Computing and Informatics, Vol. 45, 2026, 410–438, doi: 10.31577/cai_2026_2_410

FEDERATED GRAPH LEARNING VIA STRUCTURAL KNOWLEDGE SHARING FOR HYBRID BRAIN-COMPUTER INTERFACE

KaiZheng Li, HongFei Ji*

Translational Research Center
Shanghai Yangzhi Rehabilitation Hospital
(Shanghai Sunshine Rehabilitation Center)
&
School of Computer Science and Technology
Tongji University, Shanghai 201804, China
e-mail: `likaizheng@tongji.edu.cn, jhf@tongji.edu.cn`

Wei Yang

College of Economics and Management
Shanghai Ocean University, Shanghai 201306, China
e-mail: `wyang@shou.edu.cn`

Jie Li

Translational Research Center
Shanghai Yangzhi Rehabilitation Hospital
(Shanghai Sunshine Rehabilitation Center)
&
School of Computer Science and Technology
Tongji University, Shanghai 201804, China
e-mail: `nijanice@163.com`

Abstract. Deep Learning-based electroencephalogram (EEG) decoding in brain-computer interface (BCI) faces many challenges: First, EEG has low spatial reso-

* Corresponding author

lution, which makes it difficult to accurately identify specific brain regions; Second, deep learning requires large amounts of data, while EEG datasets are typically small. Merging small datasets raises privacy concerns because EEG signals contain a wealth of personal information. Finally, while federated learning (FL) offers a potential solution for addressing the issues of limited data and privacy, it faces the challenge of non-independent and identically distributed (non-IID) data, as the local data on the client is often derived from datasets of different tasks or subjects. To solve these problems, this paper proposes a novel Federated Graph Learning (FGL) framework based on Electroencephalogram and Electromyography (EEG-EMG), named FSDFGL. During the construction of the graph data, EMG data is incorporated into EEG data through SPMI, as EMG signals provide distinct features of muscle activity. In the local Graph Neural Network (GNN) process, the structural information is decoupled from the node features. Then, the structure encoder is used to capture the structural information and share it globally during FL. This allows FSDFGL to capture more commonly shared structural information while avoiding degradation of local model performance due to sharing heterogeneous features. The framework was evaluated on the TAN dataset, and experimental results demonstrate the superiority of FSDFGL in identifying complex motion intentions.

Keywords: Brain-computer interface, graph neural network, federated learning, non-independent and identically distributed

1 INTRODUCTION

Motor imagery (MI) refers to the mental process of envisioning movement without actually physically executing it [1, 2, 3, 4, 5]. BCIs based on motor imagination have become one of the hottest topics in the field of rehabilitation medicine [6, 7]. Through motor imagination training, the motor network of the brain can be activated [8] to promote brain plasticity.

Studies have shown that specific EEG patterns are generated in different brain regions during specific movement imagination, and these patterns can be analyzed to capture movement intentions. Patients can control external devices or virtual scenes through EEG-based MI to achieve better rehabilitation outcomes [9]. However, the spatial resolution of EEG signals is relatively low, and EEG signals cannot provide enough information to accurately identify complex motor intentions [10].

EMG signals have a higher spatial resolution than EEG. They can capture rapid changes in muscle activity, provide more detailed movement features, and enable more precise identification of movement intentions [11, 12, 13, 14, 15]. Therefore, identifying complex motor intentions based on EEG and EMG signals by placing electrodes on the scalp and skin surface is a more comprehensive and persuasive biological approach. In recent years, complex motor intention recognition based on multi-source signals has attracted a lot of attention; some studies utilize EEG to detect motor intentions and then analyze the actual execution using EMG [16, 17]. Others aim to improve classification accuracy by integrating EEG and EMG models at the model level [18, 19]. There are also studies that use EMG to remove motion artefacts from EEG, improving signal quality for better classification [20].

GNN is a deep learning model used to process graph-structured data. EEG-EMG data can be viewed as a graph structure, where each node represents an electrode or brain region, and edges represent connections between electrodes. GNN effectively captures and models the connections and interactions between these electrodes to better understand and analyze the information in EEG-EMG data. The application of GNNs in the field of BCI is increasing [21]. Complex graph-structured data is often difficult to analyze using traditional methods; however, it can be efficiently processed by GNNs. Demir et al. [22] demonstrated the superiority of the GNN model in modeling functional neural connections. Aung et al. [23] proposed the EEG Graph Lottery Ticket (EEG_GLT) algorithm to optimize the construction of EEG channel adjacency matrices, achieving a performance improvement of 13.39 % over the PCC method.

GNNs based on EEG and EMG have been widely used in BCI systems. However, a large-scale dataset is highly needed to achieve good generalization and performance before constructing a deep learning model. Due to the expensive data collection process, EEG-BCI data is often available as multiple small datasets that are widely distributed. Additionally, privacy concerns make it challenging to create a sufficiently large dataset by aggregating these small datasets. As EEG signals contain individuals' electroencephalographic activity, which may involve personal privacy, according to privacy regulations such as GDPR [24], EEG data should be de-identified or anonymized before sharing to protect personal identity and privacy information. Motivated by the privacy-preserving characteristics of FL [25], it allows for model training on local devices without transmitting raw data to a central server. Therefore, FGL is constructed based on the FL framework to realize the co-training of the GNN model.

One of the biggest challenges in FL is the non-IID problem; each participant may possess datasets related to different tasks or subjects, which exhibit distinct feature distributions. For example, people with neurological diseases such as epilepsy, Parkinson's disease, or Alzheimer's disease, will have different electrical activity in their brains compared to normal individuals. In addition, different cognitive or motor tasks activate different regions of the brain, resulting in differences in EEG features. However, traditional FL methods like Fed-Avg still rely on feature-based message passing schemes for communication and aggregation. In FGL procedures, feature heterogeneity may lead to conflict or incompatibility of model parameters of different participants during aggregation, ultimately leading to a degradation of the model's performance on the overall data. Therefore, addressing the non-IID problem while ensuring data security is a worthwhile research topic.

Although non-IID data distributions are inevitable in FL, there are indeed some general structural properties in EEG and EMG signals [26, 27, 28], such as small-world properties, degree distributions of nodes, and centrality features. These characteristics inspire us to exploit a model relying on structure-based message passing schemes. As shown in Figure 1 b), graphs generated by different subjects have high structural similarity among each other, while graphs generated by different tasks also show some similarity compared to random graphs. Therefore, a novel FGL framework is proposed based on structural knowledge sharing. It is called FSDFGL for motor intentions recognition. Our goal is to improve the performance of local clients by extracting and sharing universal knowledge across multiple clients while preserving data privacy.

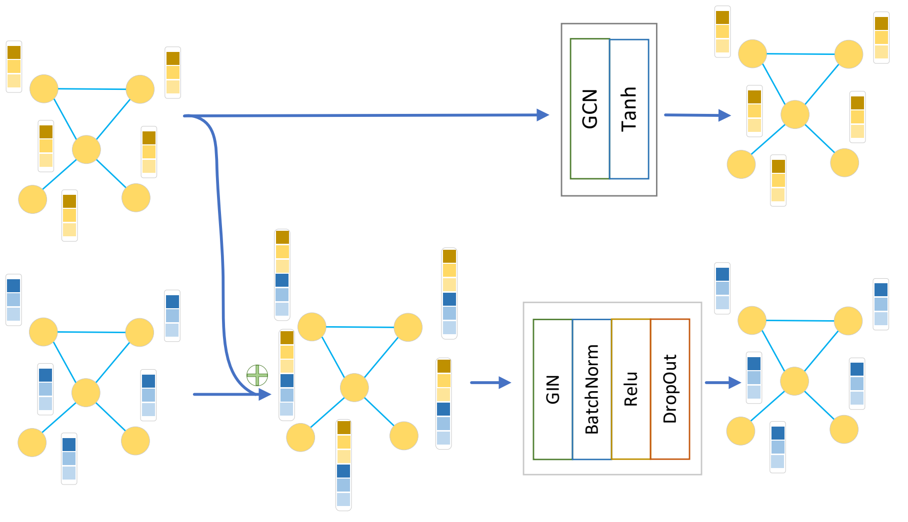
The main contributions of this article can be summarized as follows:

1. By combining EEG and EMG signals, we address the limitations of EEG signals in spatial resolution. Extensive experiments on the TAN dataset [30] were conducted to compare the performance of single-source and multi-source signals. The results show that using EEG and EMG improves the average accuracy by 11.6 % compared to EEG alone and by 16.3 % compared to EMG alone.
2. We extracted node degree and PageRank (PR) from graph structure data to enhance node understanding and provide structured information for FL. Extensive experiments on the TAN dataset showed that models using either degree, PR, or both features, outperformed those with none, with the model using both achieving the highest average accuracy.
3. We designed a dual-channel GNN with Graph Convolutional Network (GCN) for node structural understanding and Graph Isomorphism Network (GIN) for node features. GCN parameters are used in FL to avoid sharing heterogeneous information, while the hidden embeddings from the previous GCN layer and the GIN layer serve as inputs for the next GIN layer, enriching the feature representation. Extensive experiments show that the dual-channel GNN achieves 2.24 % higher accuracy on the TAN dataset compared to the single-channel GNN, demonstrating the effectiveness of our approach.
4. We propose a novel FGL framework that exploits structural knowledge sharing to improve model accuracy while preserving data privacy. To verify the robustness of our method in the non-IID scenario of the EEG-EMG domain, we conduct comprehensive experiments on the TAN and DEAP [31] datasets. The results show that FSDFGL significantly outperforms the baseline methods.

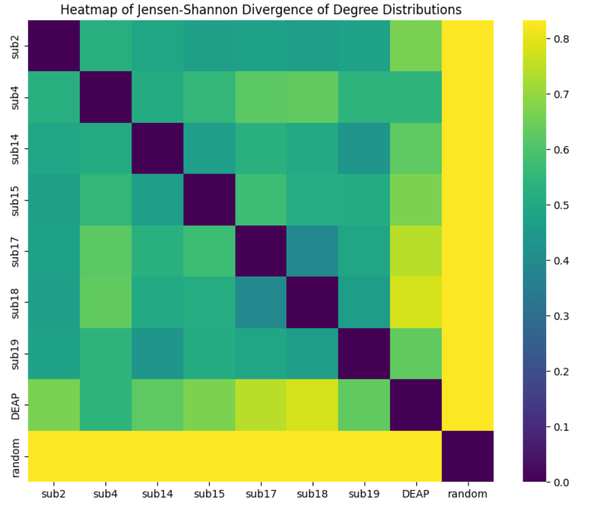
2 PRELIMINARIES

2.1 Graph Neural Networks for BCIs

GNNs have received a lot of attention in recent years due to their ability to capture complex relationships and dependencies in many domains, such as natural language processing, social networks, and biological research. Recently researchers have started using GNNs on BCI. Many GNN models map EEG electrodes to nodes in a graph, with node features representing the signal samples collected from EEG



a)



b)

Figure 1. a) The specific process of dual-channel GNN, b) Heatmaps of Jensen-Shannon divergence of degree distributions for seven subject datasets from TAN, one subject dataset from DEAP, and random graphs. The random graphs are generated using the Erdős-Rényi model [29]

channels. The connections between nodes are defined according to strategies designed by neuroscientists, typically based on the functional correlation or signal synchrony between electrodes. Through this graph structure, GNNs can effectively capture the complex connectivity patterns between brain regions, thereby enhancing the classification performance of motor intention and cognitive tasks, surpassing traditional deep learning methods. For example, Ding et al. [32] successfully developed the Local-Global-Graph network by incorporating prior knowledge from neuroscience into neural network design. Massa et al. [33] applied GNNs to high-density EMG data to recognize action intentions and achieved a low classification error rate of only 8.75 % in 65 gestures. Li et al. [34] applied dynamic GNN and attention mechanism to single-channel EEG signals to identify epileptic seizures, and the average accuracy reached 99.83 %.

2.2 Federated Learning for BCI

FL is an emerging machine learning approach that can be combined with deep learning to achieve the goal of powerful model training and learning on large-scale distributed data while preserving data privacy. Wei and Faisal [35] proposed MF-SCSN, combining privacy protection and deep transfer learning, achieving a 3 % performance improvement over the baseline decode. Liu et al. [36] proposed a hierarchical personalized Federated Learning EEG decoding (FLEEG) framework, which improved overall performance in MI classification tasks using nine EEG datasets collected from different devices. However, the main challenge faced by FL is non-IID. Different subjects or tasks usually have different feature distributions. This means that when using FL, we need to think about how to handle heterogeneity and ensure the robustness of the model.

2.3 Federated Graph Learning

FGL is a method that combines FL with GNN, which aims to deal with the problem of privacy protection and distributed processing of graph data. Currently, the research on FGL can be divided into three main directions: inter-graph FGL, intra-graph FGL, and graph-structured FGL. In inter-graph FGL, each client holds a set of graph data and jointly participates in model training through FGL to improve the modeling ability of local data. In intra-graph FGL, where each client owns only a subset of the graph, the aim is to solve the problem of missing connections or discovering communities [37]. On the other hand, graph-structured FGL considers the topological relationship between different clients and uses graph neural networks to aggregate local models, which is mainly applied in fields such as federal traffic flow prediction [38].

Although there is no research on applying FGL to BCI, EEG and EMG data can be naturally represented as graph structures, and faced with the challenges of small data volume and privacy protection, so FGL has significant potential in the application of EEG and EMG-based BCI. FGL also faces the non-IID problem in

BCI, that is, there may be heterogeneity in graph structure and node features under different subjects or tasks, which makes the effect of the global model less than expected. Therefore, how to balance the performance of global and local models in this scenario of inconsistent data distribution has become a key research direction of FGL in BCI. Tang et al. (2024) [39] applied the personalized Federated Graph Learning Framework (PEARL) to the electronic health record (EHR), proposing an effective solution for non-IID data to improve the performance of the client model through self-supervised learning and local fine-tuning schemes. This study shows that even when dealing with non-IID data, FGL can achieve significant results as long as appropriate strategies are adopted. This paper focuses on inter-graph FGL methods. The aim is to protect data privacy while training better local models for EEG and EMG-based BCI.

3 METHODS

3.1 Overall Framework

This paper presents a study based on an FGL architecture designed to recognize complex motion intentions. The overall framework of the proposed FSDFGL is shown in Figure 2. As shown in the flowchart, on the left side of client M shows the construction of the graph, which includes the EEG-EMG heterogeneous graph as well as the EEG graph. First, the EEG signal and EMG signal were aligned to ensure synchronization and timeliness. After alignment, the signal is sliced at a rate of once per second to provide a finer temporal resolution. Finally, to construct the graph data, PSD is used to extract node features, and then SPMI is used to obtain edges between nodes. In the middle of client M, the structural information of the graph is extracted, and it is fed into the two-channel GNN at the same time as the node features. The channels used to process the structure information will not only share knowledge through FL, but also supplement the channels used to process the feature information, so that the global structure information and local node features can be used to better capture the characteristics of the graph. On the right side of client M, the outputs of these two channels are merged together and then fed into the classifier for the final classification task.

3.2 EEG-EMG Heterogeneous Graphs Construction

The functional connectivity of the brain can reveal the coordination and communication mode between different brain regions during the generation and execution of motor intention. Graph data provides an effective way to represent and study the functional connectivity of the brain, helping us to gain a deep understanding of the complex structure and function of the brain, as well as the role of the brain in cognition and behavior.

Traditional BCI typically focus on analyzing graph data constructed from EEG. However, in recognizing complex motor intentions, traditional EEG-based methods

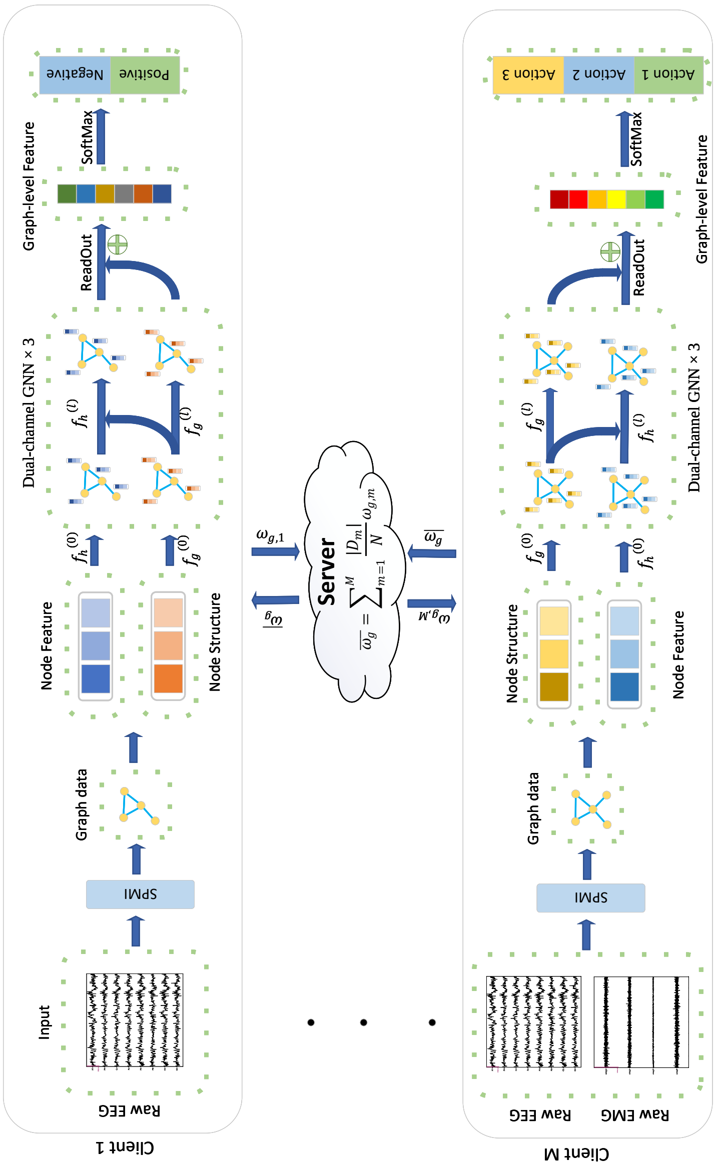


Figure 2. Overall framework of FSDFGL. The raw data is input for constructing graph-structured data, where SPMI is used to calculate the coherence between nodes. Then, structural embeddings and node features are separately extracted and fed into a dual-channel GNN, with the parameters of the structural encoder shared through federated learning. The graph-level features obtained through the readout function are used for final classification.

may exhibit limitations. To address this issue, researchers have introduced EMG signals as supplementary inputs, constructing EEG-EMG heterogeneous graphs. This approach allows for a more comprehensive understanding of the coordination mechanisms between the brain and muscles, thereby improving the accuracy of motor intention recognition.

Let $G = (V, E)$ be a graph consisting of a set of nodes V and a set of edges E connecting these nodes. Each node $v \in V$ represents the channel of EEG and EMG. Each $e \in E$ represents the degree of correlation between channels.

PSD is used to calculate the node features. Although both EEG and EMG signals are time series signals, the frequency components related to the physiological activity of interest can be extracted after proper preprocessing and filtering. To extract features in the frequency domain, the PSD of different frequency bands can be calculated, and then node numbers can be added as position encoding to these features. The relevant physiological states and movement intentions can then be inferred from the features and changes in these frequencies.

SPMI is used to construct the edges of graphs. It is a nonlinear measure of functional connectivity that was initially developed to describe the coherence between two EEG signal channels during anesthesia, thereby assessing communication between cortical regions. TAN et al. found significant discontinuities in the α/β coherence distribution between EMG and EMG electrodes. However, the distribution of SPMI among all electrode pairs shows a smooth and continuous feature. Therefore, SPMI is suitable for quantifying the coherence between EEG and EMG signals to effectively assess the functional connectivity between different brain regions or muscle regions.

In order to calculate the coherence between signals, the permutation entropy of the two channel vectors X and Y of the signal first needs to be calculated.

$$PE_X = -\sum_{i=1}^{k!} P_X(i) \log(P_X(i)), \tag{1}$$

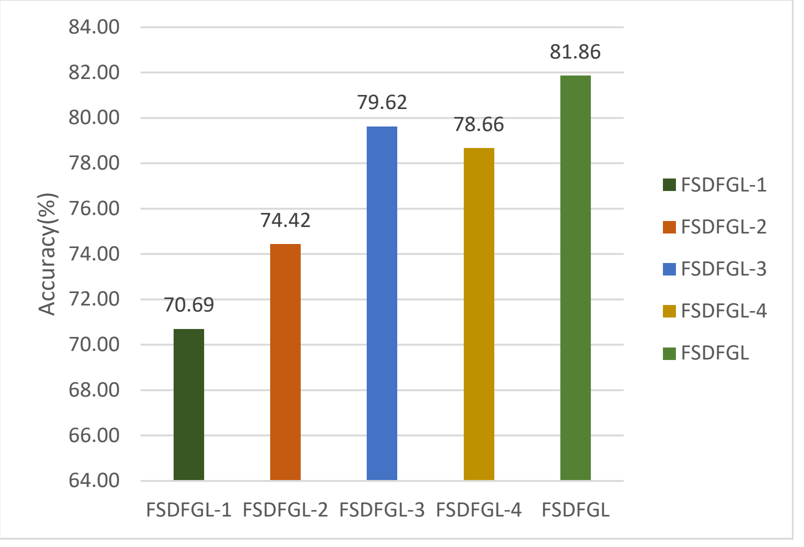
where $P_X(i)$ is the permutation probability of X and k is the dimension of X.

$$PE_{(X,Y)} = -\sum_{i=1}^{k!} \sum_{j=1}^{k!} P_{XY}(i,j) \log(P_{XY}(i,j)), \tag{2}$$

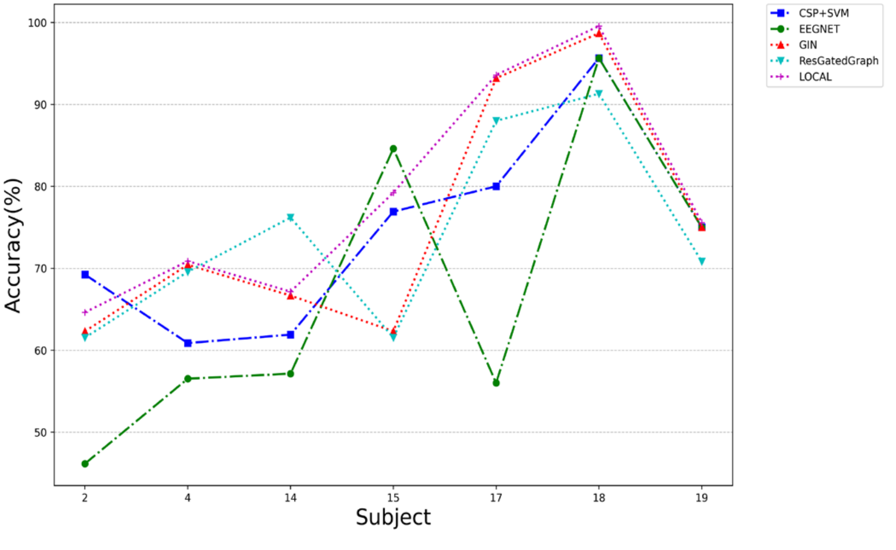
where $P_{XY}(i,j)$ is the joint probability of permutation of X and Y. Finally, how to calculate SPMI is described as follows:

$$SPMI_{(X,Y)} = \frac{(PE_X + PE_Y - PE_{(X,Y)})}{PE_{(X,Y)}}. \tag{3}$$

SPMI values range from 0 to 1. By comparing the SPMI values with a predefined threshold, channels exhibiting a degree of correlation above the threshold can be identified. These channels can be connected to form the edge of the graph.



a)



b)

Figure 3. a) Contribution comparison of different novel parts in improving performance. b) Average accuracies comparison of the proposed method with other different methods on motor intention classification tasks for each subject in TAN dataset.

3.3 Structure Embedding Initialization

In the non-IID FGL scenario, datasets from different clients usually involve different subjects or tasks, leading to feature heterogeneity and making it difficult to share useful information among clients. There are some general structural properties in EEG and EMG signals, but in GNNs, structural information is implicitly encoded to nodes along with feature information through the feature aggregation process, making it potentially difficult for local models to distinguish and extract this information during training and inference. To address this issue, we propose a strategy that encodes structural information separately as structural embeddings and uses structural embeddings instead of features for message-passing and aggregation during FL.

To extract structural information from the graph, we use Degree Structural Encoding (DSE) and PR to capture local and global knowledge, respectively. DSE is represented using one-hot encoding. For a given node v, its DSE is represented as follows:

$$s_v^{DSE} = [I(d_v = 1), I(d_v = 2), \ldots, I(d_v \geq k_1)] \in \mathbb{R}^{k_1}, \tag{4}$$

where d_v represents the degree of the node, I represents the identity function, and k represents the dimension. Using DSE to represent local structural knowledge has the following advantages: firstly, the degree distribution of a node is a fundamental graph property that can be obtained through simple computations, without requiring complex preprocessing or computational resources. Secondly, in brain networks, the degree distribution typically follows a power-law distribution, where a few nodes with a large number of connections are referred to as hub nodes. By capturing anomalies in these hub nodes, it is possible to identify various neurological diseases.

Meanwhile, PR is used to capture global structural knowledge. The algorithm is based on a random walk model and determines the relative importance of each node by calculating the link relationships between nodes in the network. The PR value of node v is represented as follows:

$$s_v^{PR} = \frac{(1-q)}{N} + q \sum_{j \in B(V)} \frac{PR(j)}{L(j)}, \tag{5}$$

where s_v^{PR} represents the PR value of electrode v_i. The term $1 - q$ denotes the probability that a random walker jumps to any electrode, q is the damping factor, and N is the total number of electrodes. $B(V)$ represents the set of all electrodes connected to v, which influence v_i through electrical activity signals. $L(j)$ denotes the connection count of electrode v, indicating the number of outgoing electrical signals from j. PR evaluates the relative importance of different brain regions by calculating the probability of a walker staying in the network for a long time to determine the ranking of nodes. This analysis is able to reveal information flow patterns in functional brain networks, identify key brain regions, and help understand the role of these brain regions in cognitive tasks. Unlike node degree, which only provides a simple measure of the number of connections between nodes, PR takes into account the global structure of the graph and the link relationship between nodes through iterative calculation.

$$s_v = \text{concat}\left[s_v^{DSE}, s_v^{PR}\right]. \tag{6}$$

Finally, the results of the DSE and PR were aggregated to obtain a structural encoder that provides a richer and more comprehensive representation of the structure.

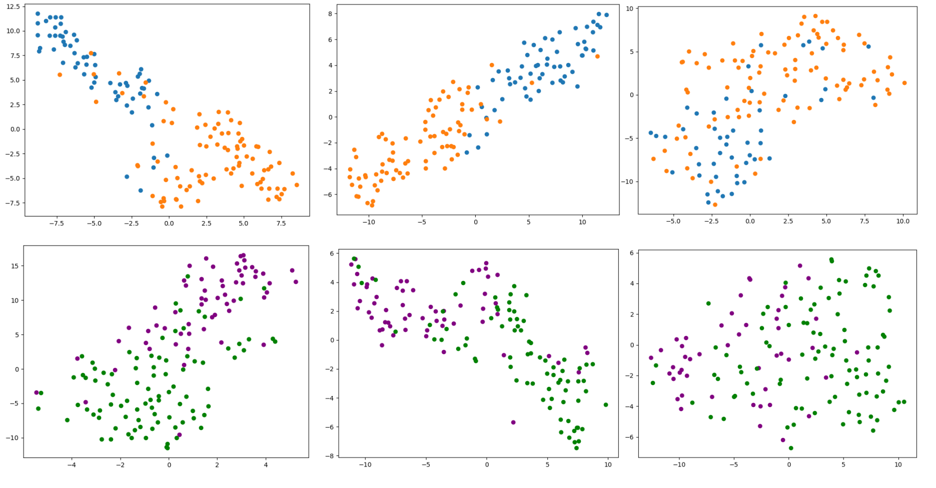


Figure 4. t-SNE visualizations of FSDFGL and FedAvg methods for tasks 17, 18, and 19. a) FSDFGL method, b) FedAvg method.

3.4 Local Dual-Channel GNN

When performing FGL, a single-channel GNN is usually used. However, when the datasets of clients come from different subjects or tasks, there may be differences in feature distributions and data characteristics. In such cases, directly using a single-channel GNN to build and share feature encoders may impact personalization performance due to inconsistencies in feature representation space. Even if the model performs well on one client, it may not generalize effectively to other clients. Additionally, although structural embeddings containing shared structural knowledge have been obtained, single-channel GNNs cannot simultaneously handle structural and feature information. To address these issues, a dual-channel GNN is adopted.

In this module, the graph structure data is used as the original input. For a given node, we extract the structural embedding and the node feature. Since these two vectors may have different dimensions, we use two parallel linear layers $f_h(0)$ and $f_g(0)$ to convert them into vectors with a consistent dimension.

Then, as shown in Figure 1 a), the structural embeddings are fed into the GCN with L layers. In the L-layer GCN, only the hidden structure embeddings of the previous layer are accepted as input, which enables the structure encoder to focus on learning structural knowledge and avoid the adverse effects of heterogeneous features on other client structure encoders when messages are shared. Choosing GCN to process structure embedding has the following advantages: first, the model structure of GCN is relatively simple, and the calculation process is more efficient than some complex GNN models. Second, GCN updates the representation of a node by aggregating the features of the node and its neighbors. This method can effectively capture local structure information, so that the model can adapt to different graph structures and have better generalization ability.

$$H^{(l+1)} = \sigma\left(D^{-\frac{1}{2}}\tilde{A}D^{-\frac{1}{2}}H^{(l)}W^{(l)}\right), \tag{7}$$

where D represents the diagonal node degree matrix of the graph, A represents the adjacency matrix, $H^{(l)}$ represents the node feature matrix of the l^{th} layer, $W^{(l)}$ represents the weight matrix of the l^{th} layer, $\tilde{D}$ represents the normalized diagonal node degree matrix, $\tilde{A}$ represents the normalized adjacency matrix, and σ represents the activation function.

As shown in Figure 1 a), the feature embeddings are fed into GIN with L layers. In GIN with l layers, the concatenation of hidden structure embeddings and hidden feature embeddings from the previous layer is used as input, which enables the feature encoder to make full use of the global structure knowledge learned by the structure encoder to supplement. GIN is chosen to handle feature embedding because it can effectively model graph isomorphism through complex feature aggregation operations and multi-layer perceptron (MLP), thus generating rich global feature representations.

$$h_v^{(k)} = \mathrm{MLP}^k\left(\left(1+\epsilon^{(k)}\right)\cdot h_v^{(k-1)} + \sum_{u\in N(v)} h_u^{(k-1)}\right), \tag{8}$$

where MLP represents a Multi-Layer Perceptron, $h_v^{(k)}$ represents the feature representation of node v at the k^{th} layer, and $\epsilon^{(k)}$ represents a learnable parameter.

$$h^G = \mathrm{concat}\left(\mathrm{readout}\left(\{h_v^{(k)} | v \in G\}\right) | k = 0, 1, \ldots, K\right), \tag{9}$$

$$\omega^*_{(h,m)} = \arg\min_{\omega_{(h,m)},\omega_{(g,m)}} L_m\left(\omega_m; D_m\right). \tag{10}$$

In the last layer, the hidden structure embeddings output by GCN and the hidden feature embeddings output by GIN are concatenated together as input to the read function. Through this function, these concatenated embeddings are transformed into graph-level features to support subsequent classification tasks. Finally, the cross-entropy loss function L_m is used to optimize the local model, and the parameters of the structure encoder $\omega_{(h,m)}$ and feature encoder $\omega_{(g,m)}$ are updated during the training process.

3.5 Federated Learning Layer

EEG and EMG datasets are usually small in size and involve sensitive, privacy-related information. In addition, these datasets often come from different subjects and tasks, which makes the data characteristics heterogeneous and negatively impacts the model's performance. To address these challenges, a FL approach has been proposed that focuses on sharing the parameters of local structural encoders to share universal structural knowledge. Since FL only shares structural encoder model parameters, it can leverage distributed data for model training and improvement while preserving data privacy.

In this approach, the central server aggregates these local structure encoder parameters to construct a global model using the following aggregation formula:

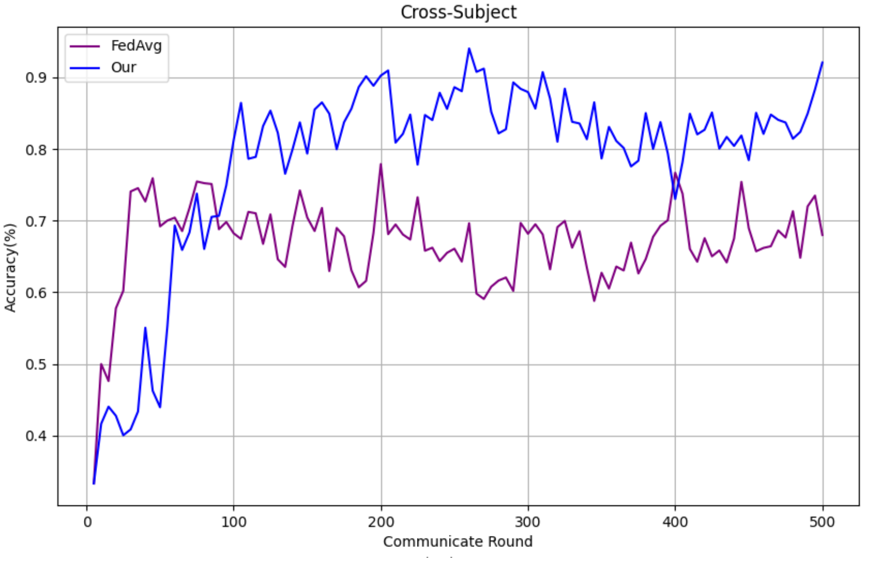
$$\bar{\omega}_g = \sum_{m=1}^{M} \frac{|D_m|}{N} \omega_{(g,m)}. \tag{11}$$

Here, ω_g represents the global model parameters, $\omega_{(g,m)}$ denotes the parameters of the local structure encoder at client m, $|D_m|$ is the data size of client m, and N is the total data size across all clients. $\frac{|D_m|}{N}$ represents the proportion of data size $|D_m|$ of each client m relative to the total data size N. The aim is to ensure that each client's contribution to the global model is proportional to the amount of data it has. Doing so helps balance the contributions of each client and allows for a fairer aggregation of global model parameters, especially when the data is unevenly distributed.

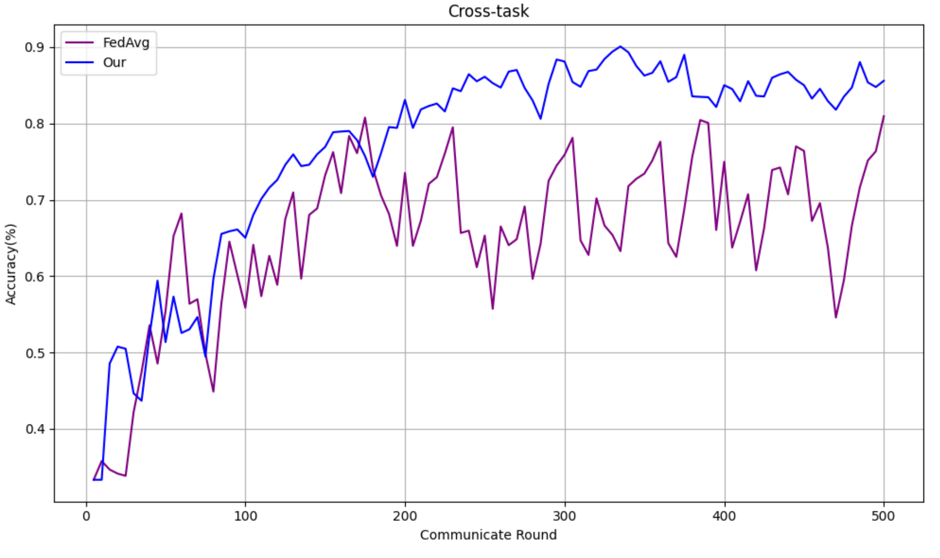
Then, $\bar{\omega}_g$ is sent back to all clients to update the local structural encoder model parameters and initiate the next round of local training. By only sharing the parameters of the structural encoder, this approach effectively protects data privacy. Additionally, the structural encoder focuses on processing structural information and performs message aggregation on the server side, allowing FSDFGL to more accurately capture the universally shared structural knowledge in EEG and EMG data. This method can also effectively improve the performance of local models, even in experimental settings involving cross-subjects or tasks.

4 EXPERIMENT RESULTS AND ANALYSIS

In this section, first, two publicly well-known datasets, namely TAN and DEAP are introduced. These datasets are used to record physiologically based signals. Experiments conducted on these two datasets will be described in detail, including



a)



b)

Figure 5. Test accuracy curves of FSDFGL and FedAvg methods over communication rounds. a) Cross-subject experimental settings, b) cross-task experimental settings.

hyperparameter settings and network architectures. Finally, the experimental results are displayed and analyzed in multiple dimensions to verify the performance of the proposed framework.

4.1 Datasets

In our experiments, TAN and DEAP are used. TAN is a motion intention dataset, while DEAP is an emotion recognition dataset.

1. TAN: The TAN dataset is a multimodal dataset containing EEG-EMG signals. The dataset included five healthy volunteers and two stroke patients. Each participant performed static movements of equal length, such as pushing and pulling, lasting 3 seconds. The dataset contained a total of 516 trials of healthy individuals and 174 trials of stroke patients. For EEG, 32 electrodes were placed for recording using the international 10-10 system with a sampling rate of 500 Hz. At the same time, eight EMG electrodes were placed on the arm, each at the following locations:

   (a) flexor digitorum superficialis (FDS),
   (b) flexor carpi ulnaris (FCU),
   (c) flexor carpi radialis (FCR),
   (d) extensor carpi ulnaris (ECU),
   (e) extensor carpi radialis longus (ECRL),
   (f) biceps brachii short head (BBS),
   (g) triceps brachii long head (TBL) and
   (h) lateral deltoid (LD).

   The sampling rate of EMG data is 1 000 Hz. In this article, EMG data was down sampled to 500 Hz and aligned with EEG data. The EEG and EMG signals were then bandpass filtered at 10–200 Hz and noise was removed using independent component analysis (ICA) methods.

2. DEAP: The DEAP dataset is a publicly available dataset for the study of emotion analysis and emotion recognition. The dataset collected both EEG and physiological signal data. In the DEAP dataset, an international 10-20 system was used for electrode placement, where the first 32 channels were used to collect EEG signals and the rest were used to collect other physiological signals. The sampling frequency of EEG data was 512 Hz. The emotional data in the DEAP dataset was obtained by having subjects watch 40 music video clips. Participants were asked to rate the video clips based on their subjective experience on four dimensions (1-9): valence, arousal, dominance, and arousal. In this paper, The EEG data was down sampled to 128 Hz, and the first 32 channels were selected to remove other irrelevant artifacts, such as EOG. Each data trial was 63 seconds long, including a baseline time of 3 seconds.

4.2 Experiments Setup

In this paper, our main focus is on the data heterogeneity scenario that exists in FL. To simulate this scenario, two non-IID settings are performed:

1. a cross-dataset setting based on different experimental paradigms and
2. a same-dataset setting based on different subjects.

In each setting, the data is split into an 80 % training set and a 20 % test set. In our experiment, the Adam optimizer was used with a learning rate of 0.005 and a dropout rate of 0.5. The dimensions of both DSE and PR are set to four. For the N-layer local dual-channel GNN, a three-layer GCN was used as the structure encoder. Additionally, a three-layer GIN was used as the feature encoder, with a hidden layer size of 64. In all experimental settings based on FL, the number of local training rounds was set to 3, the batch size was set to 128, and a total of 500 federated shares were executed. In the non-federated experimental setting, local training consisted of 300 rounds with a batch size of 128. All methods were implemented in PyTorch, and the experiments were conducted on an NVIDIA RTX 5000.

To evaluate the performance of the trained local model, the label of each subaction is predicted, and the score is added according to the quality of the action. After completing a set of actions, the action with the highest score is selected as the final predicted label, and then compared with the actual label.

4.3 Experiments Results

The proposed FSDFGL contains four innovative parts: integrating EMG into EEG signals through SPMI, extracting structural information using DSE and PR, using a dual-channel GNN to process structure embedding and node features, respectively, and realizing FL based on structural knowledge sharing. To evaluate the contribution of each innovation part, four frameworks were designed by removing part of the components and compared with the proposed framework.

1. FSDFGL-1: The framework uses only EEG data as input, without utilizing SPMI to integrate EMG signals.
2. FSDFGL-2: The framework uses an all-ones vector to represent structural embedding.
3. FSDFGL-3: The framework uses a single-channel GNN that receives both node feature and structure embedding as input.
4. FSDFGL-4: The model without federated layers for sharing structural knowledge.

In Figure 3 a), the average classification accuracy histograms of various variant frameworks and the proposed framework compared on the TAN dataset are shown. It is observed that different innovative parts contribute differently to the classification accuracy. Specifically, FSDFGL has the highest classification accuracy improvement compared with FSDFGL-1, reaching 11.17 %. This verifies that by incorporating EMG signals into EEG signals using SPMI, more comprehensive and rich information can be obtained, thus improving the classification accuracy. Additionally, FSDFGL increased by 7.42 %, 2.24 %, and 3.20 % compared to FSDFGL-2, FSDFGL-3, and FSDFGL-4, respectively. This verifies that classification accuracy can be improved to some extent by extracting structural and feature information separately. By putting this information into a dual-channel GNN and utilizing FL to share structural knowledge, further improvements can be achieved. Taken together, the proposed framework integrates the advantages of each innovative part and achieves better classification accuracy.

Method	2	4	14	15	17	18	19	Mean
LOCAL	64.62	70.87	67.14	79.23	93.60	99.57	75.58	78.66
FedAvg	70.00	70.43	70.48	73.08	94.40	98.70	80.00	79.58
Ours	66.15	73.48	67.62	87.69	96.00	100.00	82.08	81.86

Table 1. Accuracy comparison of different federated learning methods in cross-subject experimental settings on the TAN dataset

Method	2	4	14	15	17	18	19	Mean
LOCAL	64.62	70.87	67.14	79.23	93.60	99.57	75.58	78.66
FedAvg	63.85	61.74	56.19	46.15	86.40	83.04	71.67	67.01
Ours	66.92	71.30	69.52	77.69	98.80	99.57	83.33	81.02

Table 2. Accuracy comparison of different federated learning methods in cross-task experimental settings on the TAN dataset

Signal	2	4	14	15	17	18	19	Mean
EEG	49.57	50.43	66.67	71.54	86.40	84.78	85.42	70.69
EMG	61.54	76.09	56.19	65.38	72.80	67.39	66.67	65.58
EEG + EMG	66.15	73.48	67.62	87.69	96.00	100.00	82.08	81.86

Table 3. Accuracy comparison of single-source and multi-source signals

To further verify the superiority of FSDFGL, the local model of FSDFGL will be compared with traditional machine learning and deep learning models on the TAN dataset, as shown in Figure 3 b). This comparison involved benchmark methods like SVM + CSP and EEGNet. Additionally, two GNN models, GIN and ResGatedGraph, and a standard FL model, FedAvg, were evaluated. It can be observed that the local model significantly outperforms CSP + SVM on subjects 4, 14, 15, 17, 18, and 19. This is because CSP + SVM relies on manual feature engineering and cannot automatically extract complex patterns and relationships from the data. At the same time, the local model is better than EEGNet on subjects 2, 4, 14, 17, 18, and 19. Although EEGNet can automatically extract features, its network structure is designed for processing EEG data and may not effectively capture the specific patterns and characteristics of EMG data, resulting in some limitations when processing multimodal data. Meanwhile, while the other two GNN methods perform better compared to CSP + SVM and EEGNet, they almost completely underperform when compared to local models. This proves that GNN can better learn and analyze the topological relationships in the graph structure and capture the information transfer and interaction in the brain network. However, single-channel GNN models can only process node characteristics and structural information through a single network channel, which limits their performance in capturing and fusing these two types of information. At the same time, these models lack the advantages of FL and cannot utilize distributed data for model training while protecting data privacy.

In addition, in order to verify the effectiveness of structure sharing FL, FSDFGL is compared with the traditional FL method FedAvg, as well as the local model, as shown in Tables 1 and 2. Although in Table 1 the average accuracy of FedAvg is higher than the local method, in Table 2, the accuracy of all subjects is lower than the local model. Specifically, the average accuracy is 11.65 % lower than the local method. Despite FedAvg performing well in cross-subject experiments, its performance declines when dealing with more complex cross-task experiments. This indicates the limitations of FedAvg in dealing with the effects of non-IID, is easily affected by the heterogeneity of traditional EEG-EMG datasets. On the contrary, FSDFGL outperforms the local model and FedAvg in both cross-subject experimental Settings and cross-task experimental subjects. This indicates that in traditional FL, due to the heterogeneous characteristics and distribution of data among different clients, it may lead to different degrees of interference when the model is trained locally. Structure-based FL methods have better ability to deal with non-independent and identical distributions. By sharing structural knowledge, it can better avoid the influence of heterogeneous features on the local model and improve the overall learning effect.

As shown in Figure 4, we compared the performance of the two methods under t-SNE by selecting data from three high-performing subjects (17, 18, and 19) in the cross-task experimental paradigm. The upper three plots represent the results of FSDFGL, while the lower three plots correspond to FedAvg. The results indicate that our model achieves a clearer separation between the two categories, demonstrating more distinct differentiation. In contrast, FedAvg shows more overlap between the two colors, with less defined category boundaries. Furthermore, our method excels in latent space alignment by suppressing dataset or subject-specific features, allowing the model to focus more on task-related common features. This results in better generalization across different datasets or subjects and clearer category separation.

5 DISCUSSION

5.1 Effects of EEG-EMG Signal Fusion

As shown in Figure 3 a), EEG-EMG data significantly improves the accuracy of movement intention recognition by more than 11 %. To verify its superiority over a single signal, the signals were divided into three groups: EEG, EMG, and EEG + EMG. These experiments were carried out under the same configuration. The results are shown in Table 3, where the model with EEG-EMG achieves the best average accuracy.

In the experiments, data from five subjects 2, 14, 15, 17 and 18 were analyzed, and the EEG-EMG fusion signal significantly outperformed EEG or EMG signals alone. Meanwhile, in the data of subjects 4 and 19, although the EEG-EMG fusion signal is slightly lower than the single EEG or EMG signal, it still provides a relatively reliable recognition result. This indicates that EEG-EMG multi-source signal fusion can provide more comprehensive information and has significant advantages for movement intention recognition.

In addition, the results also showed that EMG signals alone performed poorly in stroke patients, which may be due to the fact that EMG signals in stroke patients are usually weaker. Therefore, EEG-EMG signal fusion shows more robustness and accuracy in these cases, which can better compensate for the deficiency of a single signal and thus provide more reliable movement intention recognition results.

5.2 Effects of Decoupling and Sharing Mechanisms

As detailed in Section 3.5, message sharing mechanisms in FL are worth discussing, in particular the impact of different message sharing mechanisms on local models and which mechanism performs best in enhancing local model performance. The message sharing mechanisms are feature-based, structure-based and shared-all parameter. All experiments were performed using the same local model, and the results are shown in Tables 4 and 5, and are discussed for the two non-IID experimental setups across subjects and tasks.

In the experimental results in Tables 4 and 5, the structure-based message sharing mechanism is significantly better than other mechanisms, and the accuracy is improved by more than 2 % compared with the local model. This shows that the structure-based message mechanism can effectively avoid the influence of non-independent and identical distributions on the local model and can share useful structural knowledge.

While the feature-based message sharing mechanism also performed well in the cross-subject experimental setup, with better results than the local model, it significantly lagged behind the local model in the cross-task experimental setup. This means that the feature-based message sharing mechanism is more suitable for datasets with similar feature distributions, but in scenarios with less similar features, the extracted feature information cannot adapt well to the

Method	2	4	14	15	17	18	19	Mean
None	64.62	70.87	67.14	79.23	93.60	99.57	75.58	78.66
Fe	65.38	69.13	69.05	79.23	97.20	99.57	75.83	79.34
ALL	63.08	67.83	65.71	73.85	96.80	99.57	78.33	77.88
Stru (Ours)	66.15	73.48	67.62	87.69	96.00	100.00	82.08	81.86

Table 4. Effects of different decoupling and sharing mechanisms in cross-subject experimental settings

Method	2	4	14	15	17	18	19	Mean
None	64.62	70.87	67.14	79.23	93.60	99.57	75.58	78.66
Fe	60.77	57.83	58.57	66.92	83.60	94.35	64.58	69.52
ALL	52.31	52.17	48.57	54.62	66.40	76.96	68.33	59.91
Stru (Ours)	66.92	71.30	69.52	77.69	98.80	99.57	83.33	81.02

Table 5. Effects of different decoupling and sharing mechanisms in cross-task experimental settings

dataset clients with heterogeneous features due to the existence of heterogeneous features.

In contrast, the message passing mechanism with full parameter sharing lags behind the local model in both cross-task and cross-subject experimental Settings. This suggests that too much sharing of model parameters may introduce noise and irrelevant information, adversely affecting the performance of the local model. Although full parameter sharing attempts to comprehensively integrate the information of all participants, due to the heterogeneity of the data distribution, this approach is prone to model overfitting or bias, thus reducing the overall performance.

5.3 Convergence Analysis

Figure 5 illustrates the performance of the proposed framework and the baseline method as a function of the number of communication rounds, and Figures 5 a) and 5 b) correspond to the experimental Settings across subjects and tasks, respectively. It can be observed that compared with FSDFGL, FedAvg usually converges faster, but has lower accuracy and less stability. FedAvg adopts a simple model parameter averaging method and synchronously updates the model parameters of all clients in each round of communication, which makes the global model reflect the data distribution of all clients immediately after each round, so as to achieve fast overall convergence. However, when there is a significant difference in data distribution across clients, updates from different clients may conflict with each other, which will impair the overall model performance. In addition, FedAvg does not have a special mechanism to deal with the data heterogeneity between clients (non-IID), which makes the model difficult to adapt to the data characteristics of all clients when the data distribution is different, resulting in low accuracy and poor stability.

In contrast, FSDFGL uses a structured knowledge sharing mechanism to capture widely shared structural knowledge, so as to effectively deal with the problem of dataset feature heterogeneity, improve the accuracy and enhance the stability of the model.

5.4 Effects of Different Structure Embeddings

As shown in Figure 3 a), extracting structural knowledge effectively improves the accuracy of motion intention recognition. However, the composition of the node structure deserves to be explored in depth, whether all components are beneficial to the recognitioon results. The components of the node structure are the DSE, the PR, the DSE and the PR (DSE + PR), and the all-one vector (none) as a contrast. All experiments were performed using the same model, and the results are listed in Tables 6 and 7. Two non-IID experimental setups across subjects and tasks are discussed. In Tables 6 and 7, the combination of DSE and PR comprehensively outperforms none and other single component node structures, which indicates that PR and DSE as components of nodes can significantly improve the accuracy of the model.

In Tables 6 and 7, the average accuracy of PR in Table 6 shows a significant improvement compared to "none", while in Table 7, the accuracy improvement of DSE is more pronounced. PR evaluates the importance of nodes through random walks, which better handles the heterogeneity of cross-subject data. On the other hand, DSE reflects local structural features through node degrees and is more adaptable to task variations (such as in cross-task experiments). The combination of DSE and PR outperforms "none" and other single-component node structures, indicating that they complement each other in different data scenarios and jointly enhance the model's accuracy. This further validates that DSE and PR, as components of nodes, can significantly improve the classification performance of the model. Although the average accuracy of the structure composed of only DSE and PR has achieved a good improvement, the effect lags behind none in some subjects. This is attributed to the fact that DSE only considers the degree of nodes, while PR obtains information based on the random walk path between nodes. Both of these methods only capture local information or certain structural properties, and may not cover all key information. In some subjects, this local information may not be sufficient to fully describe the complex structure or characteristics of the data, resulting in poor performance.

5.5 Varying Local Epochs

The setting of local rounds has been a deeply concerned issue in FL. Properly set local rounds directly determine the contribution of each client to the global model and the degree of performance improvement during training. If the local rounds are

DSE	PR	2	4	14	15	17	18	19	Mean
×	×	64.62	73.04	67.14	57.69	95.60	97.83	65.00	74.42
×	✓	63.85	71.30	65.71	77.69	95.60	98.70	75.83	78.38
✓	×	66.92	71.74	64.76	80.00	95.60	100.00	82.50	80.22
✓	✓	66.15	73.48	67.62	87.69	96.00	100.00	82.08	81.86

Table 6. Effects of different structure embeddings in cross-subject experimental settings

DSE	PR	2	4	14	15	17	18	19	Mean
×	×	62.31	70.43	64.29	69.23	96.00	98.26	78.33	76.98
×	✓	65.38	70.87	63.33	76.15	94.00	98.26	80.42	78.35
✓	×	63.08	70.00	61.64	69.23	97.60	100.00	80.42	77.42
✓	✓	66.92	71.30	69.52	77.69	98.80	99.57	83.33	81.02

Table 7. Effects of different structure embeddings in cross-task experimental settings

set too few, the local model of the client may not be able to fully utilize its local data characteristics, thus failing to significantly improve the performance of the global model. On the contrary, too many local rounds may increase the computational cost and communication overhead, and even lead to model overfitting or performance degradation. Furthermore, given the heterogeneity of data per client, some clients may have more or more complex data distributions, and thus may need more local rounds to learn and update the model efficiently.

EPOCHS	1	2	3	4
FedAvg	79.58	81.04	80.16	81.70
Ours	80.56	79.35	81.86	79.28

Table 8. The performance on cross-subject experimental settings with varying local epochs

To explore these issues, we set different number of local rounds (1, 2, 3, 4) on the two methods FSDFGL and FedAvg. All experiments use the same model and conduct in-depth analysis under two non-IID experimental settings: cross-subject and cross-task. No obvious rule is observed from the experimental results in Tables 8 and 9 indicating that more local rounds have a greater impact on the local model.

In Table 8, in the cross-subject experimental setup, FedAvg outperforms FSDFGL when the local rounds are 2 and 4. However, in Table 9, the overall performance of FSDFGL is better than FedAvg in the experimental setup across tasks. This indicates that the feature-based message sharing mechanism is not the cause of the performance degradation of the local model due to inadequate training.

5.6 Varying Batch Size

To test the most appropriate Batch Size, we set the number of client samples to 32, 64, 128 and 256 for both methods, FSDFGL and FedAvg, and conducted an

EPOCHS	1	2	3	4
FedAvg	67.01	65.94	65.51	65.54
Ours	80.56	81.25	81.02	80.65

Table 9. The performance on cross-task experimental settings with varying local epochs

in-depth analysis under two non-IID experimental settings: cross-subject and cross-task. The experimental results in Tables 10 and 11 show that the model performs best when the Batch Size is set to 128.

BATCH SIZE	32	64	128	256
FedAvg	77.88	79.34	80.16	78.66
Ours	79.35	81.02	81.86	79.28

Table 10. The performance on cross-subject experimental settings with varying batch size

BATCH SIZE	32	64	128	256
FedAvg	63.85	64.62	65.51	63.08
Ours	78.35	79.34	81.02	80.65

Table 11. The performance on cross-task experimental settings with varying batch size

6 CONCLUSION

This paper proposes a novel FGL framework for hybrid brain-computer interfaces, named FSDFGL. The framework aims to address the limitations of EEG in spatial resolution and the challenges of complex motor intention recognition due to the lack of large datasets. Through an effective structured knowledge-sharing method, FSDFGL successfully solves the non-IID problem in FL when the local client data sets come from different subjects and tasks. To evaluate the performance of the proposed method under non-IID scenarios, experiments were conducted on two publicly available datasets. The performance of the proposed method was assessed using various metrics and compared with baseline methods. The results indicate that FSDFGL achieved accuracies of 81.86 % and 81.02 % in cross-subject and cross-task settings, respectively, significantly outperforming the baseline methods. Additionally, extensive ablation studies were performed, demonstrating the superiority of the model. The model shows potential in recognizing complex motor intentions and effectively addresses the performance limitations of traditional brain-computer interfaces related to the spatial resolution of EEG and the scarcity of large EEG datasets.

Funding

This work was supported in part by the National Natural Science Foundation of China under Grant No. 62472319, in part by the Second Round of "Three-Year Action Plan to Promote Clinical Skills and Clinical Innovation in Municipal Hospitals" Research Physician Innovation Transformation Capability Training Program under Grant No. SHDC2023CRT001, in part by the National Clinical Key Specialty Construction Project of China under Grant No. Z155080000004, in part by the Shanghai Research Center of Rehabilitation Medicine (Top Priority Research Center of Shanghai) under Grant No. 2023ZZ02027, in part by the 2023 Annual Innovative Medical Device Application Demonstration Project "Weight Reduction Gait Training System Application Demonstration" under Grant No. 23SHS05500, and in part by the Shanghai Disabled Persons' Federation Key Laboratory of Intelligent Rehabilitation Assistive Appliance and Technology.

REFERENCES

[1] Jeannerod, M. J.: Mental Imagery in the Motor Context. Neuropsychologia, Vol. 33, 1995, No. 11, pp. 1419–1432, doi: 10.1016/0028-3932(95)00073-C.

[2] Müller-Putz, G. R.—Scherer, R.—Pfurtscheller, G.—Rupp, R.: EEG-Based Neuroprosthesis Control: A Step Towards Clinical Practice. Neuroscience Letters, Vol. 382, 2005, No. 1–2, pp. 169–174, doi: 10.1016/j.neulet.2005.03.021.

[3] Zhao, Q.—Zhang, L.—Cichocki, A.: EEG-Based Asynchronous BCI Control of a Car in 3D Virtual Reality Environments. Chinese Science Bulletin, Vol. 54, 2009, No. 1, pp. 78–87, doi: 10.1007/s11434-008-0547-3.

[4] Meng, J.—Zhang, S.—Bekyo, A.—Olsoe, J.—Baxter, B.—He, B.: Noninvasive Electroencephalogram Based Control of a Robotic Arm for Reach and Grasp Tasks. Scientific Reports, Vol. 6, 2016, No. 1, Art. No. 38565, doi: 10.1038/srep38565.

[5] Foong, R.—Ang, K. K.—Quek, C.—Guan, C.—Phua, K. S.—Kuah, C. W. K.—Deshmukh, V. A.—Yam, L. H. L.—Rajeswaran, D. K.—Tang, N.—Chew, E.—Chua, K. S. G.: Assessment of the Efficacy of EEG-Based MI-BCI with Visual Feedback and EEG Correlates of Mental Fatigue for Upper-Limb Stroke Rehabilitation. IEEE Transactions on Biomedical Engineering, Vol. 67, 2020, No. 3, pp. 786–795, doi: 10.1109/TBME.2019.2921198.

[6] Murphy, S. M.: Imagery Interventions in Sport. Medicine and Science in Sports and Exercise, Vol. 26, 1994, No. 4, pp. 486–494, doi: 10.1249/00005768-199404000-00014.

[7] Sharma, N.—Pomeroy, V. M.—Baron, J. C.: Motor Imagery: A Backdoor to the Motor System after Stroke? Stroke, Vol. 37, 2006, No. 7, pp. 1941–1952, doi: 10.1161/01.STR.0000226902.43357.fc.

[8] Kraeutner, S.—Gionfriddo, A.—Bardouille, T.—Boe, S. G.: Motor Imagery-Based Brain Activity Parallels That of Motor Execution: Evidence from Magnetic Source Imaging of Cortical Oscillations. Brain Research, Vol. 1588, 2014, pp. 81–91, doi: 10.1016/j.brainres.2014.09.001.

[9] Johnson, S. H.: Imagining the Impossible: Intact Motor Representations in Hemiplegics. Neuroreport, Vol. 11, 2000, No. 4, pp. 729–732, doi: 10.1097/00001756-200003200-00015.

[10] Babiloni, C. C.—Pizzella, V.—Del Gratta, C.—Ferretti, A.—Romani, G. L.: Fundamentals of Electroencefalography, Magnetoencefalography, and Functional Magnetic Resonance Imaging. International Review of Neurobiology, Vol. 86, 2009, pp. 67–80, doi: 10.1016/s0074-7742(09)86005-4.

[11] Kipf, T. N.—Welling, M.: Semi-Supervised Classification with Graph Convolutional Networks. International Conference on Learning Representations, 2017, doi: 10.48550/arXiv.1609.02907.

[12] Tavakoli, M.—Benussi, C.—Lopes, P. A.—Osorio, L. B.—de Almeida, A. T.: Robust Hand Gesture Recognition with a Double Channel Surface EMG Wearable Armband and SVM Classifier. Biomedical Signal Processing and Control, Vol. 46, 2018, pp. 121–130, doi: 10.1016/j.bspc.2018.07.010.

[13] Chen, L.—Fu, J.—Wu, Y.—Li, H.—Zheng, B.: Hand Gesture Recognition Using Compact CNN via Surface Electromyography Signals. Sensors, Vol. 20, 2020, No. 3, Art. No. 672, doi: 10.3390/s20030672.

[14] Côté-Allard, U.—Fall, C. L.—Drouin, A.—Campeau-Lecours, A.—Gosselin, C.—Glette, K.—Laviolette, F.—Gosselin, B.: Deep Learning for Electromyographic Hand Gesture Signal Classification Using Transfer Learning. IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 27, 2019, No. 4, pp. 760–771, doi: 10.1109/tnsre.2019.2896269.

[15] Toro-Ossaba, A.—Jaramillo-Tigreros, J.—Tejada, J. C.—Peña, A.—López-González, A.—Castanho, R. A.: LSTM Recurrent Neural Network for Hand Gesture Recognition Using EMG Signals. Applied Sciences, Vol. 12, 2022, No. 19, Art. No. 9700, doi: 10.3390/app12199700.

[16] Ruhunage, I.—Mallikarachchi, S.—Chinthaka, D.—Sandaruwan, J.—Lalitharatne, T. D.: Hybrid EEG-EMG Signals Based Approach for Control of Hand Motions of a Transhumeral Prosthesis. 2019 IEEE 1st Global Conference on Life Sciences and Technologies (LifeTech), 2019, pp. 50–53, doi: 10.1109/LifeTech.2019.8883865.

[17] Sarasola-Sanz, A.—Irastorza-Landa, N.—López-Larraz, E.—Bibián, C.—Helmhold, F.—Broetz, D.—Birbaumer, N.—Ramos-Murguialday, A.: A Hybrid Brain-Machine Interface Based on EEG and EMG Activity for the Motor Rehabilitation of Stroke Patients. 2017 International Conference on Rehabilitation Robotics (ICORR), 2017, pp. 895–900, doi: 10.1109/ICORR.2017.8009362.

[18] Cincotti, F.—Pichiorri, F.—Aricò, P.—Aloise, F.—Leotta, F.—de Vico Fallani, F.—del R. Millán, J.—Molinari, M.—Mattia, D.: EEG-Based Brain-Computer Interface to Support Post-Stroke Motor Rehabilitation of the Upper Limb. 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2012, pp. 4112–4115, doi: 10.1109/EMBC.2012.6346871.

[19] Zhang, J.—Wang, B.—Zhang, C.—Xiao, Y.—Wang, M. Y.: An EEG/EMG/EOG-Based Multimodal Human-Machine Interface to Real-Time Control of a Soft Robot Hand. Frontiers in Neurorobotics, Vol. 13, 2019, Art. No. 7, doi: 10.3389/fnbot.2019.00007.

[20] Hooda, N.—Das, R.—Kumar, N.: Fusion of EEG and EMG Signals for Clas-

sification of Unilateral Foot Movements. Biomedical Signal Processing and Control, Vol. 60, 2020, Art. No. 101990, doi: 10.1016/j.bspc.2020.101990.

[21] Klepl, D.—Wu, M.—He, F.: Graph Neural Network-Based EEG Classification: A Survey. IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 32, 2024, pp. 493–503, doi: 10.1109/TNSRE.2024.3355750.

[22] Demir, A.—Koike-Akino, T.—Wang, Y.—Haruna, M.—Erdogmus, D.: EEG-GNN: Graph Neural Networks for Classification of Electroencephalogram (EEG) Signals. 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021, pp. 1061–1067, doi: 10.1109/embc46164.2021.9630194.

[23] Aung, H. W.—Li, J. J.—Shi, B.—An, Y.—Su, S. W.: EEG_GLT-Net: Optimising EEG Graphs for Real-Time Motor Imagery Signals Classification. Biomedical Signal Processing and Control, Vol. 104, 2025, Art. No. 107458, doi: 10.1016/j.bspc.2024.107458.

[24] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data and Repealing Directive 95/46/EC (General Data Protection Regulation). 2016, `https://eur-lex.europa.eu/eli/reg/2016/679/oj`.

[25] Gahlan, N.—Sethia, D.: Federated Learning in Emotion Recognition Systems Based on Physiological Signals for Privacy Preservation: A Review. Multimedia Tools and Applications, Vol. 84, 2025, No. 13, pp. 12417–12485, doi: 10.1007/s11042-024-19467-3.

[26] Achard, S.—Salvador, R.—Whitcher, B.—Suckling, J.—Bullmore, E.: A Resilient, Low-Frequency, Small-World Human Brain Functional Network with Highly Connected Association Cortical Hubs. Journal of Neuroscience, Vol. 26, 2006, No. 1, pp. 63–72, doi: 10.1523/jneurosci.3874-05.2006.

[27] Barabási, A. L.—Albert, R.: Emergence of Scaling in Random Networks. Science, Vol. 286, 1999, No. 5439, pp. 509–512, doi: 10.1126/science.286.5439.509.

[28] Eguíluz, V. M.—Chialvo, D. R.—Cecchi, G. A.—Baliki, M.—Apkarian, A. V.: Scale-Free Brain Functional Networks. Physical Review Letters, Vol. 94, 2005, Art. No. 018102, doi: 10.1103/physrevlett.94.018102.

[29] Erdős, P.—Rényi, A.: On the Evolution of Random Graphs. Publicationes Mathematicae Instituti Hungarici Academiae Scientiarum, Vol. 5, 1960, No. 1–2, pp. 17–60, `https://renyi.hu/~p_erdos/1960-10.pdf`.

[30] Tan, G.—Wang, J.—Liu, J.—Sheng, Y.—Xie, Q.—Liu, H.: A Framework for Quantifying the Effects of Transcranial Magnetic Stimulation on Motor Recovery from Hemiparesis: Corticomuscular Network. Journal of Neural Engineering, Vol. 19, 2022, No. 2, Art. No. 026053, doi: 10.1088/1741-2552/ac636b (Data Set: TAN, Data Available from the Author).

[31] Koelstra, S.—Muhl, C.—Soleymani, M.—Lee, J. S.—Yazdani, A.—Ebrahimi, T.—Pun, T.—Nijholt, A.—Patras, I.: DEAP: A Database for Emotion Analysis Using Physiological Signals. IEEE Transactions on Affective Computing, Vol. 3, 2012, No. 1, pp. 18–31, doi: 10.1109/T-AFFC.2011.15.

[32] Ding, Y.—Robinson, N.—Tong, C.—Zeng, Q.—Guan, C.: LGGNet: Learning from Local-Global-Graph Representations for Brain-Computer Interface. IEEE Transactions on Neural Networks and Learning Systems, Vol. 35, 2024, No. 7, pp. 9773–9786, doi: 10.1109/TNNLS.2023.3236635.

[33] Massa, S. M.—Riboni, D.—Nazarpour, K.: Explainable AI-Powered Graph Neural Networks for HD EMG-Based Gesture Intention Recognition. IEEE Transactions on Consumer Electronics, Vol. 70, 2024, No. 1, pp. 4499–4506, doi: 10.1109/TCE.2023.3333421.

[34] Li, Y.—Yang, Y.—Zheng, Q.—Liu, Y.—Wang, H.—Song, S.—Zhao, P.: Dynamical Graph Neural Network with Attention Mechanism for Epilepsy Detection Using Single Channel EEG. Medical and Biological Engineering and Computing, Vol. 62, 2024, No. 1, pp. 307–326, doi: 10.1007/s11517-023-02914-y.

[35] Wei, X.—Faisal, A. A.: Federated Deep Transfer Learning for EEG Decoding Using Multiple BCI Tasks. 2023 11$^{\text{th}}$ International IEEE/EMBS Conference on Neural Engineering (NER), 2023, pp. 1–4, doi: 10.1109/NER52421.2023.10123713.

[36] Liu, R.—Chen, Y.—Li, A.—Ding, Y.—Yu, H.—Guan, C.: Aggregating Intrinsic Information to Enhance BCI Performance Through Federated Learning. Neural Networks, Vol. 172, 2024, Art. No. 106100, doi: 10.1016/j.neunet.2024.106100.

[37] Baek, J.—Jeong, W.—Jin, J.—Yoon, J.—Hwang, S. J.: Personalized Subgraph Federated Learning. In: Krause, A., Brunskill, E., Cho, K., Engelhardt, B., Sabato, S., Scarlett, J. (Eds.): Proceedings of the 40$^{\text{th}}$ International Conference on Machine Learning. Proceedings of Machine Learning Research (PMLR), Vol. 202, 2023, pp. 1396–1415, `https://proceedings.mlr.press/v202/baek23a.html`.

[38] Meng, C.—Rambhatla, S.—Liu, Y.: Cross-Node Federated Graph Neural Network for Spatio-Temporal Data Modeling. Proceedings of the 27$^{\text{th}}$ ACM SIGKDD Conference on Knowledge Discovery & Data Mining (KDD '21), 2021, pp. 1202–1211, doi: 10.1145/3447548.3467371.

[39] Tang, T.—Han, Z.—Cai, Z.—Yu, S.—Zhou, X.—Oseni, T.—Das, S. K.: Personalized Federated Graph Learning on Non-IID Electronic Health Records. IEEE Transactions on Neural Networks and Learning Systems, Vol. 35, 2024, No. 9, pp. 11843–11856, doi: 10.1109/TNNLS.2024.3370297.

KaiZheng Li is a Master's student in the School of Computer Science and Technology, Tongji University. His research interests include machine learning, brain-computer interface, and cognitive neuroscience.

Hongfei Ji is Associate Professor at the School of Computer Science and Technology, Tongji University, Shanghai, China. He received his Ph.D. degree in software engineering from the Tongji University, Shanghai, China, in 2015. His research interests include machine learning, pattern recognition, brain computer interface, and cognitive neuroscience.

Wei Yang is Associate Professor with the Department of Economics and Management, Shanghai Ocean University, Shanghai, China. She received her Ph.D. degree from Fudan University, Shanghai, China. Her research interests include technological innovation and policy evaluation.

Jie Li is Associate Professor in the School of Computer Science and Technology, Tongji University, Shanghai, China. She received her Ph.D. degree in computer science from the Shanghai Jiao Tong University, Shanghai, China, in 2010. Her research interests include machine learning, brain-computer interface, brain-like computation, and cognitive neuroscience.