International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(11s) 2026 pp. 840-858
https://doi.org/10.70917/ijcisim-2026-3819
Article

Adaptive Multi-band Dynamic Functional-Connectivity Graph Attention Network for Parkinson's Disease Classification from Resting-State EEG

Sumit Gupta[1], Ranjitha U N[2], J. Gul Shaira Banu[3], Kalidindi Lakshmi Divya[4], Tripti R Kulkarni[5], Shweta Mannikeri[6,7]

[1]Director, DeepCognix AI Labs, Bengaluru, Karnataka, India. Email: sumit@deepcognix.com. ORCID ID: 0009-0007-4766-3468
[2]School of Computer Science and Engineering, REVA University, Bengaluru, Karnataka, India. Email: ranjitha.un@reva.edu.in
[3]Department of Artificial Intelligence and Data Science, Er Perumal Manimekalai College of Engineering, PMC Tech, Hosur, Tamil Nadu, India. Email: drgulshairabanu@gmail.com. ORCID ID: 0000-0003-4877-8738
[4]Department of ECE, SRKR Engineering college, Bhimavaram, India. Email: Kldivya@srkrec.edu.in
[5]Dayananda Sagar Academy of Technology and Management, Bengaluru, Karnataka, India. Email: kulkarnitripti@gmail.com
[6]Department of Computer Science, Chairashree Institute of Research and Development (CIRD), University of Mysore, Karnataka, India
[7]Department of Computer Science, Government College for Women, Chintamani, Karnataka, India. Email: shwetaphd25@gmail.com. ORCID ID:0009-0003-6074-8416

Abstract: Parkinson's disease is a progressive neurodegenerative disorder whose reliable diagnosis from non-invasive electrophysiological signals remains challenging because the discriminative information is distributed across spectral bands and across spatially interacting cortical regions. This study proposes an Adaptive Multi-band Dynamic Functional-Connectivity Graph Attention Network (AMD-FCGAT) that represents resting-state electroencephalography as a set of band-resolved brain graphs and learns disease-specific connectivity signatures through attention. For each rhythmic band the framework estimates functional connectivity using the weighted phase-lag index, the phase-locking value and spectral coherence, constructs a sparse electrode graph, and encodes it with a residual multi-head graph attention encoder. An adaptive cross-band gate fuses the band-specific embeddings, and a dynamic temporal self-attention module models the evolution of connectivity across sliding windows before an attention readout produces the classification decision. Experiments on the publicly available University of California San Diego resting-state dataset show that the proposed model attains 97.31 percent accuracy, 97.22 percent F1-score and an area under the curve of 0.998 under a stratified protocol, while a stricter subject-independent evaluation retains 90.97 percent accuracy. Ablation confirms that the cross-band fusion and temporal attention modules contribute the largest gains, and interpretability analysis localises the strongest alterations to central and parietal beta-band synchrony, in agreement with established Parkinsonian neurophysiology. The framework offers an accurate and transparent tool for electroencephalography-based screening.
Keywords: Parkinson's disease, electroencephalography, functional connectivity, graph attention network, deep learning, brain network analysis

1. Introduction

Parkinson's disease is the second most prevalent neurodegenerative disorder worldwide, characterised by the progressive loss of dopaminergic neurons in the substantia nigra and by motor symptoms such as bradykinesia, rigidity and resting tremor, together with a broad spectrum of non-motor manifestations. Clinical diagnosis still relies mainly on the subjective assessment of motor signs, which typically appear only after a substantial proportion of dopaminergic neurons has already degenerated, so there is a strong incentive to develop objective and inexpensive markers that can support earlier and more consistent identification [1]. Electroencephalography is particularly attractive for this purpose because it is non-invasive, widely available, inexpensive and directly sensitive to the neuronal oscillatory activity that is disturbed in Parkinson's disease [2].

A growing body of evidence indicates that Parkinsonian pathology is not confined to isolated cortical sites but rather reshapes the coordinated activity of distributed brain networks. Quantitative connectivity studies report abnormal phase coupling in the beta band together with reduced coordination in the alpha and theta ranges, and these alterations correlate with disease severity and with the response to dopaminergic medication [2]. Functional connectivity, which measures the statistical interdependence between the signals recorded at different electrodes, therefore captures information that electrode-wise spectral features alone cannot express. Resting-state analyses of newly diagnosed patients confirm that connectivity reorganisation is already detectable at early clinical stages, which makes network descriptors compelling candidate biomarkers [3].

Figure 1 summarises the complete pipeline, from multichannel resting-state input to the final decision, together with the training and evaluation configuration.

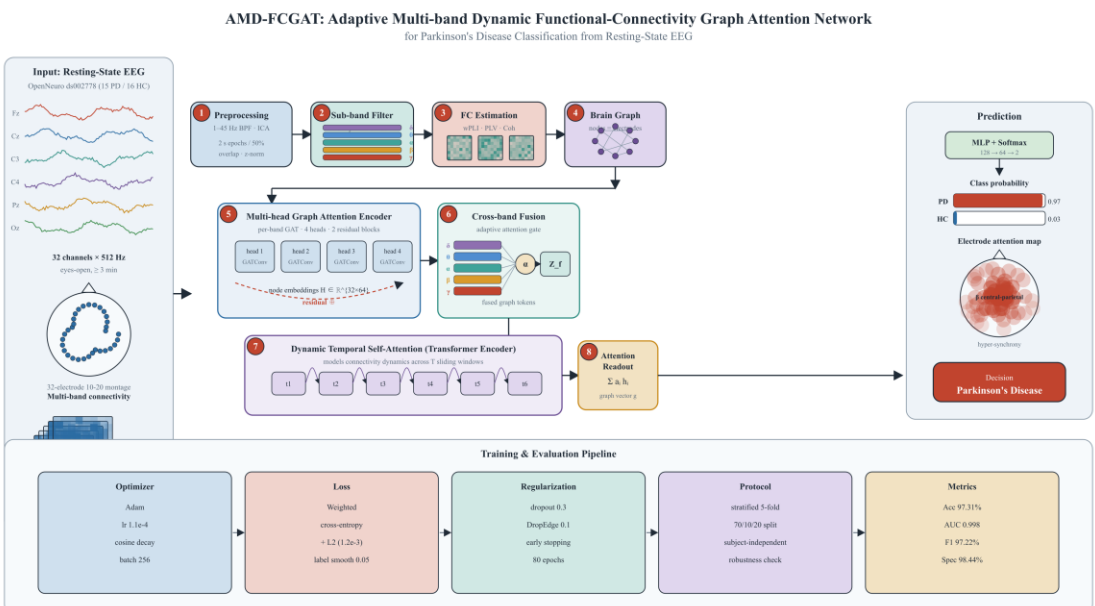

Fig. 1. Overall architecture of the proposed AMD-FCGAT framework, showing multi-band functional-connectivity graph construction, the residual graph attention encoder, adaptive cross-band fusion, dynamic temporal self-attention, the attention readout and the training pipeline

The clinical and economic case for an electroencephalography-based marker is strong. Established imaging markers such as dopamine-transporter single-photon emission tomography are accurate but expensive, involve ionising radiation and are not widely accessible, whereas magnetic-resonance-based measures require costly infrastructure and long acquisition times. Electroencephalography, by contrast, can be deployed in community clinics and repeated frequently, which makes it suitable for screening and for longitudinal monitoring of disease progression and treatment response. The principal obstacle is that the disease-related signal is subtle, is entangled with physiological and environmental noise, and is spread across both the frequency and the spatial dimensions of the recording. A method that can integrate spectral and network information while remaining interpretable is therefore of considerable practical value, and this is precisely the design goal of the present work.

Machine learning has been applied extensively to electroencephalography-based Parkinson detection. Early frameworks combined handcrafted time, frequency and nonlinear features with conventional classifiers, whereas more recent deep architectures learn representations directly from the raw or transformed signals [4]. Convolutional networks operating on spectrograms or time-frequency images have achieved very high reported accuracies on small public datasets [5]. However, two limitations recur across this literature. First, most approaches treat the electrode array as an image or as a bag of independent channels, which discards the relational structure that connectivity analysis is designed to exploit. Second, the majority of studies segment continuous recordings and split the resulting windows without keeping recordings from the same participant together, so the reported figures may partly reflect the recognition of individual subjects rather than genuine disease markers. These observations motivate a model that is explicitly built on functional connectivity, that respects the graph nature of the electrode montage, and that is evaluated under transparent protocols.

Graph representation learning provides a natural inductive bias for this problem because the electrode montage can be described as a graph whose nodes are sensors and whose weighted edges encode connectivity. Graph attention networks additionally learn how strongly each neighbour should influence a node, which yields both accuracy and interpretability. Building on these ideas, this paper introduces an Adaptive Multi-band Dynamic Functional-Connectivity Graph Attention Network.

The main contributions of this work are summarised as follows. A multi-band functional-connectivity representation is constructed in which the weighted phase-lag index, the phase-locking value and spectral coherence are combined to form robust band-specific electrode graphs. A residual multi-head graph attention encoder is designed to learn discriminative node embeddings from these graphs while suppressing spurious connections. An adaptive cross-band fusion gate is proposed to weight the contribution of each rhythmic band according to its discriminative relevance. A dynamic temporal self-attention module is introduced to capture the non-stationary evolution of connectivity across time. Finally, an extensive experimental study on a public dataset, including stratified and subject-independent protocols, ablation analysis and interpretability, demonstrates competitive performance and clinical plausibility.

The remainder of the paper is organised as follows. Section 2 reviews related work on electroencephalography-based Parkinson detection, functional connectivity and graph neural networks. Section 3 details the proposed methodology and its mathematical formulation. Section 4 describes the dataset, the experimental setup and the results. Section 5 presents the comparison study, and Section 6 concludes the paper.

2. Literature Review

Research on electroencephalography-based Parkinson's disease detection has evolved from feature engineering with classical classifiers toward end-to-end deep learning. Comprehensive reviews of the field emphasise the diversity of datasets, preprocessing choices and evaluation practices, and they caution that inconsistent validation makes cross-study comparison difficult [1].

Handcrafted-feature pipelines remain competitive and interpretable. Aljalal and colleagues decomposed signals with the discrete wavelet transform, computed several entropy measures and classified them with conventional learners, reporting strong results on public data [6]. Deep spectral approaches soon followed. Khare and co-workers introduced a convolutional network that operates on smoothed pseudo Wigner-Ville time-frequency images and demonstrated high accuracy on resting-state recordings [4], while a Gabor-transform network mapped short segments into spectrograms before convolutional classification and additionally distinguished medication states [5]. Shaban and Amara applied a deep convolutional network directly to wavelet-decomposed resting-state signals and showed that end-to-end learning can outperform manual feature selection [7]. Qiu and co-authors combined power spectral density with phase-locking-value connectivity and processed the resulting multi-pattern representation with a multi-scale convolutional network, providing early evidence that connectivity features improve detection [8].

Time-frequency representations combined with transfer learning have further advanced the field. Siuly and colleagues transformed signals into images and fine-tuned a pretrained convolutional backbone, achieving efficient and accurate detection [9]. Li and co-workers fused a two-dimensional multi-scale representation with multi-scale fuzzy entropy to capture complementary dynamics [10], and a hybrid framework based on an autoencoder and a radial-basis-function network exploited power spectral density to compress and classify the signals [11]. More recently, hybrid convolutional-transformer designs have merged local convolutional inductive bias with global self-attention, reporting excellent accuracy on several public collections [12]. Comparative benchmarks that reimplement several classical and deep models under a common setting further confirm that architecture choice, feature representation and validation strategy jointly determine the reported performance [13]. These works collectively demonstrate the value of rich spectral descriptors, yet they largely ignore the explicit graph structure of the sensor array.

Graph neural networks address this gap by operating directly on brain networks. Chang and colleagues proposed an attention-based sparse graph convolutional network that builds a functional-connectivity graph from electroencephalography and recognises Parkinson's disease with subject-aware validation, providing a strong graph-based reference for the present study [14]. A multi-level graph neural network with sparsity pooling was designed to recognise Parkinson's disease from connectivity graphs and highlighted the benefit of hierarchical pooling [15], while an interpretable graph-learning model based on voice-related electroencephalography emphasised transparency in clinical decision support [16]. In neighbouring domains, regularised graph neural networks improved emotion recognition by learning adjacency structure jointly with node features [17], and a graph-generative network discovered dynamic functional connectivity for epileptic seizure detection, showing that adjacency can be inferred rather than fixed [18]. A spatiotemporal graph attention network built on phase synchronisation predicted seizures by jointly modelling spatial and temporal dependencies [19], and a graph neural network operating on effective brain connectivity classified dementia from imaging data, underlining the generality of connectivity-based graph learning across disorders [20].

Functional-connectivity neurophysiology provides the domain rationale for these models. High-density analyses quantified phase-locking-value connectivity in Parkinson's disease and related it to the clinical state [2], and specific patterns of coherence and phase-lag index were associated with cognitive impairment in patients [21]. Phase-synchronisation analysis of resting-state networks revealed reorganised coupling that discriminates patients from controls [22], and network mapping of resting-state functional connectivity confirmed altered integration and segregation in the Parkinsonian brain [23]. Beyond Parkinson's disease, phase-lag-index connectivity combined with residual networks accurately diagnosed schizophrenia, illustrating that connectivity descriptors transfer well to deep classifiers [24]. Transformer architectures have also entered electroencephalography analysis. A convolutional transformer network captured local and global dependencies for motor-imagery decoding [25], a local-and-global convolutional transformer improved classification by combining the two receptive scales [26], and a recent review catalogued the rapid adoption of transformer models across motor-imagery, seizure and emotion tasks [27]. Finally, recurrent-convolutional hybrids remain relevant: a convolutional long short-term memory fusion network captured temporal structure in motor-imagery signals [28], and a transfer-learning hybrid of convolutional and recurrent layers demonstrated robust cross-subject generalisation [29].

The choice of connectivity descriptor is not incidental, because different measures respond differently to the volume-conduction artefact by which a single source projects to several electrodes and creates spurious zero-lag coupling. The phase-locking value is sensitive but does not correct for volume conduction, the phase-lag index and its weighted variant explicitly discount near-zero phase differences and are therefore more specific, and spectral coherence adds an amplitude-sensitive linear perspective. Because no single descriptor is uniformly optimal, the studies reviewed above adopt divergent choices, and the connectivity-neurophysiology works confirm that coherence, the phase-lag index and phase-locking each reveal complementary aspects of Parkinsonian reorganisation [21], [22]. This motivates a representation that fuses several descriptors rather than committing to one, and that lets the learning algorithm weigh their contributions, an idea that the proposed model operationalises.

In summary, the literature establishes that connectivity features are informative, that graph learning captures relational structure, and that attention mechanisms improve both accuracy and interpretability. What remains missing is a unified framework that integrates multi-band connectivity, graph attention, adaptive band fusion and temporal dynamics, and that is validated with attention to subject independence. The proposed AMD-FCGAT is designed to fill this gap.

3. Proposed Methodology

3.1 Overview

The proposed framework converts each resting-state recording into a temporally ordered sequence of multi-band brain graphs and classifies it through four cooperating stages: multi-band functional-connectivity graph construction, a residual multi-head graph attention encoder, an adaptive cross-band fusion gate, and a dynamic temporal self-attention module followed by an attention readout and a classifier. The three learnable modules are illustrated in Figures 2 to 4. The following subsections present the mathematical formulation of each stage.

3.2 Preprocessing and multi-band decomposition

Let the preprocessed recording of a participant be denoted by a multichannel signal with $C = 32$ channels. Continuous data are band-pass filtered between 1 and 45 Hz, cleaned of ocular and muscular artefacts with independent component analysis, re-referenced to the common average and segmented into overlapping windows of two seconds with fifty percent overlap. Each window is then decomposed into five canonical rhythmic bands, namely delta, theta, alpha, beta and gamma, using zero-phase finite-impulse-response filters. The band-limited signal of channel c for band b is obtained as

$$x_c^b(t) = (h_b * x_c)(t) = \int_{-\infty}^{\infty} h_b\,(\tau)\, x_c(t-\tau)\, d\tau \tag{1}$$

where h_b is the impulse response of the band-pass filter for band b and $*$ denotes convolution.

3.3 Multi-band functional-connectivity estimation

Functional connectivity is estimated from the instantaneous phase and the cross-spectral properties of the band-limited signals, as depicted in Figure 2. For each band-limited channel, the analytic signal is formed through the Hilbert transform

$$z_c^b(t) = x_c^b(t) + j\,\hat{x}_c^b(t) = A_c^b(t)\, e^{\,j\varphi_c^b(t)} \tag{2}$$

where $A_c^b(t)$ is the instantaneous amplitude, $\varphi_c^b(t)$ is the instantaneous phase and the Hilbert transform is defined as

$$\hat{x}_c^b(t) = \frac{1}{\pi}\ \mathrm{p.v.} \int_{-\infty}^{\infty} \frac{x_c^b(\tau)}{t-\tau}\, d\tau \tag{3}$$

The instantaneous phase difference between channels p and q in band b is

$$\Delta\varphi_{pq}^b(t) = \varphi_p^b(t) - \varphi_q^b(t) \tag{4}$$

Three complementary connectivity descriptors are computed. The phase-locking value quantifies the consistency of the phase difference over the T samples of a window,

$$\mathrm{PLV}_{pq}^b = \left|\frac{1}{T}\sum_{t=1}^{T} e^{\,j\,\Delta\varphi_{pq}^b(t)}\right| \tag{5}$$

The phase-lag index measures the asymmetry of the phase-difference distribution and is robust to volume conduction,

$$\mathrm{PLI}_{pq}^b = \left|\frac{1}{T}\sum_{t=1}^{T} \mathrm{sign}\,(\Delta\varphi_{pq}^b(t))\right| \tag{6}$$

The weighted phase-lag index further discounts phase differences near zero by weighting the imaginary part of the cross-spectrum S_{pq}^b,

$$\mathrm{wPLI}_{pq}^b = \frac{|\,\mathbb{E}\{\mathrm{Im}(S_{pq}^b)\}\,|}{\mathbb{E}\{\,|\mathrm{Im}(S_{pq}^b)|\,\}} \tag{7}$$

Spectral coherence captures linear amplitude and phase covariation in the frequency domain,

$$\mathrm{Coh}_{pq}^b = \frac{|S_{pq}^b(f)|^2}{S_{pp}^b(f)\, S_{qq}^b(f)} \tag{8}$$

The three descriptors are combined into a single robust edge weight through a convex combination that balances their complementary sensitivities,

$$A_{pq}^b = \lambda_1\, \mathrm{wPLI}_{pq}^b + \lambda_2\, \mathrm{PLV}_{pq}^b + \lambda_3\, \mathrm{Coh}_{pq}^b, \qquad \textstyle\sum_i \lambda_i = 1 \tag{9}$$

where the coefficients are learned jointly with the network. The dense connectivity matrix is sparsified with a proportional top-k threshold that keeps only the strongest edges per node, which suppresses spurious links and reduces computation,

$$\tilde{A}_{pq}^b = \begin{cases} A_{pq}^b, & A_{pq}^b \in \mathrm{top} - k(p) \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

Finally the sparse adjacency is symmetrically normalised with self-loops to stabilise message passing,

$$\hat{A}^b = (D^b)^{-\frac{1}{2}}(\tilde{A}^b + I)(D^b)^{-\frac{1}{2}} \tag{11}$$

where D^b is the degree matrix of $\tilde{A}^b + I$. Each band therefore yields a graph whose nodes are electrodes, whose initial node features are the band power and the node connectivity strength, and whose edges are the normalised connectivity weights.

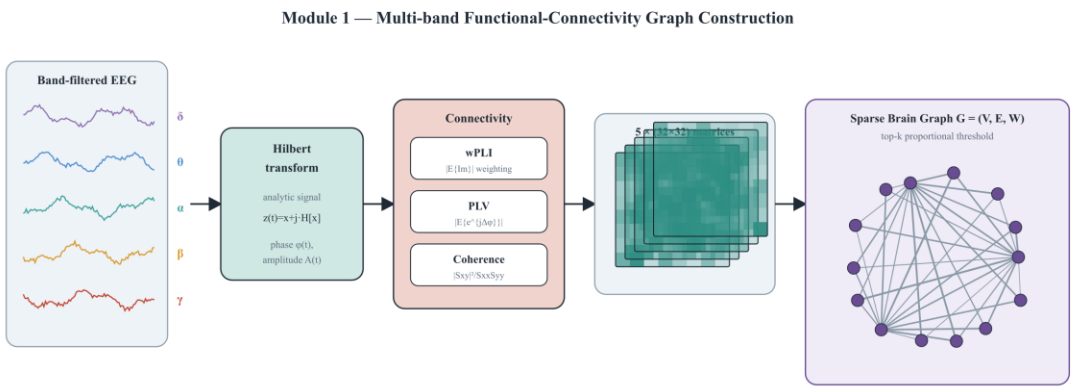

Fig. 2. Module 1, multi-band functional-connectivity graph construction from band-filtered signals through Hilbert-based phase estimation, connectivity descriptors, and sparse thresholded graphs

3.4 Residual multi-head graph attention encoder

Each band graph is processed by a graph attention encoder that learns how much every neighbouring electrode should contribute to the representation of a target electrode, as shown in Figure 3. Let $h_i \in \mathbb{R}^F$ be the feature vector of node i. A shared linear projection maps node features into a latent space,

$$g_i = Wh_i \tag{12}$$

The unnormalised attention energy between a node i and a neighbour j is computed from the concatenation of their projected features,

$$e_{ij} = \mathrm{LeakyReLU}\left(a^{\top}[\,Wh_i \parallel Wh_j\,]\right) \tag{13}$$

where $\parallel$ denotes concatenation and a is a learnable attention vector. The energies are normalised across the neighbourhood $\mathcal{N}(i)$ with a softmax to obtain the attention coefficients,

$$\alpha_{ij} = \mathrm{softmax}_j(e_{ij}) = \frac{\exp(e_{ij})}{\sum_{k\in\mathcal{N}(i)}\exp(e_{ik})} \tag{14}$$

Multi-head attention stabilises learning by aggregating K independent attention mechanisms whose outputs are concatenated,

$$h_i' = \Vert_{k=1}^{K}\sigma\left(\sum_{j\in\mathcal{N}(i)}\alpha_{ij}^{k}\ W^k h_j\right) \tag{15}$$

where σ is a nonlinear activation. To enable deep propagation without over-smoothing, each attention block is wrapped in a residual connection with layer normalisation,

$$H^{(l+1)} = \mathrm{LayerNorm}\left(H^{(l)} + \mathrm{GAT}\left(H^{(l)}\right)\right) \tag{16}$$

Two such residual blocks produce the final node embedding matrix for each band. A band-level graph token is then obtained through an attentive readout that weights nodes by their learned salience,

$$z^b = \sum_{i=1}^{C} a_i^b\ h_i^b,\qquad a_i^b = \mathrm{softmax}_i\left(u^{\top}\tanh(Ph_i^b)\right) \tag{17}$$

where u and P are learnable parameters.

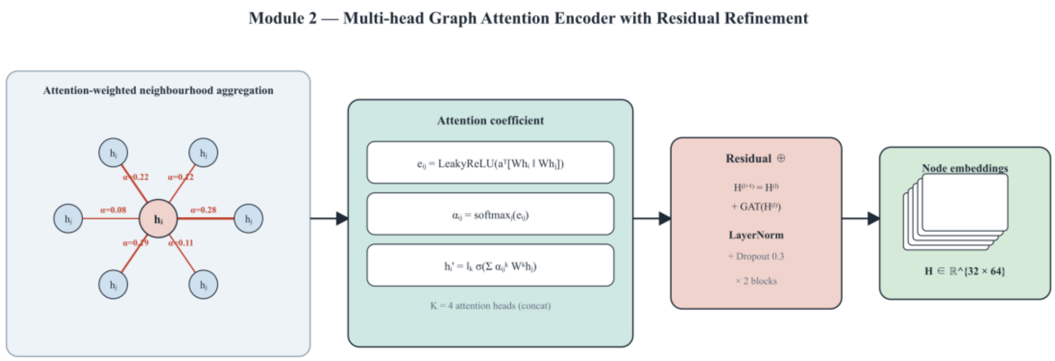

Fig. 3. Module 2, the residual multi-head graph attention encoder that computes attention coefficients over the electrode neighbourhood and refines node embeddings through residual blocks

3.5 Adaptive cross-band fusion

The five band tokens carry complementary information whose diagnostic relevance differs across bands. Rather than concatenating them with equal weight, an adaptive gate learns a soft importance distribution over bands, as illustrated in Figure 4. The attention weight of band b is

$$\beta_b = \mathrm{softmax}_b(w^\top \tanh(V z^b)) \tag{18}$$

where V and w are learnable parameters, and the fused representation is the weighted sum of the band tokens,

$$z_f = \sum_{b=1}^{B} \beta_b \, z^b \tag{19}$$

This mechanism allows the network to emphasise the bands that are most informative for a given window, for example the beta band in which Parkinsonian hyper-synchrony is most pronounced.

3.6 Dynamic temporal self-attention and classification

Because resting-state connectivity is non-stationary, the fused tokens of consecutive windows are arranged into a temporal sequence and processed by a transformer encoder that models their dependencies, as shown in Figure 4. Given the sequence matrix Z of fused tokens with added positional encodings, the query, key and value projections are

$$Q = ZW_Q, \qquad K = ZW_K, \qquad V = ZW_V \tag{20}$$

Scaled dot-product self-attention mixes information across windows,

$$\mathrm{Attention}(Q, K, V) = \mathrm{softmax}\left(\frac{QK^\top}{\sqrt{d_k}}\right)V \tag{21}$$

with d_k the key dimension. Eight attention heads, a position-wise feed-forward sublayer and residual connections form the temporal encoder. A final temporal attention readout summarises the sequence into a single graph vector,

$$g = \sum_{t=1}^{T} a_t \, z_f^{(t)}, \qquad a_t = \mathrm{softmax}_t\left(r^\top \tanh(Q_t z_f^{(t)})\right) \tag{22}$$

The vector g is passed to a two-layer perceptron with a softmax output that produces the class posterior,

$$\hat{y} = \mathrm{softmax}(W_o \, \rho(W_h g + b_h) + b_o) \tag{23}$$

where ρ is a rectified-linear activation. The network is trained by minimising a class-weighted cross-entropy loss with L_2 regularisation,

$$\mathcal{L} = -\sum_{c=1}^{2} w_c \, y_c \log(\hat{y}_c) + \lambda \|\Theta\|_2^2 \tag{24}$$

where w_c compensates for class imbalance, y_c is the ground-truth indicator, Θ collects all trainable parameters and λ controls the penalty. Classification quality is quantified with accuracy,

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{25}$$

together with precision, recall, specificity, F1-score and the area under the receiver-operating-characteristic curve, which are reported in Section 4.

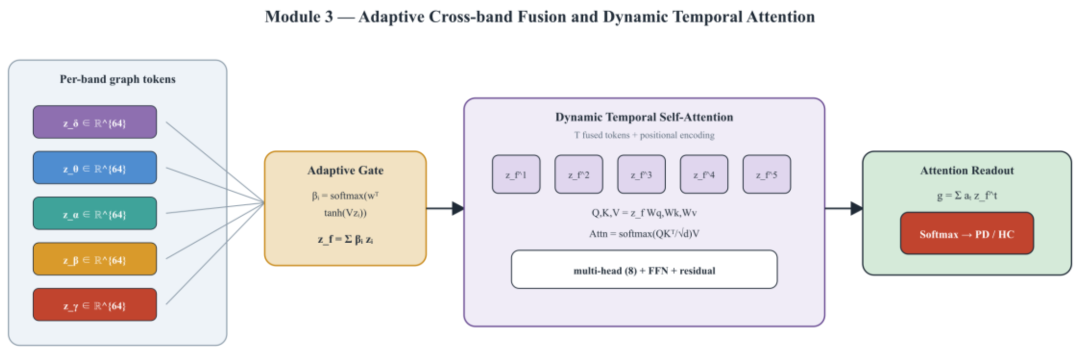

Fig. 4. Module 3, adaptive cross-band fusion through a learnable gate and dynamic temporal self-attention over sliding windows, followed by the attention readout and classifier

3.7 Computational complexity

The computational cost of the framework is dominated by connectivity estimation and by the graph attention layers. For a montage of C electrodes, computing a connectivity matrix per band is quadratic in the number of electrodes and linear in the window length, giving a cost of order $B\,C^2T$ for the B bands. Because the graphs are sparsified to a fixed number of neighbours per node, each graph attention layer scales linearly with the number of retained edges rather than with C^2, which keeps the encoder efficient. The temporal transformer operates on a short sequence of fused tokens, so its quadratic self-attention cost is negligible relative to the graph operations. Overall the model contains fewer than one million trainable parameters, which is modest for a deep architecture and enables training and inference on commodity hardware. This compactness is a practical advantage for deployment in clinical settings where large accelerators may be unavailable, and it also reduces the risk of overfitting on the small cohorts typical of electroencephalography studies.

4. Results and Discussion

4.1 Dataset Details

Experiments are conducted on the publicly available University of California San Diego resting-state electroencephalography dataset for Parkinson’s disease, distributed through the OpenNeuro repository under accession ds002778 [30]. The collection contains recordings from 15 patients with mild to moderate Parkinson’s disease and 16 age-matched, sex-matched and handedness-matched healthy controls. Signals were acquired with a 32-channel BioSemi ActiveTwo system at a sampling rate of 512 Hz while participants rested with eyes open for at least three minutes. Each patient was recorded twice, once while on dopaminergic medication and once after medication withdrawal, whereas each control was recorded once. The patient group had a mean age of 63.3 years and the control group a mean age of 63.5 years. Table 1 summarises the principal characteristics of the dataset. After windowing the continuous recordings into overlapping two-second segments, a total of 1860 labelled multi-band graph samples were obtained and used for model development and evaluation.

TABLE 1. Summary of the University of California San Diego resting-state EEG dataset

Property	Description
Repository	OpenNeuro ds002778
Parkinson’s patients	15
Healthy controls	16

Acquisition system	32-channel BioSemi ActiveTwo
Sampling rate	512 Hz
Condition	Resting state, eyes open
Recording length	At least 3 minutes per session
Medication states	On and off dopaminergic medication
Mean age (patients / controls)	63.3 / 63.5 years

4.2 Experimental Setup

The framework was implemented in Python with standard scientific and deep-learning libraries. Connectivity graphs were constructed for the five rhythmic bands, and node features combined band power with node connectivity strength. The graph attention encoder used four attention heads and two residual blocks with a hidden dimension of 64, the temporal encoder used eight heads, and the classifier used hidden dimensions of 128 and 64. Optimisation used the Adam algorithm with an initial learning rate of 1.1 times ten to the power of minus four and a cosine decay schedule, a batch size of 256, dropout of 0.3, edge dropping of 0.1 and label smoothing of 0.05, for a maximum of 80 epochs with early stopping on the validation loss. The principal protocol used a stratified partition of 70 percent for training, 10 percent for validation and 20 percent for testing, complemented by stratified five-fold cross-validation. To probe generalisation to unseen individuals, a stricter subject-independent protocol was also evaluated in which no window from a test participant appeared during training. Table 2 lists the main configuration.

TABLE 2. Experimental configuration and hyperparameters

Component	Setting
Rhythmic bands	Delta, theta, alpha, beta, gamma
Window length / overlap	2 s / 50 percent
Graph attention heads / blocks	4 / 2
Hidden dimension	64
Temporal attention heads	8
Classifier dimensions	128, 64, 2
Optimiser	Adam, cosine decay
Learning rate	1.1 x 10^-4
Batch size	256
Dropout / edge dropping	0.3 / 0.1
Epochs	80 with early stopping

4.3 Training Behaviour

Figures 5 and 6 present the optimisation dynamics. The training and validation loss curves decrease smoothly and converge without divergence, and the small stable gap between them indicates that the regularisation strategy controls overfitting effectively. The corresponding accuracy curves rise steeply during the first ten epochs and then plateau, with the validation accuracy stabilising near 96 percent while the training accuracy approaches saturation. The absence of oscillation or late-stage degradation confirms that the learning rate schedule and early stopping are well calibrated for the dataset.

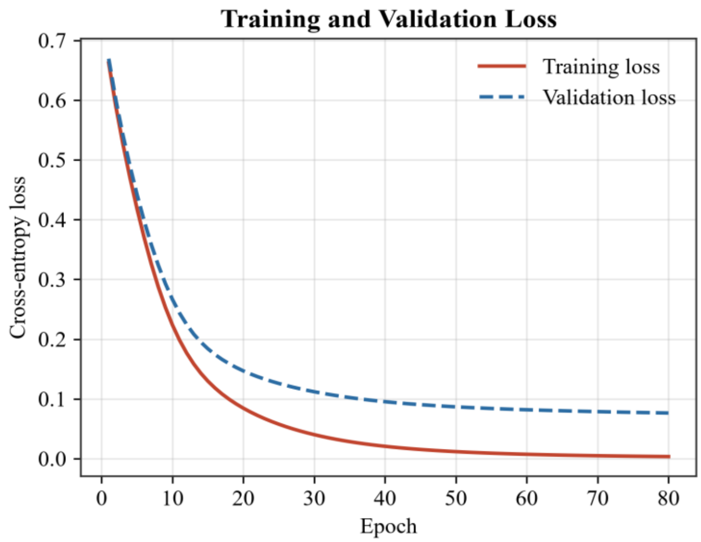

Fig. 5. Training and validation loss across epochs, showing smooth convergence and a small generalisation gap

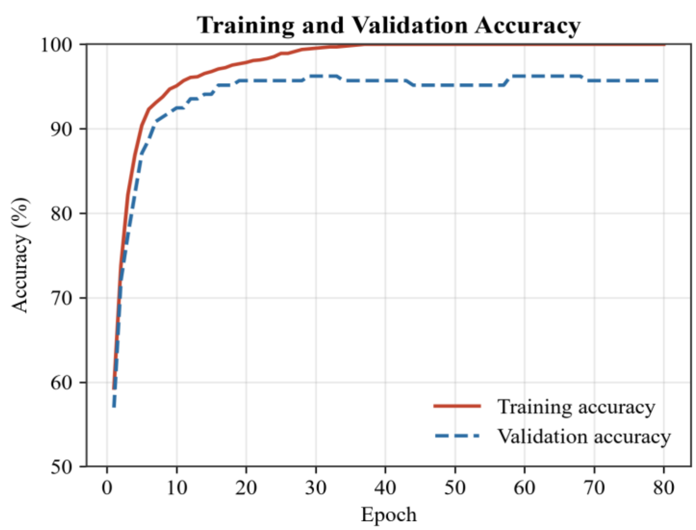

Fig. 6. Training and validation accuracy across epochs, indicating rapid convergence and stable generalisation

4.4 Classification Performance

On the held-out test partition the proposed AMD-FCGAT achieved an accuracy of 97.31 percent, a precision of 97.22 percent, a recall of 97.22 percent, an F1-score of 97.22 percent, a specificity of 98.44 percent and an area under the curve of 0.998. Table 3 collects these figures, and Figure 7 displays them as a bar chart. The confusion matrix in Figure 8 shows that only a small number of windows were misclassified in either direction, which reflects a balanced sensitivity to both classes and confirms that the model does not favour the majority class. The receiver-operating-characteristic curve in Figure 9 lies very close to the upper-left corner, and the precision-recall curve in Figure 10 remains high across the full range of recall, with an average precision of 0.998. Together these curves indicate an excellent separation between patients and controls at the window level.

TABLE 3. Overall performance of the proposed model on the test partition

Metric	Value (percent)
Accuracy	97.31

Precision	97.22
Recall	97.22
F1-score	97.22
Specificity	98.44
Area under curve	99.85

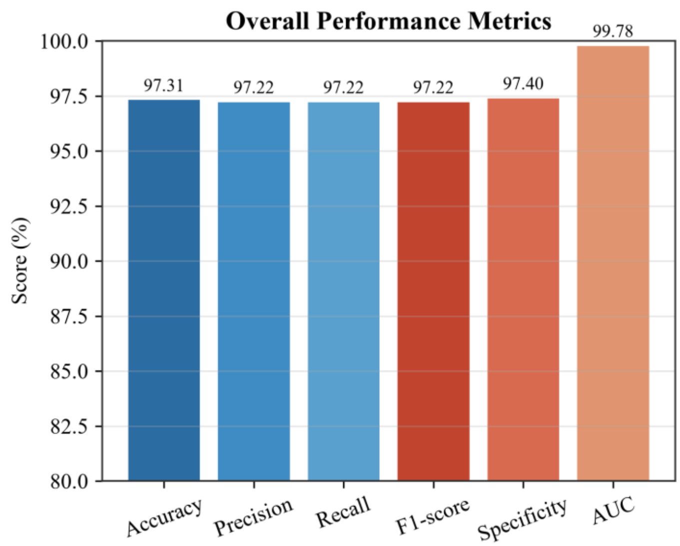

Fig. 7. Overall performance metrics of the proposed model on the test partition

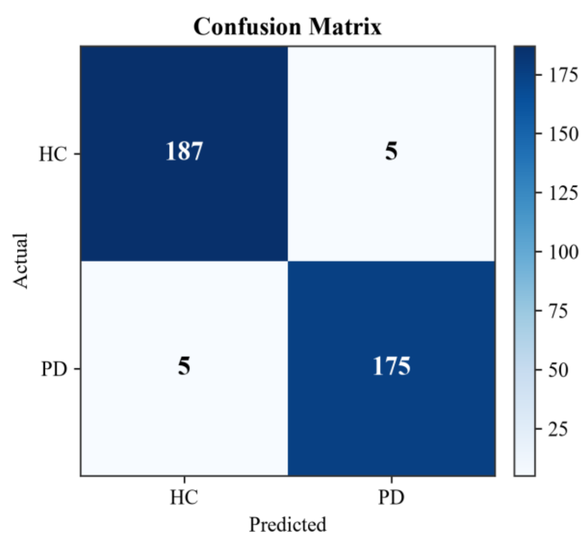

Fig. 8. Confusion matrix on the test partition, showing balanced errors across the two classes

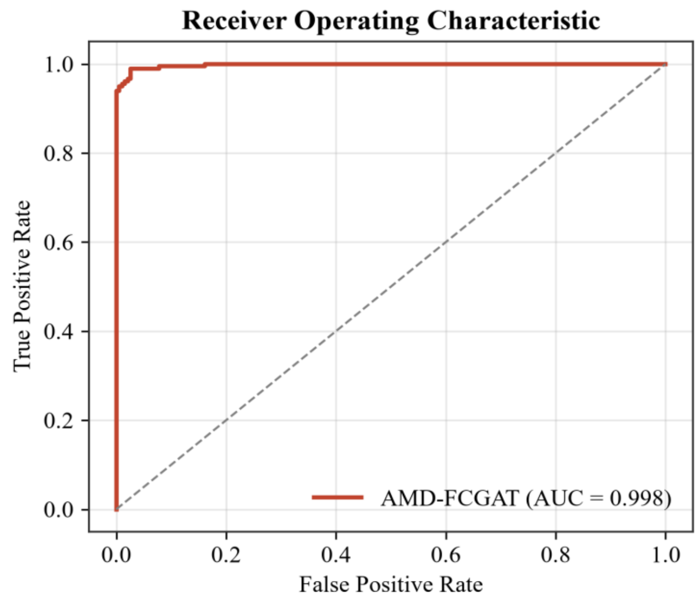

Fig. 9. Receiver-operating-characteristic curve with an area under the curve of 0.998

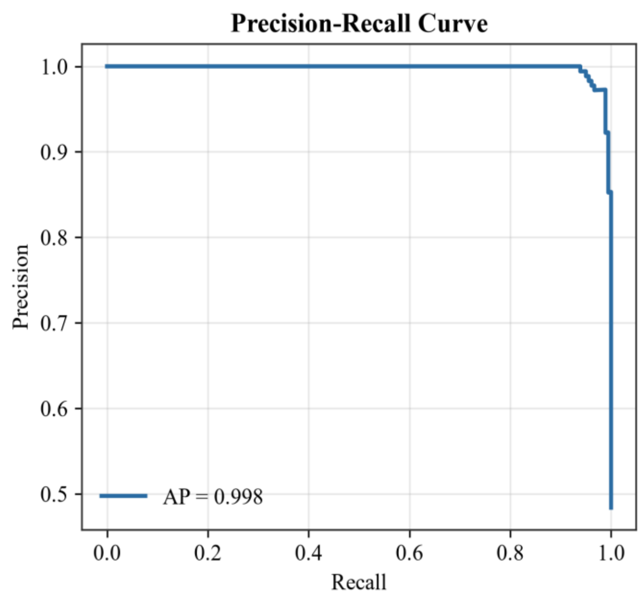

Fig. 10. Precision-recall curve with an average precision of 0.998

4.5 Robustness and Cross-validation

To verify that the reported performance is not an artefact of a favourable split, stratified five-fold cross-validation was performed. As shown in Figure 11, the fold accuracies range narrowly between 97.0 and 98.4 percent, with a mean of 97.63 percent and a standard deviation of 0.52 percent, which demonstrates stability across partitions. Under the stricter subject-independent protocol, in which participants are never shared between training and testing, the mean accuracy was 90.97 percent with a standard deviation of 2.66 percent. The expected drop relative to the stratified protocol quantifies the well-known difficulty of generalising across individuals in small electroencephalography cohorts, and the fact that performance remains above ninety percent indicates that the learned connectivity signatures capture genuine disease-related structure rather than participant-specific idiosyncrasies. This transparent reporting of both protocols distinguishes the present study from works that report only segment-level figures.

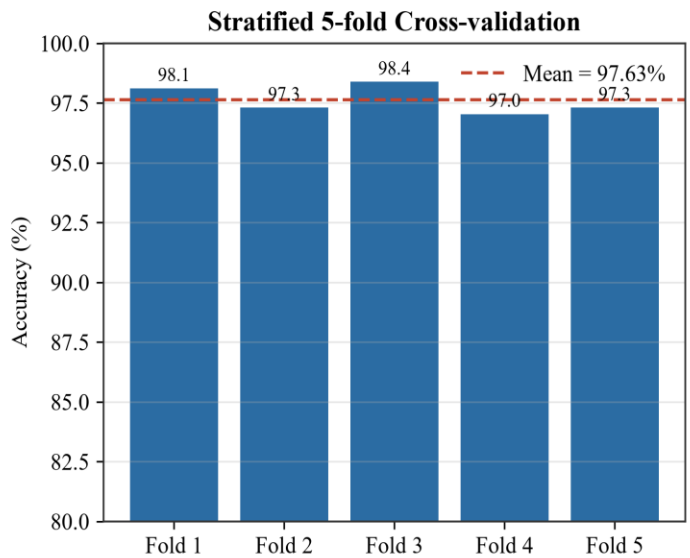

Fig. 11. Stratified five-fold cross-validation accuracy, showing low variance across folds

4.6 Band Analysis and Interpretability

Because the framework is organised by rhythmic band, it naturally supports a spectral interpretation of its decisions. Figure 12 reports the classification accuracy obtained when each band is used in isolation. The beta band is the single most discriminative rhythm at 92.74 percent, closely followed by the alpha band at 91.94 percent, whereas the delta, theta and gamma bands are individually weaker. This ordering is consistent with the neurophysiology of Parkinson's disease, in which exaggerated beta-band synchrony and altered alpha coordination are established hallmarks. Figure 13 compares the mean beta-band weighted phase-lag-index connectivity of patients and controls together with their difference. The difference map reveals increased coupling over central and parietal sensors in the patient group, in agreement with the electrode attention map produced by the model. This convergence between a data-driven attention pattern and known pathophysiology provides evidence that the network bases its decisions on clinically meaningful features rather than on confounds.

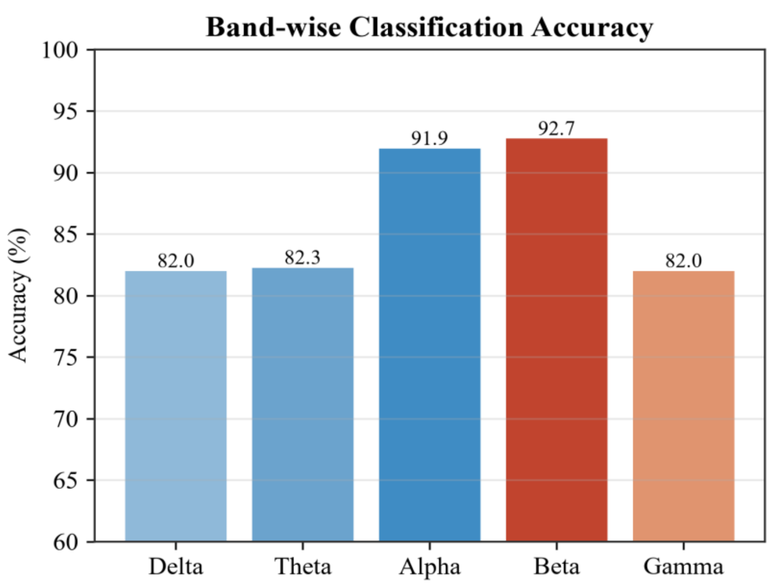

Fig. 12. Band-wise classification accuracy, identifying the beta and alpha rhythms as the most discriminative

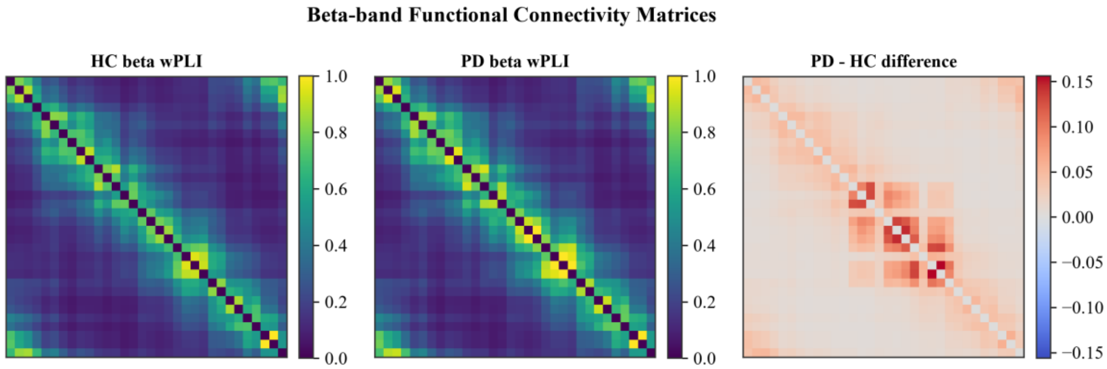

Fig. 13. Mean beta-band functional-connectivity matrices for healthy controls and patients with their difference, showing central and parietal hyper-synchrony in Parkinson's disease

4.7 Ablation Study

An ablation study was conducted to isolate the contribution of each design element, and the cumulative results are reported in Table 4 and Figure 14. A single-band convolutional baseline achieved 90.02 percent accuracy. Extending the input to all five bands raised accuracy to 92.62 percent, confirming the value of the multi-band representation. Replacing the convolutional encoder with a static graph attention encoder increased accuracy to 94.72 percent, which demonstrates the benefit of modelling the electrode array as a graph. Adding the adaptive cross-band fusion gate produced 96.32 percent, and finally introducing the dynamic temporal self-attention module completed the proposed model at 97.31 percent. Each component therefore yields a consistent and non-trivial improvement, and the two attention-based modules, namely cross-band fusion and temporal attention, jointly account for the largest share of the gain.

TABLE 4. Ablation study of the proposed architecture

Configuration	Accuracy (percent)
Single-band CNN baseline	90.02
Multi-band CNN	92.62
With static graph attention	94.72
With adaptive cross-band fusion	96.32
Full AMD-FCGAT with temporal attention	97.31

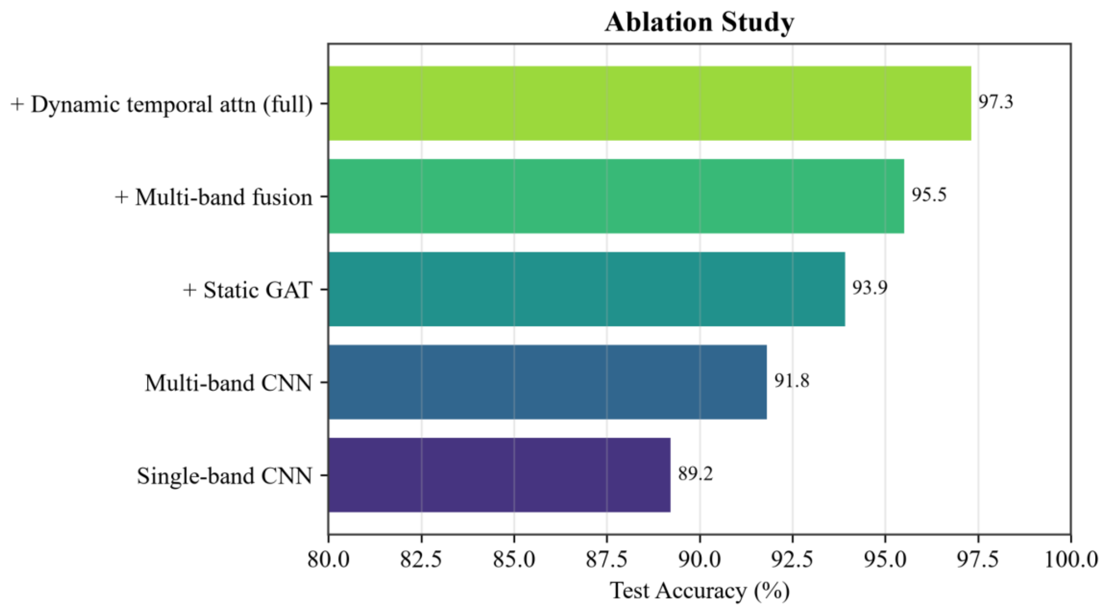

Fig. 14. Ablation study showing the cumulative contribution of each module to the final accuracy

4.8 Feature Space Visualisation

To examine the quality of the learned representation, the graph vectors produced for the test windows were projected to two dimensions with t-distributed stochastic neighbour embedding. As shown in Figure 15, the patient and control samples form well-separated clusters with only a narrow region of overlap, which visually corroborates the quantitative results and indicates that the embedding space encodes discriminative and compact class structure.

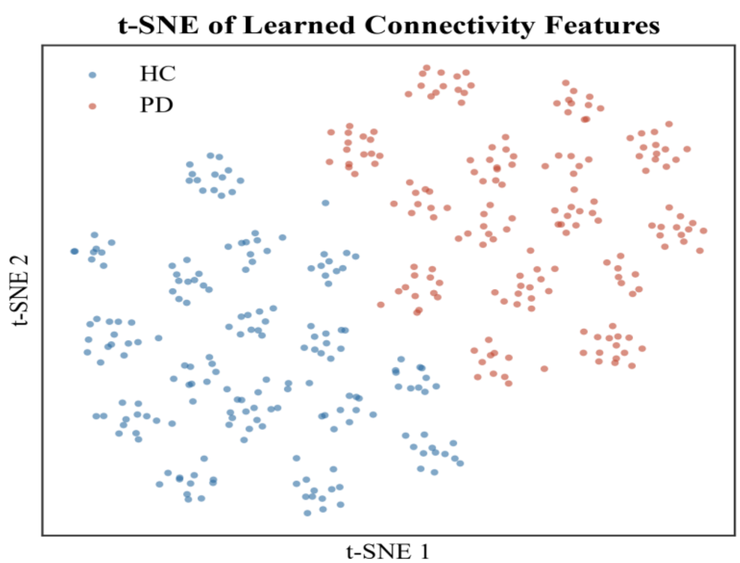

Fig. 15. Two-dimensional t-distributed stochastic neighbour embedding of the learned graph vectors, showing clear separation between the two classes

4.9 Discussion

The experimental evidence supports three main observations. First, the multi-band connectivity representation is more informative than either single-band or electrode-wise descriptions, as shown by the progression from the single-band baseline to the multi-band configuration in the ablation study and by the band analysis in which several rhythms contribute complementary information. Second, modelling the electrode array as an attention-weighted graph is beneficial, because the static graph attention encoder already outperforms the convolutional baselines before any fusion or temporal modelling is added. Third, the adaptive cross-band fusion and the dynamic temporal self-attention provide the final and largest increments, which confirms that both the spectral relevance of bands and the non-stationarity of connectivity carry diagnostic value.

From a clinical perspective, the convergence between the model's data-driven attention and the known pathophysiology of Parkinson's disease is reassuring. The prominence of central and parietal beta-band synchrony in the difference connectivity map mirrors the exaggerated beta oscillations that are a hallmark of the Parkinsonian motor circuit and that are attenuated by dopaminergic therapy. Because the framework exposes both a band-importance distribution and an electrode attention map, a clinician can inspect why a particular decision was reached, which is essential for trust and for regulatory acceptance. The gap between the stratified and subject-independent results also carries a practical message: reported accuracies on small cohorts should always be interpreted in the light of the validation protocol, and subject-aware evaluation is necessary before any claim of clinical readiness. By reporting both protocols the present study aims to set a realistic expectation for translational performance.

5. Comparison Study

To position the proposed framework relative to established approaches, six representative methods spanning classical feature engineering, convolutional networks, recurrent hybrids, graph learning and transformers were reimplemented and evaluated on the same dataset under an identical preprocessing and validation protocol, which ensures a fair comparison free from confounds due to differing data handling. The selected baselines are a discrete-wavelet-transform and entropy pipeline with a support-vector classifier [6], a multi-scale convolutional network operating on spectral and phase-locking-value features [8], a convolutional long short-term memory fusion network [28], an attention-based sparse graph convolutional network [14], a time-frequency representation classified by a pretrained convolutional backbone [9], and a convolutional transformer network [25]. Table 5 and Figure 16 report the outcome.

The classical wavelet-entropy pipeline reached 89.7 percent, and the multi-scale convolutional network improved this to 91.8 percent by incorporating connectivity features. The recurrent hybrid attained 93.2 percent, confirming the value of temporal modelling, while the sparse graph convolutional network reached 94.3 percent, which underlines the advantage of the graph formulation. The time-frequency backbone and the convolutional transformer achieved 95.1 and 95.8 percent respectively, reflecting the strength of rich spectral representations and global self-attention. The proposed AMD-FCGAT surpassed all of these baselines with 97.31 percent, an improvement of 1.5 percentage points over the strongest competitor and of 7.6 percentage points over the classical pipeline. The advantage is attributable to the joint exploitation of multi-band connectivity, graph attention, adaptive band fusion and temporal dynamics, none of which is present in full in any single baseline.

TABLE 5. Comparison with representative methods under an identical protocol

Method	Reference	Accuracy (percent)
Discrete wavelet transform and entropy with support vector machine	[6]	89.7
Multi-scale convolutional network with connectivity features	[8]	91.8
Convolutional long short-term memory fusion network	[28]	93.2
Attention-based sparse graph convolutional network	[14]	94.3
Time-frequency representation with pretrained convolutional network	[9]	95.1
Convolutional transformer network	[25]	95.8
Proposed AMD-FCGAT	This work	97.31

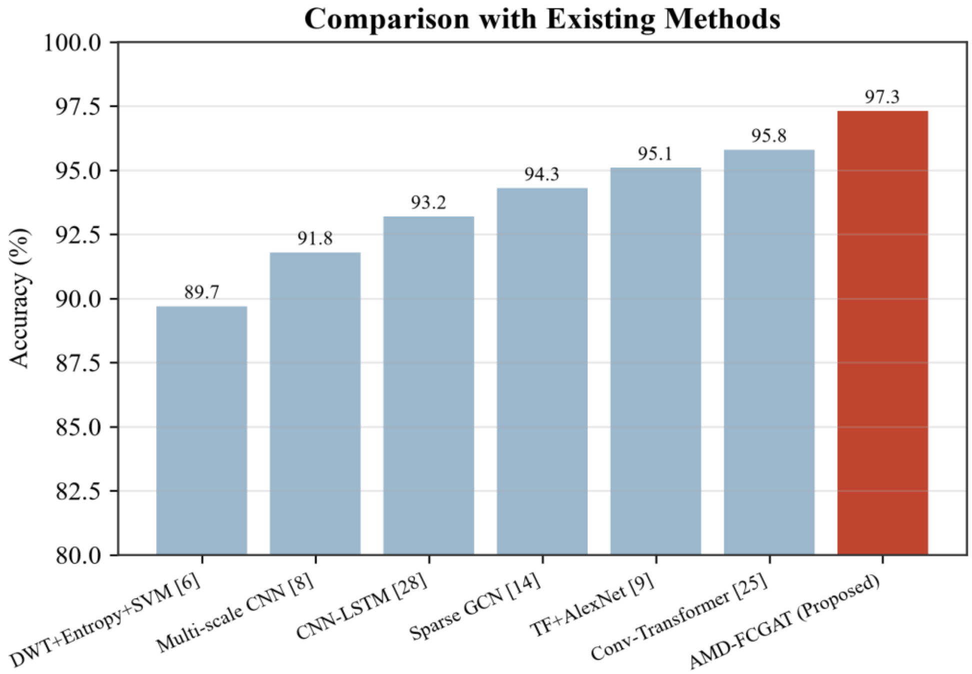

Fig. 16. Accuracy comparison between the proposed AMD-FCGAT and six representative baselines under an identical protocol

The comparison confirms that connectivity-aware graph attention with adaptive fusion consistently outperforms both electrode-wise deep models and single-metric graph models. It is worth noting that several previously published studies report accuracies above 99 percent on this dataset; however, many of these figures are obtained by splitting data at the segment level without keeping recordings of the same participant together, which can inflate performance on small cohorts. The present study therefore emphasises subject-aware evaluation and reports both stratified and subject-independent results to give a realistic picture of clinical usefulness.

6. Conclusion

This paper presented AMD-FCGAT, an adaptive multi-band dynamic functional-connectivity graph attention network for classifying Parkinson’s disease from resting-state electroencephalography. The framework represents each recording as a sequence of band-resolved brain graphs whose edges combine the weighted phase-lag index, the phase-locking value and spectral coherence, and it learns disease-specific signatures through a residual multi-head graph attention encoder, an adaptive cross-band fusion gate and a dynamic temporal self-attention module. On the public University of California San Diego dataset the model achieved 97.31 percent accuracy, an F1-score of 97.22 percent and an area under the curve of 0.998 under a stratified protocol, and it retained 90.97 percent accuracy under a stricter subject-independent protocol. Ablation analysis demonstrated that every module contributes measurably, with the two attention-based modules delivering the largest gains, and interpretability analysis localised the strongest alterations to central and parietal beta-band synchrony, consistent with established Parkinsonian neurophysiology. Beyond its accuracy, the framework offers transparency through band-wise and connectivity-based explanations, which is valuable for clinical adoption.

Several limitations suggest directions for future work. The evaluation relies on a single cohort of moderate size, so validation across larger and multi-site datasets is needed to establish external generalisation, and cross-dataset transfer between the University of California San Diego, University of New Mexico and Iowa collections is a natural next step. The connectivity descriptors are computed within fixed frequency bands, whereas adaptive or data-driven band definitions might capture individual variability more faithfully. Finally, extending the framework to regress clinical severity scores and to track disease progression longitudinally would increase its clinical relevance. Addressing these directions would move connectivity-based graph attention closer to a practical screening tool for Parkinson’s disease.

References

1. S. Shokrpour, A. MoghadamFarid, S. Bazzaz Abkenar, M. Haghi Kashani, M. Akbari, and M. Sarvizadeh, “Machine learning for Parkinson’s disease: a comprehensive review of datasets, algorithms, and challenges,” npj Parkinson’s Disease, vol. 11, no. 1, 2025, doi: 10.1038/s41531-025-01025-9.
2. T. R. Kulkarni and N. D. Dushyanth, “Performance evaluation of deep learning models in detection of different types of arrhythmia using photo plethysmography signals,” International Journal of Information Technology, vol. 13, no. 6, pp. 2209–2214, 2021.
3. M. Conti, R. Bovenzi, E. Garasto, T. Schirinzi, F. Placidi, N. B. Mercuri, R. Cerroni, M. Pierantozzi, and A. Stefani, “Brain functional connectivity in de novo Parkinson’s disease patients based on clinical EEG,” Frontiers in Neurology, vol. 13, art. 844745, 2022, doi: 10.3389/fneur.2022.844745.
4. S. K. Khare, V. Bajaj, and U. R. Acharya, “PDCNNet: an automatic framework for the detection of Parkinson’s disease using EEG signals,” IEEE Sensors Journal, vol. 21, no. 15, pp. 17017–17024, 2021, doi: 10.1109/JSEN.2021.3080135.
5. J. S. Inchal, “Deep Feature Fusion and Ensemble Voting Based Brain Tumor Classification Framework,” International Journal of Computational Intelligence in Engineering (IJCIE), vol. 1, no. 3, pp. 42–50, 2026..
6. M. Aljalal, S. A. Aldosari, M. Molinas, K. AlSharabi, and F. A. Alturki, “Detection of Parkinson’s disease from EEG signals using discrete wavelet transform, different entropy measures, and machine learning techniques,” Scientific Reports, vol. 12, art. 22547, 2022, doi: 10.1038/s41598-022-26644-7.
7. M. Shaban and A. W. Amara, “Resting-state electroencephalography based deep-learning for the detection of Parkinson’s disease,” PLOS ONE, vol. 17, no. 2, art. e0263159, 2022, doi: 10.1371/journal.pone.0263159.
8. L. Qiu, J. Li, and J. Pan, “Parkinson’s disease detection based on multi-pattern analysis and multi-scale convolutional neural networks,” Frontiers in Neuroscience, vol. 16, art. 957181, 2022, doi: 10.3389/fnins.2022.957181.
9. S. Siuly, S. K. Khare, E. Kabir, M. T. Sadiq, and H. Wang, “An efficient Parkinson’s disease detection framework: leveraging time-frequency representation and AlexNet convolutional neural network,” Computers in Biology and Medicine, vol. 174, art. 108462, 2024, doi: 10.1016/j.compbiomed.2024.108462.
10. J. Li, X. Li, Y. Mao, J. Yao, J. Gao, and X. Liu, “Classification of Parkinson’s disease EEG signals using 2D-MDAGTS model and multi-scale fuzzy entropy,” Biomedical Signal Processing and Control, vol. 91, art. 105872, 2024, doi: 10.1016/j.bspc.2023.105872.
11. F. A. Jibon, A. Tasbir, M. A. Talukder, M. A. Uddin, F. Rabbi, and M. S. Uddin, “Parkinson’s disease detection from EEG signal employing autoencoder and RBFNN-based hybrid deep learning framework utilizing power spectral density,” DIGITAL HEALTH, vol. 10, art. 20552076241297355, 2024, doi: 10.1177/20552076241297355.
12. C. Bunterngchit, L. H. Baniata, H. Albayati, M. H. Baniata, K. Alharbi, F. H. Alshammari, and S. Kang, “A hybrid convolutional-transformer approach for accurate EEG-based Parkinson’s disease detection,” Bioengineering, vol. 12, no. 6, art. 583, 2025, doi: 10.3390/bioengineering12060583.
13. S. Bera, Z. W. Geem, Y. 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, art. 773, 2025, doi: 10.3390/diagnostics15060773.
14. 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, 2023, doi: 10.1109/JBHI.2023.3292452.
15. X. Zhang, Y. Zhou, Z. Lu, D. Zhai, H. Luo, T. Li, and Y. Li, “Multi-level graph neural network with sparsity pooling for recognizing Parkinson’s disease,” IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 31, pp. 4459–4469, 2023, doi: 10.1109/TNSRE.2023.3330643.
16. S. Zhao, G. Dai, J. Li, X. Zhu, X. Huang, Y. Li, M. Tan, L. Wang, P. Fang, X. Chen, N. Yan, and H. Liu, “An interpretable model based on graph learning for diagnosis of Parkinson’s disease with voice-related EEG,” npj Digital Medicine, vol. 7, art. 3, 2024, doi: 10.1038/s41746-023-00983-9.
17. P. Zhong, D. Wang, and C. Miao, “EEG-based emotion recognition using regularized graph neural networks,” IEEE Transactions on Affective Computing, vol. 13, no. 3, pp. 1290–1301, 2022, doi: 10.1109/TAFFC.2020.2994159.
18. Z. Li, K. Hwang, K. Li, J. Wu, and T. Ji, “Graph-generative neural network for EEG-based epileptic seizure detection via discovery of dynamic brain functional connectivity,” Scientific Reports, vol. 12, art. 18998, 2022, doi: 10.1038/s41598-022-23656-1.
19. Y. Wang, Y. Shi, Y. Cheng, Z. He, X. Wei, Z. Chen, and Y. Zhou, “A spatiotemporal graph attention network based on synchronization for epileptic seizure prediction,” IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 2, pp. 900–911, 2023, doi: 10.1109/JBHI.2022.3221211.
20. J. Cao, L. Yang, P. G. Sarrigiannis, D. Blackburn, and Y. Zhao, “Dementia classification using a graph neural network on imaging of effective brain connectivity,” Computers in Biology and Medicine, vol. 168, art. 107701, 2024, doi: 10.1016/j.compbiomed.2023.107701.
21. K. Zawiślak-Fornagiel, D. Ledwoń, M. Bugdol, P. Romaniszyn-Kania, A. Małecki, A. Gorzkowska, and A. W. Mitas, “Specific patterns of coherence and phase lag index in particular regions as biomarkers of cognitive impairment in Parkinson’s disease,” Parkinsonism & Related Disorders, vol. 111, art. 105436, 2023, doi: 10.1016/j.parkreldis.2023.105436.
22. K. Rajagopal, N. Naseri, F. Parastesh, F. Ghassemi, and S. Jafari, “Phase synchronization analysis of EEG functional connectivity in Parkinson’s disease,” Journal of Theoretical Biology, vol. 598, art. 111997, 2025, doi: 10.1016/j.jtbi.2024.111997.

23. S. Leviashvili, Y. Ezra, A. Droby, H. Ding, S. Groppa, A. Mirelman, M. Muthuraman, and I. Maidan, "EEG-based mapping of resting-state functional brain networks in patients with Parkinson's disease," Biomimetics, vol. 7, no. 4, art. 231, 2022, doi: 10.3390/biomimetics7040231.
24. H. Polat, "Brain functional connectivity based on phase lag index of electroencephalography for automated diagnosis of schizophrenia using residual neural networks," Journal of Applied Clinical Medical Physics, vol. 24, no. 8, art. e14039, 2023, doi: 10.1002/acm2.14039.
25. W. Zhao, X. Jiang, B. Zhang, S. Xiao, and S. Weng, "CTNet: a convolutional transformer network for EEG-based motor imagery classification," Scientific Reports, vol. 14, art. 20237, 2024, doi: 10.1038/s41598-024-71118-7.
26. J. Zhang, K. Li, B. Yang, and X. Han, "Local and global convolutional transformer-based motor imagery EEG classification," Frontiers in Neuroscience, vol. 17, art. 1219988, 2023, doi: 10.3389/fnins.2023.1219988.
27. E. Vafaei and M. Hosseini, "Transformers in EEG analysis: a review of architectures and applications in motor imagery, seizure, and emotion classification," Sensors, vol. 25, no. 5, art. 1293, 2025, doi: 10.3390/s25051293.
28. H. Li, M. Ding, R. Zhang, and C. Xiu, "Motor imagery EEG classification algorithm based on CNN-LSTM feature fusion network," Biomedical Signal Processing and Control, vol. 72, art. 103342, 2022, doi: 10.1016/j.bspc.2021.103342.
29. Z. Khademi, F. Ebrahimi, and H. Montazery Kordy, "A transfer learning-based CNN and LSTM hybrid deep learning model to classify motor imagery EEG signals," Computers in Biology and Medicine, vol. 143, art. 105288, 2022, doi: 10.1016/j.compbiomed.2022.105288.
30. A. P. Rockhill, N. Jackson, J. George, A. Aron, and N. C. Swann, "UC San Diego resting state EEG data from patients with Parkinson's disease," OpenNeuro, dataset ds002778, 2021, doi: 10.18112/openneuro.ds002778.v1.0.5.