# ST-GF: Graph-based Fusion of Spatial and Temporal Features for EEG Motor Imagery Decoding

Xuhui Wang
*School of Computer Science*
*Northwestern Polytechnical University*
Xi'an China
xuhuiwang@mail.nwpu.edu.cn

Kui Zhao
*School of Computer Science*
*Northwestern Polytechnical University*
Xi'an China
kui.zhao@mail.nwpu.edu.cn

Enze Shi
*School of Computer Science*
*Northwestern Polytechnical University*
Xi'an China
ezshi@mail.nwpu.edu.cn

Sigang Yu
*School of Computer Science*
*Northwestern Polytechnical University*
Xi'an China
sgyu@mail.nwpu.edu.cn

Geng Chen
*School of Computer Science*
*Northwestern Polytechnical University*
Xi'an China
geng.chen.cs@gmail.com

Shu Zhang
*School of Computer Science*
*Northwestern Polytechnical University*
Xi'an China
shu.zhang@nwpu.edu.cn

***Abstract*—The Motor Imagery (MI) decoding based on electroencephalogram (EEG), has promising applications. However, most current methods face two main issues: (1) They usually rely on convolutional neural networks to extract temporal features of MI signals without fully considering the brain's functional connectivity during MI tasks. (2) They lack analysis and recognition of MI features slices and non-tasks slices within EEG signals, leading to poor generalization and robustness. To address these problems, we propose a novel deep learning model based on graph neural network to learn spatial features between multiple electrode channels and integrate the brain's functional connectivity features. Additionally, it restructures time slices features segmented by the sliding time window algorithm to enhance MI temporal features in EEG signal. Therefor our model achieves the fusion of spatial and temporal features. To enhance the convergence effect of the model, we introduce electrode channel spatial positions as prior knowledge to initialize the parameters of the graph convolutional network parameters. Experimental evaluations on the publicly available EEG MI dataset from BCI Competition IV 2a show that our model achieves a four-class cross-session classification accuracy of 82.38%. Compared with other methods, our model yields the best results, demonstrating its superiority. Furthermore, the results indicate that the spatial feature obtained through our model bears resemblance to the brain functional connectivity patterns identified during MI tasks. To conclude, the fusion of spatial and temporal features with graph model shows the great application potential for EEG MI signals decoding and other EEG analysis.**

***Keywords—Motor Imagery, electroencephalogram, graph neural network, brain functional connectivity, fusion of spatial and temporal features***

## I. INTRODUCTION

Electroencephalography (EEG) provides neurophysiological signals that can directly record brain activity with millisecond-level temporal resolution [1]. It is widely used in Brain-Computer Interface (BCI) as a medium for interaction between the human brain and the external world. EEG Motor Imagery (MI) signals are produced by the brain when simulating a specific movement through imagination without actual motion represents. It's a typical type of self-induced mental activity. Due to its advantages of noninvasive, low clinical risk, economy, convenience and independent of stimulus targets, it has been widely researched for various fields [2]. For example, EEG-MI BCI has been applied to control intelligent mechanical equipment and motor recovery such as robot [3], drones [4], and Parkinson's disease patients [5].

With the development of machine learning technologies, by leveraging their strong capabilities in solving nonlinear problems, significant progress has been made in decoding EEG signals. H. Lu *et al.* [6] uses the Common Spatial Patterns (CSP) algorithm to establish a common spatial filter for binary data classification. However, traditional machine learning models have limited ability to extract spectro-spatio-temporal features and they are mainly optimized for binary classification. Therefor they are unable to achieve end-to-end training and testing. The development of deep learning resolves this problem. Vernon J. Lawhern *et al.* [8] proposed the EEGNet model based on deep learning and deep Separable convolutions, which demonstrated excellent classification performance in a four-class task. This model operates end-to-end without requiring traditional signal analysis on EEG data. Based on Convolutional Neural Network (CNN) Wang J *et al.* [2] introduced the Interactive Frequency Convolutional Neural Network (IFNet) based on CNN, which divides EEG signals into multiple frequency bands to extract features separately, followed by a feature fusion analysis to further enhance classification performance. These methods research from the perspective of computer algorithms, lacking guidance from neuroscience. Neuroscientific studies suggest that cooperation between different brain functional regions drives high-level cognitive processes. Traditional CNN-based methods are not suitable for capturing the rich spatial information in multi-electrode EEG signals and the collaborative relationships between different brain functional regions. To address this issue, Yi Ding *et al.* [9] proposed the Local Global Graph Network (LGGNet), which combines graph convolution with medical prior knowledge of brain functional connectivity. LGGNet demonstrates great potentials of Graph Neural Networks (GNNs) in accurate EEG emotion classification tasks.

Although these models have achieved advancements in EEG signal classification, they still grapple with two limitations: Firstly, few studies regarding EEG MI signals use graph to make good use of multi electrode feature and some graph based model usually relies on medical prior knowledge This problem makes it difficult for the model to have superior performance and transfer from other classification tasks to EEG MI decoding tasks. Secondly, The current researches lack analysis and recognition of continuous MI features in EEG MI signals. It leads to EEG signals containing some non-task state noises and weakening task state features. These two limitations limit model performance and reduce generalization.

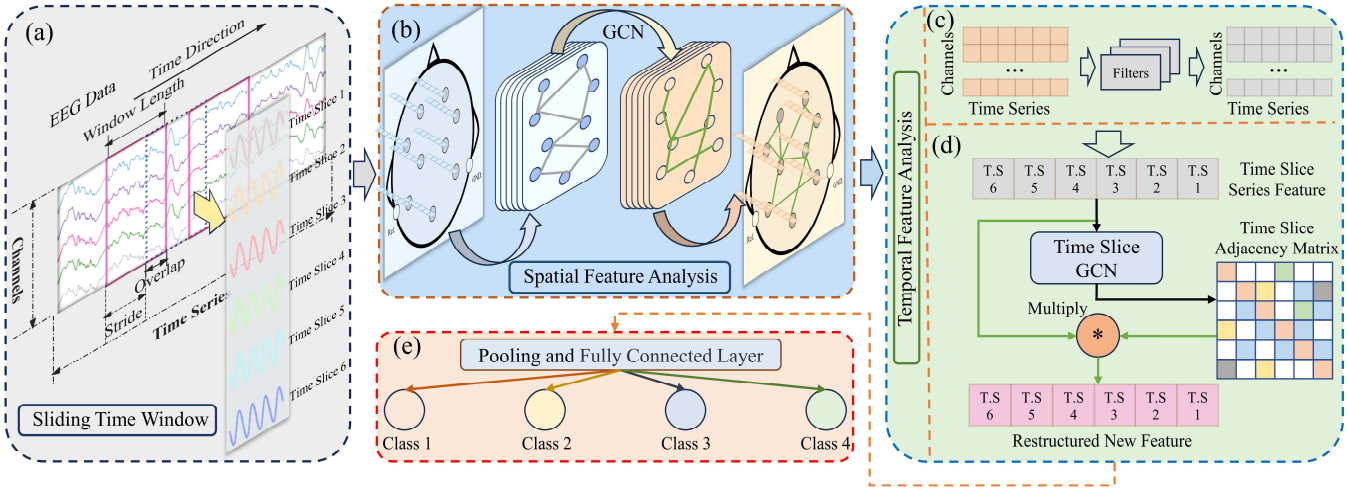


Fig. 1. The architecture of proposed ST-GF. (a) represents the sliding time window algorithm. (b) represents the Channel Correlation GCN (CCG) module to analysis spatial feature. (c) represents the EEG time series feature extraction. (d) represents the Time Slice GCN (TSG) module to restructure time feature. (c) and (d) represemts the temporal feature analysis module. (e) represents the classfication module.

To overcome these limitations, we propose **S**patial and **T**emporal **G**raph **F**usion (ST-GF), which integrates spatial and temporal features with a graph-based fusion model for EEG-MI decoding. We use GNNs to structure Channel Correlation Graph (CCG) and Time Slice Graph (TSG) module. The CCG module can learn EEG multi electrode channels spatial feature. The electrode channels spatial positions are introduced as the prior knowledge to initialize adjacency matrix instead of medical prior knowledge with poor generalization. The TSG module can restructure EEG signals to achieve more continuous MI features and strengthen task state temporal feature which is fused with spatial feature. The main contributions of our work are summarized as follow:

• We propose a GNN-based module to analyze spatial feature connections among multiple electrode channels. This allows a visual analysis of brain functional connectivity features during MI tasks using neuroscience theories.

• We design a sliding time window algorithm to segment EEG signals and restructure the time segments using GNNs to strengthen MI features within the signals.

• We introduce a simple prior knowledge—the electrode channels spatial positions as the adjacency matrix initialization method to strengthen the model's generalization capabilities.

Our code is publicly available at: https://github.com/guadawcz/ST-GF

## II. Methods

### A. Overview

The architecture of the proposed **S**patial and **T**emporal **G**raph **F**usion model is illustrated in Fig. 1: In general it involves five sections. (1) We segment the input EEG signal using the sliding time window algorithm, as shown in Fig. 1. (a). (2) The segmented data are input into the CCG module to learn and extract the spatial relationship features between multiple electrode channels, as shown in Fig. 1. (b). (3) We utilize a Depthwise Separable convolutional neural network [10] to extract EEG temporal features of the signal, as shown in Fig. 1. (c). To strengthen the model's ability to solve nonlinear problems, introduce the GELU activation function after extracting the temporal features. (4) Once the time slices with fused spatial and temporal features are obtained, as shown in Fig. 1. (d), we use the TSG module to restructure, strengthen the features within the time slices, acquire more continuous features of MI tasks. (5) We compresses the features using an average pooling layer, and then reduce the dimensionality of the feature tensor using fully connected layers to obtain the final classification results, as shown in Fig. 1. (e).

### B. EEG Spatial Feature Analysis

We utilize GCN to structure the CCG module, which can integrate the spatial feature of electrode channels into the classification model to improve robustness and performance. For EEG data, the input data $H^0$ is $X \in R^{N\times(C\times T)}$, where $N$ is the number of trail, $C$ is the number of electrode channels and $T$ is the number of time point, each electrode channel is treated as a node, and signals on each channel are treated as the feature vector. The graph convolution formula for one layer is shown below:

$$H_C^{(l+1)} = \sigma\left(D_C^{-\frac{1}{2}} A_C D_C^{-\frac{1}{2}} H_C^l W_C^{(l)}\right) \tag{3}$$

where $\sigma$ is the activation function. $A_{ij}$ is set as a trainable parameter and optimize it through backpropagation. The degree matrix $D_{ii} = \sum_j A_{ij}$. The connections between brain region's strength decreases with the physical distance increases. Therefor the edges between two spatially neighboring electrode channels are set to 1, which represent that the edges exists. The edges that are not between two spatially neighboring electrode channels are set to 0, which represents that the edge doesn't exit. The neighboring electrodes are the top, bottom, left, right, top left, bottom left, top right, and bottom right electrodes. Therefor the values $A_{ij}$ in the adjacency matrix is shown below:

$$A_{ij} = \begin{cases} 1, & if\ n_i\ and\ n_j\ are\ neighbours\ or\ i = j. \\ 0, & if\ n_i\ and\ n_j\ are\ not\ neighbours. \end{cases} \tag{6}$$

### C. EEG Temporal Feature Analysis

#### 1) EEG Time Series Feature Extraction

While GCNs are effective in capturing the spatial features of EEG signals, CNNs have an advantage in capturing temporal features. In sequential data, neighboring time steps exhibit high correlation, and local features are of significant

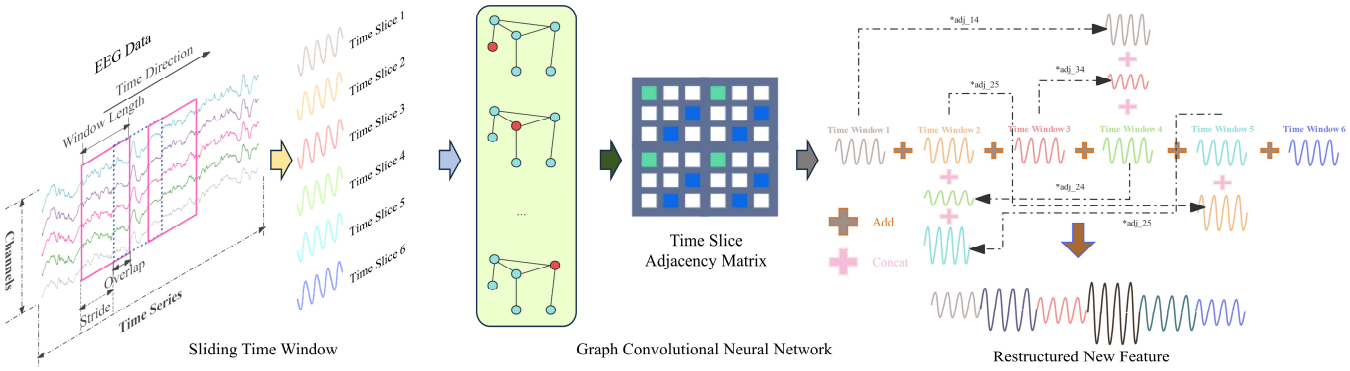


Fig. 2. Illustraiton of the TSG module procedure. The feature is restructured a new feature by GCN.

importance. Through convolutional operations, CNNs through their convolutional operations can extract these local features. We employ Depthwise Separable convolution to separate the convolutional operation into Depthwise convolution and pointwise convolution. This reduces the number of parameters and the computational complexity of the model. Additionally, pointwise convolution facilitates information exchange between channels, making the model more lightweight, efficient, and capable of generalizing well. Fig. 1. (c) illustrates the Time Series Feature Extraction module.

*2) Time Slice GCN*

In a complete segment of EEG MI signals, there may be multiple MI tasks present. Therefore, segmenting the signal into multiple time slices and performing analysis within each slice allows for increased data quantity and improved extraction of local patterns and trends. This approach captures the temporal features and dynamic changes of the EEG signal, which is beneficial for analyzing Event-Related Potentials (ERPs) and frequency variations, among other brainwave activities. We utilize the sliding time window algorithm to segment EEG data, which is illustrated in Fig. 1. (a). For EEG data $X \in R^{N\times(C\times T)}$, it can be segmented into $X \in R^{N\times W\times(C\times T)}$, where $W$ represents the number of time windows. The calculation formula for $W$ is as shown below:

$$W = \frac{len(X) - W_n}{Stride} + 1 \tag{7}$$

where $len(X)$ represents the number of time point, $W_n$ represents the length of each window, $Stride$ represents the length of each sliding step of the sliding time window.

However, when using the sliding time window algorithm to segment signals, it may disrupt the complete EEG MI features. If these segmented signals are directly input into a neural network, it may affect the classification performance and model robustness. To address this issue, we propose the TSG module. Specifically, we abstract multiple temporal slices of data into graph data, where each temporal slice represents a node, and the data within each temporal slice represents the node's feature vector. After learning the temporal features within the temporal slices using CNN, we further utilize GCN to learn the connections between the abstracted graph data of multiple temporal slices and restructure the features of the temporal slices. This feature enhancement is beneficial for the classification performance of the model. Fig. 2 illustrates the procedure of TSG module. The graph convolution formula for one layer is shown below:

$$H_T^{(l+1)} = \sigma\left( D_T^{-\frac{1}{2}} A_T D_T^{-\frac{1}{2}} H_T^l W_T^{(l)} \right) \tag{8}$$

where $H_T^{(0)} = X$, $X \in R^{W\times N\times(C\times T)}$.

## III. Experiments and Results

### *A. Dataset and Data Processing*

To evaluate our method, we use one publicly available dataset named BCI Competition IV 2a (BCIC-IV-2a) [11].

*1) Dataset*

The BCIC-IV-2a dataset consists of EEG MI data from 9 healthy subjects. It includes 22 Ag/AgCl electrode channels sampled at 250 Hz, with a bandpass filter applied between 0.5 Hz and 100 Hz to record EEG signals, and 3 monopolar electrode channels for recording EOG signals. The amplifier sensitivity was set to 100 μV, and an additional 50 Hz notch filter was used to suppress line noise. The dataset features four different MI tasks: imagination of left-hand movement (Class 1), right-hand movement (Class 2), both feet (Class 3), and tongue (Class 4). Data were recorded across two sessions on different days, comprising training and test sets. Each session contains 288 trials, with 72 trials per class.

*2) Data Processing*

We extract the 4 seconds motion imagination component from the raw signal as the experimental data. Thus there are 1000 time points for each trail. 0.5 Hz~40 Hz bandpass filter is utilized to preserve the alpha and beta frequency bands, which are most relevant to MI tasks. We adopt repeated trail

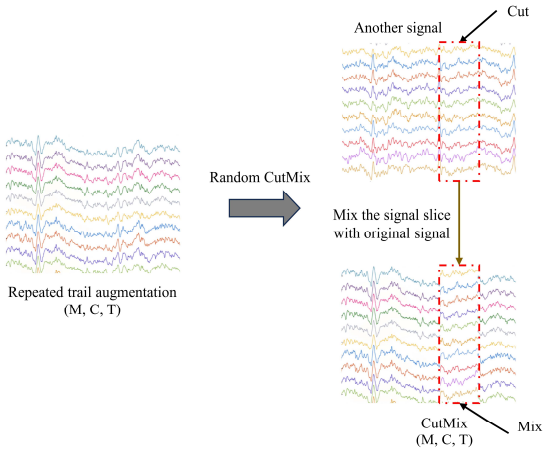


Fig. 3. Illustration of data augmentation. M: number of repeated trails, C: number of EEG channels, T: number of time points.

augmentation consisting of random cutmix [12] to relieve the overfitting problem caused by small sample EEG dataset. Specifically, for each sample in a batch, we generate $M = 5$ multiple instances through random cutmix which mixes the same regions of two signals of the same category together with p=0.5 probability. Fig. 3 shows the detailed illustration of data augmentation.

### *B. Experiment settings*

In our experiments, we utilize cross-entropy loss and Adam optimizer with a learning rate of $2^{-12}$ and weight decay of 0.01 to update network parameters during training. The batch size is set to 32 and epoch of each experiment is set to 2000. Throughout the whole train and test procedure a random seed is set to 2024. The training and testing program runs using NVIDIA RTX 3090 with PyTorch 1.13.3 and CUDA 11.2. The training time for 2000 epochs of each subjects and testing time for each epoch of one subject respectively takes approximately 19 minutes and less than 0.15s, which indicate that our model has the potential for real-time EEG-MI signal decoding. In the CCG module and TSG module, the number of graph convolution layers is set to 2. We set $W_n = 125$ which represents 0.5s-long EEG characteristics and $Stride = 100$ to segment data. Therefor each trail is segmented into 9 slices i.e. $W = 9$. The size of pooling layer is set to 125.

### *C. Experimental Results*

To evaluate the performance of our proposed approach, we compare the strategy with some advanced models on BCIC-IV-2a dataset, including one machine learning model (FBCSP-SVM [7] ), and five deep learning models (SCCNet [13], EEGNet [8], FBCNet [14], ConvNet [15] and IFNet [2]). Table I shows the classification results of different models on BCIC-IV-2a dataset. T-test is used for significance analysis and the discrepancy in results is statistically significant. Different adjacency matrix initialization methods have an impact on the performance of the model. Therefore, we research other two CCG module adjacency matrix initialization methods as the comparative experiment: one method named Gaussian Random without prior knowledge, another method named Pearson Correlation Coefficient with simple prior knowledge. We adopt the same training procedure and data augmentation to evaluate on BCIC-IV-2a dataset. Table II shows the classification results with different initialization method. GR stands for the Gaussian Random. PCC stands for the Pearson Correlation Coefficient. CSL stands for the Channel Spatial Location.

From Table I and Table II we have the following observations: (1) In Table I, ST-GF achieves a classification of 82.38% which consistently outperforms the other models. Notably, ST-GF and other deep-learning-based model also outperform FBCSP-SVM. These result demonstrate that the deep learning has the more powerful non-linear problem fitting ability to learn EEG complex feature. (2) In Table II, the adjacency matrices initialization method based on Channel Spatial Location achieves the highest classification accuracy, which indicates that integrating prior knowledge of channel spatial location in the classification model of EEG MI tasks can improve the classification performance of the model, and this prior knowledge is a relatively high-quality information compared to Pearson correlation coefficient. Meanwhile, the adjacency matrices initialization method based on Pearson Correlation Coefficient achieves higher accuracy than Gaussian Random. It shows that simple prior knowledge is beneficial for classification model.

From CCG module (as shown in Fig. 1. b), our model can learn electrode channels correlation. To research the functional area connections of the brain during MI tasks, we visualize the channel correlation of 9 subjects of BCIC-IV-2a to brain regions. The Fig. 4 shows the visualization results. The 22 nodes stands for the 22 electrode channels and the locations refer to the actual position of the electrode channels. The size of nodes represents the different weight, and the thickness and color depth of edges between nodes represent the strength of connections. To make the results clearer, we removed edges with very small connections, because this part of the edge has a negligible impact on the result due to its low weight. From Fig. 4, we have the following observations: (1) The channel weights in the prefrontal cortex region of the brain are relatively high. The prefrontal cortex contains the motor cortex region, which is involved in planning, controlling, and executing autonomous movements and plays an important role in controlling human movement, (2) The channels in the temporal lobe and occipital lobe region of the brain also have significant weight. The temporal lobe is mainly responsible for processing auditory information and conducting auditory perception. The occipital lobe is mainly responsible for processing visual information and forming visual perception. When collecting the dataset, there will be a brief prompt sound, and the subjects need to perform MI tasks or rest according to the prompt sound. Meanwhile, when the subjects start the experiment, they need to observe the computer screen and perform corresponding MI tasks. Therefore, it is necessary to activate the functional areas of auditory and visual perception to process information. (3) The

TABLE I. CLASSIFICATION ACCURACY AND STANDARD DEVIATION (IN PERCENTAGE %) OF DIFFERENT ALGORITHMS ON DATASET.

| **Method** | **Subject** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Sbj 01** | **Sbj 02** | **Sbj 03** | **Sbj 04** | **Sbj 05** | **Sbj 06** | **Sbj 07** | **Sbj 08** | **Sbj 09** | **Avg Acc±Std** |
| FBCSP-SVM [7] | 81.60 | 52.78 | 84.38 | 65.28 | 56.25 | 44.44 | 89.24 | 81.94 | 72.92 | 69.87$^{**}$±15.90 |
| SCCNet [13] | 78.82 | 59.38 | 89.93 | 68.40 | 67.71 | **64.93** | 80.56 | 80.56 | 76.74 | 74.11$^{**}$±9.05 |
| EEGNet [8] | 83.90 | 54.26 | 91.56 | **76.52** | 64.29 | 48.31 | 86.45 | 82.93 | 82.20 | 74.49$^{**}$±14.36 |
| FBCNet [14] | 85.80 | 56.01 | 89.67 | 70.87 | 65.70 | 57.17 | 90.70 | 84.48 | 83.06 | 75.94$^{**}$±13.73 |
| ConvNet [15] | 76.39 | 55.21 | 89.24 | 74.65 | 56.94 | 54.17 | 91.27 | 77.08 | 76.39 | 72.37$^{**}$±13.19 |
| IFNet [2] | 88.47 | 56.35 | 91.77 | 73.78 | 69.72 | 60.42 | 89.24 | 85.42 | **88.72** | 78.21$^{*}$±13.51 |
| **Ours** | **90.57** | **72.10** | **93.72** | 76.10 | **77.10** | 62.85 | **91.31** | **90.03** | 87.64 | **82.38**±10.90 |

The *:p<0.05 and **:p<0.01 represent the T-test difference between the performance of Our Model and given model. "Avg Acc±Std" means average accuracy and standard deviation.

TABLE II. THE CLASSIFICATION ACCURACY AND STANDARD DEVIATION (IN PERCENTAGE %) WITH DIFFERENT INITIALIZATION METHOD OF CHANNEL CORRELATION GRAPH ADJACENCY MATRIX.

| Method | Subject | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Sbj 01** | **Sbj 02** | **Sbj 03** | **Sbj 04** | **Sbj 05** | **Sbj 06** | **Sbj 07** | **Sbj 08** | **Sbj 09** | **Avg Acc±Std** |
| GR | 87.71 | 69.06 | 93.55 | 73.96 | 71.63 | 61.86 | **91.67** | 87.67 | 86.05 | 80.33±10.69 |
| PCC | 89.00 | 69.57 | 93.12 | 74.12 | 72.71 | 62.64 | 90.22 | 87.50 | 87.56 | 80.71±10.35 |
| CSL | **90.57** | **72.10** | **93.72** | **76.10** | **77.10** | **62.85** | 91.31 | **90.03** | **87.64** | **82.38**±10.90 |

GR is short for Gaussian Random. PCC is short for Pearson Correlation Coefficient. CSL is short for Channel Spatial Location.

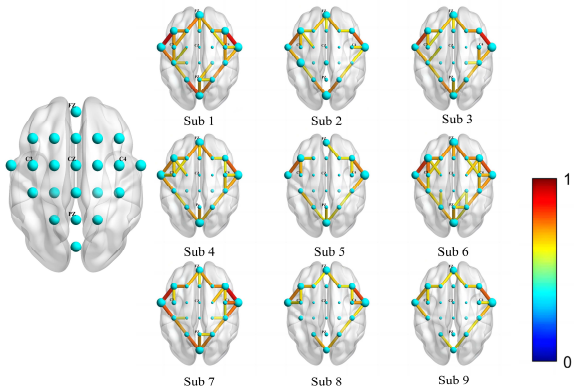


Fig. 4. The channel correlation of 9 subjects of BCIC-IV-2a on brain regions visualization results. The left of the figure shows the channel locatrion in brain region. The 22 nodes stands for the 22 electrode channels The size of nodes represents the different weight. The thickness and color depth of edges between nodes represent the strength of connections, which varies from 0 to 1.

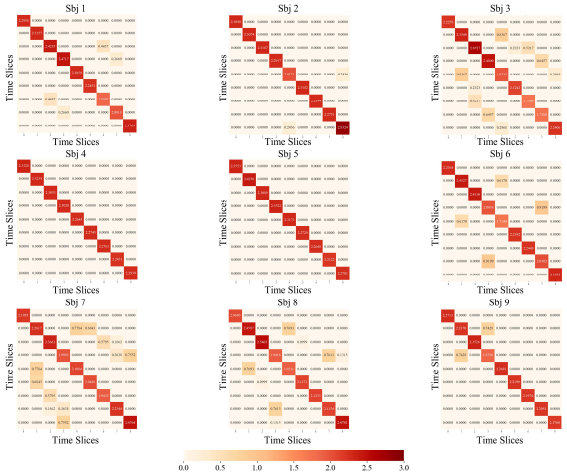


Fig. 5. The heatmap of the relationships between different time slices. The diagonal elements represent the weight of the time slice. The remaining elements represent the connections between different time slices.

edges with higher weights connect the electrode channels with higher weights in the frontal, occipital, and temporal regions, indicating strong functional connection between these areas. More importantly the connections of the 9 individuals are similar. The results are consistent with the cognitive understanding of brain functional regions working together to perform tasks in neuroscience, proving that the channel correlation obtained by our model are to some extent consistent with the brain functional connectivity network.

With TSG module, it can learn the relationships between different time slices sliced in one same time series through GNNs and restructure them. Fig. 5 visualizes the relationship on BCIC-IV-2a dataset. From the observation of Fig. 5, we have the following observations: (1) The diagonal elements represent the weights of the time slices themselves with each time slice being assigned a distinct weight. Consequently, it diminishes the influence of time slices exhibiting ambiguous MI features while enhancing the influence of time slices characterized by vivid and consistent MI features on classification outcomes. Enhanced vivid and consistent MI features prove advantageous in improving the performance of the model. (2) The non-diagonal elements represent the connections between different time slices and different weights shows different connections. Through these connections, it can restructure discontinuous or segmented MI features into more coherent features, amalgamating them into an cohesive feature set alongside weighted time slice features. It serves to boost the classification efficacy of the model. Given the diverse cognitive habits exhibited by different subjects, the 9 subjects results shown in Fig. 5 are different with each other. These time slice weights and connections improve the classification performance of the model and support CCG module in obtaining more pertinent brain functional connectivity patterns during MI tasks.

### *D. Ablation Study*

In this subsection, to validate the effectiveness of graph convolution module and data augmentation in ST-GF, we conduct the ablation study using cross-session experiment on the BCIC-IV-2a dataset. We adopt the same training procedure for all ablation experiments.

#### *1) Ablation on Graph Convolution Module*

Table III shows the individual and average classification results achieved with removing different graph convolution module. The experimental results indicate that the model with CCG and TSG module achieve the best performance with improves the overall accuracy improvement of 6.46%, compared with the model only using EEG time series feature extraction module. CCG module improves the overall accuracy by 2.17%, which indicates that incorporating the spatial feature of electrode channels into the model is beneficial for classification tasks. Meanwhile, it proves that the functional area connections of the brain during MI tasks which achieved by our model is consistent with the cognition of neuroscience. TSG module improves the overall accuracy by 4.32%, which demonstrates that obtaining more continuous MI features by restructuring time-slice features contributes to the improvement in accuracy.

#### *2) Ablation on Data Augmentation*

Table IV shows the individual and average classification accuracy achieved with removing data augmentation. RC stands for Random Cutmix. Data augmentation improves the overall accuracy by 5.02%, indicating the effectiveness of this method in the context of MI- EEG decoding. The result can attribute to two aspects: (1) Repeated trail augmentation increases the dataset. This helps the model learn a wider range

TABLE III. CONTRIBUTION OF EACH GRAPH MODULE IN OUR MODEL ON DATASET.

| Module | Subject | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Sbj 01** | **Sbj 02** | **Sbj 03** | **Sbj 04** | **Sbj 05** | **Sbj 06** | **Sbj 07** | **Sbj 08** | **Sbj 09** | **Avg Acc±Std** |
| No Graph | 85.48 | 57.93 | 91.80 | **78.11** | 61.80 | 53.35 | 82.27 | 86.60 | 85.90 | 75.92*±13.48 |
| CCG | 86.80 | 68.73 | 93.30 | 71.38 | 68.25 | 61.73 | 84.73 | 83.67 | 84.26 | 78.09**±10.10 |
| TSG | 89.17 | 70.84 | 92.70 | 73.86 | 71.12 | 59.17 | 91.00 | 87.87 | 86.41 | 80.24**±11.07 |
| CCG&TSG | **90.57** | **72.10** | **93.72** | 76.10 | **77.10** | **62.85** | **91.31** | **90.03** | **87.64** | **82.38**±10.90 |

The *: p<0.05 and **:p<0.01 represent the T-test difference between the performance of each module and the whole model.
CCG is short for Channel Correlation Graph. TSG is short for Time Slice Graph

TABLE IV. CONTRIBUTION OF EACH DATA AUGMENTATION IN OUR MODEL ON DATASET.

| Module | Subject | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Sbj 01** | **Sbj 02** | **Sbj 03** | **Sbj 04** | **Sbj 05** | **Sbj 06** | **Sbj 07** | **Sbj 08** | **Sbj 09** | **Avg Acc±Std** |
| No RC | 86.91 | 55.36 | 92.06 | **78.91** | 64.66 | 56.17 | 89.81 | 84.98 | 87.42 | 77.36*±13.82 |
| RC | **90.57** | **72.10** | **93.72** | 76.10 | **77.10** | **62.85** | **91.31** | **90.03** | **87.64** | **82.38**±10.90 |

The *: p<0.05 represent the T-test difference between the performance of data augmentation and no data augmentation. RC is short for Random Cutmix

of feature representations, and improves its robustness and generalization. (2)Through the random blending of partial regions from the same class sample, it increases the model's robustness to local variations in the dataset. It's beneficial in handling noise present in EEG data.

## IV. CONCLUSION

In this work, we propose ST-GF model to further improve EEG MI decoding ability. In ST-GF model, the CCG module is structured to integrate the spatial feature of electrode channels. To enhance the convergence of the algorithm, we introduce electrode channel spatial positions as the simple prior knowledge to initialize the adjacency matrix. The TSG module is structured to restructure time slices and enhance MI features in EEG signal. The results on the BCIC-IV-2a dataset show that our proposed model achieves superior performance compared to other models. We discovered that the spatial feature obtained through our model bears resemblance to the brain functional connectivity patterns identified in neuroscience during MI tasks. Future work will focus on upgrading our model with attention mechanism to achieve more accurate brain functional connectivity patterns and better classification performance of other BCI cross-subject tasks with their brain functional connectivity patterns.

## ACKNOWLEDGMENT

This work was supported by Foundational Research Project in Specialized Discipline (Grant No. G2024WD0146); Faculty Construction Project (Grant No. 24GH0201148).

## V. REFERENCES

[1] A. Subasi and E. Erçelebi, "Classification of EEG signals using neural network and logistic regression," *Computer Methods and Programs in Biomedicine*, vol. 78, no. 2, pp. 87–99, May 2005.

[2] J. Wang, L. Yao, and Y. Wang, "IFNet: An Interactive Frequency Convolutional Neural Network for Enhancing Motor Imagery Decoding From EEG," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 1900–1911, 2023.

[3] Jd. R. Millan, F. Renkens, J. Mourino, and W. Gerstner, "Noninvasive brain-actuated control of a mobile robot by human EEG," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1026–1033, Jun. 2004.

[4] P. M. Jonassen and H. T. Lønvik, "An Exploration of Techniques for Electroencephalography-Based Motor Imagery Classification for Real-Time Drone Control," Master thesis, NTNU, 2023.

[5] A. Miladinović *et al.*, "Evaluation of Motor Imagery-Based BCI methods in neurorehabilitation of Parkinson's Disease patients," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Jul. 2020, pp. 3058–3061.

[6] H. Lu, H.-L. Eng, C. Guan, K. N. Plataniotis, and A. N. Venetsanopoulos, "Regularized Common Spatial Pattern With Aggregation for EEG Classification in Small-Sample Setting," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 12, pp. 2936–2946, Dec. 2010.

[7] K. K. Ang, Z. Y. Chin, H. Zhang, and C. 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)*, Jun. 2008, pp. 2390–2397.

[8] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, p. 056013, Jul. 2018.

[9] Y. Ding, N. Robinson, C. Tong, Q. Zeng, and C. Guan, "LGGNet: Learning From Local-Global-Graph Representations for Brain–Computer Interface," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 7, pp. 9773–9786, Jul. 2024.

[10] A. G. Howard *et al.*, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," Apr. 16, 2017, *arXiv*: arXiv:1704.04861.

[11] C. Brunner, R. Leeb, G. Müller-Putz, A. Schlögl, and G. Pfurtscheller, "BCI Competition 2008–Graz data set A," *Institute for knowledge discovery (laboratory of brain-computer interfaces), Graz University of Technology*, vol. 16, pp. 1–6, 2008.

[12] S. Yun, D. Han, S. Chun, S. J. Oh, Y. Yoo, and J. Choe, "CutMix: Regularization Strategy to Train Strong Classifiers With Localizable Features," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Oct. 2019, pp. 6022–6031.

[13] C.-S. Wei, T. Koike-Akino, and Y. Wang, "Spatial Component-wise Convolutional Network (SCCNet) for Motor-Imagery EEG Classification," in *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, Mar. 2019, pp. 328–331.

[14] R. Mane *et al.*, "FBCNet: A Multi-view Convolutional Neural Network for Brain-Computer Interface," Mar. 17, 2021, *arXiv*: arXiv:2104.01233.

[15] R. T. Schirrmeister *et al.*, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Human Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.