

Name of Participating Students: Troy William Deng

High School: Concordia International School Shanghai

Province: Shanghai

Country/Region: China

Name of Advisor(s): Dr. Xin Lou

Affiliation of Advisor(s): ShanghaiTech University

Title of Research Report: A Synergistic Framework Integrating
Dynamic Graph Attention Network and Multiscale Entropy for
EEG-Based Alzheimer's Disease Classification

A Synergistic Framework Integrating Dynamic Graph Attention Networks and Multiscale Entropy for EEG-Based Alzheimer's Disease Classification

Troy William Deng

Concordia International School Shanghai troywdeng@gmail.com

Abstract. Accurate identification of Alzheimer's disease (AD) from non-invasive brain signals remains an important challenge. Electroencephalography (EEG) provides a cost-effective approach for monitoring brain activity, yet existing methods often overlook temporal dynamics and rely on simplified handcrafted features that poorly characterize nonlinear neural activity. We propose a Dynamic Graph Attention Network (DGAT) with Multiscale Entropy (MSE) for EEG-based AD classification. MSE quantifies nonlinear EEG dynamics within sliding temporal windows, while phase-locking value (PLV) constructs dynamic functional connectivity graphs, enabling DGAT to jointly model spatial brain network topology and temporal variations. GNNExplainer further identifies brain regions and temporal windows most relevant to model predictions. On an OpenNeuro public dataset, the proposed method achieves 85.7% classification accuracy, outperforming conventional approaches based on static connectivity or power-spectrum features, while ablation studies confirm the contributions of MSE and dynamic graph modeling. Interpretability analysis identifies the frontopolar (Fp2, Fp1) and frontal midline (Fz) regions as the most influential, followed by posterior temporal (T6) and occipital (O1) regions, providing biologically plausible candidates for further investigation. Cross-dataset evaluation demonstrates above-chance zero-shot transfer ($AUC = 0.751$) to an unseen eyes-open photic-stimulation paradigm, while revealing substantial domain shift from eyes-closed resting-state EEG. Fine-tuning improves accuracy from 66.2% to 84.6% and AUC from 0.751 to 0.950, demonstrating the potential of domain adaptation to improve robustness across acquisition conditions. Overall, the findings support the value of combining nonlinear EEG complexity with dynamic graph learning for interpretable AD classification and provide a potentially generalizable framework for other neurological disorders.

Keywords: Alzheimer's Disease · EEG · Dynamic Graph Neural Network · Multiscale Entropy · Brain Functional Connectivity.

Outline

- 1 Introduction
- 2 Related Work
 - 2.1 Traditional Machine Learning Methods
 - 2.2 Deep Learning Methods
 - 2.3 Graph Neural Network Methods
 - 2.4 Entropy Analysis Methods
- 3 Proposed Method
 - 3.1 Problem Overview
 - 3.2 Overall Framework
 - 3.3 Multiscale Entropy Feature Extraction
 - 3.4 Dynamic Functional Connectivity Construction
 - 3.5 Dynamic Graph Attention Network
 - 3.6 Training Strategy
 - 3.7 Interpretability Analysis
- 4 Experiments and Results Analysis
 - 4.1 Experimental Settings
 - 4.2 Comparison with Existing Methods
 - 4.3 GNNExplainer Interpretability Analysis
 - 4.4 Ablation Study
 - 4.5 Cross-Dataset Generalization to OpenNeuro ds006036
 - 4.6 Fine-Tuning on OpenNeuro ds006036
- 5 Discussion
 - 5.1 Methodological and Scientific Significance
 - 5.2 Limitations
 - 5.3 Future Work
- 6 Acknowledgments
 - 6.1 Research Background and Topic Source
 - 6.2 Division of Labor
 - 6.3 Detailed Research Process
 - 6.4 AI Usage Statement
 - 6.5 External Assistance Statement
 - 6.6 Ethics and Compliance Statement
- References

1 Introduction

Alzheimer’s disease (AD) is a progressive neurodegenerative disorder characterized by gradual cognitive decline and irreversible impairment of brain function. Traditional diagnosis mainly relies on clinical symptom assessment and neuropsychological testing, which often suffers from strong subjectivity and limited sensitivity for early-stage detection. In recent years, electroencephalography (EEG) has attracted increasing attention for AD early-stage classification due to its non-invasive nature, high temporal resolution, and low acquisition cost [4]. However, effectively modeling the nonlinear dynamics and time-varying functional connectivity of EEG signals remains a significant challenge.

Existing EEG/fMRI-based AD classification methods can generally be categorized into traditional machine learning, deep learning, graph neural networks (GNNs), and entropy analysis methods. Traditional methods (e.g., support vector machines (SVMs)), rely on handcrafted features and have limited representation capability [1, 2]. Deep learning methods (e.g., convolutional neural networks (CNNs)) can automatically learn features but often ignore the intrinsic topological structure of brain functional networks [5–7]. GNN-based methods introduce graph-based representation learning for brain connectivity analysis and have shown promising performance [8, 11, 16], but most existing GNN approaches are built on static brain graphs and fail to capture the temporal evolution of functional connectivity networks. On the other hand, entropy-based methods, such as multiscale entropy (MSE), can effectively characterize the complexity and irregularity of neural signals [3, 9, 13], but their discriminative ability is limited when used alone.

To address these limitations, this paper proposes a novel Dynamic Graph Attention Network with Multiscale Entropy (DGAT-MSE) framework for EEG-based AD classification. Specifically: MSE is first used to extract nonlinear complexity features from EEG signals across multiple temporal scales [13]; meanwhile, dynamic functional connectivity graphs are constructed using phase-locking value (PLV), where nodes represent EEG channels and edges characterize inter-channel synchronization [10, 14]; then, a DGAT is introduced to jointly learn spatial topological dependencies and temporal evolution patterns [12]; finally, GNNExplainer is incorporated to improve model interpretability by identifying important brain regions and critical time windows associated with AD [15].

Experimental results demonstrate that our framework achieves superior performance compared with several baseline methods, validating the effectiveness of integrating nonlinear complexity analysis, dynamic functional connectivity modeling, and graph attention learning for EEG-based AD classification. Recent studies on brain functional connectivity and graph-theoretical analysis further support the feasibility of exploiting dynamic neural interaction patterns for AD diagnosis [17–19].

2 Related Work

2.1 Traditional Machine Learning Methods

Traditional machine learning methods have been widely applied in AD early classification. Researchers typically extract handcrafted features from neuroimaging or electrophysiological data and subsequently perform classification using conventional machine learning algorithms. For example, Wee et al. [1] constructed brain functional connectivity networks and used SVMs for AD classification; Rabeh et al. [2] utilized SVMs and decision trees to distinguish AD, mild cognitive impairment (MCI), and healthy controls. In EEG-based studies, Ieracitano et al. [4] projected EEG signals into the time-frequency domain and combined spectral representations with logistic regression for dementia classification.

These methods have relatively low computational complexity and strong interpretability, but their performance heavily depends on manually designed features and prior expert knowledge. Consequently, they often struggle to capture the highly nonlinear and complex spatiotemporal patterns of brain activity.

2.2 Deep Learning Methods

Deep learning methods can automatically learn discriminative representations from raw data and have shown remarkable potential in AD early classification. For instance, Li and Liu [5] proposed a multiple-cluster DenseNet architecture that effectively captures complementary multi-scale features from MRI data, improving the discriminative power and robustness of AD classification compared with conventional CNN approaches. Farooq et al. [6] designed deep CNN architectures for multi-class AD diagnosis using MRI, demonstrating the effectiveness of deep feature learning.

Although deep learning methods have strong feature extraction capabilities, most of them process neuroimaging data in Euclidean space and ignore the intrinsic graph topology of brain functional networks. Therefore, they are insufficient for modeling complex interactions among brain regions and EEG channels.

2.3 Graph Neural Network Methods

GNN methods can directly model the topological structure of brain functional networks and have attracted widespread attention in AD research. Zhang et al. [8] utilized graph convolutional networks (GCNs) to classify brain functional connectivity networks and achieved competitive results. Recent studies have shown that abnormal functional connectivity and neural synchronization are important pathological characteristics of AD. Smailovic et al. [21] demonstrated that EEG functional connectivity disruptions are strongly associated with regional brain glucose metabolism abnormalities in AD and amyloid-positive MCI patients. These findings support the importance of graph-based brain network analysis for AD diagnosis.

However, most existing GNN-based approaches rely on static brain graphs and cannot capture the dynamic temporal evolution of functional connectivity. Moreover, the interpretability of GNNs remains a significant challenge in medical applications. Inspired by GNNExplainer research [20], incorporating dynamic graph modeling and explainability mechanisms into GNN frameworks is highly important for reliable AD classification.

2.4 Entropy Analysis Methods

Entropy analysis methods evaluate brain functional states by measuring the complexity and irregularity of physiological signals. Costa et al. [13] proposed MSE analysis to characterize signal complexity across multiple temporal scales. Furthermore, Grieder et al. [9] applied MSE to resting-state fMRI data and found that AD patients exhibit significantly reduced brain functional complexity compared with healthy controls. Entropy methods can effectively characterize nonlinear dynamic properties of brain activity and provide complementary information to functional connectivity analysis.

Nevertheless, entropy analysis methods are usually used alone and lack the ability to model spatial interactions and temporal dependencies within brain functional networks. Therefore, combining entropy analysis with dynamic graph representation learning may provide a more comprehensive framework for AD early-stage classification.

3 Proposed Method

3.1 Problem Overview

Traditional graph structure models focus on pairwise relationships between two nodes. This modeling approach has significant limitations: it can only capture simple pairwise correlations, ignoring the complexity and multi-dimensionality of real-world relationship networks. In recent years, graph networks have gradually been applied in the medical field, but many challenges remain to be addressed promptly. Once the graph topology is constructed, it remains fixed throughout the entire model training cycle, lacking the flexibility for dynamic adjustment. This static graph structure cannot flexibly adapt to dynamic changes in the data or the evolution of internal data structures, nor can it capture the temporal evolution characteristics of brain networks during cognitive tasks. Can a dynamic spatiotemporal graph neural network that integrates multiscale EEG complexity and time-varying functional connectivity improve the classification of Alzheimer’s disease from cognitively normal controls, and what spatiotemporal EEG patterns contribute to its predictions?

3.2 Overall Framework

Our proposed DGAT-MSE framework models the nonlinear, spatial, and temporal characteristics of EEG signals in a unified manner. The overall pipeline is

illustrated in Figure 1. Given preprocessed EEG signals, we first extract nonlinear node features using MSE. Next, dynamic functional connectivity graphs are constructed based on PLV, where nodes represent EEG channels and edges encode inter-channel synchronization. Finally, the resulting graph sequence is fed into a DGAT to jointly learn spatial topology and temporal dynamics for classification.

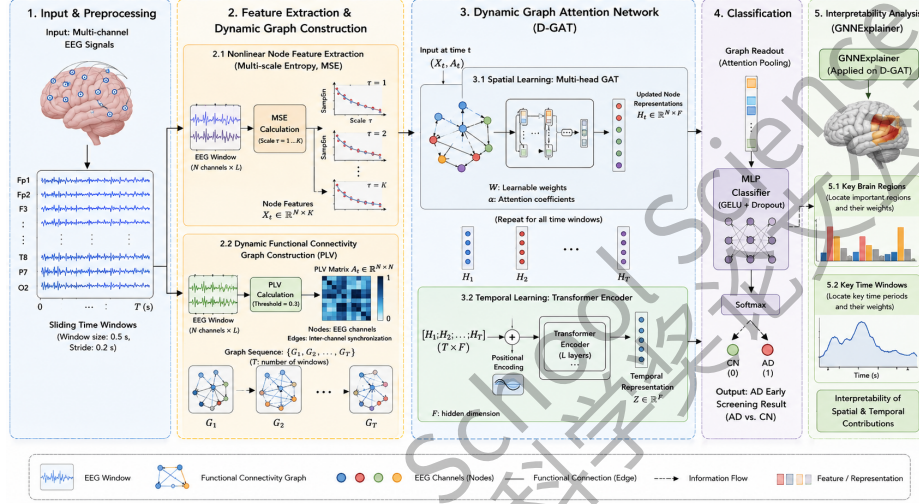


Fig. 1. Proposed DGAT-MSE Architecture

3.3 Multiscale Entropy Feature Extraction

To characterize the nonlinear dynamics of EEG signals, we employ MSE as node features. For a time series $\{x_i\}_{i=1}^L$ of length L , MSE is computed via coarse-graining followed by sample entropy estimation.

For a scale factor τ , the coarse-grained time series is defined as:

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, \quad j = 1, 2, \dots, \left\lfloor \frac{L}{\tau} \right\rfloor. \quad (1)$$

Sample entropy is defined as:

$$\text{SampEn}(m, r) = -\ln \frac{A}{B}, \quad (2)$$

where $m = 2$ is the embedding dimension, $r = 0.15 \times \text{std}(x)$ is the tolerance, and A and B denote the number of matched patterns of length $m + 1$ and m , respectively.

MSE is obtained across multiple scales:

$$\text{MSE} = \{\text{SampEn}(y^{(\tau)}) \mid \tau = 1, \dots, K\}. \quad (3)$$

To capture temporal variations, a sliding window strategy is adopted with a window length of 0.5 s and a step size of 0.1 s. For each time window t , node features are represented as:

$$\mathbf{X}_t \in \mathbb{R}^{N \times K}, \quad (4)$$

where N is the number of EEG channels and K is the number of scales.

3.4 Dynamic Functional Connectivity Construction

To model interactions between EEG channels, we construct dynamic functional connectivity using PLV. After obtaining instantaneous phases $\phi_i(t)$ and $\phi_j(t)$ via the Hilbert transform, PLV is defined as:

$$\text{PLV}_{ij} = \left| \frac{1}{T} \sum_{k=1}^T e^{i(\phi_i(t_k) - \phi_j(t_k))} \right|, \quad (5)$$

where T is the number of windows. The adjacency matrix $\mathbf{A}_t$ is then constructed as:

$$\mathbf{A}_t(i, j) = \begin{cases} \text{PLV}_{ij}, & \text{if } \text{PLV}_{ij} > \theta \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

where θ is a predefined threshold and set to 0.3 in this study [21, 22]. This value ensures that only moderately strong phase synchronization is retained as functional connections while filtering out spurious low-level coupling.

Thus, we obtain a sequence of dynamic graphs $\{\mathcal{G}_t = (\mathbf{X}_t, \mathbf{A}_t)\}_{t=1}^T$.

3.5 Dynamic Graph Attention Network

The constructed graph sequence is fed into the DGAT model to learn spatiotemporal representations.

Spatial Encoding. For each graph $\mathcal{G}_t$, a GAT is applied to capture spatial dependencies. GAT uses an attention mechanism to compute an unnormalized attention score e_{ij} between node i and node j :

$$e_{ij} = \text{LeakyReLU}(\mathbf{a}^\top [\mathbf{W}\mathbf{h}_i \parallel \mathbf{W}\mathbf{h}_j]), \quad (7)$$

where $e_{ij} \in \mathbb{R}$, $\mathbf{h}_i, \mathbf{h}_j \in \mathbb{R}^d$ are the input feature vectors of node i and its neighbor j (with d being the feature dimension), $\mathbf{W} \in \mathbb{R}^{d' \times d}$ is a learnable shared linear transformation matrix projecting each node feature from dimension d to d' , $\parallel$ denotes concatenation of the two projected vectors, $\mathbf{a} \in \mathbb{R}^{2d'}$ is a learnable attention weight vector, and $\text{LeakyReLU}(\cdot)$ is the leaky rectified linear unit with negative slope 0.2.

The scores are normalized across all neighbors of node i via softmax to obtain attention coefficients α_{ij} :

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k \in \mathcal{N}_i} \exp(e_{ik})}, \quad (8)$$

where $\alpha_{ij} \in [0, 1]$ is the normalized attention coefficient representing the relative importance of neighbor j to node i (satisfying $\sum_{j \in \mathcal{N}_i} \alpha_{ij} = 1$), $\exp(\cdot)$ is the exponential function, and $\mathcal{N}_i$ denotes the set of first-order neighbors of node i (including i itself via a self-loop).

The updated representation $\mathbf{h}'_i$ of node i is then computed as a weighted combination of the projected neighbor features:

$$\mathbf{h}'_i = \sigma \left(\sum_{j \in \mathcal{N}_i} \alpha_{ij} \mathbf{W} \mathbf{h}_j \right), \quad (9)$$

where $\mathbf{h}'_i \in \mathbb{R}^{d'}$ is the output feature vector of node i after one GAT layer (the prime symbol denotes the updated representation, equivalent to $\mathbf{h}_i^{(l+1)}$ at layer $l+1$), $\sigma(\cdot)$ is the ELU (Exponential Linear Unit) nonlinear activation function applied element-wise, α_{ij} serves as the attention weight for neighbor j , and the summation runs over all neighbors in $\mathcal{N}_i$.

After the GAT layer, we stack the updated feature vectors of all N nodes at each time step t to form the spatial feature matrix:

$$\mathbf{H}_t = \begin{bmatrix} \mathbf{h}'_{t,1}^\top \\ \mathbf{h}'_{t,2}^\top \\ \vdots \\ \mathbf{h}'_{t,N}^\top \end{bmatrix} \in \mathbb{R}^{N \times d'}, \quad (10)$$

where $\mathbf{H}_t$ is the spatial feature matrix of graph snapshot $\mathcal{G}_t$, and $\mathbf{h}'_{t,i}$ is the GAT-updated feature vector of node i at time step t . Stacking multiple GAT layers yields the final per-snapshot spatial encoding $\mathbf{H}_t^{\text{GAT}}$, producing the sequence $\{\mathbf{H}_t^{\text{GAT}}\}_{t=1}^T$.

Temporal Encoding After aggregating node features at each time step into a global vector, we feed the sequence into a Transformer encoder to capture temporal dependencies. Mean pooling at each time step is applied to obtain a compact global representation: $\mathbf{z}_t = \frac{1}{N} \sum_{i=1}^N \mathbf{h}_{t,i}^{\text{GAT}} \in \mathbb{R}^{d_{\text{out}}}$, which forms the temporal sequence $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_T] \in \mathbb{R}^{T \times d_{\text{out}}}$. The core operation is the scaled dot-product self-attention:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax} \left(\frac{\mathbf{Q} \mathbf{K}^\top}{\sqrt{d_k}} \right) \mathbf{V}, \quad (11)$$

where $\mathbf{Q} \in \mathbb{R}^{T \times d_k}$ is the query matrix, $\mathbf{K} \in \mathbb{R}^{T \times d_k}$ is the key matrix, $\mathbf{V} \in \mathbb{R}^{T \times d_k}$ is the value matrix, d_k is the dimensionality of the query/key vectors,

$\mathbf{Q}\mathbf{K}^\top \in \mathbb{R}^{T \times T}$ computes pairwise similarity between all time steps, the scaling factor $\sqrt{d_k}$ prevents large dot-product values that would push softmax into regions of vanishing gradients, and $\text{softmax}(\cdot)$ normalizes each row to a probability distribution over the T positions. In self-attention, $\mathbf{Q}$, $\mathbf{K}$, and $\mathbf{V}$ are all linear projections of the same input sequence $\mathbf{Z}$. Stacking L Transformer layers, the output of the final layer is $\mathbf{Z}^{(L)} \in \mathbb{R}^{T \times d_{\text{out}}}$. The final spatiotemporal representation $\mathbf{h}$ is extracted by reading out the last time step: $\mathbf{h} = \mathbf{Z}_T^{(L)} \in \mathbb{R}^{d_{\text{out}}}$, which aggregates information from the entire temporal sequence through the self-attention mechanism.

Classification The learned spatiotemporal representation $\mathbf{h}$ is passed through a linear layer and softmax function to obtain predicted probabilities for each class:

$$\hat{\mathbf{p}} = \text{softmax}(\mathbf{W}_o \mathbf{h} + \mathbf{b}_o), \quad (12)$$

where $\hat{\mathbf{p}} \in [0, 1]^C$ denotes the predicted probability vector over C classes. The predicted class label is obtained as:

$$\hat{y} = \arg \max_c \hat{p}_c. \quad (13)$$

3.6 Training Strategy

The proposed DGAT-MSE model is trained in a supervised manner using cross-entropy loss for AD classification. The cross-entropy loss function is adopted as:

$$\mathcal{L} = - \sum_{c=1}^C y_c \log(\hat{y}_c), \quad (14)$$

where y_c and $\hat{y}_c$ denote the ground truth label and predicted probability for class c , respectively.

Model optimization is performed using the Adam optimizer with an initial learning rate of 0.001. A batch size of 16 is employed during training, where each sample corresponds to a graph sequence derived from one subject.

To improve generalization and prevent overfitting, we use early stopping, in which training is stopped if the validation loss does not decrease for 15 consecutive epochs. Moreover, a data augmentation strategy is applied by using a sliding window with a step size of 0.2 seconds. This effectively increases the number of training samples and enhances the robustness of the model. Dropout regularization is also utilized to mitigate over-fitting.

3.7 Interpretability Analysis

To enhance model transparency and clinical interpretability, we employ GNNExplainer to identify the most informative brain regions and temporal segments contributing to the model's predictions.

Given a trained DGAT model and an input graph sequence $\{\mathcal{G}_t\}_{t=1}^T$, GNNExplainer aims to learn a compact subgraph and a subset of node features that are most influential for a specific prediction. This is achieved by optimizing a mask over the adjacency matrix and node features.

Specifically, let $\mathbf{A}_t$ and $\mathbf{X}_t$ denote the adjacency matrix and node features at time step t , respectively. GNNExplainer learns a mask $\mathbf{M}_A$ over edges and a mask $\mathbf{M}_X$ over node features by maximizing the mutual information between the model prediction $\hat{y}$ and the masked input:

$$\max_{\mathbf{M}_A, \mathbf{M}_X} I(\hat{y}, (\mathbf{A}_t \odot \mathbf{M}_A, \mathbf{X}_t \odot \mathbf{M}_X)), \quad (15)$$

where $\odot$ denotes element-wise multiplication.

In practice, this objective is approximated by minimizing the following loss:

$$\mathcal{L}_{\text{expl}} = -\log P_\theta(y \mid \mathbf{A}_t \odot \mathbf{M}_A, \mathbf{X}_t \odot \mathbf{M}_X) + \lambda_1 \|\mathbf{M}_A\|_1 + \lambda_2 \|\mathbf{M}_X\|_1, \quad (16)$$

where λ_1 and λ_2 control the sparsity of the edge and feature masks.

By applying GNNExplainer to each time window, we obtain importance scores for both spatial connections and temporal segments. The edge mask highlights critical functional connections between EEG channels, while the feature mask identifies the most informative MSE scales.

This enables the identification of key brain regions and time periods associated with cognitive processing, providing neurophysiological insights into AD-related alterations and improving the interpretability of the proposed model.

4 Experiments and Results Analysis

4.1 Experimental Settings

Dataset The experiments were conducted on two complementary and publicly available EEG datasets obtained from the OpenNeuro repository. The first, ds004504 [23], contains resting-state scalp EEG recorded with the eyes closed; the second, ds006036 [24], contains recordings of the same participants acquired under an eyes-open condition with photic stimulation (photomark). Both datasets include 88 participants, divided into three clinical groups: Alzheimer’s disease (AD, $n=36$), cognitively normal controls (CN, $n=29$), and frontotemporal dementia (FTD, $n=23$). For this study, we restricted the analysis to the AD and CN groups, excluding the 23 FTD participants to maintain a specific focus on AD classification; this yielded a final cohort of 65 participants (36 AD and 29 CN) in each dataset.

Dataset Partitioning To ensure the reliability of the model evaluation and avoid data leakage, the datasets were partitioned as follows: The 65 participants from the OpenNeuro dataset (36 AD and 29 CN) were split into a training set and a testing set with a ratio of 8 : 2, and 5-fold cross-validation was further performed on the training set. Critically, both the initial partitioning and cross-validation were conducted at the subject level, ensuring that all EEG segments from a single participant were assigned exclusively to the same subset or fold.

Evaluation Metric To quantitatively evaluate the classification performance of the proposed framework for AD classification, several standard metrics are employed, including Accuracy, Precision, Recall, F1-score, and the Area Under the Receiver Operating Characteristic curve (AUC-ROC).

Let TP , TN , FP , and FN represent the number of True Positives, True Negatives, False Positives, and False Negatives, respectively. The specific definitions of these metrics are as follows:

1. **Accuracy:** The proportion of correctly predicted instances (both AD and CN) among the total number of cases.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (17)$$

2. **Precision:** The ratio of correctly predicted positive observations to the total predicted positives, reflecting the reliability of the AD detection.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (18)$$

3. **Recall (Sensitivity):** The ratio of correctly predicted positive observations to all observations in the actual class, which is crucial in clinical classification to avoid missing potential patients.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (19)$$

4. **F1-score:** The harmonic mean of Precision and Recall, providing a balanced measure especially when the class distribution is uneven.

$$F1\text{-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (20)$$

5. **AUC-ROC:** The Area Under the Receiver Operating Characteristic curve. The ROC curve is generated by plotting the True Positive Rate (TPR) against the False Positive Rate (FPR) at various threshold settings. The AUC value provides an aggregate measure of performance across all possible classification thresholds, where a value closer to 1.0 indicates superior discriminative capability.

4.2 Comparison with Existing Methods

To evaluate the effectiveness of the proposed DGAT-MSE framework, we compare its performance with several representative methods for AD classification based on the OpenNeuro dataset ds004504. Table 1 summarizes the comparison results in terms of accuracy, precision, recall, F1-score, and AUC-ROC.

As shown in Table 1, the proposed DGAT-MSE framework achieves the best overall performance among all compared methods. Specifically, our method attains the highest classification accuracy of 85.71%, outperforming Dynamic

Table 1. Performance comparison with existing AD-EEG classification methods

Method	Acc.	Prec.	Recall	F1	AUC
DGAT-MSE (Ours)	0.8571	1.0000	0.8571	0.9231	0.9048
Dynamic Graph + Power Spectrum	0.7692	0.7000	1.0000	0.8235	0.7619
Dynamic GCN	0.6923	1.0000	0.4286	0.6000	0.6429
Static Graph + Power Spectrum	0.6718	0.8000	0.5714	0.6667	0.6190
CNN-LSTM	0.5385	0.5385	1.0000	0.7000	0.5238

Graph + Power Spectrum, Dynamic GCN, and CNN-LSTM by 8.79%, 16.48%, and 31.86%, respectively. Moreover, DGAT-MSE achieves the highest F1-score (0.9231) and AUC-ROC (0.9048), indicating superior classification consistency and discriminative capability.

Compared with traditional static graph-based methods, the proposed framework demonstrates the advantage of modeling dynamic functional connectivity. Static graph representations fail to capture temporal variations in neural interactions, resulting in relatively lower classification performance. By constructing time-varying functional connectivity graphs, DGAT-MSE is able to better characterize dynamic brain activity patterns associated with AD.

Compared with CNN-LSTM, graph-based approaches achieve significantly better performance, highlighting the importance of explicitly modeling brain network topology. Although CNN-LSTM achieves high recall, its low precision indicates a high false-positive rate and poor classification stability. In contrast, the proposed method achieves perfect precision while maintaining high recall, demonstrating a more balanced and reliable classification performance.

Furthermore, compared with Dynamic GCN, the proposed DGAT-MSE framework achieves substantial improvements in accuracy, F1-score, and AUC-ROC. This improvement can be attributed to two key factors. First, MSE effectively captures nonlinear temporal complexity in EEG signals across multiple scales. Second, the graph attention mechanism adaptively assigns importance weights to functional connections, enabling the model to focus on disease-relevant brain regions and discriminative neural interactions.

Overall, these results demonstrate that the integration of multiscale nonlinear feature extraction and dynamic graph attention learning provides a robust and effective framework for EEG-based AD classification.

4.3 GNNExplainer Interpretability Analysis

To elucidate the decision-making mechanism of the DGAT-MSE model, we employed GNNExplainer to quantify the contribution of each EEG channel and temporal window to the classification outcome. The explainer was applied to representative AD and CN subjects, generating node-level importance masks and temporal importance profiles.

Brain Region Importance We conducted interpretability analysis on all AD and cognitively normal (CN) subjects in the test set. We employed GNNExplainer [15] to quantify the contribution of each EEG channel to the model's

classification decision on a per-window basis. For each subject, the explainer was applied to 30 consecutive time windows, yielding a channel-wise importance vector per window. These vectors were then averaged across windows to obtain a subject-level importance profile over the 19-channel montage, as well as a window-level importance score reflecting the temporal informativeness of each segment. The resulting channel importance distributions were compared between the AD and CN groups to identify neurophysiological divergences captured by the model. A summary of the group-level findings is presented in Table 2, followed by detailed per-subject analyses in Figure 2, Figure 3 and Figure 4.

Channel-level Importance As shown in Table 2 (Panel A), the AD and CN subjects exhibit markedly different channel importance profiles. For AD, the top three contributing channels are Fp2 (0.351), Fp1 (0.343), and Fz (0.336), all located in the prefrontal region. This strong frontal dominance suggests that the model relies primarily on prefrontal EEG patterns to identify AD-related abnormalities, which aligns with established evidence of prefrontal cortical dysfunction in Alzheimer’s disease. In contrast, the CN subject displays a mixed temporal-frontal profile, with T6 (0.378), F4 (0.376), and T5 (0.364) as the top three channels. The emergence of temporal lobe electrodes (T6, T5) as the most informative channels in the healthy brain is consistent with the known role of temporal regions in normal resting-state EEG dynamics. Notably, F4, the least important channel in the AD subject (rank 19, 0.277), rises to rank 2 in the CN subject (0.376), representing the most prominent divergence between the two profiles ($\Delta = 0.099$, Panel E).

Lobe-level Analysis Aggregating channel importance by cortical lobe (Panel C) reveals a systematic shift in the spatial distribution of informative activity. The CN subject yields higher mean importance than the AD subject across all five lobes, with the largest gaps observed in the temporal ($\Delta = 0.042$), parietal ($\Delta = 0.037$), and occipital ($\Delta = 0.040$) regions. This global reduction in importance for the AD subject suggests that pathological EEG signals are not confined to a single region but reflect a diffuse degradation of cortical information, consistent with the widespread nature of AD-related neurodegeneration. At the lobe level, the frontal lobe ranked first in both groups, with a mean importance of 0.327 in AD and 0.359 in CN. This result confirms that the frontal lobe serves as the dominant and most discriminative brain region for EEG-based AD classification, consistent with its established role as the primary locus of AD-related neurodegeneration and compensatory recruitment.

Temporal Dynamics The window-level statistics (Panel D) reveal that the mean window importance is lower for AD (0.307) than for CN (0.342), indicating that individual time segments in AD contribute less discriminative evidence on average. However, both subjects exhibit comparable temporal variability (standard deviation of 0.062 vs. 0.061), suggesting that fluctuation in signal informativeness over time is a shared characteristic rather than a disease-specific effect. The similar temporal variance also rules out a simple signal-to-noise interpretation, as the AD signal is not merely noisier but rather differs in its spatial-channel composition (Panels A–C).

Table 2. Channel-wise importance analysis via GNNExplainer for representative AD and CN subjects. Importance scores are averaged over 30 consecutive time windows per subject. Higher values indicate stronger contribution to the model’s classification decision.

Panel A: Top-5 contributing channels				
Rank	AD Channel Importance		CN Channel Importance	
1	Fp2	0.351	T6	0.378
2	Fp1	0.343	F4	0.376
3	Fz	0.336	T5	0.364
4	T6	0.336	Fp1	0.358
5	O1	0.331	Fp2	0.358
Panel B: Bottom-5 contributing channels				
Rank	AD Channel Importance		CN Channel Importance	
1	F4	0.277	T4	0.280
2	T3	0.279	Cz	0.311
3	C3	0.279	C3	0.318
4	Cz	0.282	C4	0.320
5	T4	0.283	T3	0.321
Panel C: Lobe-level mean importance				
Lobe	AD		CN	
Frontal (Fp1, Fp2, F3, F4, Fz)	0.327		0.359	
Central (C3, C4, Cz)	0.296		0.316	
Temporal (F7, F8, T3–T6)	0.296		0.338	
Parietal (P3, P4, Pz)	0.306		0.343	
Occipital (O1, O2)	0.313		0.353	
Panel D: Window-level statistics				
Metric	AD		CN	
Mean window importance	0.307		0.342	
Range	[0.241, 0.399]		[0.262, 0.404]	
Standard deviation	0.062		0.061	
Panel E: Largest cross-group differences (Δ)				
Channel	AD	CN	Δ	Direction
F4	0.277	0.376	0.099	CN $\uparrow$
T5	0.288	0.364	0.076	CN $\uparrow$
O2	0.296	0.355	0.059	CN $\uparrow$
F7	0.292	0.343	0.050	CN $\uparrow$
T6	0.336	0.378	0.043	CN $\uparrow$

Clinical Interpretation The observation that all top-difference channels (Panel E: F4, T5, O2, F7, T6) show higher importance in CN than AD implies that the model leverages normal functional patterns in CN that are attenuated or distorted in AD. This is consistent with the “disconnection hypothesis” of Alzheimer’s disease, wherein pathological processes disrupt the coordinated activity of distributed cortical networks. The shift from temporal-frontal balance in CN to prefrontal dominance in AD (visualized in Figure 2) may reflect a compensatory mechanism or a consequence of progressive temporal lobe dysfunction. From a model interpretation perspective, these results demonstrate that the proposed DGAT-MSE model attends to neurophysiologically plausible features, enhancing confidence in its clinical applicability.

Moreover, Figure 2 shows the channel-wise importance for the AD subject. The GNNExplainer node importance reveals that prefrontal channels dominate the model’s decision for the AD subject. The top five contributing channels are Fp2 (0.351), Fp1 (0.343), Fz (0.336), T6 (0.336), and O1 (0.331). The three highest-ranked electrodes are all located in the prefrontal region, indicating that frontal EEG patterns are the primary discriminative markers leveraged by the model for AD classification.

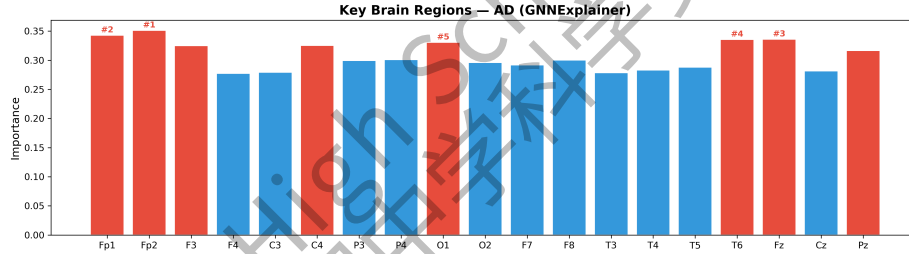


Fig. 2. Channel-wise importance for the AD subject

Figure 3 shows the channel-wise importance for the CN subject. For the cognitively normal control, the importance distribution exhibits a mixed temporal-frontal pattern. The leading channels are T6 (0.378), F4 (0.376), T5 (0.364), Fp1 (0.358), and Fp2 (0.358). Compared to the AD case, temporal lobe electrodes (T6, T5) gain higher relative importance while frontal involvement becomes more bilateral.

Figure 4 shows the comparison of channel importance between AD and CN. Direct comparison highlights the most divergent regions. The largest differences are observed at F4 ($\Delta = 0.099$, CN > AD), T5 ($\Delta = 0.076$, CN > AD), O2 ($\Delta = 0.059$, CN > AD), and T6 ($\Delta = 0.043$, CN > AD). Notably, all top-difference channels exhibit higher importance in CN than in AD, suggesting that these regions carry discriminative information that the model attenuates when classifying AD subjects. This pattern implies that AD pathology may manifest

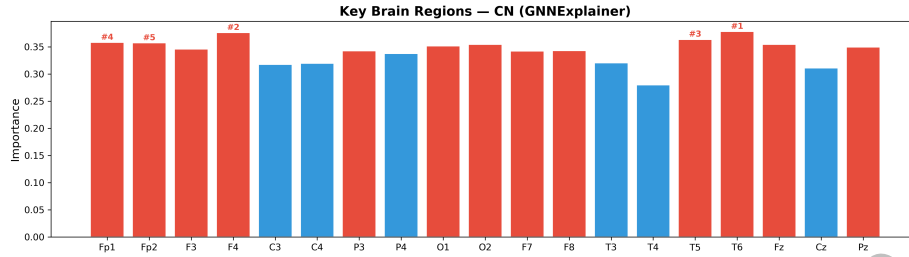


Fig. 3. Channel-wise importance for the CN subject

as a disruption of the normal functional contribution patterns seen in temporal and right-frontal regions.

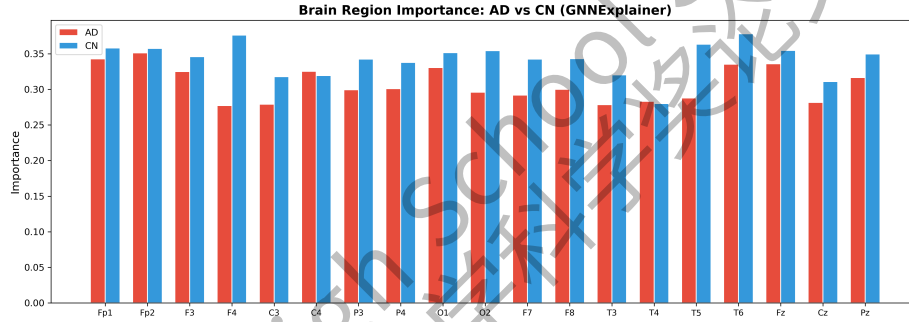


Fig. 4. Comparison of channel-wise importance between AD and CN

Temporal Window Importance Figure 5 illustrates the temporal importance profiles across 30 consecutive sliding windows for representative AD and CN subjects. Both profiles exhibit oscillatory patterns, reflecting the dynamic nature of neural information processing over time. However, distinct characteristics differentiate the two groups.

For the AD subject, temporal importance exhibited pronounced fluctuations with a mean score of 0.307 (SD = 0.062). Several windows demonstrated markedly elevated importance (windows 6, 9, 13, 16, 21 and 28, all exceeding 0.39), punctuated by intervals of reduced importance (windows 1, 2, with scores below 0.25). This alternating pattern of high-salience and low-salience windows suggests that AD-related discriminative features are concentrated in specific temporal segments, which may reflect transient spontaneous oscillatory patterns or time windows with greater discriminative relevance in resting-state EEG.

The CN subject displayed a higher baseline temporal importance (mean = 0.342, SD = 0.061) with a more evenly distributed profile. While periodic fluctuations were still present, the troughs were shallower relative to the AD profile,

indicating that CN neural activity maintains consistently informative features across a broader temporal span. This pattern aligns with the neuroscience literature documenting reduced trial-to-trial variability and more sustained task-related activation in healthy brain function [9].

The contrasting temporal dynamics between AD and CN subjects suggest that AD pathology not only alters the spatial distribution of discriminative EEG features but also compresses the temporal window during which diagnostically relevant neural activity can be reliably detected. This finding reinforces the importance of joint spatiotemporal modeling, as implemented in the DGAT-MSE framework, for capturing the full complexity of AD-related EEG alterations.

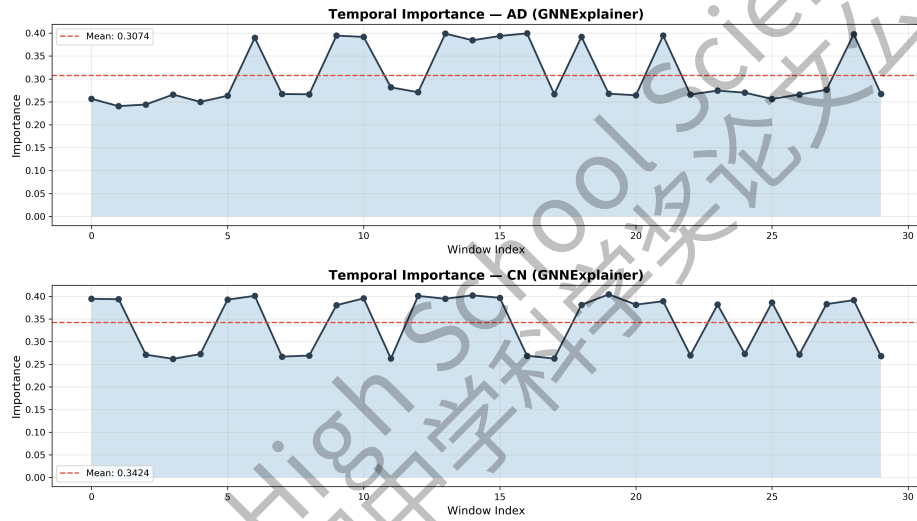


Fig. 5. Temporal importance profiles across sliding windows for AD (top) and CN (bottom) subjects. The dashed red line indicates the mean importance. Note the more pronounced peaks and deeper troughs in the AD profile.

4.4 Ablation Study

To systematically investigate the individual contribution of each key component in our proposed **DGAT-MSE** framework, we conducted a series of ablation experiments. Specifically, several model variants were constructed by removing or replacing core modules of the proposed architecture.

- **w/o MSE:** The MSE features are replaced with traditional Power Spectral Density (PSD) features to evaluate the contribution of nonlinear temporal complexity.

- **w/o Dynamic Graph:** The dynamic functional connectivity graph is replaced with an identity adjacency matrix, thereby removing inter-channel connectivity modeling and retaining only node-wise features.
- **w/o Temporal Transformer:** The Transformer-based temporal evolution modeling is replaced with simple mean pooling to evaluate the effectiveness of temporal dependency learning.

The experimental results are summarized in Table 3. As shown in Table 3, the complete DGAT-MSE framework achieves the best overall performance, demonstrating the effectiveness of jointly modeling nonlinear EEG complexity, dynamic functional connectivity, and temporal evolution.

When MSE features are replaced with conventional power spectrum features, the classification accuracy drops significantly from 85.71% to 61.54%. This substantial degradation indicates that nonlinear temporal complexity plays a critical role in characterizing AD-related neural activity. Traditional spectral features are insufficient to capture the irregular and multi-scale dynamics of EEG signals associated with neurodegenerative processes.

Removing the dynamic graph structure also leads to a noticeable performance decline. Specifically, the accuracy decreases to 82.63%, suggesting that modeling inter-channel functional connectivity is essential for capturing disease-related brain network alterations. The dynamic graph representation enables the model to characterize time-varying neural interactions that cannot be effectively represented by independent channel features alone.

In addition, replacing the Temporal Transformer with simple mean pooling reduces the classification accuracy to 84.48%. This result demonstrates that temporal evolution modeling contributes substantially to the final performance. The Transformer module effectively captures long-range temporal dependencies and dynamic transitions in EEG patterns, which are important for identifying AD-related cognitive abnormalities.

Overall, the ablation results confirm that each component of the proposed framework contributes positively to the final classification performance. In particular, the combination of multi-scale entropy features, dynamic graph modeling, and Transformer-based temporal learning provides complementary information for robust EEG-based AD classification.

Table 3. Ablation study of the proposed DGAT-MSE framework

Method	Accuracy	Precision	Recall	F1	AUC-ROC
w/o Dynamic Graph (identity adj)	0.8263	0.8000	0.8571	0.8276	0.8968
w/o MSE (Power Spectrum only)	0.6154	0.6154	1.0000	0.7619	0.7571
w/o Temporal Transformer (mean pool)	0.8448	0.8333	0.8571	0.8450	0.9161
DGAT-MSE (Full Model)	0.8571	1.0000	0.8571	0.9231	0.9048

Table 4. Subject-level zero-shot transfer performance of DGAT-MSE on ds006036 (36 AD / 29 CN).

Aggregation	Acc	Precision	Recall	F1	AUC
Majority vote	0.6308	0.8750	0.3889	0.5385	0.7510
Probability threshold (0.5)	0.6615	0.85	0.47	0.6071	0.7510

4.5 Cross-Dataset Generalization to OpenNeuro ds006036

To assess the generalizability of the proposed DGAT-MSE model beyond the distribution on which it was trained, we conducted a zero-shot cross-dataset transfer experiment on the OpenNeuro dataset ds006036 [24]. Unlike the training set ds004504, which contains eyes-closed resting-state EEG, the recordings in ds006036 were acquired under an eyes-open condition with photic stimulation (photomark). All test recordings share the same acquisition settings as the training data: 19 EEG channels in the Fp1...Pz montage, a sampling rate of 500 Hz, and EEGLAB .set format, with the label encoding A = AD and C = CN.

Feature Extraction and Experimental Settings To ensure a fair comparison, the feature extraction pipeline was kept identical to the one used during training. Raw EEG was band-pass filtered to 0.5 Hz to 40 Hz, re-referenced to the average, and resampled to 256 Hz. Only the first 60 s of each recording were used. Sliding windows of length $W = 128$ samples (0.5 s) were extracted with a stride of 51 samples (0.2 s). For each window, three complementary feature representations were computed: (i) multiscale entropy (MSE) across 5 scales, (ii) a dynamic functional connectivity graph based on the phase-locking value (PLV) thresholded at 0.3, and (iii) band-power features over five frequency bands (delta 0.5 Hz to 4 Hz, theta 4 Hz to 8 Hz, alpha 8 Hz to 13 Hz, beta 13 Hz to 30 Hz, and gamma 30 Hz to 40 Hz) with a $\log(1 + x)$ transform. The normalization statistics for the MSE and power features were those fitted on the source domain (ds004504), following the standard zero-shot protocol in which the target domain is processed without any adaptation.

For inference, windows were aggregated into sequences of length 10 with a stride of 5, and the pre-trained model *DGAT_MSE_viz.pth* was applied to each sequence without any fine-tuning. Subject-level predictions were obtained by aggregating the sequence-level outputs via majority voting and via a probability threshold of 0.5 on the mean predicted probability of AD.

Results and Analysis Table 4 reports the subject-level classification performance on the 65 test participants. The corresponding confusion matrix under majority voting is given in Table 5.

Several observations emerge from these results. First, the model retains a non-trivial discriminative ability on the target domain: the area under the ROC curve is 0.751, markedly above the chance level of 0.5. This indicates that the

Table 5. Confusion matrix of the majority-vote aggregation on `ds006036`.

	Predicted CN	Predicted AD
True CN	27	2
True AD	22	14

Table 6. Subject-level fine-tuning performance of DGAT-MSE on `ds006036`.

Method	Acc	Precision	Recall	F1	AUC
DGAT-MSE (fine-tuned)	0.8462	0.8750	0.8750	0.8750	0.9500

representations learned from eyes-closed resting-state EEG still carry information that transfers to an eyes-open photic-stimulation paradigm. Second, the model exhibits a pronounced sensitivity-specificity imbalance. It achieves high specificity (27/29 CN correctly classified, 93 %) but substantially lower sensitivity for AD (only 14/36 AD correctly classified under majority voting, 39 %). The probability-threshold aggregation alleviates this somewhat (recall improves from 0.3889 to 0.4722, with F1 rising from 0.5385 to 0.6071), but a large fraction of AD participants are still assigned to the CN class. This bias toward predicting CN is a typical signature of domain shift: the AD-related spectral and connectivity patterns captured during eyes-closed rest drift when the same participants are measured under eyes-open photic stimulation.

In summary, the transfer experiment demonstrates that DGAT-MSE generalizes above chance to a previously unseen acquisition paradigm, while the observed domain shift highlights the sensitivity of EEG-based AD markers to the recording condition (eyes-closed rest versus eyes-open photic stimulation) and motivates domain-adaptation strategies as a direction for future work.

4.6 Fine-Tuning on OpenNeuro `ds006036`

The zero-shot results above indicate that the domain gap between eyes-closed rest and eyes-open photic stimulation degrades the sensitivity of the model toward the AD class. To quantify how much of this degradation can be recovered with a modest amount of target-domain supervision, we fine-tuned the pre-trained `DGAT_MSE_viz.pth` model directly on `ds006036`. Because the test cohort contains only 65 participants, the fine-tuning was carried out under a k -fold cross-validation protocol, keeping the feature-extraction pipeline and sequence construction identical to the zero-shot setting described in Sec. 4.5. The pre-trained weights were adapted with a reduced learning rate, and each held-out fold was evaluated in the same way as before.

Table 6 reports the performance of the fine-tuned model.

Fine-tuning yields a marked and consistent improvement over the zero-shot baseline. Accuracy rises from 0.6615 to 0.8462, and the AUC increases from 0.7510 to 0.95. Most importantly, the sensitivity-specificity imbalance observed

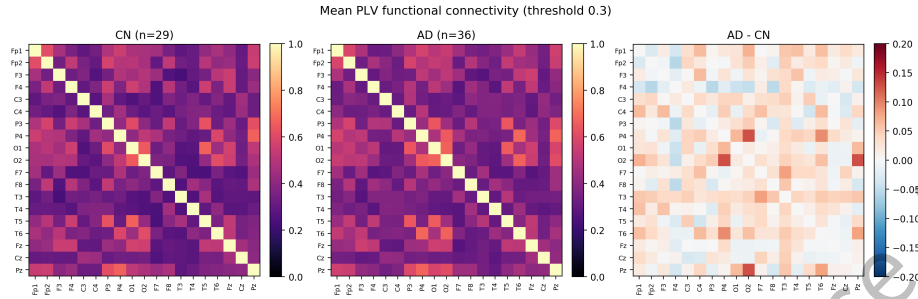


Fig. 6. Group-averaged PLV functional connectivity on ds006036: CN ($n = 29$, left), AD ($n = 36$, middle), and AD–CN difference (right). Connectivity below the 0.3 threshold is set to zero and self-loops are fixed to one, matching the graph construction used by the model.

under domain shift is essentially resolved: recall climbs from 0.47 to 0.875, and the balanced precision, recall, and F1 (all 0.875) indicate that the model no longer collapses toward the majority CN class. This confirms that the spectral and connectivity features are inherently discriminative for AD, and that the zero-shot degradation stemmed primarily from a shifted decision boundary rather than from a loss of information content in the representation.

Figure 6 visualizes the mean dynamic functional connectivity estimated from the PLV features, averaged over all 0.5 s windows of each participant. The AD group shows a modest overall increase in connectivity relative to CN, with the largest difference concentrated over parieto-occipital channels (notably P4–O2).

These results strongly suggest that a small amount of target-domain adaptation substantially recovers cross-paradigm generalization performance, and reinforces domain adaptation as a practical strategy for deploying EEG-based AD classifiers on previously unseen acquisition protocols.

5 Discussion

5.1 Methodological and Scientific Significance

The significance of the proposed DGAT-MSE framework extends beyond its classification performance and lies in the way it represents AD-related alterations in brain dynamics. From a neuroscience perspective, the framework treats EEG not merely as a collection of independent temporal signals, but as a dynamically evolving network in which the complexity of local neural activity and the interactions among brain regions change over time. MSE provides a nonlinear description of the temporal complexity of individual EEG channels, while dynamic functional connectivity captures time-varying relationships among channels. Therefore, their integration offers a representation that simultaneously reflects local neural complexity, inter-regional coordination, and temporal evolution.

The ablation results support the complementary nature of these components: replacing MSE with conventional power-spectrum features reduced accuracy from 85.71% to 61.54%, while removing dynamic graph modeling reduced it to 82.63%. These findings suggest that biomedical time-series problems may benefit from architectures that preserve the multiple scales at which biological systems operate, rather than forcing heterogeneous information into a single static representation. More importantly, the cross-dataset experiment provides additional methodological insight: although the model was trained on eyes-closed resting-state EEG, it retained above-chance discriminative information when directly transferred to eyes-open EEG with photic stimulation ($AUC = 0.751$), suggesting that part of the learned representation captures AD-related structure that is not entirely specific to one recording paradigm. Thus, the contribution of this work is not simply a more accurate classifier, but a framework for investigating how nonlinear neural dynamics and evolving brain connectivity jointly encode disease-related information.

5.2 Limitations

Several limitations should be considered. First, the relatively small and heterogeneous public cohorts may not fully capture the clinical and demographic variability of AD, while differences in recording paradigms between datasets introduce substantial domain shift and limit generalizability. Second, the framework assumes that PLV-based connectivity and MSE-derived complexity provide stable representations of disease-related dynamics across acquisition conditions, which may not hold under different preprocessing procedures, recording devices, or cognitive states. Finally, EEG functional connectivity reflects statistical synchronization rather than direct causal communication; therefore, model attention weights and discriminative connectivity patterns should not be interpreted as definitive physiological mechanisms. The improvement from zero-shot transfer ($AUC = 0.751$) to target-domain fine-tuning ($AUC = 0.950$) further demonstrates both the strength of the learned representation and its sensitivity to distributional differences. Thus, the present findings demonstrate a promising computational representation of AD-related EEG dynamics rather than evidence of clinical diagnostic validity.

5.3 Future Work

Future work will investigate whether the proposed representation captures robust and reproducible neurobiological signatures of AD. First, EEG will be integrated with structural and functional MRI to link electrophysiological dynamics with complementary anatomical and functional information. Second, longitudinal studies will examine whether changes in MSE, functional connectivity, and spatiotemporal representations can track individual cognitive decline and progression from cognitively normal states to mild cognitive impairment and AD. Third, systematic domain adaptation and generalization across datasets and acquisition protocols will be explored to separate disease-related patterns

from recording-specific variation. Together, these directions aim to determine whether the learned nonlinear complexity and dynamic connectivity patterns reflect stable, subject-independent, and acquisition-independent signatures of neurodegeneration.

6 Acknowledgments

6.1 Research Background and Topic Source

This research focuses on classification of Alzheimer’s disease using EEG signals. The motivation initially stemmed from learning about the severe impact of neurodegenerative diseases on the health of older adults during a school science fair. During the seven-month volunteer service program, "Caring for Alzheimer’s," in 2025, frequent contact with Alzheimer’s patients inspired the author to alleviate the suffering of patients and their families through effective AD classification. With the advisor’s introduction to recent advances in graph neural networks for healthcare, it is noticed that existing methods mostly ignore the nonlinear dynamics of EEG signals and the time-varying nature of brain connectivity. Therefore, we decided to explore a new method combining dynamic graph learning and entropy analysis. The specific research topic was formulated by the author and the advisor after reading papers on multiscale entropy [13], GNNExplainer [15], and EEG AD classification [4].

6.2 Division of Labor

Advisor Contributions Dr. Xin Lou. The advisor’s role was to introduce basic EEG analysis methods, guide the theoretical foundation of graph neural networks, help debug programming issues, and review the scientific accuracy of the paper draft. The advisor received no compensation, and the relationship is a standard teacher-student relationship.

Student Contributions

- **Literature Review and Methodology Formation:** Summarized the related works, gradually formed own method for AD classification, and finally proposed the DGAT-MSE framework.
- **Model Construction and Code Implementation:** Responsible for implementing the DGAT model, extracting MSE features, implementing PLV graph construction, conducting all model training, hyperparameter tuning, ablation studies, and interpretability analysis.
- **Data Processing and Paper Writing:** Downloaded and preprocessed the OpenNeuro dataset, performed data augmentation and train/test splitting, and wrote the research paper in LaTeX.

The advisor provided directional guidance and review but did not perform any experiments or writing on behalf of the students. All code, experiments, and the first draft were completed independently by the student.

6.3 Detailed Research Process

Data Acquisition The author downloaded the ds004504 and ds006036 datasets from the OpenNeuro website (<https://openneuro.org/>). The datasets have been public for years and contain anonymized EEG recordings; informed consent was obtained during the original data collection. We strictly followed the data-use agreement and did not interact directly with any human subjects.

Data Analysis and Computing Environment All experiments were run on a laptop with an Intel i7-10750H CPU, 16GB RAM, and an NVIDIA GTX 1660 Ti GPU. Software environment: Windows 10 + Python 3.9, with major libraries including MNE (0.24), NumPy (1.21), SciPy (1.7), PyTorch (1.10), and PyTorch Geometric (2.0). Preprocessing included band-pass filtering (0.5–40 Hz), resampling to 256 Hz, retaining the first 60 seconds, and sliding window segmentation with a window length of 0.5 seconds and a step of 0.2 seconds.

Experimental Design and Implementation Five experiments were designed: 1) comparison with four baseline methods on ds004504 dataset; 2) ablation studies (removing dynamic graph, removing MSE, removing Transformer); 3) GNNExplainer interpretability analysis; 4) subject-level cross-validation on ds006036 dataset; 5) fine-tuning on ds006036 dataset.

Major difficulties were encountered during various experiment stages: 1) In the graph construction stage, The PLV threshold significantly affected the results. A small parameter search was conducted to find the optimal threshold value; 2) In the training stage, the GAT attention mechanism sometimes produced NaN losses. The issue was solved by gradient clipping and ELU activation; 3) In the interpretability stage, slow GNNExplainer computation was overcome by random window sampling. Finally, we successfully reproduced and validated the results.

Writing of the Thesis The author independently wrote the first draft of the thesis in LaTeX in the format of Springer research paper. Then the advisor reviewed the draft and provided the comments. After a few iterations, the author finally completed and formatted the entire document. The writing process took about three weeks.

6.4 AI Usage Statement

During the preparation of this research report, Grammarly (Web Version) was used solely as a language-editing and proofreading tool. Grammarly was used several times during the writing and revision process, from August 2026 to September 2026, to check grammar, spelling, punctuation, and overall language structure, and to suggest improvements in organization where appropriate. Grammarly was not used to generate the research question, scientific methodology,

experimental design, data processing, code, or interpretation of experimental results.

All research ideas, experimental procedures, data analysis, implementation, interpretation, and discussions were developed and completed by the student. All suggestions provided by Grammarly were reviewed critically by the student, and only appropriate language corrections were incorporated into the final report.

6.5 External Assistance Statement

No one else but the author and the advisor provided substantial help. The author referred to the official PyTorch Geometric examples for the GAT implementation and modified it for our dynamic graph sequence. No paid paper-writing or code-writing services were used.

6.6 Ethics and Compliance Statement

This research does not violate any Chinese laws or regulations. It does not involve food, tobacco, alcohol, drugs, or medical instruments that could cause physical or psychological harm. The EEG data used are from a public and anonymized dataset. The dataset's original collection followed informed consent procedures and we did not interact with any human subjects directly. This research adheres to internationally recognized ethical standards for data science and medical informatics.

The similarity detection report indicates an overall similarity rate of less than 5%.

References

1. C. Wee, P. Yap, D. Zhang, et al., "Identification of MCI individuals using structural and functional connectivity networks," *Neuroimage*, vol. 59, no. 3, pp. 2045–2056, 2012.
2. A. B. Rabeh, F. Benzarti, and H. Amiri, "Diagnosis of alzheimer diseases in early step using svm (support vector machine)," in 2016 13th International Conference on Computer Graphics, Imaging and Visualization (CGiV), Beni Mellal, Morocco, 2016, pp. 364–367.
3. Q. Jiang and J. Shi, "Sparse kernel entropy component analysis for dimensionality reduction of neuroimaging data," in 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2014, pp. 3366–3369.
4. C. Ieracitano, N. Mammone, A. Hussain, et al., "A novel multi-modal machine learning based approach for automatic classification of eeg recordings in dementia," *Neural Networks*, vol. 123, pp. 176–190, 2020.
5. F. Li and M. Liu, "Alzheimer's disease diagnosis based on multiple cluster dense convolutional networks," *Computerized Medical Imaging and Graphics*, vol. 70, pp. 101–110, 2018.
6. A. Farooq, S. Anwar, M. Awais, et al., "A deep cnn based multi-class classification of alzheimer's disease using mri," in 2017 IEEE International Conference on Imaging Systems and Techniques (IST), 2017, pp. 1–6.
7. Y. AbdulAzeem, W. Bahgat, and M. Badawy, "A cnn based framework for classification of alzheimer's disease," *Neural Computing and Applications*, vol. 33, no. 16, pp. 10415–10428, 2021.
8. Y. Zhang, L. Xue, S. Zhang, et al., "A novel spatiotemporal graph convolutional network framework for functional connectivity biomarkers identification of Alzheimer's disease," *Alzheimer's Research & Therapy*, vol. 16, no. 60, 2024.
9. M. Grieder, D. Wang, T. Dierks, et al., "Default mode network complexity and cognitive decline in mild Alzheimer's disease," *Frontiers in Neuroscience*, vol. 12, no. 770, 2018.
10. J. D. Power, A. L. Cohen, S. M. Nelson, et al., "Functional network organization of the human brain," *Neuron*, vol. 72, no. 4, pp. 665–678, 2011.
11. T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *International Conference on Learning Representations*, 2017.
12. A. Vaswani, N. Shazeer, N. Parmar, et al., "Attention is all you need," *Advances in Neural Information Processing Systems*, pp. 5998–6008, 2017.
13. M. Costa, A. L. Goldberger, and C. Peng, "Multiscale entropy analysis of complex physiologic time series," *Physical Review Letters*, vol. 89, p. 068102, 2002.
14. S. Wu, P. Zhan, G. Wang, et al., "Changes of brain functional network in Alzheimer's disease and frontotemporal dementia: A graph-theoretic analysis," *BMC Neuroscience*, vol. 25, no. 1, p. 30, 2024.
15. Z. Ying, D. Bourgeois, J. You, M. Zitnik, and J. Leskovec, "GNNExplainer: Generating explanations for graph neural networks," *Advances in Neural Information Processing Systems*, vol. 32, pp. 9240–9251, 2019.
16. J. Wang, J.-E. Ding, F. Liu, et al., "Flexible and explainable graph analysis for EEG-based Alzheimer's disease classification," *arXiv preprint arXiv:2504.01329*, 2025.
17. V. P. Oikonomou, K. Georgiadis, I. Lazarou, et al., "Exploring functional brain networks in Alzheimer's disease using resting state EEG signals," *Journal of Dementia and Alzheimer's Disease*, vol. 2, no. 2, p. 12, 2025.

18. A. F. Parker, C. Szoëke, H. Kwan, et al., “Functional connectivity alterations in women with subjective cognitive decline,” *Journal of Alzheimer’s Disease Reports*, vol. 9, p. 25424823251328340, 2025.
19. F. Miraglia, F. Vecchio, C. Pappalettera, et al., “Brain connectivity and graph theory analysis in Alzheimer’s and Parkinson’s disease: the contribution of electrophysiological techniques,” *Brain Sciences*, vol. 12, no. 3, p. 402, 2022.
20. M. Zhdanov, S. Steinmann, and N. Hoffmann, “Investigating brain connectivity with graph neural networks and GNNExplainer,” in *26th International Conference on Pattern Recognition (ICPR)*, pp. 5155–5161, 2022.
21. U. Smailovic, T. Koenig, I. Savitcheva, et al., “Regional disconnection in Alzheimer dementia and amyloid-positive mild cognitive impairment: Association between EEG functional connectivity and brain glucose metabolism,” *Brain Connectivity*, vol. 10, no. 10, pp. 555–565, 2020.
22. J. P. Gurja, S. P. Muthukrishnan, M. Tripathi, et al., “Reduced resting-state cortical alpha connectivity reflects distinct functional brain dysconnectivity in Alzheimer’s disease and mild cognitive impairment,” *Brain Connectivity*, vol. 12, no. 2, pp. 134–145, 2022.
23. A. Miltiadous, et al., “A dataset of EEG recordings from: Alzheimer’s disease, Frontotemporal dementia and Healthy subjects,” *OpenNeuro*, 2026. doi: 10.18112/open-neuro.ds004504.v1.0.9.
24. A. Ntetska, et al., “A complementary dataset of open-eyes EEG recordings in a photo-stimulation setting from: Alzheimer’s disease, Frontotemporal dementia and Healthy subjects,” *NEMAR*, 2026. doi: 10.82901/nemar.on006036.