TEMPORAL CROSS-GRAPH NETWORK FOR BRAIN FUNCTIONAL ACTIVITY PREDICTION

Xinyu Yuan[1], *Wenhan Wang*[1], *Youyong Kong*[1,2,⋆], *Jiasong Wu*[1,2], *Guanyu Yang*[1,2], *Huazhong Shu*[1,2]

[1]Jiangsu Provincial Joint International Research Laboratory of Medical Information Processing, School of Computer Science and Engineering, Southeast University, China
[2]Key Laboratory of Computer Network and Information Integration, Southeast University, Ministry of Education, Nanjing, China
⋆Corresponding Author. E-mail:kongyouyong@seu.edu.cn

ABSTRACT

Prediction of brain functional activity is of great significance for neuroscience research. The brain functional activities at different regions are highly related, and their relationships can be captured with functional connectivity and structural connectivity. The existing works are challenging to integrate two connectivity information for functional activity prediction. In this paper, we propose a Temporal Cross-Graph Network (TCGN) for predicting brain functional activity, which can comprehensively exploit multi-modal spatial dependence and temporal patterns. In particular, a novel cross-graph convolution module is developed to capture the spatial features of brain structural and functional connectivity. A temporal fusion module is designed to learn the pattern of dynamic functional connectivity to guide the prediction. Specially, a multi-task loss function is proposed to incorporate functional activity and dynamic functional connectivity. Extensive experiments on the Human Connectome Project dataset demonstrate the effectiveness of the proposed framework.

Index Terms— Brain functional activity prediction, structural connectivity, graph convolutional network

1. INTRODUCTION

Brain functional activity analysis plays a vital role in understanding individual differences in human cognitive and behavioral performance[1, 2]. Resting state functional magnetic resonance imaging (rsfMRI) has become an essential technique to capture brain functional activity with blood oxygen level-dependent (BOLD) signals [3]. Plenty of evidence has demonstrated the reproducibility of intrinsic functional activity and dynamic pattern over time [4, 5]. Therefore, it is interesting to explore accurate functional activity prediction for the brain cognition and disease understanding.

This work is supported by the National Key Research and Development Program of China (No. 2021ZD0113202) and grant 62171125, 61876037, 31800825 National Natural Science Foundation of China.

In the past decade, great effort have been made to investigate functional activity prediction with functional connectivity (FC)[6, 7]. Indeed, the functional activity is not only related to FC, and is also highly related to the structural connectivity (SC). Recent evidence suggests that functional activity highly correlates with SC at an aggregate level [8, 9]. These findings provide strong evidence that functional interactions between brain regions are shaped by higher-level indirect anatomical connections and motivated relative research into combining FC and SC to predict functional activity [10].

A large number of approaches have been developed for temporal prediction in recent years. Machine learning methods such as Support Vector Regression (SVR) [11] and Auto-Regressive Integrated Moving Average (ARIMA) [12] rely on the stationary assumption. Long Short-term Memory (LSTM) [13] can better capture the long-distance temporal dependencies but ignore the spatial information. Thus, graph convolutional network (GCN) [14] model such as T-GCN [15] have been proposed to capture both the spatial and temporal features. However, these methods are challenging to integrate the multi-modal graph information for the functional activity prediction.

In this paper, we propose a novel **T**emporal **C**ross-**G**raph **N**etwork(**TCGN**) for forecasting brain functional activity consists of BOLD signals and dynamic FC matrices. We use temporal fusion of multi-modal spatial features to extract patterns of dynamic FC, and design a loss function for multi-scale assessment of brain functional activity prediction. To the best of our knowledge, this is the first study to forecast brain functional activity by integrating SC and FC.

2. METHODOLOGY

In this section, we give the problem statement and introduce the proposed network. The architecture of our model is shown in Figure 1.

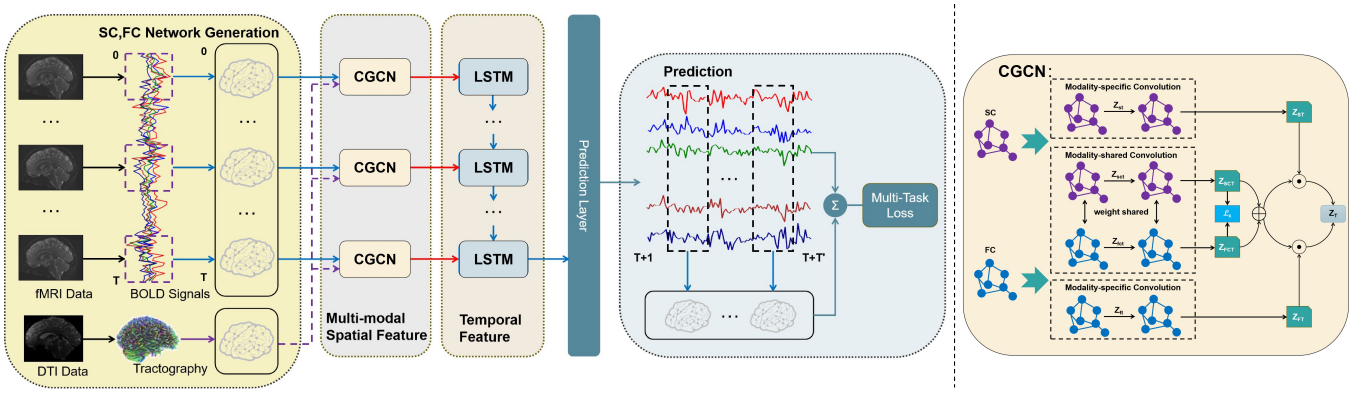

Fig. 1. The architecture of TCGN model. The brain SC and FC are considered as input to extract multi-modal spatial information and temporal feature through CGCN and LSTM respectively. The fully connected layer is designed to predict the BOLD signals from time $T+1$ to $T+T'$. Future dynamic FC is calculated by predicted BOLD signals. Fusion loss is applied as the overall model constraint.

2.1. Problem statement

An adjacency matrix $A_s \in \Re^{|V|*|V|}$ is used to describe the SC extracted from Diffusion Tensor Imaging (DTI) using fiber tracking [16]. Dynamic FC, computed by fMRI [17], can be defined as a sequence of matrices snapshots $A_f = A_{f_1}, ..., A_{f_\tau}$ and for the snapshot of time slice t, we use $A_{f_t} \in \Re^{|V|*|V|}$ to describe the corresponding dynamic FC.

After then we obtained the FC matrices and BOLD signals feature matrices of previous T time slices and current time slice $A_{f_{t-T}}, ..., A_{f_t}$ and $X_{t-T}, ..., X_t$. The feature matrix X_t includes the BOLD signals in the sliding window corresponding to its time slice t. Also we can get SC matrix as A_s. The goal of our task is to predict the BOLD signals of the next T' time slices $(t+1, ..., t+T')$, which can be formally described by:

$$X_{t+1}, ..., X_{t+T'} = f(X_{t-T}, ..., X_t; A_{f_{t-T}}, ..., A_{f_t}; A_s) \tag{1}$$

where $f(\cdot)$ is the desirable model for prediction while X_{t+1}, ..., $X_{t+T'}$ represent the prediction results.

2.2. Multi-modal Spatial Dependence Module

The existing brain functional activity prediction [6, 7] only utilized the FC information, which ignores the important SC. A number of recent work have demonstrated the strong relationships between the FC and SC [8, 9]. However, the existing works are challenging to integrate two connectivity information from these two modalities for functional activity prediction.

To overcome this issue, we propose a novel multi-modal spatial dependence module to integrate the FC and SC. Cross-graph convolutional network (CGCN) is proposed to extract the topological structure feature in two modals. The overall framework of CGCN is shown in the right of Figure 1. The key idea is to obtain the shared and specific features in topological space of SC and FC. The graph convolutional network (GCN) [14] is used for capture information of brain connectivity. Our CGCN includes the modality-shared convolution and modality-specific convolution. Each input will pass modality-specific convolutional layers to learn two specific embedding Z_s and Z_f, respectively. Further, considering that the information in these two spaces have common characteristics, a modality-shared convolution module with parameter sharing strategy is designed to learn common embedding Z_{st} and Z_{ft}. In order to enhance the similarity of the shared features of the two modalities, a consistency constraint $\mathcal{L}_c$ is adopted. As FC predictions are related to SC and FC, Hadamard product is utilized to integrate specific embedding with shared embedding. Therefore we can obtain the final embedding information Z.

Modality-specific Convolution With A_s as the brain SC network, we have the input graph $G_{st} = (A_{st}, X_t)$ where $A_{st} = A_s$ at time t. The output Z_{st} can be represented by a GCN layer as:

$$Z_{st} = ReLU(\tilde{D}_{st}^{-1/2} \tilde{A}_{st} \tilde{D}_{st}^{-1/2} X_t W_s) \tag{2}$$

where W_s is the weight matrix of the layer in GCN, $ReLU(\cdot)$ is the Relu activation function. Specifically, we have $\tilde{A}_{st} = A_{st} + I_{st}$ and $\tilde{D}_{st}$ is the diagonal degree matrix of $\tilde{A}_{st}$. We represent the output embedding of the layer as Z_{st}. In this way, we can learn specific spatial features from brain SC network.

As for the brain FC space, we have the original graph input $G_{ft} = (A_{ft}, X_t)$ at time t. The learned embedding output Z_{ft} can be then extracted in the same way as in brain SC network represented as:

$$Z_{ft} = ReLU(\tilde{D}_{ft}^{-1/2} \tilde{A}_{ft} \tilde{D}_{ft}^{-1/2} X_t W_f) \tag{3}$$

Modality-shared Convolution Previous studies [8, 9]

provide strong evidence that the functional interactions between brain regions are formed by indirect anatomical connections. Therefore, we cannot only extract the topological features in SC and FC spaces separately, but also need to capture the common embedding features shared from these two modals. In this way, we need to know which part of the shared information is more relevant for the task. In order to solve this problem, we design a modality-shared convolution with parameter sharing strategy to get the embedding shared in these two brain connectivities.

First, we use modality-shared convolution to extract the node embedding Z_{st} from SC graph (A_{st}, X_t) at time t as follows:

$$Z_{sct} = ReLU(\tilde{D}_{st}^{-1/2}\tilde{A}_{st}\tilde{D}_{st}^{-1/2}X_tW_c) \tag{4}$$

where W_c is the weight matrix of modality-shared convolution and Z_{sct} is the output embedding in this layer. The same weight matrix W_c is shared to learn spatial features from FC, in order to extract the shared information, as follows:

$$Z_{fct} = ReLU(\tilde{D}_{ft}^{-1/2}\tilde{A}_{ft}\tilde{D}_{ft}^{-1/2}X_tW_c) \tag{5}$$

The shared weight matrix can filter out the shared features from two modals and the shared embedding Z_{ct} of the two brain connectivities can be calculated by structural and functional embedding as:

$$Z_{ct} = (Z_{sct} + Z_{fct})/2 \tag{6}$$

Feature Fusion CGCN extracts the shared and specific features for each modality. We denote the shared spatial features as Z_{ct}, and represent the specific SC and FC features as Z_{st} and Z_{ft} at time t. The Hadamard product is used to combine these three embedding, which can integrate the influence of modality-specific characteristics and modality-shared information on the connection of each pair of brain regions. The final embedding Z_t at time t is obtained as follows:

$$Z_t = Z_{st} \odot Z_{ct} + Z_{ft} \odot Z_{ct} \tag{7}$$

2.3. Temporal Dependence Module

The brain functional activity has been demonstrated to have intrinsic dynamic change patterns [7]. Therefore, the temporal dependence module is proposed to capture the dynamic change pattern using the popular long short-term memory (LSTM). It uses gated mechanism to memorize as much long-term information as possible and is effective for various tasks. We obtain a hidden status for each time step, which will be used to control the information flow to the next time step and serve as the output at the current moment. It is worth noting that the LSTM operation is applied to each time node separately, and the parameters of LSTMs for all nodes are shared with each other. The output of the CGCN layer which represents the multi-modal spatial embedding is fed into the LSTM layer to capture temporal information. For each node at time t, the operation of the LSTM can be expressed as follows:

$$\begin{aligned} f_t &= \sigma(W_f \cdot [H_{t-1}, Z_t] + b_f) \\ i_t &= \sigma(W_i \cdot [H_{t-1}, Z_t] + b_i) \\ \tilde{C}_t &= tanh(W_C \cdot [H_{t-1}, Z_t] + b_C) \\ C_t &= f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \\ o_t &= \sigma(W_o \cdot [H_{t-1}, Z_t] + b_o) \\ H_t &= o_t \odot tanh(C_t) \end{aligned} \tag{8}$$

where f_t, i_t, o_t, and C_t are the output of the forget gate, input gate, output gate and the input memory cell. Z_t is the output of the CGCN layer and the input of the LSTM layer at time t, the hidden state H_t is the output of the current time step, which also serves as the part of input to the next time step. W_f, W_i, W_c, W_o are the parameters to be learned.

2.4. Loss Function

For the two output embeddings Z_{sct} and Z_{fct} in modality-shared convolution, a consistency constraint is designed to enhance their resemblance. We set the consistency constraint as $\mathcal{L}_c$ where:

$$\mathcal{L}_c = \|Z_{sct} - Z_{fct}\|_2^2 \tag{9}$$

After the LSTM layer, we use a fully connected layer to predict the BOLD signals in the future periods. Then we calculate the future FC matrices from the predicted BOLD signals. The mean squared error is used as the loss both for BOLD signals and FC prediction, which can be expressed as $\mathcal{L}_b$ and $\mathcal{L}_f$. A multi-task loss $\mathcal{L}_{mt}$ is used for two prediction tasks, the $\mathcal{L}_{mt}$ can be expressed as:

$$\mathcal{L}_{mt} = \mathcal{L}_b + \lambda\mathcal{L}_f \tag{10}$$

where λ is the parameter to balance the MSE loss of BOLD signals and FC matrices.

Combing the two prediction tasks and constraints, we have the following overall loss function:

$$\mathcal{L} = \mathcal{L}_{mt} + \gamma\mathcal{L}_c \tag{11}$$

where γ is parameter to balance consistency terms and MSE loss. We can train our predictive model by minimizing the loss of Eq.(11).

3. EXPERIMENT

3.1. Dataset

The proposed approach is evaluated on the real dataset from the publicly available Human Connectome Project (HCP) [18]. The dataset comprises of MRI imaging derived from 100 healthy subjects. The DTI and rs-fMRI imaging were processed to obtain the SC and dynamic FC for each subject.

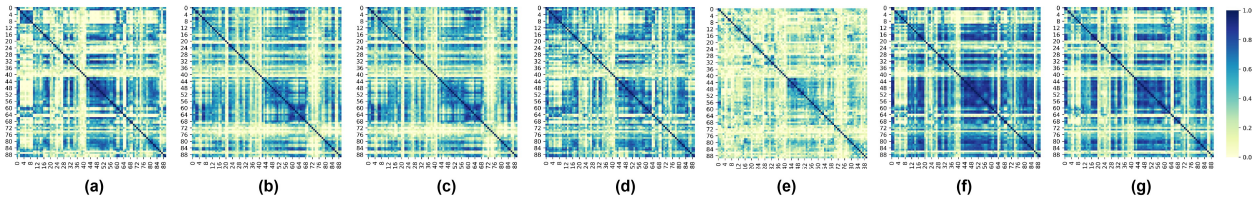

Fig. 2. (a)Ground Truth (b)ARIMA (c)SVR (d)FNN (e)LSTM (f)T-GCN (g)TCGN

Table 1. Performance of different methods for brain activity (including dynamic FC and BOLD signals) prediction

Steps	Metric(%)		Methods					
			ARIMA	SVR	FNN	LSTM	T-GCN	TCGN
10	FC	RMSE	8.52	6.20	8.72	4.65	4.95	**4.36**
		MAE	5.79	3.99	4.46	2.80	2.89	**2.57**
	BOLD	RMSE	7.17	6.56	6.89	6.79	6.53	**6.25**
		MAE	5.22	4.97	5.30	4.72	4.86	**4.62**
40	FC	RMSE	17.42	15.78	11.56	11.56	11.99	**10.88**
		MAE	13.29	11.53	10.75	7.15	7.74	**7.05**
	BOLD	RMSE	7.11	**6.61**	7.15	6.78	6.90	6.66
		MAE	5.32	5.01	5.55	5.08	5.16	**4.91**

Table 2. Comparison of prediction in TCGN without FC or SC modal or multi-task loss (ml)

Steps	Metrics(%)		w/o FC	w/o SC	w/o ml	TCGN
10	FC	RMSE	4.52	4.94	4.39	**4.36**
		MAE	2.69	2.88	2.59	**2.57**
	BOLD	RMSE	6.41	6.49	6.29	**6.25**
		MAE	4.72	4.88	4.64	**4.62**
40	FC	RMSE	11.19	12.00	11.07	**10.88**
		MAE	7.25	7.81	7.15	**7.05**
	BOLD	RMSE	6.88	6.98	6.73	**6.66**
		MAE	5.04	5.14	4.94	**4.91**

The commonly Automated Anatomical Labeling (AAL) template [19, 20] was utilized to define brain regions in both the FC and SC networks. Accordingly, both networks have 90 regions of interest (ROIs).

3.2. Experimental Settings

The RMSE and MAE metrics are utilized to evaluate the performance. For each subject, 80% of the data is used for training and remaining 20% for testing. It is worth noting that for each individual we use the training set and the testing set to construct sliding windows separately, which avoids the possibility that the test time points are involved in any training samples. After splitting, we compute the dynamic FC as sequence samples and generate them by sliding a window of width $T + T'$ as defined in 2.1. Especially, each sample sequence consists of 80 time steps, where the first $T = 40$ time steps are treated as the input while the rest $T' = 40$ are considered as the ground truth. We train the GCN and LSTM layer with the hidden dimension 32 and set batch size to 4. Only one GCN layer is used in our model. λ and γ are set to 0.5 and 1.0, respectively. Adam optimizer is employed and the max training iteration is set to 100. The initial learning rate is set to 0.01. All the models are implemented using Python 3.6 and tensorflow 1.12.0 on NVIDIA RTX 2080 Ti GPU.

3.3. Experiment Results

Contrast experiment Table 1 shows the quantitative evaluations of different methods for 10 as short time steps and 40 as long time steps prediction of FC and BOLD on HCP dataset. A lower values of RMSE and MAE represent a better performance. The proposed TCGN model obtains the best overall performance compared to other methods. The deep learning methods (LSTM & T-GCN) obtain better performance than the traditional approaches (ARIMA & SVR).

We further visualize one representative sample at 40 time steps of ground truth FC and the predicted FCs of different methods, shown as Figure 2. The results are consistent with quantitative assessment. The proposed TCGN obtains the best performance and the deep learning methods is better than the traditional approaches.

Ablation experiment Ablation studies are conducted to validate the effects of major components of TCGN model by removing a certain component. w/o FC represents removing the functional connectivity, w/o SC denotes removing the structural connectivity and w/o ml represents without the multi-task loss. From the results shown in Table 2, we can observe a drop in the performance of TCGN variants, which demonstrates the effectiveness of those components.

4. CONCLUSION

In this paper, we propose a temporal cross-graph network framework for brain functional activity prediction. Extensive experiments demonstrate the effectiveness of the proposed framework. In particular, for other spatio-temporal prediction problems of multi-modal graph structure, our model can also be used as a general framework for research.

5. REFERENCES

[1] William R Shirer, Srikanth Ryali, Elena Rykhlevskaia, Vinod Menon, and Michael D Greicius, "Decoding subject-driven cognitive states with whole-brain connectivity patterns," *Cerebral cortex*, vol. 22, no. 1, pp. 158–165, 2012.

[2] Andrew T Reid, Drew B Headley, Ravi D Mill, Ruben Sanchez-Romero, Lucina Q Uddin, et al., "Advancing functional connectivity research from association to causation," *Nature neuroscience*, vol. 22, no. 11, pp. 1751–1760, 2019.

[3] Andrew T Drysdale, Logan Grosenick, Jonathan Downar, Katharine Dunlop, Farrokh Mansouri, Yue Meng, et al., "Resting-state connectivity biomarkers define neurophysiological subtypes of depression," *Nature medicine*, vol. 23, no. 1, pp. 28–38, 2017.

[4] Robert Kim and Terrence J Sejnowski, "Strong inhibitory signaling underlies stable temporal dynamics and working memory in spiking neural networks," *Nature Neuroscience*, vol. 24, no. 1, pp. 129–139, 2021.

[5] Jae-Joong Lee, Hong Ji Kim, Marta Čeko, Bo-yong Park, Soo Ahn Lee, et al., "A neuroimaging biomarker for sustained experimental and clinical pain," *Nature medicine*, vol. 27, no. 1, pp. 174–182, 2021.

[6] Carlton Chu, Yizhao Ni, Geoffrey Tan, Craig J Saunders, and John Ashburner, "Kernel regression for fmri pattern prediction," *NeuroImage*, vol. 56, no. 2, pp. 662–673, 2011.

[7] R Cameron Craddock, Michael P Milham, and Stephen M LaConte, "Predicting intrinsic brain activity," *Neuroimage*, vol. 82, pp. 127–136, 2013.

[8] Xuegang Song, Feng Zhou, Alejandro F Frangi, Jiuwen Cao, Xiaohua Xiao, Yi Lei, et al., "Graph convolution network with similarity awareness and adaptive calibration for disease-induced deterioration prediction," *Medical Image Analysis*, vol. 69, pp. 101947, 2021.

[9] Tabinda Sarwar, Ye Tian, BT Thomas Yeo, Kotagiri Ramamohanarao, and Andrew Zalesky, "Structure-function coupling in the human connectome: A machine learning approach," *NeuroImage*, vol. 226, pp. 117609, 2021.

[10] Yang Li, Rasoul Shafipour, Gonzalo Mateos, and Zhengwu Zhang, "Mapping brain structural connectivities to functional networks via graph encoder-decoder with interpretable latent embeddings," in *2019 IEEE GlobalSIP*. IEEE, 2019, pp. 1–5.

[11] Yuting Sun, Shifei Ding, Zichen Zhang, and Weikuan Jia, "An improved grid search algorithm to optimize svr for prediction," *Soft Computing*, vol. 25, no. 7, pp. 5633–5644, 2021.

[12] Mingzhu Liao, "Rmb exchange rate prediction based on arima and bp neural network fusion model," *World Scientific Research Journal*, vol. 6, no. 4, pp. 252–262, 2020.

[13] Weicong Kong, Zhao Yang Dong, Youwei Jia, David J Hill, Yan Xu, and Yuan Zhang, "Short-term residential load forecasting based on lstm recurrent neural network," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 841–851, 2017.

[14] Thomas N Kipf and Max Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907*, 2016.

[15] Ling Zhao, Yujiao Song, Chao Zhang, Yu Liu, Pu Wang, Tao Lin, Min Deng, and Haifeng Li, "T-gcn: A temporal graph convolutional network for traffic prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 9, pp. 3848–3858, 2019.

[16] Aina Frau-Pascual, Jean Augustinack, Divya Varadarajan, Anastasia Yendiki, et al., "Detecting structural brain connectivity differences in dementia through a conductance model," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2019, pp. 591–595.

[17] Youyong Kong, Shuwen Gao, Yingying Yue, Zhenhua Hou, Huazhong Shu, Chunming Xie, et al., "Spatio-temporal graph convolutional network for diagnosis and treatment response prediction of major depressive disorder from functional connectivity," *Human brain mapping*, vol. 42, no. 12, pp. 3922–3933, 2021.

[18] David C Van Essen, Stephen M Smith, Deanna M Barch, Timothy EJ Behrens, Wu-Minn HCP Consortium, et al., "The wu-minn human connectome project: an overview," *Neuroimage*, vol. 80, pp. 62–79, 2013.

[19] Zhinan Yu, Jiaolong Qin, Xinyuan Xiong, Fengguo Xu, Jun Wang, et al., "Abnormal topology of brain functional networks in unipolar depression and bipolar disorder using optimal graph thresholding," *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 96, pp. 109758, 2020.

[20] Siyuan Liu, Jakob Seidlitz, Jonathan D Blumenthal, Liv S Clasen, and Armin Raznahan, "Integrative structural, functional, and transcriptomic analyses of sex-biased brain organization in humans," *Proceedings of the National Academy of Sciences*, vol. 117, no. 31, pp. 18788–18798, 2020.