



Prediction of Cognitive Scores by Joint Use of Movie-Watching fMRI Connectivity and Eye Tracking via Attention-CensNet

Jiaxing Gao¹, Lin Zhao², Tianyang Zhong¹, Changhe Li¹, Zhibin He¹, Yaonai Wei¹,
Shu Zhang³, Lei Guo¹, Tianming Liu², Junwei Han¹, and Tuo Zhang¹(✉)

¹ School of Automation, Northwestern Polytechnical University, Xi'an, China
tuozhang@nwpu.edu.cn

² Cortical Architecture Imaging and Discovery Lab, Department of Computer Science and
Bioimaging Research Center, The University of Georgia, Athens, GA, USA

³ School of Computer Science, Northwestern Polytechnical University, Xi'an, China

Abstract. Brain functional connectivity under the naturalistic paradigm has been demonstrated to be better at predicting individual behaviors than other brain states, such as rest and task. Nevertheless, the state-of-the-art methods are difficult to achieve desirable results from movie-watching paradigm fMRI(mfMRI) induced brain functional connectivity, especially when the datasets are small, because it is difficult to quantify how much useful dynamic information can be extracted from a single mfMRI modality to describe the state of the brain. Eye tracking, becoming popular due to its portability and less expense, can provide abundant behavioral features related to the output of human's cognition, and thus might supplement the mfMRI in observing subjects' subconscious behaviors. However, there are very few works on how to effectively integrate the multimodal information to strengthen the performance by unified framework. To this end, an effective fusion approach with mfMRI and eye tracking, based on Convolution with Edge-Node Switching in Graph Neural Networks (CensNet), is proposed in this article, with subjects taken as nodes, mfMRI derived functional connectivity as node feature, different eye tracking features used to compute similarity between subjects to construct heterogeneous graph edges. By taking multiple graphs as different channels, we introduce squeeze-and-excitation attention module to CensNet (A-CensNet) to integrate graph embeddings from multiple channels into one. The experiments demonstrate the proposed model outperforms the one using single modality, single channel and state-of-the-art methods. The results suggest that brain functional activities and eye behaviors might complement each other in interpreting trait-like phenotypes. Our code will make public later.

Keywords: Functional Connectivity · Naturalistic Stimulus · Eye Movement · CensNet · Attention

1 Introduction

There is a growing interest in leveraging brain imaging data to predict non-brain-imaging phenotypes in individual participants, since brain functional activity could intrinsically serve as an “objective” observer of a subject given that the emergence of behavior and cognition were widely attributed to the orchestration of local and remote cortical areas by means of a densely connected brain network [1]. In this context, the movie-watching paradigm has been widely demonstrated to provide richer life-like scenarios [2–5], better subject compliance in contrast to rest and task states, and were suggested to be better at predicting emotional/cognitive reaction [6–9] that could be more easily and saliently evoked by naturalistic input load, making movie watching the upper bound of current paradigms.

However, in recent works [8, 10, 11], the individual trait prediction from movie-watching paradigm fMRI (mfMRI) induced brain activity/functional connectivity can only achieve an accuracy around 0.40 (Pearson/Spearman’s r). The accuracy drops even dramatically when dataset size is small. Regarding the reality that it is a challenging task to increase mfMRI data size, at least two strategies can be intuitively proposed to improve the performance on the basis of limited number of subjects: 1) Incorporation of other data modalities, such as behavior. Conceptually, a joint use of what a subject “thinks” and what the subject “reacts” to a stimulus could help to increase the accuracy of prediction of “who” he/she is. Eye movement behaviors [12], as one example, have been related to subjects’ cognitive and phenotypical measures [12–14], and might supplement the fMRI derived brain activities in monitoring subjects’ attention and task compliance and observing subjects’ subconscious traits [15]; 2) Increase the number of different video clips watched by the same group of subjects.

Integration of multimodal data has been realized by using a graph to present the relation between subjects, where subjects are defined as nodes, node feature is fMRI connectivity, and edge is estimated by thresholding the similarity of behaviors, including eye movement [16]. This graph convolution networks could realize an embedding of a cohort of subjects’ brain activity features according to their behavior similarity, and estimate a mapping of these embedded features to cognitive scores, and further propagate the mapping to other nodes as a prediction of their scores. However, behavior in this model only provides the topology of the graph but are not fully involved in the process of embeddings. Also, different definitions of edges, such as eye trajectory and pupil size in this work, yield a set of graphs with different topologies on the same set of nodes. Therefore, we aim to solve the following two technique problems: 1) how to integrate the different edge features to node ones and learn graph embedding for both node and edge; and 2) how to integrate the embeddings of heterogeneous graphs with the same set of nodes to fulfill classification or regression. Based on Convolution with Edge-Node Switching graph neural network (CensNet) [17], we proposed Attention-CensNet (A-CensNet for short), where subjects are nodes, with the mfMRI derived functional connectivity as nodal features. Eye tracking derived gaze trajectory and temporal pupil size variation were respectively used to measure the similarity between subjects and to construct a set of heterogeneous edges. Each of these heterogeneous graphs was taken as an independent channel, where CensNet was used to alternatively learn both node embeddings and edge embeddings. Then, Squeeze-and-Excitation attention module (SENet)

[18] was used to integrate the node-edge embeddings from multiple channels into one hybrid graph on which the final round of node embedding was performed. Note that mfMRI and eye tracking data from the same cohort exposed to different movies inputs also yield additional channels in this work.

In the following sections, we firstly introduce the dataset and preprocessing steps. Basics of CensNet and SENet, and proposal of A-CensNet are introduced followed by its application to our task. Comparative and ablation studies regarding to prediction accuracy (AUC) were presented in the Results to demonstrate the effectiveness of multiple channel integration strategy, and the better performance of A-CensNet than others.

2 Materials and Methods

2.1 Dataset

In the Human Connectome Project (HCP) 7T release [19], movie-watching fMRI and resting-state fMRI data were acquired on a 7 T Siemens Magnetom scanner [20]. Among the four scan sessions (MOVIE 1~4), MOVIE 2&3 were used as a testbed in this work. Important imaging parameters are as follows: TR = 1000 ms, TE = 22.2 ms, flip angle = 45 deg, FOV = 208×208 mm, matrix = 130×130 , spatial resolution = 1.6mm^3 , number of slices = 85, multiband factor = 5. During MOVIE 2 runs, 4 video clips (818 time points, TRs), separated by five 20s rest sessions (100 time points in total), were presented to subjects. In MOVIE 3 runs, 5 video clips (789 time points, TRs), separated by five 20s rest sessions (120 time points in total), were presented to subjects. Eye tracking data were acquired during MOVIE runs using an EyeLink S1000 system with a sampling rate of 1000 Hz. HCP provides many phenotypic measures from a variety of domains. As suggested by [8], we focus on measures in the cognition domain in this work. After a quality control of data modalities, 81 subjects are selected.

2.2 Preprocessing

fMRI data have been preprocessed by the minimal preprocessing pipeline for the Human Connectome Project [20, 21]. The signals were mapped to the grayordinate system, which includes 64k vertices on the reconstructed cortical surface plus 30k subcortical voxels for an individual. The within-subject cross-modal registration and cross-subject registration are adopted to warp the grayordinate vertices and voxels to the same space, such that the associated fMRI signals have cross-subject correspondence.

In this study, subcortical regions are not our major interest and not included. Destrieux atlas [22] is applied to cortical surfaces to yield 75 cortical areas on each hemisphere. We compute the mean fMRI signal averaged over vertices within a cortical area, and construct a 150-by-150 functional connectivity matrix, by means of Pearson correlation between these average signals (blue panel in Fig. 1(a)). We zero the negative correlation and 90% of the lowest positive correlation. The upper triangular matrix is converted to a vector and used as the functional feature.

For eye tracking data, time stamps are used to extract the effective data points and synchronize the eye behavior features across subjects. Blink session is not considered.

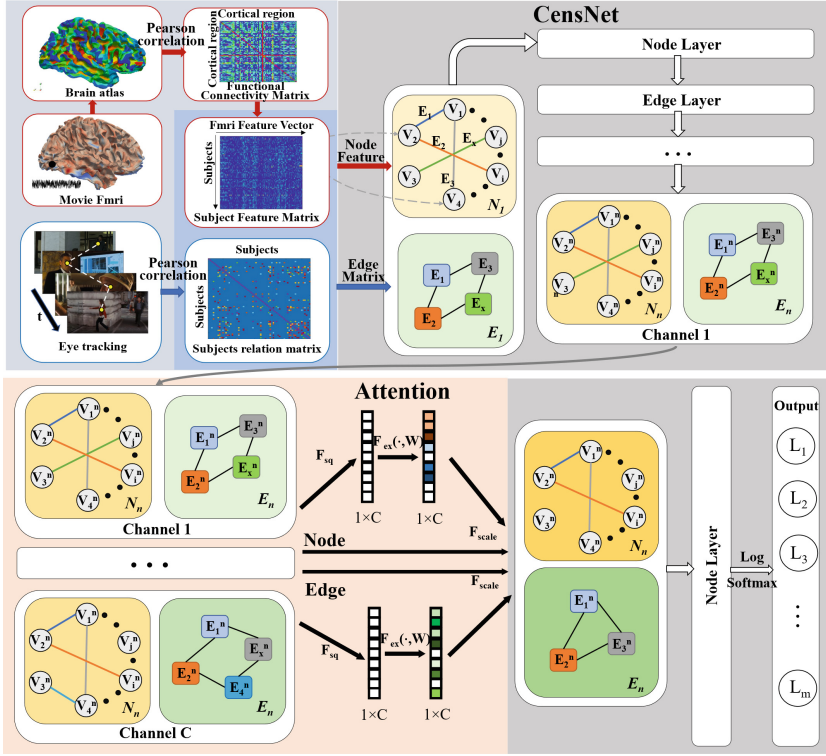


Fig. 1. The preprocessings and flowchart of A-CensNet. Data preprocessings are in the shallower blue panel. Construction of the graph is in the darker blue panel. Node feature generation is highlighted by red arrows and edge generation blue arrows. The graph (white box) is presented by its node-center version (yellow box) and edge-center version (green box) and was fed to CensNet in each channel. Gray frame highlights a typical CensNet flow. Attention module (SENet) is highlighted by carneose panel. Note that SENet was inserted in the middle of a CensNet procedure, before the last round of node convolution. (Color figure online)

Since many phenotypic measures within a domain could be correlated with one another, we perform principal components analysis (PCA) to measures classified in “cognition” domains [8]. The first principal component is used to classify the subjects to four groups by their scores. Subject number balance among groups is considered. Each group is assigned a label $l \in \mathcal{L}$.

2.3 Classification of Population via Attention-CensNet

Basics of CensNet. Supposing we have a dataset of M subjects, our objective is to assign each subject a cognitive group label l . We construct a graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, A\}$ to represent the entire cohort as shown in Fig. 1(a), where $v \in \mathcal{V}$ is a node of the graph, the subjects in this work. Edges \mathcal{E} s as well as the adjacent matrix A encode the similarity between subjects. The implementation of CensNet was detailed elsewhere [17], and we provide a summarized version as follows:

For spectral graph convolution, normalized graph Laplacian of a graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, A\}$ is computed: $\mathcal{L} = I_N - D^{-1/2}AD^{-1/2}$ where I_N is the identity matrix and D is the diagonal degree matrix. One of the important steps is the layer-wise propagation rule based on an approximated graph spectral kernel as follows:

$$H^{l+1} = \sigma(\tilde{D}^{-1/2}\tilde{A}\tilde{D}^{-1/2}H^lW^l) \quad (1)$$

where $\tilde{A} = A + I_N$ and \tilde{D} is the degree matrix, H^l and W^l are the hidden feature matrix and learnable weight of the l^{th} layer.

On this basis, the CensNet is proposed to have both node and edge convolution layers. For convenience, the graph (white box) in Fig. 1 (a) consists of a node-center version (yellow box) and an edge-center one (green box). In the node convolution layer, the embedding of nodes in the white box is updated, while the edge adjacency matrix and edge features in the green box are intact. Then, a similar update is implemented in the edge layer. Such a node-edge switching is implemented by the following equations:

$$H_v^{l+1} = \sigma(T\Phi(H_e^lP_e)T^T \odot \tilde{A}_vH_v^lW_v^l) \quad (2)$$

$$H_e^{l+1} = \sigma(T^T\Phi(H_v^lP_v)T \odot \tilde{A}_eH_e^lW_e^l) \quad (3)$$

where $\tilde{A}_v = \tilde{D}_v^{-\frac{1}{2}}(A_v + I_{N_v})\tilde{D}_v^{-\frac{1}{2}}$, $\tilde{A}_e = \tilde{D}_e^{-\frac{1}{2}}(A_e + I_{N_e})\tilde{D}_e^{-\frac{1}{2}}$, $T \in \mathbb{R}^{N_v \times N_e}$ is a binary matrix that indicates whether an edge connects a node. P_e is a learnable weight vector, Φ denotes the diagonalization operation. \odot denotes the element-wise product.

The loss function is defined as:

$$\mathcal{L}(\Theta) = -\sum_{l \in Y_L} \sum_{f=1}^F Y_{lf} \log M_{lf} \quad (4)$$

where Y_L is the subset of nodes with labels, M is the softmax results of the last node layer where node feature map has F dimensions.

Squeeze-and-Excitation Attention Block. The SENet [18] is introduced to integrate multiple graphs that share the same nodes but have different node features and edges. Note that a typical CensNet includes 1) the node-and-edge switching embedding plus 2) an additional round of node embedding (for node classification). The SENet is inserted between 1) and 2) (lower panel in Fig. 1(a)). The process (node as same as edge) is expressed by:

$$z_{node} = F_{sq}(\mathcal{F}_{node}) = \frac{1}{M \times H_n} \sum_{i=1}^M \sum_{j=1}^{H_n} \mathcal{F}_{node}(i, j) \quad (5)$$

$$s_{node} = F_{ex}(z_{node}, W) = \sigma(g(z_{node}, W)) = \sigma(W_2\delta(W_1z_{node})) \quad (6)$$

It is important to note that previous works [9] adopted regression scheme to predict cognitive scores, and the prediction accuracy was quantified by Pearson/Spearman correlation. However, the small dataset size significantly reduces the prediction accuracy [9], especially when sample size is below 100. We followed the suggestion in [16] to use

the classification strategy instead and the accuracy of class label prediction was quantified by the AUC. Accordingly, the regression layer of some of state-of-the-art methods [8, 9, 16, 23] are replaced by classification one. We set the dimension of the final output to 4 (the number of classes) and pass it through logsoftmax function. The loss function and the AUC calculation is the same as that used in A-CensNet, while keeping the other layers by their default configuration.

3 Results

3.1 Implementation Details

It was demonstrated in He et al., 2020 that the behavior score prediction accuracy via regression drops dramatically when subject number is below 100 (no more than $r = 0.1$ via Spearman correlation). Since we only have 81 subjects, a low regression accuracy cannot be a trustworthy to be used to compare with state-of-the-arts. As a compromise solution, we adopted classification scheme to demonstrate the effectiveness of our proposed framework.

In our application, subjects are divided to four cognitive groups (around 20 subjects in each one, a total of 81 subjects). Then, we randomly split the dataset to training, validation and testing sets, respectively. We randomly selected 10 subjects from each group, a total of 40 subjects (about 50% of the total) to form the training set. Among the remaining 41 people, we randomly selected 20 people to be the verification group and 21 people to be the testing group (each accounting for 25% of the total). Such a random division of training/validation/testing subsets was repeated 100 times, independently. The results (AUCs) were the average of 100 independent replicates.

We experiment on preserving {10%, 15%, 20%} top graph and edges by their weights, and find that preserving 10% node feature and 10% edges yields the best prediction performance. We try different settings of learning rate from {0.05, 0.01, 0.005, 0.001}, dropout {0.2, 0.3, 0.4, 0.5}, hidden {16, 32, 64, 128, 512, 1024}, and found that the best performance is yielded by learning rate to 0.005, dropout to 0.2, hidden to 1024. We implement the A-CensNet structure by adding the Attention mechanism based on CensNet, which empirically produce the best performance. AUC is adopted for each independent replication experiment to evaluate the prediction performance.

3.2 Ablation Study

The prediction accuracy measured by AUC of 100 repeated experiments are reported in Table 1. As a comparison, attention module is removed, as shown in the “Attention-No” rows in Table 1. In our method, the attention block is added after the node-edge updates in each channel (see Fig. 1(a)). The results are reported in “Attention-Middle” section. We also move the attention module ahead of CensNet. The results are reported in “Attention-Before” section. It is noted that data from different datasets are taken as different channels for attention module.

Table 1. Ablation Studies. The prediction accuracy is measured by AUC. Graph structure is represented by {Node, Edge}. “+” denotes a temporal concatenation of two features. C: channel. Trj: eye movement trajectory. Ppl: pupil size variation. The index after mfMRI, Trj and Ppl indicates which movie dataset it comes from. **Red**: the highest AUC. **Blue**: the second-highest one.

Models	Graph Structure			
Attention No (CensNet)	{mfMRI2, Trj2}	{mfMRI2, Ppl2}	{mfMRI3, Trj3}	{mfMRI3, Ppl3}
	50.36±0.71	52.91±0.73	51.49±0.75	51.83±0.85
	{mfMRI2+mfMRI3, Trj2+Trj3}			
	51.94±0.81			
	{mfMRI2+mfMRI3, Ppl2+Ppl3}			
	45.45±0.43			
Attention Middle	C1: {mfMRI2, Trj2} C2: {mfMRI2, Ppl2}		C1: {mfMRI3, Trj3} C2: {mfMRI3, Ppl3}	
	50.66±0.67		51.79±0.67	
	C1: {mfMRI2, Trj2} C2: {mfMRI3, Trj3}			
	49.35±0.69			
	C1: {mfMRI2, Ppl2} C2: {mfMRI3, Ppl3}			
	50.38±0.62			
	C1: {mfMRI2+mfMRI2, Trj2+Trj3} C2: {mfMRI2+mfMRI2, Ppl2+Ppl3}			
	46.21 ± 0.55			
Attention Before	C1: {mfMRI2, Trj2} C2: {mfMRI2, Ppl2}		C1: {mfMRI3, Trj3} C2: {mfMRI3, Ppl3}	
	49.00±0.52		48.49±0.78	
	C1: {mfMRI2, Trj2} C2: {mfMRI2, Ppl2} C3: {mfMRI3, Trj3} C4: {mfMRI3, Ppl3}			
	50.24±0.67			
	54.63±0.65			

To evaluate whether multiple stimulus loads to the same subjects enhance the performance, we only use two channels (trajectory and pupil variation as two sets of edges) within a single movie. The results are in the first row in “Attention-Top” and “Attention-Middle” sections. We use the format {Node, Edge} to describe the graph structure.

In the non-attention algorithms, CensNet on movie2 dataset ({mfMRI2, Pupil2}) yields the best performance. Concatenation of node feature or edge feature from two movie datasets even decreases the accuracy. When attention module is added (Attention-Middle), two-channel models (gray rows) do not significantly increase the accuracy. Within one movie dataset, when node feature is fixed and eye trajectory and pupil size variation are taken as two channels, the performance is not better than that on single-channel model on movie3 dataset. When two movie datasets are considered (carneose&green rows), they are integrated by means of channels when models with attention are used. In contrast, in non-attention models, the two movie datasets are integrated by means of feature concatenation. It is seen that the channel integration outperforms feature concatenation when pupil size is used as edges (green). But this does not hold when eye trajectory was used as edges (carneose). When both trajectory and pupil size are both considered as two channels, but features from two dataset are concatenated (white rows),

the prediction performance is worst (46.21 ± 0.55) in all attention models. Finally, when both trajectory and pupil size from two dataset are used as four channels (blue row) in our algorithm, the best performance was yielded.

These comparisons suggest that integration of edge features by channel attention does not always improves the performance than single use of an edge (gray rows), while integration of datasets by channel attention significantly outperforms the integration by feature concatenation. This observation still holds for “Attention-Before” models. But their accuracy is not as high as that via “Attention-Middle” models (see the cross-comparison between blue rows and gray rows), suggesting that a node-edge embedding could yield latent features more sensitive to individual variations, such that a channel attention works better at this deep feature space than being applied immediately after the original shallow features (“Attention-Before”). Finally, in addition to the 4-group classification, we evenly divided the subjects into 6 and 8 groups, respectively. AUCs via our model are 56.34 ± 0.65 and 56.95 ± 0.48 , demonstrating the robustness of the algorithm to class numbers.

3.3 Comparison with State-of-the-Arts

We compare our results with the ones by state-of-the-art methods listed in Table 2. The results via linear model and GCN were applied on Movie2 and have been reported in [16], and is thus listed for reference. Note that the linear, FNN and BrainNetCNN are not GNN-related methods. They have no edges but only rely on mfMRI features.

Table 2. Comparison with state-of-the-arts. Red font highlights the highest AUC and blue font highlights the second-highest one. “+” denotes a temporal concatenation of two features.

Models	Data		AUC
	Edge	Node	
Linear ^[8]	-	mfMRI2	41.94±0.81
GCN ^[16]	Trj2	mfMRI2	48.51±0.94
CensNet ^[17]		mfMRI2	49.75±0.65
CensNet ^[17]		Trj2	50.36±0.71
	Ppl2	52.91±0.73	
FNN ^[9]	-	mfMRI2	50.81±0.66
		mfMRI3	49.45±0.85
		mfMRI2+3	50.54±1.01
BrainNet CNN ^[23]	-	mfMRI2	53.42±0.63
		mfMRI3	49.81±0.78
		mfMRI2+3	50.66±0.82
A-CensNet	All	All	54.63±0.65

All nonlinear deep neural networks yield significant improvement in contrast to the linear method (41.94 ± 0.81). A concatenation of mfMRI features (mfMRI2 +3 rows) does not improve the prediction accuracy in contrast to that on a single dataset, suggesting the importance of the strategy selection for concatenating features from multiple datasets. Within a single movie dataset (unshaded rows), models integrating mfMRI and eye tracking outperform the models (FNN and BrainNetCNN) that used single modality

(mfMRI). After integrating mfMRI and eye behavior from multiple datasets, our results outperform all the-state-of-arts. These results demonstrate the effectiveness of integration of brain activity and eye behavior to one framework in cognition prediction. Also, given the limited subjects, multiple loads of stimuli integrated via attention modules could significantly improve the prediction performance.

4 Conclusion

We propose A-CensNet to predict subjects' cognitive scores, with subjects taken as nodes, mfMRI derived functional connectivity as node feature, different eye tracking features are used to compute similarity between subjects to construct heterogeneous graph edges. These graphs from different dataset are all taken as different channels. The proposed model integrates graph embeddings from multiple channels into one. This model outperforms the one using single modality, single channel and state-of-the-art methods. Our results suggest that the brain functional activity patterns and the behavior patterns might complement each other in interpreting trait-like phenotypes, and might provide new clues to studies of diseases with cognitive abnormality [24].

Acknowledgement. This work was supported by the National Natural Science Foundation of China (31971288, U1801265, 61936007, 62276050, 61976045, U20B2065, U1801265 and 61936007); the National Key R&D Program of China under Grant 2020AAA0105701; High-level researcher start-up projects (Grant No. 06100-22GH0202178); Innovation Foundation for Doctor Dissertation of Northwestern Polytechnical University CX2022052.

References

1. Thiebaut de Schotten, M., Forkel, S.J.: The emergent properties of the connected brain. *Science*, **378**(6619), 505–510 (2022)
2. Li, J., et al.: Global signal regression strengthens association between resting-state functional connectivity and behavior. *Neuroimage* **196**, 126–141 (2019)
3. Huijbers, W., Van Dijk, K.R.A., Boenniger, M.M., Stirnberg, R., Breteler, M.M.: Less head motion during MRI under task than resting-state conditions. *Neuroimage* **147**, 111–120 (2017)
4. Barch, D.M., et al.: Function in the human connectome: task-fMRI and individual differences in behavior. *Neuroimage* **80**, 169–189 (2013)
5. Sonkusare, S., Breakspear, M., Guo, C.: Naturalistic stimuli in neuroscience: critically acclaimed. *Trends Cogn. Sci.* **23**(8), 699–714 (2019)
6. Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R.: Intersubject synchronization of cortical activity during natural vision. *Science* **303**(5664), 1634–1640 (2004)
7. Finn, E.S., Scheinost, D., Finn, D.M., Shen, X., Papademetris, X., Constable, R.T.: Can brain state be manipulated to emphasize individual differences in functional connectivity? *Neuroimage* **160**, 140–151 (2017)
8. Finn, E.S., Bandettini, P.A.: Movie-watching outperforms rest for functional connectivity-based prediction of behavior. *Neuroimage* **235**, 117963 (2021)
9. He, T., et al.: Deep neural networks and kernel regression achieve comparable accuracies for functional connectivity prediction of behavior and demographics. *Neuroimage* **206**, 116276 (2020)

10. Gal, S., Coldham, Y., Bernstein-Eliav, M.: Act natural: functional connectivity from naturalistic stimuli fMRI outperforms resting-state in predicting brain activity. *bioRxiv* (2021)
11. He, T., et al.: Meta-matching as a simple framework to translate phenotypic predictive models from big to small data. *Nature Neurosci.* **25**, 1–10 (2022)
12. Lim, J.Z., Mountstephens, J., Teo, J.: Emotion recognition using eye-tracking: taxonomy, review and current challenges. *Sensors* **20**(8), 2384 (2020)
13. Hess, E.H., Polt, J.M.: Pupil size as related to interest value of visual stimuli. *Science* **132**(3423), 349–350 (1960)
14. Lohse, G.L., Johnson, E.J.: A comparison of two process tracing methods for choice tasks. *Organ. Behav. Hum. Decis. Process.* **68**(1), 28–43 (1996)
15. Son, J., et al.: Evaluating fMRI-based estimation of eye gaze during naturalistic viewing. *Cereb. Cortex* **30**(3), 1171–1184 (2020)
16. Gao, J., et al.: Prediction of cognitive scores by movie-watching FMRI connectivity and eye movement via spectral graph convolutions. In: 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI), pp. 1–5. IEEE (2022)
17. Jiang, X., Ji, P., Li, S.: CensNet: convolution with edge-node switching in graph neural networks. In: *IJCAI*, pp. 2656–2662 (2019)
18. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7132–7141 (2018)
19. Elam, J.: <https://www.humanconnectome.org/study/hcp-young-adult/article/first-release-of-7t-mr-image-data>. Accessed 20 June 2016
20. Griffanti, L., et al.: ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *Neuroimage* **95**, 232–247 (2014)
21. Glasser, M.F., et al., Wu-Minn HCP Consortium: The minimal preprocessing pipelines for the human connectome project. *Neuroimage*, **80**, 105–124 (2013)
22. Destrieux, C., Fischl, B., Dale, A., Halgren, E.: Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage* **53**(1), 1–15 (2010)
23. Kawahara, J., et al.: BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment. *Neuroimage* **146**, 1038–1049 (2017)
24. Tye, C., et al.: Neurophysiological responses to faces and gaze direction differentiate children with ASD, ADHD and ASD + ADHD. *Dev. Cogn. Neurosci.* **5**, 71–85 (2013)