SPATIO-TEMPORAL ATTENTION GRAPH CONVOLUTION NETWORK FOR FUNCTIONAL CONNECTOME CLASSIFICATION

Wenhan Wang[1], *Youyong Kong*[1*], *Zhenghua Hou*[2], *Chunfeng Yang*[1], *Yonggui Yuan*[2]

[1]Jiangsu Provincial Joint International Research Laboratory of Medical Information Processing, School of Computer Science and Engineering, Southeast University, Nanjing, China
[2]Department of Psychosomatic and Psychiatry, Zhongda Hospital, School of Medicine, Southeast University, Nanjing, China

ABSTRACT

Numerous evidence has demonstrated the pathophysiology of a number of mental disorders is intimately associated with abnormal changes of dysfunctional integration of brain network. Functional connectome (FC) exhibits a strong discriminative power for mental disorder identification. However, existing methods are insufficient for modeling both spatial correlation and temporal dynamics of FC. In this study, we propose a novel Spatio-Temporal Attention Graph Convolution Network (STAGCN) for FC classification. In spatial domain, we develop attention enhanced graph convolutional network to take advantage of brain regions' topological features. Moreover, a novel multi-head self-attention approach is proposed to capture the temporal relationships among different dynamic FC. Extensive experiments on two tasks of mental disorder diagnosis demonstrate the superior performance of the proposed STAGCN.

Index Terms— Mental disorder, functional connectome, graph convolutional network, attention, spatio-temporal

1. INTRODUCTION

In recent years, numerous evidence has demonstrated the pathophysiology of a number of mental disorders is intimately associated with abnormal changes of dysfunctional integration of brain network [1]. Functional magnetic resonance imaging (fMRI) has become an emerging technique to investigate the abnormal changes between disorders and controls [2]. Indeed, the functional connectome (FC) revealed by fMRI exhibits a strong discriminative power for disease identification and thus is increasingly utilized to explore the objective biomarker for mental disorder diagnosis [3]. Therefore, it is essential to develop an effective FC classification framework for accurate diagnosis.

Existing FC classification methods can be mainly divided into two streams, the static methods and the dynamic approaches. The static methods utilize one static FC matrix to represent each subject. The supervised classifiers, such as support vector machines [4] and random forest [5], have been used to extract features and classify from static FC. Moreover, FC can be represented as a graph structure naturally, including a set of brain regions of interest (ROIs) as nodes, and ROIs connections as edges [6]. Hence, graph neural networks can be extended to learn intrinsic features of the brain topology [7] from static FC. Significantly, these efforts lose sight of the time-varying feature of FC, which has been demonstrated to be an important biomarker for mental disorders [8].

Recent efforts have been made to develop dynamic approaches to capture the dynamic FC features. Several Recurrent Neural Network (RNN) based methods have been proposed to fully utilize the temporal information, such as Long Short Term Memory (LSTM) network [9][10]. However, these methods ignore the functional dependency between ROIs. Recently, a growing number of studies use dynamic FC to reveal the time-varying characteristics of brain activity [8][11]. Each subject can get a series of functional connectivity matrices, which reflect the changes of correlation level between brain regions on the time scale. Naturally, dynamic functional connectivity can be represented as a set of graphs. Based on this intuition, previous works [12][13][14] propose to leverage the advantages of Graph Convolutional Network (GCN) and LSTM to extract both spatial and temporal information. However, such a strategy cannot well solve the long-range dependency of temporal functional connectivity.

In this study, we propose a Spatio-Temporal Attention Graph Convolution Network (STAGCN) for FC classification. Firstly, we develop a novel attention-enhanced graph convolutional network to model the brain regions' topological correlations. Secondly, a temporal graph embedding attention mechanism is introduced for capturing the temporal relationships of dynamic functional connectivity. We conduct extensive experiments to evaluate the proposed framework on two

*Corresponding author. (Email: kongyouyong@seu.edu.cn)
The research is supported by grant BE2019748 Natural Science Foundation of Jiangsu Province and grant National Key Research and Development Program of China (No. 2021ZD0113202).

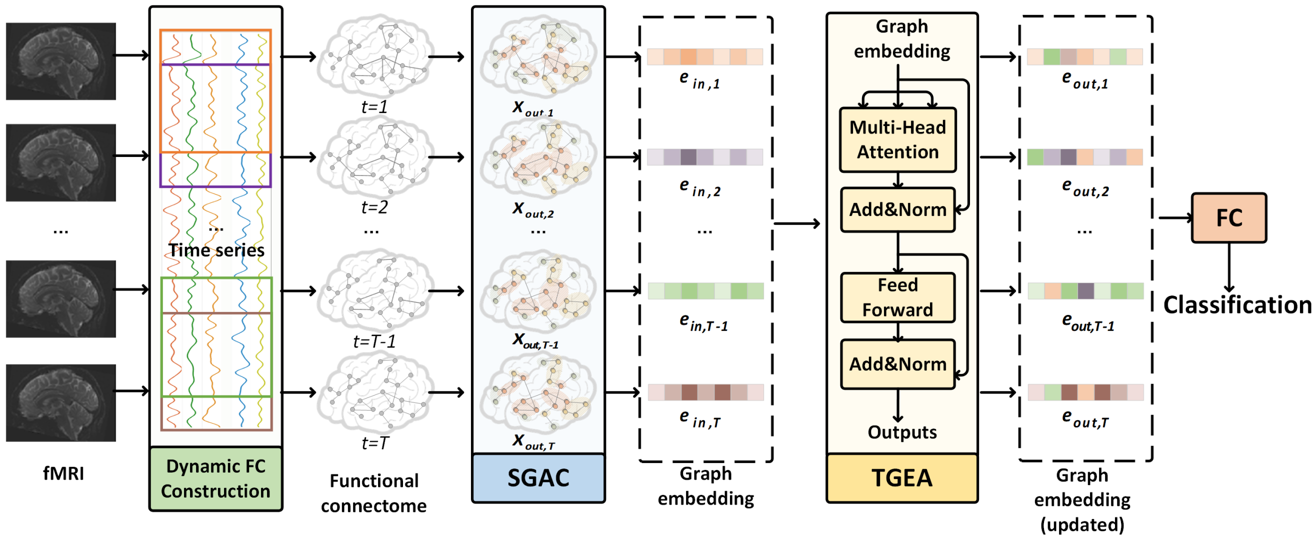

Fig. 1. Framework of Spatio-Temporal Attention Graph Convolution Network (STAGCN) for functional connectome(FC) classification.

tasks of mental disorder diagnosis, including major depressive disorder (MDD) and autistic spectrum disorder (ASD), which demonstrate the superiority of the proposed STAGCN.

2. METHODOLOGY

Fig. 1 illustrates the STAGCN framework. Firstly, the dynamic functional connectome is constructed from the resting-state fMRI. Secondly, the spatial graph attention convolution (SGAC) module is developed to capture the topological features of FC at each sliding window. Thirdly, the temporal graph embedding attention (TGEA) module is proposed to characterize the temporal features of dynamic FC among the different sliding windows. The spatio-temporal features are finally utilized for FC classification.

2.1. Dynamic Functional Connectome Construction

A set of dynamic FC is constructed from the resting-state fMRI data with the sliding window method [15]. Firstly, with a selected atlas and the preprocessed data, averaged time courses are computed for each brain region. Secondly, the entire time courses are divided into a number of overlapping sliding windows, and the Pearson correlation is utilized to calculate the functional connectivity matrix in each sliding window. Finally, the matrix is further thresholded by proportional quantization to obtain the brain network at each window.

Each subject can get T functional connectivity matrices, which can be represented as a series of graphs $\boldsymbol{G} = (\boldsymbol{V}, \boldsymbol{A}, \boldsymbol{X})$, where $\boldsymbol{V} \in \mathbb{R}^{N}, \boldsymbol{A} \in \mathbb{R}^{T\times N\times N}$ and $\boldsymbol{X} \in \mathbb{R}^{T\times N\times C}$. $\boldsymbol{V} = \{v_1, v_2, ..., v_N\}$ is the set of N nodes representing brain regions. $\boldsymbol{A} = \{A_t | A_t \in \mathbb{R}^{N\times N}, t = 1, 2, ..., T\}$ represents adjacency matrices to capture the relationships between brain regions. $\boldsymbol{X} = \{X_t | X_t \in \mathbb{R}^{N\times C}, t = 1, 2, ..., T\}$ represents C dimension features of N nodes which can be obtained by the normalized features.

2.2. Spatial Graph Attention Convolution

To learn spatial topological features from the graph at each time point, a module called spatial graph attention convolution (SGAC) is introduced. Firstly, we define a linear filter [16] to conduct a spatial-domain convolution for graphs by combining the adjacency matrix as follows,

$$H_t = h_0 I + h_1 A_t \tag{1}$$

where I is the identity matrix representing self-connections, and h_0, h_1 are scalar coefficients. A_t represents the adjacency matrix of the t-th graph in the graph sequence. To obtain a set of filter coefficients for multiple vertex features, filter H_t needs to be in $\mathbb{R}^{N\times N\times C}$ for C is the dimension of vertex features, and h_0, h_1 need to be in $\mathbb{R}^{C}$. For each feature channel, the linear filter is described as below,

$$H_t^{(c)} = h_0^{(c)} I + h_1^{(c)} A_t, \tag{2}$$

where $H_t^{(c)}$ represents an $N \times N$ slice of H_t, and $h_0^{(c)}, h_1^{(c)}$ are scalars corresponding to the given vertex feature c. The filter operation is performed as follows,

$$X_{out,t} = \sum_{c=1}^{C} H_t^{(c)} X_t^{(c)} + b \tag{3}$$

where $X_{out,t} \in \mathbb{R}^{N}$, $X_t^{(c)}$ represents the c th feature of the input, $b \in \mathbb{R}$ is the bias.

To use multiple filters, H_t can be expanded to $\mathbb{R}^{N\times N\times C\times F}$ by adding another dimension F, representing the number of filters. Thus, the output after the convolution should be $X_{out,t} \in \mathbb{R}^{N\times F}$. This operation transforms the original feature space to a new hidden space with the dimension determined by the number of filters.

Ordinary convolution operations make the weights of adjacent nodes the same. In practice, however, nodes of different brain regions are not all equally important. Identifying significant nodes and amplifying their influence are conducive to extracting effective information, which can simultaneously suppress the nodes with irrelevant and potentially confusing information. Therefore, we introduce the attention mechanism to further enhance the effective nodes in graph convolution. Different weights are assigned to the nodes and the important nodes are highlighted. With enhanced attention, the convolution filter of Equation (3) can be modified as follows,

$$X_{out,t} = \sum_{c=1}^{C}(1+att)H_t^{(c)}X_t^{(c)} + b \tag{4}$$

where $att = \{att_1, att_2, ..., att_N\}$ is a learnable parameter with an initial value of 0, representing the attention weight of each node. The brain regions with a larger value of att should be more significant for discriminating the mental disorders.

2.3. Temporal Graph Embedding Attention

Spatial graph attention convolution can extract the spatial feature of the brain topology in each graph, and obtain the high-level graph representation. Besides, it is also very important to extract the temporal feature among the FCs of different sliding windows. A novel temporal graph embedding attention module is proposed to capture the temporal relationships of dynamic functional connectivity. Features of each graph after convolution are flattened into a vector to obtain the graph embedding. Then we employ a self-attention mechanism to capture global dependencies among the graph embeddings. The input graph embedding sequence $E_{in} \in \mathbb{R}^{T\times M}$ represents the dynamic functional connectome, where T is the number of graphs, that is, the number of sliding windows, $M = N \times F$ is the feature embedding dimension. $E_{in} = \{e_{in,t} | e_{in,t} \in \mathbb{R}^{1\times M} | t = 1, 2, ..., T\}$ indicates the input feature vector of each graph.

Although the RNN models are slight and exquisite, they require a number of operations to obtain the dependency from two arbitrary input and output positions. Transformer uses self-attention, which can obtain the dependency relationship between two positions without relying on the sequence and distance [17]. Such a powerful self-attention strategy is thus transferred to capture the relationships of graph embeddings from different sliding windows.

Self-attention is calcualted among the graph embeddings of different sliding windows. For the input graph embedding sequence $E_{in} \in \mathbb{R}^{T\times M}$, using linear transformations W_Q, W_K and W_V, we create a query matrix Q, a key matrix K and a value matrix V respectively,

$$Q = E_{in}W_Q, K = E_{in}W_K, V = E_{in}W_V \tag{5}$$

where W_Q, W_K and $W_V \in \mathbb{R}^{M\times M}$ and $Q, K, V \in \mathbb{R}^{T\times M}$. The formula for calculating the output of self-attention can be expressed as,

$$Attention(Q, K, V) = Softmax(\frac{QK^{\top}}{\sqrt{d_k}})V \tag{6}$$

where d_k represents the dimension of feature embedding, here, $d_k = M$. Through this operation, the time-varying characteristics of dynamic functional connectome has been learned. In this operation, the scale factor of $\frac{1}{\sqrt{d_k}}$ is used for appropriate normalization to prevent the large value of d_k from causing a very small gradient.

Multi-head self-attention is further introduced to model the information jointly from various representation subspaces with different positions. The large space can be divided into multiple mutually exclusive small spaces, and each attention unit is calculated in different small spaces. Combining multiple small spaces can help to obtain richer features from different dimensions. The final output of multi-head attention is the concatenation of h attention head outputs.

$$e_i = Attention(Q_i, K_i, V_i), i \in [1, 2, ..., h] \tag{7}$$

$$\begin{aligned} E_{out} &= MultiHead(Q, K, V) \\ &= Concat(e_1, e_2, ..., e_h)W_o \end{aligned} \tag{8}$$

In addition to the attention layer, we also use two modules in transformer, a fully feed-forward network and a residual connection followed by layer normalization. The output keeps the same size as the input. Through the temporal graph embedding attention module, the temporal dependency of each embedding in the graph sequence is learned for better classification. Finally, the updated graph embeddings are sent to a fully connected layer for classification.

3. EXPERIMENTS

Datasets. We evaluate the proposed framework on two tasks of mental disorder diagnosis, including MDD and ASD. MDD diagnosis is evaluated on the dataset from the Second Affiliated Hospital of Xinxiang Medical University. ASD diagnosis is assessed on the public Autism Brain Imaging Data Exchange (ABIDE) [18]. The MDD dataset contains 98 MDD patients and 47 healthy controls. The ABIDE dataset is a multi-site ASD dataset and the data we use includes 295 ASD patients and 334 controls. The length of time series of Xinxiang and ABIDE is 230 and 176 respectively. The Xinxiang dataset is preprocessed with the same procedure in [13]. The preprocessed ASD data is obtained from the

Table 1. Performance of different methods for functional connectome classification on two diangosis tasks.

		MDD			ASD		
	Methods	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
Dynamic	STAGCN	**83.38 (± 1.23)**	**95.97 (± 1.06)**	54.43 (± 4.41)	**66.96 (± 0.96)**	**57.35 (± 4.22)**	**73.72 (± 3.21)**
	STGCN	78.44 (± 0.94)	86.86 (± 3.57)	50.82 (± 3.62)	60.70 (± 0.80)	47.61 (± 7.25)	68.38 (± 5.79)
	GCN-LSTM	74.36 (± 1.40)	84.08 (± 3.48)	41.10 (± 7.84)	60.77 (± 0.81)	48.30 (± 7.03)	68.40 (± 5.76)
	LSTM	68.86 (± 1.52)	81.64 (± 3.31)	40.99 (± 5.93)	60.48 (± 1.38)	55.98 (± 1.98)	64.66 (± 2.43)
Static	GCN	73.92 (± 1.81)	92.86 (± 3.34)	30.68 (± 8.95)	60.72 (± 1.10)	49.01 (± 6.69)	69.88 (± 5.65)
	GAT	70.61 (± 3.25)	86.49 (± 5.13)	36.26 (± 5.77)	61.19 (± 0.80)	50.21 (± 4.76)	69.45 (± 3.08)
	SVM	53.33 (± 0.79)	64.60 (± 1.49)	37.37 (± 1.72)	60.41 (± 0.24)	56.27 (± 0.86)	64.11 (± 0.73)
	RF	56.60 (± 0.47)	36.12 (± 1.53)	**56.00 (± 1.11)**	60.58 (± 0.58)	46.30 (± 1.97)	73.36 (± 0.88)

preprocessed connectomes project with standard processing procedure. The Anatomical Automatic Labeling atlas is utilized in the dynamic FC construction, which has 90 brain regions.

Implementation Details. The proposed method, as well as baselines, are conducted with the TensorFlow platform. All training and testing are conducted on two NVIDIA RTX 3090 GPUs. We use two layers of STAGCN to ensure the superiority of the model. We train our model using the Adam optimizer with the learning rate of 0.01 and the batch size is set to 16. The sliding window size is 100 and the stride is 2. Furthermore, we use 10-fold cross-validation to evaluate the performance and repeat 10 times to reduce the estimation uncertainty caused by stochastic gradient descent [12].

Performance Comparison. We use three common classification metrics to evaluate the classification performance: accuracy, sensitivity, and specificity. We compare our STAGCN with both static and dynamic functional connectivity based methods. SVM and Random Forest (RF) are basic machine learning methods while GCN and Graph Attention Networks (GAT) [19] are recent graph learning based approaches. For dynamic functional connectivity, LSTM is the most commonly used RNN approach. GCN-LSTM is a spatio-temporal graph neural network framework leveraging LSTM after GCN. And ST-GCN [13] embeds GCN into the operation of LSTM.

Table 2. Ablation study on MDD diagnosis.

Methods	Accuracy	Sensitivity	Specificity
STAGCN	**83.38 (± 1.23)**	**95.97 (± 1.06)**	54.43 (± 4.41)
w/o SA	82.61 (± 1.10)	94.03 (± 1.50)	**55.52 (± 6.14)**
w/o TA	66.10 (± 2.99)	85.76 (± 4.42)	25.83 (± 4.30)

Table 3. Ablation study on ASD diagnosis.

Methods	Accuracy	Sensitivity	Specificity
STAGCN	**66.96 (± 0.96)**	**57.35 (± 4.22)**	**73.72 (± 3.21)**
w/o SA	66.23 (± 0.42)	57.24 (± 4.81)	72.35 (± 4.15)
w/o TA	60.22 (± 1.54)	52.53 (± 4.78)	65.78 (± 2.94)

The performance results of different methods on two tasks are presented in Table 1. We can observe that the graph-based methods such as GCN and GAT generally perform better than traditional machine learning models such as SVM and RF, implying the significance of learning spatial correlations among brain regions. Moreover, our STAGCN method outperforms all other methods and indicates its superiority for learning spatial features and temporal dependency of dynamic functional connectome.

Ablation Analysis. To study the effects of different components, we conduct ablation studies by comparing STAGCN and its variants, in which different variants are obtained by removing spatial attention (SA) and temporal attention (TA) respectively. The comparisons of different variants on two datasets are shown in Table 2 and Table 3, demonstrating that both spatial attention and temporal attention have a positive contribution to performance.

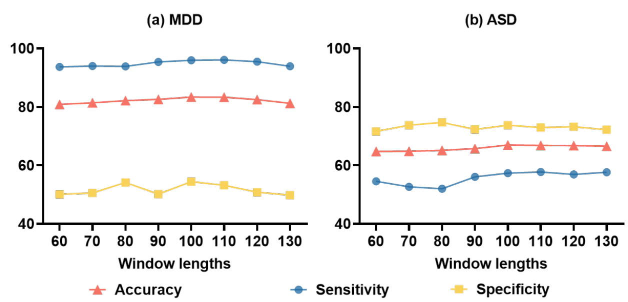

Fig. 2. Performance on two tasks with different lengths of sliding window.

Sliding Window Length. The influence of the important parameter of sliding window length on the performance is further evaluated through experiments. The results illustrated as Fig.2 indicate that the optimal window length is 100 for both the two diagnosis tasks.

4. CONCLUSION

In this work, we have proposed a novel spatio-temporal attention graph convolution network (STAGCN) for functional connectome classification. It consists of a spatial graph attention convolution module for capturing the brain's topological relationships and a temporal graph embedding attention architecture to learn the global temporal information. Extensive experiments on two diagnosis tasks demonstrate the superior performance of STAGCN.

5. REFERENCES

[1] Sartaj Ahmed Salman, Zhichao Lian, Marva Saleem, and Yuduo Zhang, "Functional connectivity based classification of adhd using different atlases," in *2020 IEEE International Conference on Progress in Informatics and Computing*, 2020, pp. 62–66.

[2] Tiago Azevedo, Luca Passamonti, Pietro Liò, and Nicola Toschi, "A deep spatiotemporal graph learning architecture for brain connectivity analysis," in *2020 IEEE Engineering in Medicine & Biology Society*, 2020, pp. 1120–1123.

[3] Li Zhang, Mingliang Wang, Mingxia Liu, and Daoqiang Zhang, "A survey on deep learning for neuroimaging-based brain disorder analysis," *Frontiers in neuroscience*, p. 779, 2020.

[4] Xin Yang, Mohammad Samiul Islam, and A M Arefin Khaled, "Functional connectivity magnetic resonance imaging classification of autism spectrum disorder using the multisite abide dataset," in *2019 IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*, 2019, pp. 1–4.

[5] Chella Kamarajan, Babak Ardekani, Ashwini Pandey, Sivan Kinreich, Gayathri Pandey, David Chorlian, Jl Meyers, Jian Zhang, Elaine Bermudez, Arthur Stimus, and Bernice Porjesz, "Random forest classification of alcohol use disorder using fmri functional connectivity, neuropsychological functioning, and impulsivity measures," *Brain Sciences*, vol. 10, pp. 1–22, 2020.

[6] Edward Bullmore and Olaf Sporns, "Complex brain networks: Graph theoretical analysis of structural and functional systems," *Nature reviews. Neuroscience*, vol. 10, pp. 186–98, 2009.

[7] Yu Zhang, Loïc Tetrel, Bertrand Thirion, and Pierre Bellec, "Functional annotation of human cognitive states using deep graph convolution," *NeuroImage*, vol. 231, pp. 117847, 2021.

[8] Maria Giulia Preti, Thomas AW Bolton, and Dimitri Van De Ville, "The dynamic functional connectome: State-of-the-art and perspectives," *NeuroImage*, vol. 160, pp. 41–54, 2017.

[9] Nicha C. Dvornek, Pamela Ventola, Kevin A. Pelphrey, and James S. Duncan, "Identifying autism from resting-state fmri using long short-term memory networks," in *Machine Learning in Medical Imaging*, 2017, pp. 362–370.

[10] Hongming Li and Yong Fan, "Brain decoding from functional mri using long short-term memory recurrent neural networks," in *MICCAI*, 2018, pp. 320–328.

[11] Ricardo Pio Monti, Peter Hellyer, David Sharp, Robert Leech, Christoforos Anagnostopoulos, and Giovanni Montana, "Estimating time-varying brain connectivity networks from functional mri time series," *NeuroImage*, vol. 103, pp. 427–443, 2014.

[12] Soham Gadgil, Qingyu Zhao, Adolf Pfefferbaum, Edith V. Sullivan, Ehsan Adeli, and Kilian M. Pohl, "Spatio-temporal graph convolution for resting-state fmri analysis," in *MICCAI*, 2020, pp. 528–538.

[13] Youyong Kong, Shuwen Gao, Yingying Yue, Zhenhua Hou, Huazhong Shu, Chunming Xie, Zhijun Zhang, and Yonggui Yuan, "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.

[14] Xiaodan Xing, Qingfeng Li, Mengya Yuan, Hao Wei, Zhong Xue, Tao Wang, Feng Shi, and Dinggang Shen, "Ds-gcns: connectome classification using dynamic spectral graph convolution networks with assistant task training," *Cerebral Cortex*, vol. 31, no. 2, pp. 1259–1269, 2021.

[15] Elena A. Allen, Eswar Damaraju, Sergey M. Plis, Erik B. Erhardt, Tom Eichele, and Vince D. Calhoun, "Tracking Whole-Brain Connectivity Dynamics in the Resting State," *Cerebral Cortex*, vol. 24, no. 3, pp. 663–676, 2012.

[16] Felipe Petroski Such, Shagan Sah, Miguel Alexander Dominguez, Suhas Pillai, Chao Zhang, Andrew Michael, Nathan D. Cahill, and Raymond Ptucha, "Robust spatial filtering with graph convolutional neural networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 6, pp. 884–896, 2017.

[17] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin, "Attention is all you need," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017, p. 6000–6010.

[18] Cameron Craddock, Yassine Benhajali, Chu Carlton, Chouinard Francois, Alan Evans, András Jakab, Budhachandra Khundrakpam, John Lewis, li Qingyang, Milham Michael, Yan Chaogan, and Pierre Bellec, "The neuro bureau preprocessing initiative: open sharing of preprocessed neuroimaging data and derivatives," *Frontiers in Neuroinformatics*, vol. 7, 2013.

[19] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio, "Graph attention networks," *6th International Conference on Learning Representations*, 2018.