kühn AA, Hotton G, Tisch S, Kupsch A, Dowsey-Limousin P, Hariz MI, Brown P, 2005. Cortico-cortical coupling in Parkinson's disease and its modulation by therapy. *Brain* 128, 1277–91. [PubMed: 15774503].
- [21] S. K. Khare, V. Bajaj, and U. R. Acharya, "Detection of Parkinson's Disease Using Automated Tunable Q Wavelet Transform Technique With EEG Signals," *Biocybernetics and Biomedical Engineering* 41, no.2 (2021): 679–689.
- [22] S. K. Khare, V. Bajaj, and U. R. Acharya, "PDCNNNet: An Automatic Framework for the Detection of Parkinson's Disease Using EEG Signals," *IEEE Sensors Journal* 21, no. 15 (2021): 17017–17024.

- [23] S. Lee, R. Hussein, R. Ward, Z. Jane Wang, and M. J. McKeown, "A Convolutional-Recurrent Neural Network Approach to Resting-State EEG Classification in Parkinson's Disease," *Journal of Neuroscience Methods* 361 (2021): 109282.
- [24] M. Aljalal, S. A. Aldosari, K. AlSharabi, A. M. Abdurraqueeb, and F. A. Alturki, "Parkinson's Disease Detection From Resting-State EEG Signals Using Common Spatial Pattern, Entropy, and Machine Learning Techniques," *Diagnostics* 12, no. 5 (2022): 1033.
- [25] Koch M, Geraedts V, Wang H, Tannemaat M, Back T, 2019. Automated Machine Learning for EEG-EEGBased Classification of Parkinson's Disease Patients, in: *Proceedings - 2019 IEEE International Conference on Big Data, Big Data 2019*. Institute of Electrical and Electronics Engineers Inc., pp. 4845–4852.
- [26] Oliveira A.P.S. de, Santana M.A. de, Andrade MKS, Gomes JC, Rodrigues MCA, Santos W.P. dos, 2020. Early diagnosis of Parkinson's disease using EEG, machine learning and partial directed coherence. *Res. Biomed. Eng.* 36, 311–331.
- [27] Cao C, Liu F, Tan H, Song D, Shu W, Li W, Zhou Y, Bo X, Xie Z, 2018. Deep Learning and Its Applications in Biomedicine. *Genomics, Proteomics Bioinforma.* 16, 17–32.
- [28] Shah SAA, Zhang L, Bais A, 2020. Dynamical system based compact deep hybrid network for classification of Parkinson disease related EEG signals. *Neural Networks* 130, 75–84. [PubMed: 32650152].
- [29] S. L. Oh, Y. Hagiwara, U. Raghavendra, et al., "A Deep Learning Approach for Parkinson's Disease Diagnosis From EEG Signals," *Neural Computing and Applications* 32 (2020): 10927–10933.
- [30] Xu S, Wang Z, Sun J, Zhang Z, Wu Z, Yang T, Xue G, Cheng C, 2020. Using a deep recurrent neural network with EEG signal to detect Parkinson's disease. *Ann. Transl. Med.* 8, 874–874. [PubMed: 32793718].
- [31] M. Aljalal, S. A. Aldosari, K. AlSharabi, A. M. Abdurraqueeb, and F. A. Alturki, "Parkinson's Disease Detection From Resting-State EEG Signals Using Common Spatial Pattern, Entropy, and Machine Learning Techniques," *Diagnostics* 12, no. 5 (2022): 1033.
- [32] S. K. Khare and V. Bajaj, "A CACDSS for automatic detection of Parkinson's disease using EEG signals," in *2021 International Conference on Control, Automation, Power and Signal Processing (CAPS)* (Jabalpur, India: IEEE, 2021).
- [33] Soojin Lee, Ramy Hussein, Rabab Ward, Z. Jane Wang, Martin J. McKeown, "A convolutional-recurrent neural network approach to resting-state EEG classification in Parkinson's disease", *Journal of Neuroscience Methods*, Volume 361, 2021, 109282, ISSN 0165-0270, <https://doi.org/10.1016/j.jneumeth.2021.109282>.

- [34] Niloufar Delfan, Mohammadreza Shahsavari, Sadiq Hussain, Robertas Damaševičius, U. Rajendra Acharya, “A Hybrid Deep Spatio-Temporal Attention-Based Model for Parkinson's Disease Diagnosis Using Resting State EEG Signals”, *International Journal of Imaging Systems and Technology*, Pages 1-15, 2024
- [35] Fatma Latifoglu, Sultan Penekli, Fırat Orhanbulucu, Muhammad E. H. Chowdhury, “A novel approach for Parkinson's disease detection using Vold-Kalman order filtering and machine learning algorithms”, *Neural Computing and Applications* (2024) 36:9297–9311.
- [36] Naveenraj Kamalakannan, Shiva Prasaath Sudha Balamurugan, Kalaivani Shanmugam, “A Novel Approach For The Early Detection Of Parkinson's Disease Using EEG Signal”, *International Journal of Electrical Engineering and Technology (IJEET)* Volume 12, Issue 5, May 2021, pp. 80-95
- [37] A. Smrdel, “Independent component analysis of oddball EEG recordings to detect Parkinson's disease,” *Sci Rep*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-07645-8.
- [38] H. Chang, B. Liu, Y. Zong, C. Lu, and X. Wang, “EEG-based Parkinson's disease recognition via attention-based sparse graph convolutional neural network,” **IEEE Journal of Biomedical and Health Informatics**, vol. 27, no. 11, pp. 5216–5224, Nov. 2023.
- [39] C. Neves, Y. Zeng, and Y. Xiao, “Parkinson's Disease Detection from Resting State EEG using Multi-Head Graph Structure Learning with Gradient Weighted Graph Attention Explanations,” Aug. 2024, [Online]. Available: <http://arxiv.org/abs/2408.00906>
- [40] Hamidi, A., Mohamed-Pour, K., & Yousefi, M. (2023). Forged Channel: A Breakthrough Approach for Accurate Parkinson's Disease Classification using Leave-One-Subject-Out Cross-Validation. *arXiv*. <https://doi.org/10.48550/arXiv.2305.02234>
- [41] S. Bera, Z. W. Geem, Y. I. Cho, and P. K. Singh, “A Comparative Study of Machine Learning and Deep Learning Models for Automatic Parkinson's Disease Detection from Electroencephalogram Signals,” *Diagnostics*, vol. 15, no. 6, Mar. 2025, doi: 10.3390/diagnostics15060773.
- [42] Zhang, Q. (2025). GEPD: GAN-enhanced generalizable model for EEG-based detection of Parkinson's disease. *arXiv*. <http://arxiv.org/abs/2508.14074>.
- [43] <https://openneuro.org/datasets/ds002778/versions/1.0.2>
- [44] <https://openneuro.org/datasets/ds003490/versions/1.1.0>
- [45] C. Liao, S. Zhao, X. Wang, J. Zhang, Y. Liao, and X. Wu, “EEG Data Augmentation Method Based on the Gaussian Mixture Model,” *Mathematics*, vol. 13, no. 5, Mar. 2025, doi: 10.3390/math13050729.

- [46] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in neural information processing systems* (pp. 5998–6008).
- [47] W. Chen, Y. Liao, R. Dai, Y. Dong, and L. Huang, “EEG-based emotion recognition using graph convolutional neural network with dual attention mechanism,” *Front Comput Neurosci*, vol. 18, 2024, doi: 10.3389/fncom.2024.1416494.
- [48] Fractal analysis of EEG signals for the classification of schizophrenia and Parkinson’s disease. *International Journal of Neural Systems*, 26(5), 1650025. <https://doi.org/10.1142/S0129065716500250>.
- [49] Ullah, I., Hussain, M., Aboalsamh, H., et al. (2022). A deep CNN-based approach using time–frequency scalogram images for Parkinson’s disease detection. *Biomedical Signal Processing and Control*, 75, 103590. <https://doi.org/10.1016/j.bspc.2022.103590>.
- [50] Bhurane, A. A., Dhok, S. B., Sharma, M., Yuvaraj, R., Murugappan, M., & Acharya, U. R. (2022). Diagnosis of Parkinson’s disease from electroencephalography signals using linear and self-similarity features. *Expert Systems*, 39(7), e12472. <https://doi.org/10.1111/exsy.12472>.
- [51] Zhang, D., Yao, L., Chen, K., Tang, Y., & Wang, Q. (2020). EEG-based classification of Parkinson’s disease using CSP and SVM. *Frontiers in Neuroscience*, 14, 578. <https://doi.org/10.3389/fnins.2020.00578>.