Biomedical Signal Processing and Control 104 (2025) 107458

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

EEG_GLT-Net: Optimising EEG graphs for real-time motor imagery signals classification☆

Htoo Wai Aung [a], Jiao Jiao Li [a], Bin Shi [b,*], Yang An [c,*], Steven W. Su [c,*]

[a] University of Technology Sydney, 15 Broadway, Ultimo, Sydney, 2007, NSW, Australia
[b] Neck-Shoulder and Lumbocrural Pain Hospital of Shandong First Medical University, Shandong First Medical University & Shandong Academy of Medical Sciences, Jinan, China
[c] College of Medical Information and Artificial Intelligence, Shandong First Medical University & Shandong Academy of Medical Sciences, Shandong, China

ARTICLE INFO

Keywords:
Brain-computer interfaces (BCIs)
Electroencephalography motor imagery (EEG MI)
Spectral graph convolutional neural networks (GCNs)
EEG_GLT (EEG graph lottery ticket)
Graph pruning

ABSTRACT

Brain-Computer Interfaces (BCIs) connect the brain to external control devices, necessitating the accurate translation of brain signals such as from electroencephalography (EEG) into executable commands. EEG MI classification has numerous applications, including neurorehabilitation for stroke patients, control of assistive robotic devices, and advancements in neurofeedback systems. Graph Neural Networks (GCN) have been increasingly applied for classifying EEG Motor Imagery (MI) signals, primarily because they incorporates the spatial relationships among EEG channels, resulting in improved accuracy over traditional convolutional methods. However, existing methods for constructing adjacency matrices, such as Geodesic distances, Pearson Correlation Coefficient (PCC), and others, often rely on predefined inter-channel relationships. These methods not only demand high computational resources during inference but often achieve limited performance accuracy, particularly for single time-point EEG MI classification where rapid interpretation is crucial.

To address this, our paper introduces the EEG Graph Lottery Ticket (EEG_GLT) algorithm, an innovative technique for constructing adjacency matrices for EEG channels. This method does not require pre-existing knowledge of inter-channel relationships, and it can be tailored to suit both individual subjects and GCN model architectures. We conducted an empirical study with 20 subjects and six different GCN architectures to compare the performance of our EEG_GLT adjacency matrix against both Geodesic and PCC adjacency matrices on time-resolved EEG MI dataset, PhysioNet dataset. Our EEG_GLT method consistently exceeded performance accuracy benchmarks. Additionally, we compared our model with state-of-the-art models, achieving superior results. EEG_GLT algorithm marks a breakthrough in development of optimal adjacency matrices, effectively boosting both computational accuracy and efficiency, making it well-suited for single time point classification of EEG MI signals that demand intensive computational resources.

1. Introduction

Brain Computer Interfaces (BCIs) form an interdisciplinary bridge between engineering and neuroscience, enabling direct communication between the human brain and control devices. Originally designed to aid those with motor impairments [1], BCIs have expanded their applications to neurofeedback, gaming, and rehabilitation. Essentially, BCIs convert neural signals into actionable commands. The primary means of brain signal acquisition include electrocorticography (ECoG) and electroencephalography (EEG). Although ECoG boasts superior spatial resolution due to directly placing electrodes on the cortex, its invasive native limits its applications [2]. In contrast, EEG uses scalp placed electrodes to capture brain activity, making it more popular due to non-invasiveness and portability. This method captures various brain signals, from event-related to spontaneous and stimulus-evoked [3].

Motor Imagery (MI) pertains to the mental simulation of motor actions, such as moving one's hands or feet, without performing the actual movement [4,5]. As highlighted by [6], action execution and its imagination share neural pathways. MI has prominent applications in rehabilitation and neuroscience. When paired with EEG, it captures neural signals generated from the intention to move. Integrating this

☆ This document is the results of the research project funded by the University of Technology Sydney and the Major Science and Technology Innovation Project of Shandong Province, China (2022CXGC020510, 2021SFGC0502)

* Corresponding authors.
E-mail addresses: htoowai.aung@student.uts.edu.au (H.W. Aung), jiaojiao.li@uts.edu.au (J.J. Li), bshi@sdfmu.edu.cn (B. Shi), an.yang1229@outlook.com (Y. An), suweidong@sdfmu.edu.cn, Steven.Su@uts.edu.au (S.W. Su).

https://doi.org/10.1016/j.bspc.2024.107458
Received 20 June 2024; Received in revised form 17 December 2024; Accepted 27 December 2024
Available online 9 January 2025

with BCIs allows decoding EEG MI signals to control external devices such as a robotic exoskeleton. This technology is pivotal for those with motor impairments, especially stroke survivors, with the potential to restore quality of life and ability to perform daily activities. By accurately decoding EEG MI signals, BCIs can provide real-time feedback and communicate with assistive devices, to facilitate patient-intended movements [7].

Convolutional Neural Networks (CNNs) have consistently showcased superior results in computer vision tasks [8–10]. However, their effectiveness is largely constrained to regular Euclidean data, such as 2-dimensional grids and 1-dimensional sequences [10]. A drop in capability is experienced with non-Euclidean data, primarily because CNN cannot accurately capture the intrinsic structure and connectivity of this data. Graphs serve as powerful tools for representing relationships among entities, and are employed in diverse application areas including traffic systems, social networks, e-commerce platforms, biological structures, and trade networks. These graphs can highlight complex structures and be variable in nature such as being might be homogeneous or heterogeneous, having weight or not, and being signed or unsigned [11]. The Graph Convolutional Neural Network (GCN) is an adaptation of CNN operations that is, tailored for graphs. GCN excel in managing non-Euclidean data, incorporating topological relationships during convolution.

With the help of GCNs, the inherent connections among electrodes can be integrated through the adjacency matrix, a capability beyond the reach of traditional CNNs. Establishing relationships between nodes is essential before deploying the GCN method. Studies [12–14] have utilised Geodesic distances between electrodes to form the adjacency matrix, while others [15–19] have employed the Pearson Coefficient Correlation (PCC) to assess correlations between EEG channels. Additionally, [20] have utilised the Phase Lag Index (PLI) in the adjacency matrix construction in their CSGNN model. Notably, [15,21] explored optimal adjacency matrices in EEG classification through a trainable matrix. [22] introduced a unified GNN sparsification technique (UGS), giving rise to a Graph Lottery Ticket (GLT) by pruning both the original adjacency matrix and GNN weights. This method decreases the Multiply Accumulate (MAC) inference, thus reducing computational overhead.

Existing methods for constructing adjacency matrices in EEG signal classification rely on prior knowledge of inter-channel relationships, which can be a limitation. This dependency is especially challenging for single time point classification of EEG motor imagery (MI) signals, where rapid signal interpretation at intervals as brief as $\frac{1}{160}$s is critical. To address this, our study proposes EEG_GLT which is a novel method for constructing adjacency matrices for GCNs specifically for EEG MI single time point classification, without requiring predefined inter-channel knowledge while enhancing both classification accuracy and computational efficiency.

2. Related work

Traditional EEG MI classifiers typically rely on machine learning techniques that classify signals based on manually crafted features, such as wavelet transforms or analytic intrinsic mode functions [23,24]. One widely used method is the filter bank common spatial pattern (FBCSP) [25], which applies common spatial patterns (CSP) across various frequency bands in EEG signals to extract discriminative features.

Deep neural networks (DNNs) have advanced EEG motor imagery (MI) classification by leveraging end-to-end architectures that combine feature extraction with classifier learning, eliminating the need for manual feature engineering. CNN-based models, such as those proposed by [26,27], excel at extracting temporal features from 1D and 2D Euclidean data, achieving high accuracy. Further refinements, as seen in [28–30], incorporate LSTM blocks to capture temporal dependencies effectively in EEG signals. The EEGProgress model [31] adopts a unique approach by applying CNN operations to individual brain regions for EEG MI signal classification, focusing on regional processing rather than all channels simultaneously. The ConTraNet model [32] combines Transformer and CNN blocks to capture both long- and short-term dependencies, fixed spatial patterns, and applies attention to non-stationary, time-varying inputs, resulting in improved performance for EEG-based emotion recognition. However, a common limitation of the methods discussed above is that they are applicable only to window-based EEG classification and not to single time-point classification.

Graph convolutional networks (GCNs) have become increasingly popular in EEG signal classification due to their ability to encode non-Euclidean data, offering flexibility in analyzing graph-structured information [11,33]. GCNs can be applied across various graph analysis tasks:

- **Node-Level Tasks**: Predicting properties of individual nodes, used for both regression and classification.
- **Edge-Level Tasks**: Predicting edge properties, mainly for classification.
- **Graph-Level Tasks**: Classifying entire graphs based on their structure and properties.

Two main categories of GCNs are the spectral method [34–36] and the spatial method [37–40]. Studies [15,41] indicate challenges associated with the spatial method, particularly for matching local neighbourhoods. GCNs have an important application in classifying EEG signals at the graph level, where EEG readings from individual electrodes are treated as node attributes. EEG feature extraction is broadly categorised into time and frequency domain features. Building on the work of [42], time-domain metrics such as Root Mean Square, skewness, minmax, variance, kurtosis, Hurst Exponent, Higuchi, and Petrosian fractal dimensions are derived within predefined time windows by [17]. Within the frequency domain, emphasis is placed on power spectral density (PSD) and power ratio (PR) across specific frequency bands: δ[0.5–4 Hz], θ[4–8 Hz], α[8–13 Hz], β[13–30 Hz], and γ[30–110 Hz]. This is supplemented by other metrics such as total power, spectral entropy, and peak frequency, all captured within chosen time windows.

The DG-HAM [43] and EEG-ARNN [44] models classify EEG tasks using raw EEG signals within a specified window length, without extracting graph-based features, such as time-domain or frequency-domain features. In contrast, [19] introduced the state-of-the-art GCNs-Net for time-point classification, which treats each time point as an independent feature, enabling a more detailed time-resolved analysis of EEG MI signals. Although GCNs-Net performs well in classifying EEG MI single time points, it only considers the functional connectivity of EEG channels during GNN operations, which can limit its accuracy. Additionally, its fully dense adjacency matrix requires high multiply-accumulate (MAC) operations, making it less efficient for single time-point EEG classification.

In this study, we propose the EEG_GLT method for adjacency matrix construction, integrated with a spectral GNN-based EEG_GLT-Net architecture, to classify EEG MI at the single-time-point level. Using the raw EEG MI single-time-point signals from the time-resolved PhysioNet dataset. The primary contributions of this study can be summarised as:

- **EEG Graph Lottery Ticket (EEG_GLT):** We present a novel method to construct an optimal adjacency matrix for EEG MI signal classification. Achieved through the iterative pruning of relationships among EEG channels, the EEG_GLT introduces a new direction in EEG adjacency matrix design.
- **Channel Relationship Optimisation:** Our approach reveals the most advantageous relationship between EEG channels. It is tailored for catering to individual subjects and the architecture of GCN models, eliminating the need for prior knowledge about the inter-relationships among EEG channels.

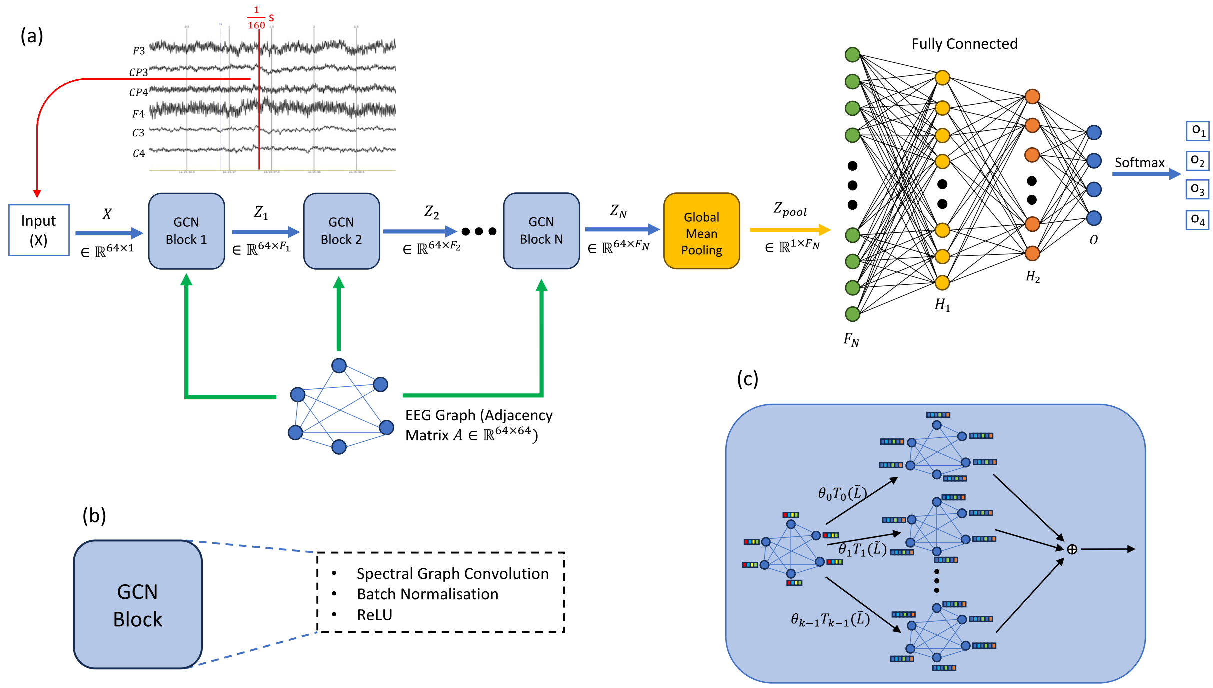


Fig. 1. Our model: (a) Overall architecture (classifying EEG MI of one time point $\frac{1}{160}s$ of signals from 64 EEG electrodes). Note that EEG Graph adjacency matrix can be $A^{Geodesic}$, A^{PCC} or A^{EEG_GLT}, (b) Components inside the spectral graph convolution block, (c) Chebyshev spectral graph convolution.

- **Computational Efficiency:** Recognising the computational intensity of classifying EEG at single time points, our strategy mitigates the high demand for computational resources, proving especially beneficial for real-time applications.
- **Performance Validation:** We benchmark the accuracy of our EEG_GLT method against two well-established techniques: the Geodesic method and the leading PCC method employed in the state-of-the-art GCNs-Net. This evaluation spans across six distinct spectral GCN models. Each model is distinguished by its unique specifications, including variations in GCN layer structures, polynomial degrees of filters, numbers of Fully Connected (FC) layers, and the amount of hidden nodes. Additionally, we compare the performance of our model with seven other state-of-the-art models to demonstrate its effectiveness.

3. Methodology

3.1. Overview

As shown in Fig. 1, the project framework was as follows:

- EEG signals from 64 channels were captured at each time point $\frac{1}{160}s$ and used as input features for the EEG_GLT-Net.
- Additionally, the EEG_GLT-Net accepted the graph representation as another form of input. This representation included the graph Laplacian, derived using three different methods: Pearson Correlation Coefficient (PCC) between EEG channels, Geodesic distance between EEG electrodes, and our newly proposed EEG Graph Lottery Ticket Adjacency Matrix Mask (m_{EEG_GLT}).
- The EEG_GLT-Net processed these inputs to decode the EEG MI time point signal, which was then categorised into one of the four MI types.

3.2. Dataset description

This paper utilised the PhysioNet EEG Motor Imagery (MI) dataset [45] encompassing over 1500 EEG recordings sourced from 109 participants. The recordings were captured using 64 EEG electrodes, consistent with the international 10-10 system, with the exclusion of F9, Nz, F10, FT9, FT10, A1, A2, TP9, TP10, P9, and P10 channels. Each participant executed 84 trials, broken down into 3 runs with, 7 trials per run, spanning 4 distinct tasks. The tasks included:

- Task 1: Imagining the act of opening and closing the left fist.
- Task 2: Imagining the act of opening and closing the right fist.
- Task 3: Imagining the act of opening and closing both fists simultaneously.
- Task 4: Imagining the act of opening and closing both feet.

Recordings in the dataset were originally sampled at 160 Hz and each recording had a duration of 4 s. Our study employed time point samples for classification, and our analysis was strictly conducted at the subject level. Although the original dataset comprised 109 participants, our study focused solely on 20 subjects, labeled S_1 to S_{20}.

3.3. Data pre-processing and feature extraction

In the initial pre-processing phase, raw signals underwent only a notch filter at the 50 Hz power line frequency, foregoing typical filtering or denoising steps to maximise data integrity. Although each task lasted for a 4-second duration, only the time period from $t = 1$ s to $t = 3$ s was considered in our experiments. This is because subjects typically exhibited greater readiness post $t = 1$ s. All 64 EEG channels were incorporated into our model. We utilised the signal values from each EEG channel at each time point as features for individual nodes. The construction methods of the adjacency matrix, which captures brain connectivity, are elaborated in Sections 3.4 and 3.5. The training data underwent normalisation, ensuring a mean $\mu = 0$ and a standard deviation $\sigma = 1$ for each channel. Following this, both the test and validation sets were adjusted in alignment with the normalisation parameters established from the training data.

3.4. Graph preliminary

3.4.1. Graph representation

Consider a directed weighted graph represented as $G = \{V, E\}$. Here, $|V| = N$ denoted the number of nodes and $|E|$ was the count of edges connecting the nodes. The node set was defined as $V = \{v_1, v_2, \dots, v_n\}$ and the node feature matrix of the entire graph was represented by $X \in \mathbb{R}^{N\times F}$. The adjacency matrix, denoted as $A \in \mathbb{R}^{N\times N}$, captured the graph's overall topology. Specifically, if an edge existed between nodes v_i and v_j (i.e., $(v_i, v_j) \in E$), then $A[i,j] \neq 0$. Otherwise, $A[i,j] = 0$.

The adjacency matrix for the PCC method was defined in Eq. (2), where I was the identity matrix and $|P|$ was the absolute PCC matrix as in Eq. (1). This PCC matrix, $|P| \in [0,1]$, captured the linear correlations among EEG channel signals.

$$P_{ij} = \frac{cov(x_i, x_j)}{\sigma_i \sigma_j} \tag{1}$$

$$A^{PCC} = |P| - I \tag{2}$$

For the Geodesic-distance adjacency matrix method, the configuration of 64 electrodes into a unit sphere acted as a stand-in for spatial brain connectivity. This allowed the computation of geodesic distances between the electrodes placed on a sphere of radius r. If two electrodes have Cartesian coordinates $(x_i,\ y_i,\ z_i)$ and $(x_j,\ y_j,\ z_j)$, the geodesic distance for the adjacency matrix was calculated using Eq. (3). These distances were standardised into the [0, 1] range.

$$A_{ij}^{Geodesic} = arcos(\frac{(x_i\ x_j + y_i\ y_j + z_i\ z_j)}{r^2}) \tag{3}$$

The degree matrix, D, was a diagonal representation of A, where the ith diagonal element of D was computed as $D_{ii} = \sum_{j=1}^{N} A_{ij}$. The combinatorial Laplacian matrix, $L \in \mathbb{R}^{N\times N}$, was described as $L = D - A$. A normalised version of this combinatorial Laplacian can be obtained using:

$$L = I_N - D^{-1/2} A D^{-1/2} \tag{4}$$

3.4.2. Spectral graph filtering

The eigenvectors of the graph Laplacian matrix can be expressed as graph Fourier modes, with $\{u_l\}_{l=0}^{N-1} \in \mathbb{R}$. The diagonal matrix of these Fourier frequencies, Λ, is given by $diag[\lambda_0, \dots, \lambda_{N-1}] \in \mathbb{R}^{N\times N}$. We defined the Fourier basis, $U = [u_0, \dots, u_{N-1}] \in \mathbb{R}^{N\times N}$, which allows for the decomposition of the Laplacian matrix, L, into $L = U\Lambda U^T$. The signal x can be transformed by graph Fourier into $\hat{x} \in \mathbb{R}^N$ using $\hat{x} = U^T x$, while the inverse graph Fourier transform is given by $x = U\hat{x}$. The convolution operation on graph G is defined as:

$$x *_G g = U((U^T x) \odot (U^T g)) \tag{5}$$

where g represents the convolutional filter and $\odot$ denotes the Hadamard product. Given that $g_\theta(\Lambda) = diag(\theta)$, where $\theta \in \mathbb{R}^N$ represents the vector of Fourier coefficients, the Graph convolution operation can be implemented as follows:

$$x *_G g_\theta = U g_\theta(\Lambda) U^T x \tag{6}$$

where g_θ is a non-parametric filter, and polynomial approximation is employed to mitigate the excessive computational complexity. Chebyshev graph convolution, a specific instance of graph convolution, utilises Chebyshev polynomials for filter approximation, thereby reducing computational complexity from $O(N^2)$ to $O(KN)$ [35]. The approximation of $g_\theta(\Lambda)$ under the Kth order Chebyshev polynomial framework is given by:

$$g_\theta(\Lambda) = \sum_{k=0}^{K-1} \theta_k T_k(\hat{\Lambda}) \tag{7}$$

$$\hat{\Lambda} = \frac{2\Lambda}{\Lambda_{max}} - I_N \tag{8}$$

Normalising Λ can be achieved by using Eq. (8), where Λ_{max} denotes the largest entry in the diagonal of Λ, and I_N represents the diagonal matrix of the scaled eigenvalues. In the equation above, θ_k refers to the Chebyshev polynomial's coefficients, and $T_k(\hat{\Lambda})$ is obtained by the following equations:

$$\{T_0(\hat{\Lambda}) = 1, T_1 = (\hat{\Lambda}), T_k(\hat{\Lambda}) = 2\hat{\Lambda}T_{k-1}(\hat{\Lambda}) - T_{k-2}(\hat{\Lambda})\} \tag{9}$$

Finally, the signal x can be convolved with the defined filter g_θ as follows:

$$x *_G g_\theta = U \sum_{k=0}^{K-1} \theta_k T_k(\hat{\Lambda}) U^T x \tag{10}$$

$$x *_G g_\theta = \sum_{k=0}^{K-1} \theta_k T_k(\widetilde{L}) x \tag{11}$$

The normalised Laplacian matrix, denoted as $\widetilde{L}$ can be computed using the following Eq. (12).

$$\widetilde{L} = \frac{2L}{\lambda_{max}} - I_N \tag{12}$$

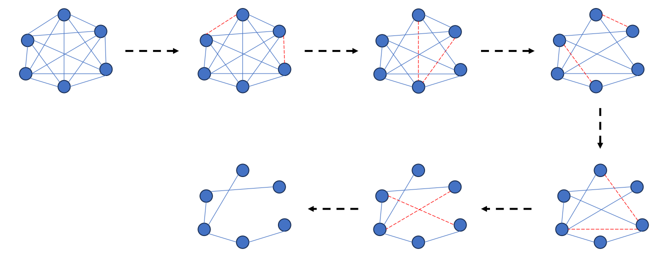


Fig. 2. EEG graph (m_g) pruning using Algorithm 1: At each N_{ep} iteration, the bottom p_g% are pruned, reducing density from 100% until the lowest density s_g%. Solid lines indicate remaining edges, while red-dashed lines depict removed edges.

3.5. EEG Graph Lottery Ticket (EEG_GLT)

In the process of executing a forward pass with the spectral GNN function, symbolised as $f(., \Theta)$, and given a graph denoted as $G = \{A, X\}$, the method presented in UGS method [22] aims to search the adjacency matrix mask $m_g \in \{0,1\}$ with the maximum sparsity that concurrently maintained the highest prediction accuracy. In our model, the original matrix $A_{original\_ij} = \{0, if \quad i = j; 1, otherwise\}$ in the shape of $|V|\times|V|$ was not trainable. The adjacency matrix mask in our model $m_g \in \mathbb{R}^{|V|\times|V|}$ was trainable.

$$A = A_{original} \odot m_g \tag{13}$$

Once the model had undergone N epochs, the lowest p_g% (p_g = 10%) of the values in the trained m_g at highest accuracy of the validation dataset were pruned. These values were set to 0, while the remaining values were set to 1 as shown in Fig. 2. Concurrently, the spectral filter weights, represented as Θ, were reset to their initial state, Θ_0. The trained m_g that yielded the highest accuracy of the validation set within the span of N epochs was designated as the Graph Lottery Ticket (GLT) and duly noted. This process continued, and a GLT was recorded for each level of graph sparsity until the sparsity of m_g fell below the pre-determined final sparsity level, s_g. The EEG_GLT was ultimately identified as the GLT that achieves the highest accuracy alongside the highest level of graph sparsity. Moreover, it delineated the optimal adjacency matrix capable of producing the highest accuracy.

3.6. General model architecture

A GCN structure was designed to classify EEG MI signals. This architecture comprised three primary blocks: the GCN block, the Global Mean Pooling Block, and the Fully Connected Block. In the GCN Block, generalised graph features for each EEG electrode were extracted. Subsequently, the features from all 64 channels were consolidated using a mean in the Global Mean Pooling Block. The Fully Connected Block was employed for the final prediction. A detailed representation of this model architecture is provided in Fig. 1 and Table 1.

Table 1
Generalised architecture of GCN model.

Layer	Type	Input size	Polynomial order	Weights	Bias	Output	Activation
Input	Input	$N \times 1$	–	–	–	–	–
Block A - GCN Block							
C1	Graph Convolution	$N \times 1$	K_1	$1 \times F_1 \times K_1$	$N \times F_1$	$N \times F_1$	–
BNC1	Batch Normalisation	$N \times F_1$	–	F_1	F_1	$N \times F_1$	ReLU
C2	Graph Convolution	$N \times F_1$	K_2	$F_1 \times F_2 \times K_2$	$N \times F_2$	$N \times F_2$	–
BNC2	Batch Normalisation	$N \times F_2$	–	F_2	F_2	$N \times F_2$	ReLU
C3	Graph Convolution	$N \times F_2$	K_3	$F_2 \times F_3 \times K_3$	$N \times F_3$	$N \times F_3$	–
BNC3	Batch Normalisation	$N \times F_3$	–	F_3	F_3	$N \times F_3$	ReLU
C4	Graph Convolution	$N \times F_3$	K_4	$F_3 \times F_4 \times K_4$	$N \times F_4$	$N \times F_4$	–
BNC4	Batch Normalisation	$N \times F_4$	–	F_4	F_4	$N \times F_4$	ReLU
C5	Graph Convolution	$N \times F_4$	K_5	$F_4 \times F_5 \times K_5$	$N \times F_5$	$N \times F_5$	–
BNC5	Batch Normalisation	$N \times F_5$	–	F_5	F_5	$N \times F_5$	ReLU
C6	Graph Convolution	$N \times F_5$	K_6	$F_5 \times F_6 \times K_6$	$N \times F_6$	$N \times F_6$	–
BNC6	Batch Normalisation	$N \times F_6$	–	F_6	F_6	$N \times F_6$	ReLU
Block B - Global Mean Pooling Block							
P	Global Mean Pool	$N \times F_6$	–	–	–	F_6	–
Block C - Fully Connected Block							
FC1	Fully Connected	F_6	–	$F_6 \times H_1$	H_1	H_1	–
BNFC1	Batch Normalisation	H_1	–	H_1	H_1	H_1	ReLU
FC2	Fully Connected	H_1	–	$H_1 \times H_2$	H_2	H_2	–
BNFC2	Batch Normalisation	H_2	–	H_2	H_2	H_2	ReLU
FC3	Fully Connected	$H_2 \times O$	–	$H_2 \times O$	O	O	–
S	Softmax Classification	O	–	–	–	O	–

N = Number of EEG Channels (i.e. 64); O = Number of EEG MI Classes (i.e. 4)

Algorithm 1 Finding EEG Graph Lottery Ticket

Input: Graph $G = \{A, X\}$, GNN $f(G, \Theta)$, GNN initialisation Θ_0, $A_{original\_ij} = \{0, if \quad i = j; 1, otherwise\}$, initial Adjacency Matrix Mask $m_g^0 = A_{original}$, learning rate $\eta = 0.01$, pruning rate $p_g = 10\%$, pre-defined lowest Graph Density Level $s_g = 13.39\%$.

Output: EEG Graph Lottery Ticket $(m_{g\_EEG\_GLT}) - m_g^{s,i}$ at the highest accuracy with the highest sparsity possible.

while $\frac{||m_g^s||_0}{||A_{original}||_0} \geq s_g$ **do**
  for for iteration $i = 0, 1, 2, \ldots, N_{ep}$ **do**
    Forward $f(., \Theta_i)$ with $G_s = \{m_g^{s,i} \odot A_{original}, X\}$ to compute Cross-Entropy Loss, L
    Backpropagate and update, Θ_i and $m_g^{s,i}$ using Adam Optimiser
  end for
  Record $m_g^{s,i}$ with the highest accuracy in validation set during the N_{ep} iteration
  Set p_g = 10% of the lowest absolute magnitude values in m_g^s to 0 and the others to 1, then obtain a new $m_g^{s+1,0}$
end while

3.7. Model setting

Let F_i represent the number of filters at each GCN level, given by $F_i \in [F_1, F_2, F_3, F_4, F_5, F_6]$. Similarly, K_i denotes the polynomial order of the filter for each ith layer, and is defined as $K_i \in [K_1, K_2, K_3, K_4, K_5, K_6]$. O indicates the number of MI classes for prediction. Due to the large volume of instances in the training set, we employed a mini-batch size B of 1024. A batch normalisation (BN) layer was incorporated after both the spectral GCN and Fully Connected layers. This BN layer re-scales and re-centres normalised signals to match the original distribution within the mini-batch, addressing the internal covariate shift issue and helping to mitigate the gradient vanishing/exploding problem. Additionally, 50% dropout layers were integrated after the ReLU layers (Eq. (14)) within the Fully Connected Block for regularisation. The details of the model settings can be found in Table 2, while the hyperparameter settings are provided in Table 3.

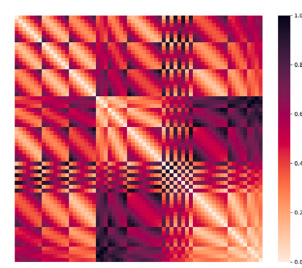


Fig. 3. Geodesic distance adjacency matrix ($A^{Geodesic}$).

$$ReLU(x) = max(0, x) \tag{14}$$

$$Softmax(\hat{y}_i) = \frac{e^{\hat{y}_i}}{\sum_{i=1}^{O} e^{\hat{y}_i}} \tag{15}$$

where $\hat{y}_i$ represent the predicted probability of an instance for each class, ranging over $\hat{y}_i \in [\hat{y}_1, \ldots, \hat{y}_O]$. O denotes the total number of classes. The loss function employed was the cross-entropy loss.

$$Loss = -\frac{1}{|B|} \sum_{b=1}^{B} \sum_{i=1}^{O} y_i.log(\hat{y}_i) \tag{16}$$

Both accuracy and F1 score evaluation metrics were employed to assess the performance of models.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{17}$$

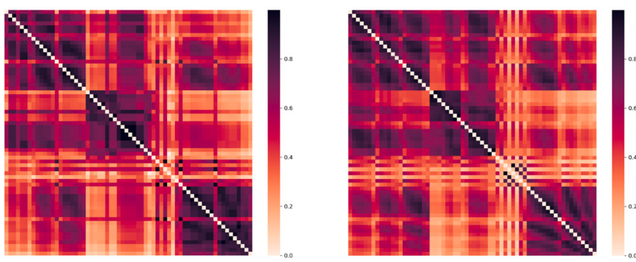
$$Recall = \frac{TP}{TP + FN} \tag{18}$$

$$Precision = \frac{TP}{TP + FP} \tag{19}$$

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{20}$$

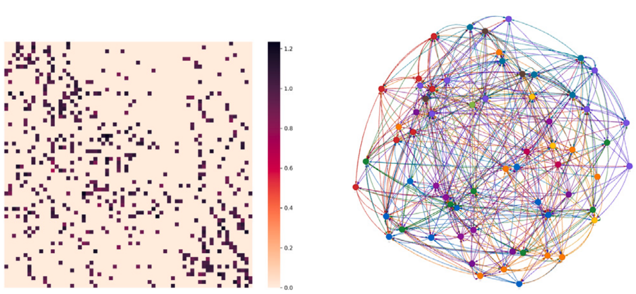
Table 2
Model setting.

Model	Model framework	Number of GCN filters	GCN filter polynomial order	Number of FC hidden nodes
A	$(C-BNC)\times 6-P-(FC-BNFC)\times 2-FC-S$	16, 32, 64, 128, 256, 512	5, 5, 5, 5, 5, 5	1024, 2048, 4
B	$(C-BNC)\times 6-P-(FC-BNFC)\times 2-FC-S$	16, 32, 64, 128, 256, 512	2, 2, 2, 2, 2, 2	1024, 2048, 4
C	$(C-BNC)\times 5-P-(FC-S)$	16, 32, 64, 128, 256	5, 5, 5, 5, 5	4
D	$(C-BNC)\times 5-P-(FC-S)$	16, 32, 64, 128, 256	2, 2, 2, 2, 2	4
E	$(C-BNC)\times 5-P-(FC-BNFC)\times 2-FC-S$	64, 128, 256, 512, 1024	5, 5, 5, 5, 5	512, 128, 4
F	$(C-BNC)\times 5-P-(FC-BNFC)\times 2-FC-S$	64, 128, 256, 512, 1024	2, 2, 2, 2, 2	512, 128, 4

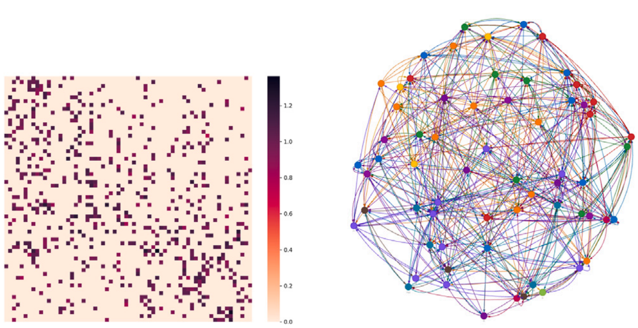


(a) Subject S_6 (b) Subject S_{14}

Fig. 4. PCC Adjacency matrix (A^{PCC}) of Subject S_6 and S_{14}.



(a) Adjacency Matrix ($m_{g\_EEG\_GLT}$) - Subject S_6, 13.39% Density, Model A (b) Graph of $m_{g\_EEG\_GLT}$ - S_6, 13.39% Density, Model A



(c) Adjacency Matrix ($m_{g\_EEG\_GLT}$) - Subject S_6, 13.39% Density, Model E (d) Graph of $m_{g\_EEG\_GLT}$ - S_6, 13.39% Density, Model E

Fig. 5. Representations of $m_{g\_EEG\_GLT}$ for Subject S_6 at 13.39% density. (a) Adjacency Matrix - Model A (Accuracy: 78.13%) (b) Graph - Model A (c) Adjacency Matrix - Model E (Accuracy: 73.55%) (d) Graph - Model E.

Table 3
Hyperparameter setting.

Hyperparameter	Value
Training Epochs (N_{ep})	1000
Batch Size (B)	1024
Dropout Rate	0.5
Optimiser	Adam
Initial Learning Rate (η)	0.01

4. Results and discussion

4.1. Geodesic vs PCC adjacency matrix construction method

The Table 4 presents the mean performance accuracy and F1 score across various models for different adjacency matrix construction methods, including Geodesic, PCC, and EEG_GLT, for each subject. Among the existing methods (PCC and Geodesic), the PCC adjacency method consistently outperformed the Geodesic method, enhancing the accuracy by 0.98%–22.60% and the F1 score by 0.99%–22.86%. Table 5 and Fig. 6 detail the mean accuracies and F1 scores for 20 subjects (S_1 - S_{20}) across different matrix construction methods for each model setting. Notably, the PCC method outperformed the Geodesic method across all model settings, improving accuracy by 9.76% and the F1 score by 9.63%. The superiority of the PCC method in EEG MI adjacency matrix construction over the Geodesic method stems a major limitation in the latter: it considers only the geodesic distance between EEG electrodes, leading to identical adjacency matrices for all 20 subjects (Fig. 3). In contrast, the PCC method produces unique matrices for each subject, offering tailored matrices that are better suited for subject-based EEG MI classification (Fig. 4). Our experiment revealed that using the relative physical distance between EEG electrodes was suboptimal due to limited accuracy. Since EEG electrodes do not have direct connections to brain tissue, electrical signals produced by large neuron groups that fire simultaneously or synchronously need to traverse multiple tissue layers such as the cerebral cortex, cerebrospinal fluid, skull, and scalp before detected by EEG electrodes. Given that the skull attenuates these signals, and causes a smearing effect [46], coupled with individual differences in skull thickness, scalp conductivity, and MI task approach, it was the most logical to use unique adjacency matrices for each individual.

In the $A^{Geodesic}$ adjacency matrix construction, we adopted a unit sphere assumption because the PhysioNet dataset lacks data on individual head shapes. Given natural variations in head structure, $A^{Geodesic}$ values could potentially differ for each subject.

4.2. EEG_GLT method vs PCC method in adjacency matrix construction

Our EEG_GLT method consistently surpassed the PCC method in both accuracy and F1 score. As shown in Table 4, EEG_GLT demonstrated substantial increase in accuracy and F1 score compared to

Table 4
Accuracy comparison across different methods of adjacency matrix construction for each subject.

Subject	Accuracy (Mean±Std)			F1 score (Mean±Std)		
	Geodesic	PCC	EEG_GLT (our method)	Geodesic	PCC	EEG_GLT (our method)
S_1	66.19% ± 4.17%	76.47% ± 9.94%	**98.51% ± 0.77%**	66.53% ± 4.36%	76.91% ± 9.78%	**98.53% ± 0.78%**
S_2	46.53% ± 1.33%	69.13% ± 7.05%	**76.18% ± 5.53%**	46.47% ± 1.46%	69.34% ± 7.37%	**76.19% ± 5.52%**
S_3	76.18% ± 4.98%	87.28% ± 9.19%	**99.17% ± 0.32%**	76.12% ± 5.00%	87.43% ± 8.97%	**99.19% ± 0.31%**
S_4	96.41% ± 1.97%	99.13% ± 1.01%	**99.97% ± 0.06%**	96.44% ± 1.98%	99.10% ± 1.12%	**99.97% ± 0.05%**
S_5	37.05% ± 1.04%	43.19% ± 3.03%	**50.95% ± 3.80%**	36.66% ± 0.97%	43.28% ± 2.73%	**50.86% ± 3.85%**
S_6	44.37% ± 1.59%	58.23% ± 5.19%	**69.60% ± 5.67%**	44.29% ± 1.65%	58.25% ± 5.49%	**69.50% ± 5.70%**
S_7	40.44% ± 1.19%	50.98% ± 3.80%	**59.45% ± 3.00%**	40.30% ± 1.23%	51.10% ± 3.49%	**59.34% ± 2.99%**
S_8	89.03% ± 7.04%	95.06% ± 5.96%	**99.95% ± 0.07%**	88.84% ± 6.88%	95.14% ± 5.81%	**99.96% ± 0.07%**
S_9	87.26% ± 14.26%	97.64% ± 3.33%	**99.95% ± 0.08%**	87.41% ± 14.49%	97.70% ± 3.78%	**99.95% ± 0.08%**
S_{10}	98.26% ± 0.31%	99.24% ± 0.19%	**99.99% ± 0.01%**	98.25% ± 0.32%	99.25% ± 0.20%	**99.99% ± 0.01%**
S_{11}	97.18% ± 1.12%	99.48% ± 0.70%	**99.99% ± 0.01%**	97.18% ± 1.13%	99.49% ± 0.74%	**99.99% ± 0.01%**
S_{12}	71.54% ± 3.44%	78.07% ± 8.95%	**99.69% ± 0.32%**	71.40% ± 3.37%	77.94% ± 8.76%	**99.70% ± 0.31%**
S_{13}	36.52% ± 0.32%	41.35% ± 1.23%	**44.50% ± 2.23%**	36.49% ± 0.45%	41.01% ± 1.34%	**44.47% ± 2.23%**
S_{14}	40.21% ± 1.80%	55.97% ± 6.47%	**72.39% ± 6.43%**	40.10% ± 1.88%	56.05% ± 6.57%	**72.71% ± 6.13%**
S_{15}	46.16% ± 1.28%	52.11% ± 3.96%	**67.55% ± 9.26%**	45.92% ± 1.93%	52.20% ± 3.66%	**67.52% ± 9.27%**
S_{16}	95.62% ± 3.87%	96.75% ± 5.00%	**99.98% ± 0.03%**	94.94% ± 5.25%	96.72% ± 5.07%	**99.98% ± 0.03%**
S_{17}	92.07% ± 8.10%	98.83% ± 2.33%	**99.98% ± 0.03%**	91.95% ± 8.31%	98.66% ± 2.76%	**99.98% ± 0.03%**
S_{18}	71.24% ± 5.96%	86.19% ± 10.95%	**99.92% ± 0.12%**	73.28% ± 3.28%	85.98% ± 11.10%	**99.93% ± 0.13%**
S_{19}	33.18% ± 0.40%	38.38% ± 2.27%	**41.41% ± 1.44%**	32.85% ± 0.32%	38.35% ± 2.32%	**41.27% ± 1.34%**
S_{20}	93.77% ± 2.08%	98.44% ± 0.68%	**99.94% ± 0.11%**	93.76% ± 2.06%	98.45% ± 0.72%	**99.95% ± 0.12%**

Table 5
Accuracy comparison across different methods of adjacency matrix construction for each model.

Model	Adj method	Avg. Accuracy	Avg. F1 score
Model A	Geodesic	70.70%	70.14%
	PCC	79.82%	79.77%
	EEG_GLT	**85.90%**	**85.89%**
Model B	Geodesic	70.70%	70.65%
	PCC	78.69%	78.32%
	EEG_GLT	**83.84%**	**83.80%**
Model C	Geodesic	65.49%	65.43%
	PCC	74.13%	74.41%
	EEG_GLT	**83.27%**	**83.28%**
Model D	Geodesic	62.97%	63.08%
	PCC	68.13%	68.05%
	EEG_GLT	**81.52%**	**81.48%**
Model E	Geodesic	69.20%	69.16%
	PCC	78.90%	78.88%
	EEG_GLT	**85.91%**	**85.88%**
Model F	Geodesic	69.34%	69.28%
	PCC	76.89%	77.26%
	EEG_GLT	**83.26%**	**83.36%**

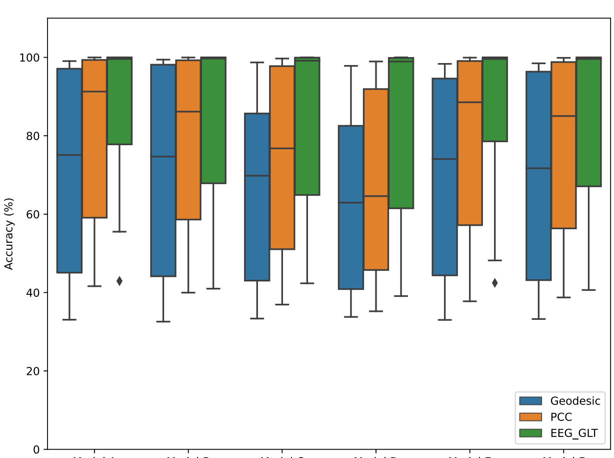


Fig. 6. Comparison of model accuracy across different adjacency matrix construction methods.

the PCC method, by 0.52%–22.04% and 0.50%–21.76%, respectively. Unlike the PCC method, our EEG_GLT adjacency matrix is dynamic with the ability to adapt to both the individual subject and the model settings of GCNs (Table 2), as shown in Fig. 5.

According to Table 5 and Fig. 6, our EEG_GLT method improved the mean accuracies and F1 scores for 20 subjects by 13.39% and 13.43%, respectively compared to the PCC method. This underscores the necessity of model-specific adjustments, in addition to subject-based tailoring in the adjacency matrix construction, to attain the best possible outcomes. Distinctly, our EEG_GLT matrix is asymmetrical due to the iterative pruning process detailed in Algorithm 1, which refines the matrix until the optimal EEG Graph Lottery Ticket is identified.

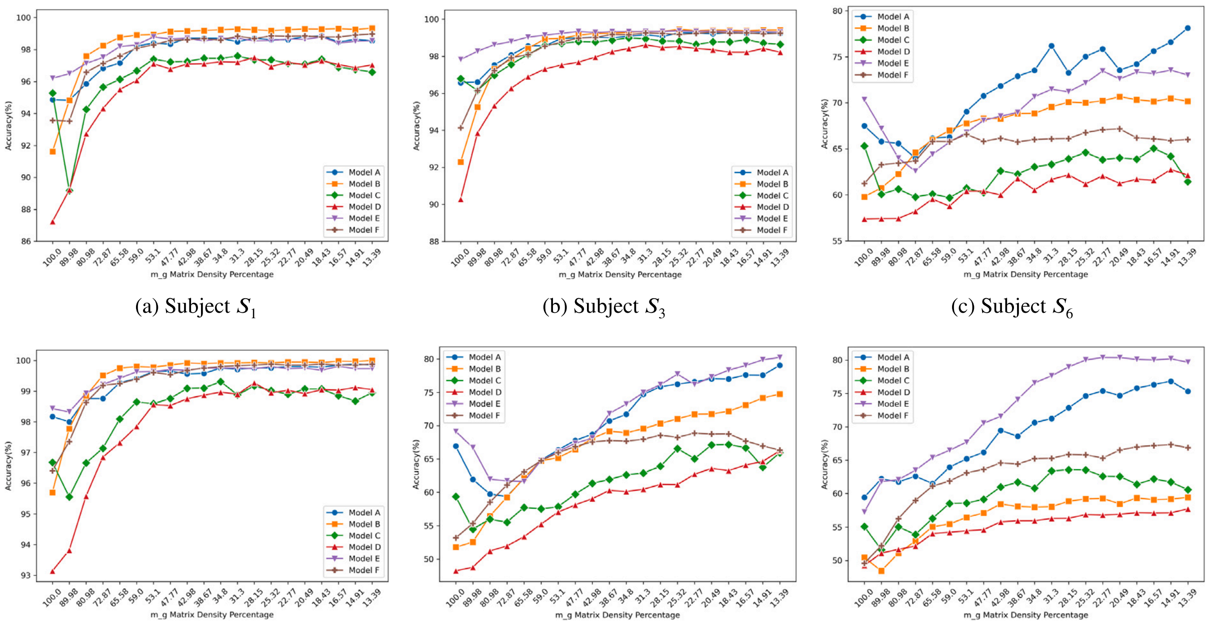
Fig. 7 presents the classification accuracy across various adjacency matrix densities for Subjects S_1, S_3, S_6, S_{12}, S_{14} and S_{15}. The data indicates an upward trend in classification accuracy with iterative pruning. Most importantly, the accuracy is notably lower at an adjacency matrix density of 100% in comparison to other densities. This observation suggests that some initial connections between EEG electrodes might be unnecessary, or even counterproductive, for achieving optimal classification. Removing these redundant links may boost the classification accuracy. Hence, a fully connected model between EEG channels may not be the most effective approach.

Table 5 displays the optimal EEG_GLT adjacency matrix ($m_{g\_EEG\_GLT}$) density for each subject. The transformation of the adjacency matrix mask m_g for the subjects S_6 and S_{14} at different densities is shown in Figs. 8 and 9 respectively. For subjects S_5, S_7, S_{13}, and S_{19}, their optimal $m_{g\_EEG\_GLT}$ were identified early at a 100% density. In contrast, other subjects attained their best results at densities below 22.53% for 2nd order models. When considering 5th order models, such as Model B, Model D, and Model F, the optimal EEG_GLTs emerged at densities of 59.00% or lower (see Table 6).

While our approach enhanced the accuracy for subjects S_5, S_7, S_{13}, and S_{19}, the results for both accuracy and F1 score lingered below 60.00%. A potential explanation is that relying on a single time point feature from EEG channels might not be adequate for MI tasks in these subjects, since there is inherent variability in the time required (or temporal dynamics) to execute the MI task among different individuals, as referenced in [47]. This variability might also explain why eliminating edges between EEG channels does not necessarily lead to improved performance accuracy for those subjects.

4.3. Model setting vs Adjacency matrix construction methods

Based on Table 5, for the Geodesic method, 2nd order GCN filters classify with higher average accuracy and F1 score than 5th order filters. However, for the PCC and EEG GLT methods, 5th order GCN filters



(a) Subject S_1 (b) Subject S_3 (c) Subject S_6

(d) Subject S_{12} (e) Subject S_{14} (f) Subject S_{15}

Fig. 7. Performance accuracy across different m_g densities using different models for Subjects S_1, S_3, S_6, S_{12}, S_{14} and S_{15} Accuracy vs m_g Densities.

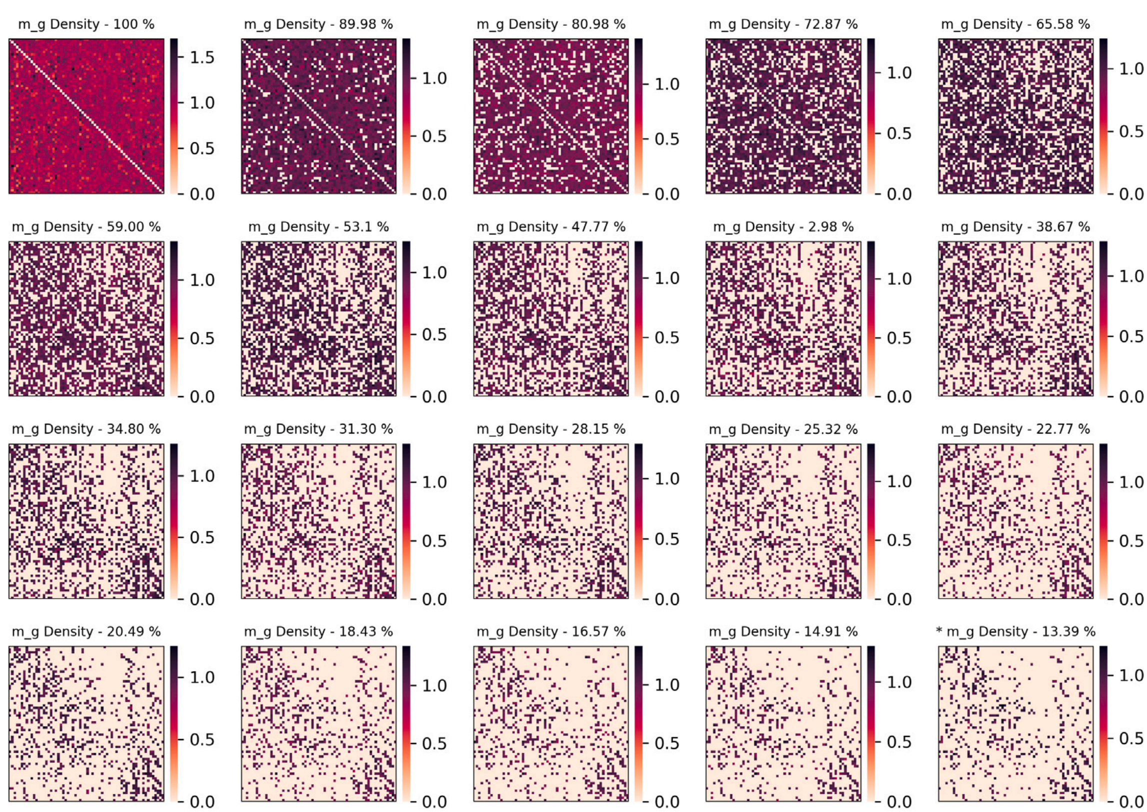


Fig. 8. EEG_GLT Adjacency matrix mask (m_g) of Subject S_6 at different densities using Model A. The m_g density at 13.39% produces the highest accuracy of 78.13%.

perform better. As highlighted in Section 4.2, our EEG_GLT method consistently achieves better accuracy than both the PCC and Geodesic methods. This remains the case even when the EEG_GLT adjacency matrix is paired with Model D, characterised by its minimal complexity, encompassing just five spectral GCN layers with 2nd order filters and a singular FC layer. These findings suggest that optimising the adjacency matrix is more importance than refining the GCN architecture when aiming for enhanced performance accuracy.

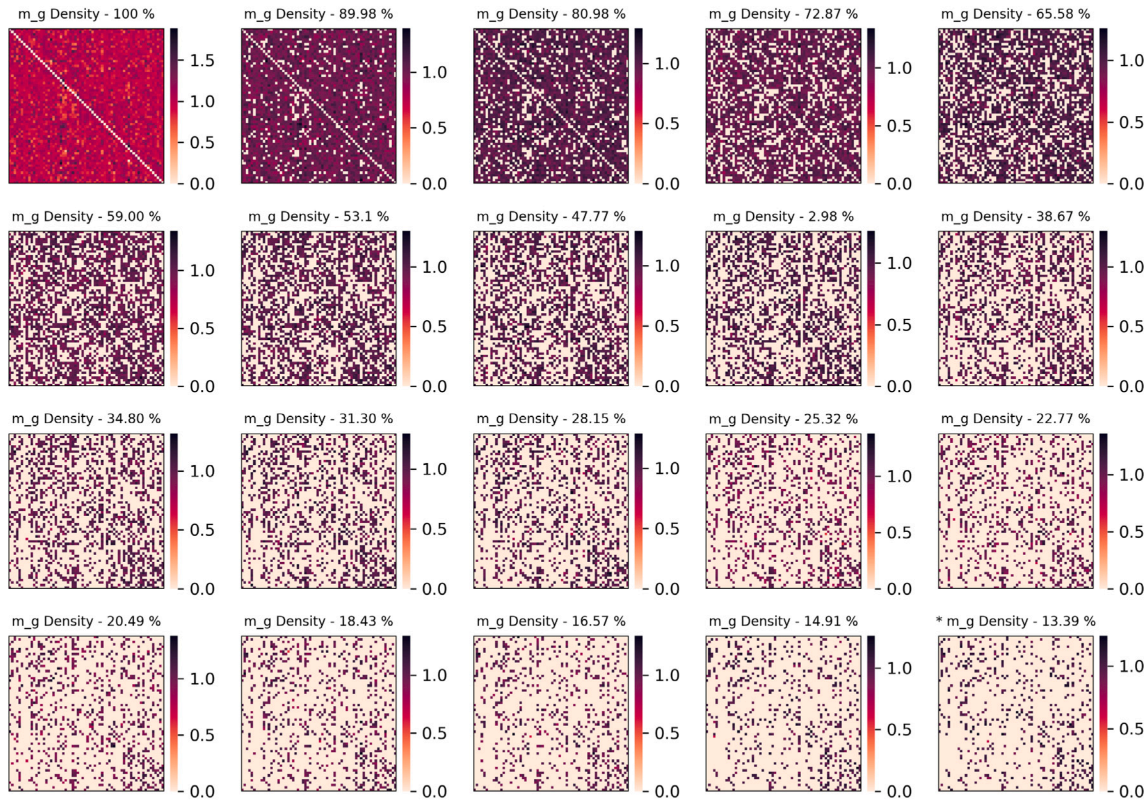


Fig. 9. EEG_GLT Adjacency matrix mask (m_g) of Subject S_{14} at different densities using Model A. The m_g density at 13.39% produces the highest accuracy of 79.06%.

Table 6
Optimal EEG_GLT adjacency matrix ($m_{g\_EEG\_GLT}$) density of each subject across models.

Subject	Model A	Model B	Model C	Model D	Model E	Model F
S_1	18.43%	13.39%	31.30%	28.15%	18.43%	13.39%
S_2	16.57%	13.39%	13.39%	28.15%	18.43%	25.32%
S_3	18.43%	25.32%	34.80%	31.30%	25.32%	20.49%
S_4	13.39%	13.39%	14.91%	20.49%	14.91%	13.39%
S_5	100.00%	31.30%	100.00%	100.00%	100.00%	100.00%
S_6	13.39%	20.49%	100.00%	14.91%	14.91%	20.49%
S_7	100.00%	28.15%	100.00%	31.30%	100.00%	59.00%
S_8	20.49%	18.43%	13.39%	14.91%	31.30%	14.91%
S_9	13.39%	16.57%	16.57%	14.91%	13.39%	13.39%
S_{10}	13.39%	13.39%	22.77%	20.49%	13.39%	13.39%
S_{11}	13.39%	13.39%	16.57%	13.39%	13.39%	13.39%
S_{12}	14.91%	13.39%	34.80%	28.15%	16.57%	13.39%
S_{13}	80.98%	34.80%	100.00%	20.49%	100.00%	22.77%
S_{14}	13.39%	13.39%	18.43%	13.39%	13.39%	22.77%
S_{15}	14.91%	13.39%	28.15%	13.39%	22.77%	14.91%
S_{16}	14.91%	13.39%	20.49%	18.43%	13.39%	13.39%
S_{17}	14.91%	13.39%	20.49%	22.77%	13.39%	13.39%
S_{18}	14.91%	13.39%	28.15%	20.49%	22.77%	31.30%
S_{19}	100.00%	59.00%	100.00%	22.77%	100.00%	31.30%
S_{20}	25.32%	22.77%	34.80%	16.57%	20.49%	34.80%

Table 7
MACs savings (%) for each subject: PCC's best model accuracy vs. EEG_GLT accuracy from models with adjacency matrix densities just surpassing PCC's best accuracy.

Subj	PCC			EEG_GLT			MACs
	Model	Acc.	MACs	Model (Adj%)	Acc.	MACs	Saving
S_1	A	87.66%	81.89M	D (13.39%)	97.04%	**8.76M**	89.30%
S_2	B	75.43%	42.26M	B (13.39%)	78.09%	**36.97M**	12.52%
S_3	A	94.89%	81.89M	D (13.39%)	98.22%	**8.76M**	89.30%
S_4	A	99.88%	81.89M	B (13.39%)	99.98%	**36.97M**	54.85%
S_5	B	46.90%	42.26M	B (13.39%)	48.73%	**36.97M**	12.52%
S_6	E	62.92%	291.62M	B (13.39%)	70.17%	**36.97M**	87.32%
S_7	E	55.04%	291.62M	B (13.39%)	57.68%	**36.97M**	87.32%
S_8	B	98.71%	42.26M	D (13.39%)	99.78%	**8.76M**	79.27%
S_9	A	99.86%	81.89M	B (13.39%)	99.98%	**36.97M**	54.85%
S_{10}	E	99.44%	291.62M	D (13.39%)	99.97%	**8.76M**	97.00%
S_{11}	E	99.90%	291.62M	D (13.39%)	99.98%	**8.76M**	97.00%
S_{12}	A	86.76%	81.89M	D (13.39%)	99.05%	**8.76M**	89.30%
S_{13}	A	42.79%	81.89M	B (13.39%)	43.57%	**36.97M**	54.85%
S_{14}	B	63.58%	42.26M	D (13.39%)	66.25%	**8.76M**	79.29%
S_{15}	E	57.01%	291.62M	D (13.39%)	57.72%	**8.76M**	97.00%
S_{16}	B	99.80%	42.26M	D (13.39%)	99.85%	**8.76M**	79.27%
S_{17}	A	99.98%	81.89M	B (13.39%)	100.00%	**36.97M**	44.93%
S_{18}	A	96.05%	81.89M	D (16.57%)	99.58%	**8.76M**	76.14%
S_{19}	A	41.62%	81.89M	A (89.98%)	41.78%	**80.67M**	1.49%
S_{20}	B	99.17%	42.26M	D (13.39%)	99.68%	**8.76M**	79.27%

4.4. MACs saving using EEG_GLT method

The MACs inference for classifying a single-time-point EEG MI signal is influenced by several model settings, including the model framework, the number and polynomial order of GCN filters, and the specifications of FC layers as the number of layers and the node count. Among these, the count and polynomial orders of GCN filters at the GCN layers are the primary determinants of the MACs requirement. Both $A^{Geodesic}$ and A^{PCC} maintain 100% densities in their adjacency matrices. Consequently, the MACs inference for a single-time-point EEG MI signal, when using models A to F, are as follows: 81.89M, 42.26M, 22.64M, 11.32M, 291.62M, and 146.10M, respectively.

Our EEG_GLT method presents varied A^{EEG_GLT} densities due to the pruning employed by Algorithm 1. As elaborated in Section 4.2, the EEG_GLT approach enhances classification accuracy through pruning, which in turn decreases the MACs. Table 7 illustrates the percentage of MACs savings for each subject, comparing the top accuracy value from the PCC method to the EEG_GLT accuracies from models with adjacency matrix densities slightly exceeding PCC's best.

For performance equivalent to or surpassing PCC's optimal accuracy, only Models D and B with the sparsest adjacency matrix density (13.39%) are necessary. The PCC method requires between 42.26M

Table 8
Performance comparisons state-of-the-art models.

Method	Avg. Accuracy	Avg. F1 score
FBCSP [25]	59.56%	60.04%
EEGNet [27]	72.20%	72.10%
CasCNN [28]	63.30%	63.18%
DG-HAM [43]	76.15%	76.08%
EEG-ARNN [44]	82.39%	82.17%
SSDA [29]	83.73%	83.24%
GCNs-Net [19]	80.16%	80.05%
Proposed EEG_GLT-Net	**86.43%**	**86.23%**

to 291.62M for one-time-point inference across 20 subjects to reach peak accuracy. In contrast, our EEG_GLT approach needs only 8.76M to 80.67M to achieve equal or better accuracy, translating to savings in MACs of up to 97.00%.

4.5. Comparison with current state-of-the-art models

In this paper, we compare our proposed method, EEG_GLT-Net, with seven other state-of-the-art (SOTA) models listed in Table 8, including FBCSP [25], EEGNet [27], CasCNN [28], DG-HAM [43], EEG-ARNN [44], SSDA [29], and GCNs-Net [19]. Our comparisons begin with the traditional FBCSP approach, which leverages CSP to extract features across multiple frequency bands and utilises SVM for classification. We then compare with EEGNet, a widely used model based solely on a CNN structure. Further, we assess CasCNN and SSDA, both of which combine CNN and LSTM networks. Finally, we evaluate our method against DG-HAM, EEG-ARNN, and GCNs-Net, which are GNN-based networks.

The traditional FBCSP method achieves 59.56%, the lowest accuracy among the SOTAs, likely due to its reliance on SVM as the classifier. The popular EEGNet achieves 72.20% accuracy, outperforming the CasCNN model, which achieves only 63.30%. Within the CNN-based SOTA models, SSDA reaches the highest accuracy at 83.73%. Among the GNN-based SOTA models, EEG-ARNN achieves the highest accuracy at 82.39%, followed by GCNs-Net and DG-HAM with accuracies of 80.16% and 76.15%, respectively.

From the perspective of adjacency matrix construction methods, the trainable adjacency matrix in EEG-ARNN outperforms the geodesic-based DG-HAM and PCC-based GCNs-Net. Our proposed EEG_GLT-Net, using single time-point classification at intervals of $(\frac{1}{160})$s, achieves the highest overall accuracy of 86.43% among all SOTAs. Notably, GCNs-Net is the only other model employing single time-point classification; however, while the GCNs-Net accuracy falls short of our EEG_GLT-Net using the EEG_GLT adjacency matrix, it surpasses our model when using a PCC-based adjacency matrix, reaching 79.82% which may be attributed to the application of pooling layer after every GNN layer within GCNs-Net.

4.6. Limitations and future works

In this paper, we introduced a novel method for constructing an adjacency matrix in GNNs to classify single time-point EEG motor imagery (MI) signals. While the single time-point classification at $(\frac{1}{160}s)$ using the EEG_GLT-Net architecture performs well for MI tasks with distinct neural patterns, such as those in the PhysioNet dataset, it is less effective for more subtle and overlapping motor imagery patterns with inherent temporal information, as seen in datasets like BCICIV_2a. In such cases, using longer signal segments typically exceeding one second and incorporating feature extraction are essential to capture meaningful temporal and spatial patterns for accurate classification.

However, the proposed EEG_GLT adjacency matrix construction method is not restricted to single time-point classification. In future work, we plan to extend EEG_GLT to other benchmark EEG MI datasets, such as BCICIV_2a, and EEG movement datasets, including the High-Gamma dataset. This extension will involve adapting the model architecture to incorporate temporal data embeddings suitable for each dataset.

5. Conclusion

Our EEG_GLT approach, developed for optimal adjacency matrix construction in EEG MI time-point signal classification, consistently outperforms both the Geodesic and PCC methods in accuracy and F1 score. It is important to note that the PCC method is currently employed in the state-of-the-art EEG time-point classification model, GCNs-Net. Specifically, our EEG_GLT method enhances accuracy and F1 score by margins ranging from 0.52% to 22.04% and 0.50% to 21.76%, respectively, compared to PCC. Furthermore, it improves the average accuracy across 20 subjects by 13.39%. With this method, optimal outcomes emerge when the adjacency matrix densities remain below 22.53%. Our study emphasises the pivotal role played by the configuration of the adjacency matrix in performance accuracy, overshadowing even model settings. In addition, our EEG_GLT approach has much higher computational efficiency, demanding between 8.76M and 80.67M MACs, which is significantly less than the 42.26M to 291.62M required by the PCC method for comparable or superior results.

While this research primarily focuses on identifying the optimal adjacency matrix, with pruning confined to the adjacency matrix, upcoming studies will explore pruning GNN and FC layers weights to further streamline computational costs. Additionally, we plan to expand the number of time points used for feature extraction, especially for subjects S_5, S_7, S_{13}, and S_{19}. In future work, we will refine Algorithm 1 to seamlessly integrate pooling layers within the GCN blocks under the EEG_GLT method, to further optimise computational efficiency. To achieve a more generalised understanding of the inter-relationships between EEG channels, it is essential to incorporate a broader range of tasks into models.

CRediT authorship contribution statement

Htoo Wai Aung: Writing – original draft, Software, Methodology, Data curation, Conceptualization. **Jiao Jiao Li:** Writing – review & editing, Validation, Supervision. **Bin Shi:** Supervision, Project administration, Funding acquisition. **Yang An:** Writing – review & editing, Methodology. **Steven W. Su:** Writing – review & editing, Supervision, Project administration, Methodology, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] Jonathan R Wolpaw, Niels Birbaumer, Dennis J McFarland, Gert Pfurtscheller, Theresa M Vaughan, Brain–computer interfaces for communication and control, Clin. Neurophysiol. 113 (6) (2002) 767–791.

[2] Mikhail A. Lebedev, Miguel A.L. Nicolelis, Brain–machine interfaces: past, present and future, Trends Neurosci. 29 (9) (2006) 536–546.

[3] Donald L. Schomer, Fernando Lopes Da Silva, Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, Lippincott Williams & Wilkins, 2012.

[4] Jason Hubbard, Atsushi Kikumoto, Ulrich Mayr, EEG decoding reveals the strength and temporal dynamics of goal-relevant representations, Sci. Rep. 9 (1) (2019) 9051.

[5] Dennis J McFarland, Laurie A Miner, Theresa M Vaughan, Jonathan R Wolpaw, Mu and beta rhythm topographies during motor imagery and actual movements, Brain Topogr. 12 (2000) 177–186.
[6] Marc Jeannerod, The representing brain: Neural correlates of motor intention and imagery, Behav. Brain Sci. 17 (2) (1994) 187–202.
[7] Andrea Biasiucci, Robert Leeb, Iñaki Iturrate, Serafeim Perdikis, Abdul Al-Khodairy, Tiffany Corbet, Armin Schnider, T Schmidlin, Huaijian Zhang, Manuela Bassolino, et al., Brain-actuated functional electrical stimulation elicits lasting arm motor recovery after stroke, Nature Commun. 9 (1) (2018) 2421.
[8] Clement Farabet, Camille Couprie, Laurent Najman, Yann LeCun, Learning hierarchical features for scene labeling, IEEE Trans. Pattern Anal. Mach. Intell. 35 (8) (2012) 1915–1929.
[9] Yann LeCun, Léon Bottou, Yoshua Bengio, Patrick Haffner, Gradient-based learning applied to document recognition, Proc. IEEE 86 (11) (1998) 2278–2324.
[10] Yann LeCun, Koray Kavukcuoglu, Clément Farabet, Convolutional networks and applications in vision, in: Proceedings of 2010 IEEE International Symposium on Circuits and Systems, IEEE, 2010, pp. 253–256.
[11] Ziwei Zhang, Peng Cui, Wenwu Zhu, Deep learning on graphs: A survey, IEEE Trans. Knowl. Data Eng. 34 (1) (2020) 249–270.
[12] Rui Zhang, Ziyang Wang, Fangmei Yang, Yu Liu, Recognizing the level of organizational commitment based on deep learning methods and EEG, in: ITM Web of Conferences, vol. 47, EDP Sciences, 2022, p. 02044.
[13] Manhua Jia, Wenjian Liu, Junwei Duan, Long Chen, CL Chen, Qun Wang, Zhiguo Zhou, Efficient graph convolutional networks for seizure prediction using scalp EEG, Front. Neurosci. 16 (2022) 967116.
[14] Neeraj Wagh, Yogatheesan Varatharajah, Eeg-gcnn: Augmenting electroencephalogram-based neurological disease diagnosis using a domain-guided graph convolutional neural network, in: Machine Learning for Health, PMLR, 2020, pp. 367–378.
[15] Guangcheng Bao, Kai Yang, Li Tong, Jun Shu, Rongkai Zhang, Linyuan Wang, Bin Yan, Ying Zeng, Linking multi-layer dynamical GCN with style-based recalibration CNN for EEG-based emotion recognition, Front. Neurorobotics 16 (2022) 834952.
[16] Weifeng Ma, Chuanlai Wang, Xiaoyong Sun, Xuefen Lin, Yuchen Wang, A double-branch graph convolutional network based on individual differences weakening for motor imagery EEG classification, Biomed. Signal Process. Control 84 (2023) 104684.
[17] Lu Meng, Jinzhou Hu, Yu Deng, Yue Hu, Electrical status epilepticus during sleep electroencephalogram waveform identification and analysis based on a graph convolutional neural network, Biomed. Signal Process. Control 77 (2022) 103788.
[18] Nastaran Khaleghi, Tohid Yousefi Rezaii, Soosan Beheshti, Saeed Meshgini, Developing an efficient functional connectivity-based geometric deep network for automatic EEG-based visual decoding, Biomed. Signal Process. Control 80 (2023) 104221.
[19] Yimin Hou, Shuyue Jia, Xiangmin Lun, Ziqian Hao, Yan Shi, Yang Li, Rui Zeng, Jinglei Lv, GCNs-net: A graph convolutional neural network approach for decoding time-resolved EEG motor imagery signals, IEEE Trans. Neural Netw. Learn. Syst. 35 (6) (2024) 7312–7323.
[20] Xuefen Lin, Jielin Chen, Weifeng Ma, Wei Tang, Yuchen Wang, EEG emotion recognition using improved graph neural network with channel selection, Comput. Methods Programs Biomed. 231 (2023) 107380.
[21] Tengfei Song, Wenming Zheng, Peng Song, Zhen Cui, EEG emotion recognition using dynamical graph convolutional neural networks, IEEE Trans. Affect. Comput. 11 (3) (2018) 532–541.
[22] Tianlong Chen, Yongduo Sui, Xuxi Chen, Aston Zhang, Zhangyang Wang, A unified lottery ticket hypothesis for graph neural networks, in: International Conference on Machine Learning, PMLR, 2021, pp. 1695–1706.
[23] Sachin Taran, Varun Bajaj, Dheeraj Sharma, Siuly Siuly, A. Sengur, Features based on analytic IMF for classifying motor imagery EEG signals in BCI applications, Measurement 116 (2018) 68–76.
[24] Tianyou Yu, Jun Xiao, Fangyi Wang, Rui Zhang, Zhenghui Gu, Andrzej Cichocki, Yuanqing Li, Enhanced motor imagery training using a hybrid BCI with feedback, IEEE Trans. Biomed. Eng. 62 (7) (2015) 1706–1717.
[25] Kai Keng Ang, Zheng Yang Chin, Haihong Zhang, Cuntai Guan, Filter bank common spatial pattern (FBCSP) in brain-computer interface, in: 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), 2008, pp. 2390–2397.
[26] Robin Tibor Schirrmeister, Jost Tobias Springenberg, Lukas Dominique Josef Fiederer, Martin Glasstetter, Katharina Eggensperger, Michael Tangermann, Frank Hutter, Wolfram Burgard, Tonio Ball, Deep learning with convolutional neural networks for EEG decoding and visualization, Hum. Brain Mapp. 38 (11) (2017) 5391–5420.
[27] Vernon J Lawhern, Amelia J Solon, Nicholas R Waytowich, Stephen M Gordon, Chou P Hung, Brent J Lance, EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces, J. Neural Eng. 15 (5) (2018) 056013.
[28] Dalin Zhang, Lina Yao, Kaixuan Chen, Sen Wang, Xiaojun Chang, Yunhao Liu, Making sense of spatio-temporal preserving representations for EEG-based human intention recognition, IEEE Trans. Cybern. 50 (7) (2020) 3033–3044.
[29] Shadi Sartipi, Mujdat Cetin, Subject-independent deep architecture for EEG-based motor imagery classification, IEEE Trans. Neural Syst. Rehabil. Eng. 32 (2024) 718–727.
[30] Jialing Wang, Shiwei Cheng, Jieming Tian, Yuefan Gao, A 2D CNN-LSTM hybrid algorithm using time series segments of EEG data for motor imagery classification, Biomed. Signal Process. Control 83 (2023) 104627.
[31] Zhige Chen, Rui Yang, Mengjie Huang, Fumin Li, Guoping Lu, Zidong Wang, EEGProgress: A fast and lightweight progressive convolution architecture for EEG classification, Comput. Biol. Med. 169 (2024) 107901.
[32] Omair Ali, Muhammad Saif ur Rehman, Tobias Glasmachers, Ioannis Iossifidis, Christian Klaes, ConTraNet: A hybrid network for improving the classification of EEG and EMG signals with limited training data, Comput. Biol. Med. 168 (2024) 107649.
[33] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, S Yu Philip, A comprehensive survey on graph neural networks, IEEE Trans. Neural Netw. Learn. Syst. 32 (1) (2020) 4–24.
[34] Joan Bruna, Wojciech Zaremba, Arthur Szlam, Yann LeCun, Spectral networks and locally connected networks on graphs, 2013, arXiv Preprint arXiv:1312.6203.
[35] Michaël Defferrard, Xavier Bresson, Pierre Vandergheynst, Convolutional neural networks on graphs with fast localized spectral filtering, Adv. Neural Inf. Process. Syst. 29 (2016).
[36] Ron Levie, Federico Monti, Xavier Bresson, Michael M Bronstein, Cayleynets: Graph convolutional neural networks with complex rational spectral filters, IEEE Trans. Signal Process. 67 (1) (2018) 97–109.
[37] Will Hamilton, Zhitao Ying, Jure Leskovec, Inductive representation learning on large graphs, Adv. Neural Inf. Process. Syst. 30 (2017).
[38] Federico Monti, Davide Boscaini, Jonathan Masci, Emanuele Rodola, Jan Svoboda, Michael M Bronstein, Geometric deep learning on graphs and manifolds using mixture model cnns, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 5115–5124.
[39] Mathias Niepert, Mohamed Ahmed, Konstantin Kutzkov, Learning convolutional neural networks for graphs, in: International Conference on Machine Learning, PMLR, 2016, pp. 2014–2023.
[40] Hongyang Gao, Zhengyang Wang, Shuiwang Ji, Large-scale learnable graph convolutional networks, in: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2018, pp. 1416–1424.
[41] David I Shuman, Sunil K Narang, Pascal Frossard, Antonio Ortega, Pierre Vandergheynst, The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains, IEEE Signal Process. Mag. 30 (3) (2013) 83–98.
[42] Difei Zeng, Kejie Huang, Cenglin Xu, Haibin Shen, Zhong Chen, Hierarchy graph convolution network and tree classification for epileptic detection on electroencephalography signals, IEEE Trans. Cogn. Dev. Syst. 13 (4) (2020) 955–968.
[43] Dalin Zhang, Lina Yao, Kaixuan Chen, Sen Wang, Pari Delir Haghighi, Caley Sullivan, A graph-based hierarchical attention model for movement intention detection from EEG signals, IEEE Trans. Neural Syst. Rehabil. Eng. 27 (11) (2019) 2247–2253.
[44] Biao Sun, Zhengkun Liu, Zexu Wu, Chaoxu Mu, Ting Li, Graph convolution neural network based end-to-end channel selection and classification for motor imagery brain–computer interfaces, IEEE Trans. Ind. Inform. 19 (9) (2023) 9314–9324.
[45] Ary L Goldberger, Luis AN Amaral, Leon Glass, Jeffrey M Hausdorff, Plamen Ch Ivanov, Roger G Mark, Joseph E Mietus, George B Moody, Chung-Kang Peng, H Eugene Stanley, PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals, Circulation 101 (23) (2000) e215–e220.
[46] Sebastianus Petrus van den Broek, F Reinders, M Donderwinkel, MJ Peters, Volume conduction effects in EEG and MEG, Electroencephalogr. Clin. Neurophysiol. 106 (6) (1998) 522–534.
[47] Catherine Tallon-Baudry, Olivier Bertrand, Oscillatory gamma activity in humans and its role in object representation, Trends in Cognitive Sciences 3 (4) (1999) 151–162.