

# Functional Decoding using Convolutional Networks on Brain Graphs

Yu Zhang and Pierre Bellec

Department of Psychology, Université de Montréal, 90 avenue Vincent d'Indy, Montreal, QC H2V 2S9  
Centre de recherche de l'Institut universitaire de gériatrie de Montréal, 4545 Queen Mary Rd, Montreal, QC H3W 1W6

## Abstract

A key goal in neuroscience is to understand brain mechanisms of cognitive functions. An emerging approach is the study of brain states dynamics using functional magnetic resonance imaging (fMRI). In this project, we applied graph convolutional networks (GCN) to decode brain activity over short time windows in a task fMRI dataset, i.e. associate a given window of fMRI time series with the task used. We investigated the performance of this GCN "cognitive state annotation" in the Human Connectome Project (HCP) database, which features 21 different experimental conditions spanning seven major cognitive domains, and high temporal resolution in task fMRI data. Using a 10-second window, the 21 cognitive states were identified with an excellent average test accuracy of 92% (chance level 4.8%). Performance remained good (60%) even at a temporal resolution of one volume (720 ms of duration). As the HCP task battery was designed to selectively activate a wide range of specialized functional networks, we anticipate the GCN annotation to be applicable over a broad range of paradigms, including resting-state.

**Keywords:** fMRI; brain decoding; graph convolutional network; deep learning

## Introduction

Modern imaging techniques, such as functional magnetic resonance imaging (fMRI), provide an opportunity to accurately map the neural substrates of human cognition. An emerging topic in the literature is the identification of brain states, characterized by a canonical spatio-temporal pattern of functional activity, and associated with specific cognitive states. A popular approach to identify these brain states, called multi-voxel pattern analysis (MVPA), uses machine learning tools to decode which task a subject performed based on recordings of brain activity in task fMRI (Norman, Polyn, et al., 2006). But the algorithm is usually limited to specific cognitive domains and relies on long acquisition of brain activity with repeated blocks to accurately decode a brain state. This paper aims at generalizing the brain decoding a wider cognitive battery and finer temporal resolution. For which, we proposed a new brain annotation pipeline based on graph convolution networks (GCNs). Three types of brain graphs were investigated. The models were validated using data from Human Connectome Project (HCP) (Van Essen, Smith, et al., 2013), which includes a large collection of fMRI data acquired from 1200 subjects, during 21 different cognitive tasks, in seven cognitive domains. Moreover, the high spatio-temporal resolution of fMRI signals (Van Essen, Ugurbil, et al., 2012), and consequently opened new avenues to characterize the dynamics of

human cognitive functions using Deep neural networks.

## Materials and Methods

To annotate the dynamics of cognitive states, we proposed a new brain decoding architecture based on graph convolutional network (Figure 1), which takes short series of functional data as input, applies information propagation among inter-connected brain regions and networks, and predicts the corresponding cognitive states based on the high order graph-level representations.

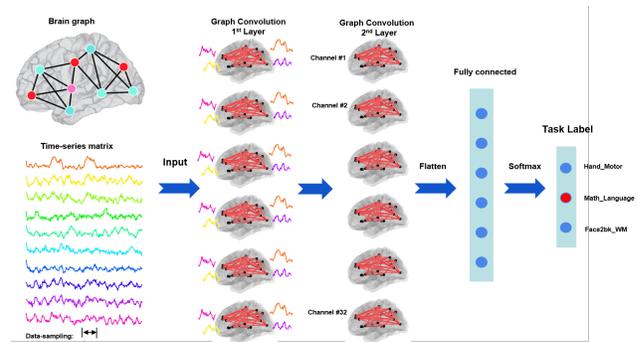


Figure 1: Pipeline of functional brain decoding using graph convolutions

## Populational brain graph

Here we used the multimodal cortical parcellation of the human brain (Glasser, Coalson, Robinson, et al., 2016), which delineates 180 functional areas per hemisphere. These brain parcels were defined as the nodes in the brain graph, while the connections between nodes were defined in different manners including: 1) spatial graph: by counting the shared vertices between two parcels on the white surface; 2) structural graph: by correlating the cortical thickness across 1096 subjects with surface curvature regressed out (Glasser et al., 2016); 3) functional graph: by calculating the group averaged functional connectivity based on 1080 minimal preprocessed resting-state fMRI data with the signals from white matter and csf regressed out and temporally bandpass filtered between 0.01 to 0.1 HZ (Glasser, Sotiropoulos, Wilson, et al., 2013). The correlation values of the structural and functional graphs were first normalized using Fisher z-transform and then weighted using a Gaussian kernel in order to scale from 0 to 1. To control for the effect of different sparsity levels in the spatial, structural and functional connectome (Figure 2), a KNN-graph was built for the brain graphs with each brain region only connected to its 8 strongest connected neighbours.

After mapping the minimal preprocessed task-fMRI data onto a set of brain regions, a sampling scheme was used on the temporal components in order to extract the task-corresponding fMRI time-series of specific length (ranging from 0.7s to 10s). We generalized the decoding models to predict all 21 task states from the seven cognitive domains, namely: emotion, gambling, language, motor, relational, social, and working memory. Within each cognitive domain, there are between-subject variations in the number of conditions and duration of task trials. Additional information on fMRI data acquisition, preprocessing, and task design can be found in (Barch, Burgess, Harms, et al., 2013) and (Glasser et al., 2013).

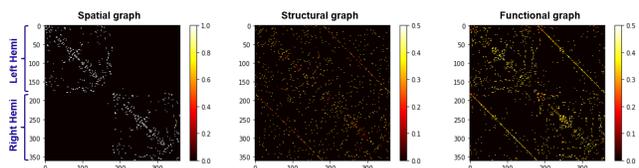


Figure 2: The sparse connectome matrix of spatial, structural and functional graphs

### Graph convolution layer

GCN originates from graph signal processing on a weighted graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$  that defines a network structure between brain regions. The set  $\mathcal{V}$  is a parcellation of cerebral cortex into  $N$  regions, and  $\mathcal{E}$  is a set of connections between each pair of brain regions, with its weights defined as  $W_{i,j}$ . Graph convolution relies on the graph Laplacian, which is a smooth operator characterizing the magnitude of signal changes between adjacent brain regions. The normalized graph Laplacian is defined as:

$$L = I - D^{-1/2} W D^{-1/2}, \quad (1)$$

where  $D$  is a diagonal matrix of node degrees and  $I$  is the identity matrix. For a signal  $x$  defined on graph, i.e. assigning a feature vector to each brain region, the convolution operation between the signal  $x$  and a filter  $g_\theta$ , based on graph  $\mathcal{G}$ , is defined as:

$$x * g_\theta = U G_\theta U^T x \quad (2)$$

where  $G_\theta = \text{diag}(U^T g_\theta)$  and  $U = (u_0, \dots, u_{N-1})$  is the matrix of Laplacian eigenvectors. With different choices of  $g_\theta$ , we can construct different graph convolutional networks. Here we use the ChebNet convolutions (Defferrard, Bresson, & Vandergheynst, 2016), which uses the Chebychev polynomial expansion of the Laplacian matrix in place of the spectral decomposition:

$$x * g_\theta = \sum_{k=0}^K \theta_k T_k(\tilde{L}) x \quad (3)$$

where  $\tilde{L}$  is a normalized version of the Laplacian, equals to  $2L/\lambda_{\max} - I$ , with  $\lambda_{\max}$  being the largest of the eigenvalues. In our models,  $\lambda_{\max} \approx 1.0$  for all three types of brain

graphs. This normalization is essential to preserve the magnitude of the graph signal  $x$  across multiple representations, especially when combining several graph convolutional layers. A simple recursive formula can be used to compute the Chebychev polynomials of order  $k$  from the previous orders,  $T_k(x) = 2T_{k-1}(x) - T_{k-2}(x)$  with  $T_0(x) = 1, T_1(x) = x$ , making the solution simple to implement. The ChebNet is computational efficiency, and scalable to large-scale graphs, as it avoids computing the full spectrum of the graph.

Up to this point, we define a parametric model for a graph convolution layer with one input channel. This can be easily generalized into multiple filters and channels. Specifically, we train a separate model for each channel (i.e. each fMRI volume per TR), and then summarize across multiple channels. This operation is repeated for each filter independently and followed by applying a nonlinear activation function, e.g.  $\text{ReLU}(\cdot) = \max(0, \cdot)$ . This multi-channel, and multi-filter structure enriches the final graph representations of fMRI time-series, and is further improved by stacking multiple layers. The final GCN architecture consists of six convolutional layers with 32 filters at each layer, followed by two fully-connected layers (256- 64 - num-of-states).

The implementation of our proposed GCN annotation model is based on Pytorch 1.1.0. The impact of  $k$  order in ChebNet is especially investigated as well as the choice of different brain graphs. The networks are trained for 100 epochs, and using the Adam optimizer which keeps separate learning rates for each weight, as well as an exponentially decaying average of previous gradients. The batch size varies from 20 to 130 depending on the chosen time windows. The entire dataset is split into training (70%), validation (10%), test (20%) sets using a subject-specific split scheme, which means that all time-series from the same subject was assigned to one set. The best model, with the highest prediction accuracy on the validation set is saved for further testing analysis.

### Representational similarity analysis

The stacked GCN layers provide a graph embedded representation of fMRI time-series corresponding to each experimental trial. In order to validate such representation includes state-specific features, we map the representations back onto the cerebral cortex, and generate a new activation pattern for each trial. Furthermore, a representational similarity analysis is conducted by calculating spatial correlations among the activation patterns. The resultant RSA matrix characterizes both within-state similarity, and the between-state dissimilarity in graph representations of fMRI time-series. We further summarize it into a state-RSA matrix by averaging the similarity values for each category of brain states. The RSA was only performed on the test set with the saved best model.

### Results

We applied the proposed GCN state annotation pipeline to infer which of the 21 task state associated with a given short time windows in a fMRI time-series. Using a 10-second window (approximately the shortest duration of all task trials), the

21 cognitive states were identified with an average test accuracy of 92%. A temporal resolution of one volume (720 ms of duration), the prediction accuracy remained good (60%). After taking into account of the delay effect of hemodynamic response function in BOLD signals, i.e. excluding fMRI volumes within 6s after task trial started from both training and test sets, the performance was improved and reached 80%. We further investigated the impact of following factors, including the order  $K$  of used Chebyshev polynomials and the choice of brain graphs.

### Impact of order $K$

The ChebNet is  $K$ -localized in space by taking up to  $K$ -th order polynomials, which means that at each layer, the information and attributes only propagate to the  $K$ -th-order neighborhood. Here we investigated the impact of order  $K$  by spanning over the list of [1,2,5,8,10]. As shown in Figure 3 A, a gap in performance of spatial-ChebNet appeared when increasing the order from 2 to 5, and reached a stable range after order 5. This is probably due to the fact that spatial brain graph is mostly consist of short-distance connections (Figure 2). Thus, the global efficiency of the brain graph is constrained by a small neighborhood, where the information is not transmitted across networks and modules. But when we enlarge this neighborhood sufficiently, this constraint disappears by rebuilding the between-network communications. In the following analysis, we fixed the order  $K = 5$  for ChebNet.

### Impact of brain graphs

In contrast to favoring short-distance connections in the spatial graph, the structural and functional graphs also include more long-distance connections (Figure 2). Thus, by increasing the order  $K$  of ChebNet, we did not see much improvement in performance especially for the structural graph. As for the functional graph, the impact only happened in the early phase of training (before 40 epochs), when we increased from 1st- to 2nd-order. The prediction accuracy plateaued after a sufficient number of epoch trials. On the other hand, when we fixed the order  $K = 5$ , similar performance was achieved from the spatial and functional graphs, with both being higher than the structural graph. This is probably due to high ratio of long-distance connections in the structural graph that the information transmits faster between-communities than within networks, which consequently confused the state predictor. Furthermore, the model training time also showed a big difference among the brain graphs, with the shortest training for spatial (350s per epoch), and the longest training for functional (480s per epoch). To summarize, with 5-th order neighborhood, the spatial-GCN achieved highest prediction accuracy with the shortest training time, and the least computational resources. In the following analysis, we focused on the spatial graph.

### Representational similarity analysis

The similarity analysis of high-order graph representations indicated a nice disassociation between different type of brain

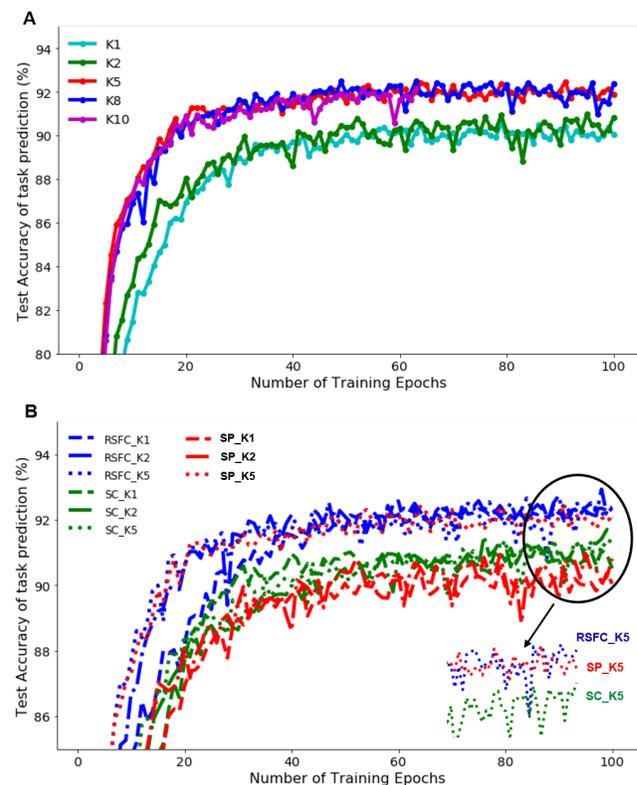


Figure 3: The effect of the order  $K$  for ChebNet using spatial graph and comparing between choices of different brain graphs

states. For instance, as shown in Figure 4, the RSA matrix of the Motor task showed high within-state similarities (0.30, 0.28, 0.27, 0.27 and 0.35 respectively), and much lower between-state similarities (average is 0.031). Moreover, we found higher similarity between left and right foot movements ( $r=0.12$ ) compared to foot versus tongue movements ( $r=0.008$ ), while the similarity between left and right hand movements was moderate ( $r=0.045$ ). In order to visualize these graph representations, the t-distributed Stochastic Neighbor Embedding (t-SNE) (Maaten & Hinton, 2008) was used to project the representations of all experimental trials down to 2 dimensions. As shown in Figure 4 C, the tongue movement was highly separated from others, while the mapping between left and right foot movements were largely overlapping. These findings was in line with the RSA above.

## Conclusion and Discussion

We proposed a GCN architecture to annotate cognitive states of the human brain. This model annotates brain activity with fine temporal resolution, and fine cognitive granularity. Using a 10s window of fMRI signals, our model identified 21 different task conditions with a test accuracy of 92%. The performance of the annotation model relies on the global effi-

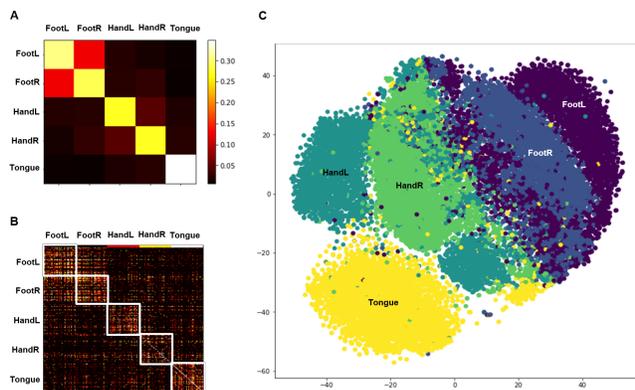


Figure 4: Representational similarity analysis of the graph representations of fMRI time-series acquired from the Motor task. The RSA matrix was shown for both brain states (A) and experimental trials (B). Visualizations of the graph representations using t-SNE (C)

ciency of information propagation on the graph. This efficiency could either come from the organization of brain graphs, or the architecture of GCN. At the level of graph organization, the global efficiency originates from the small world architecture (Bassett & Bullmore, 2006) characterized by a high local clustering coefficient, and a small shortest-path length. In this study, we compared three types of brain graphs, including the spatial graph which only connects each brain region to its direct neighbours, the functional graph which consists of dense local connections and a few long-range connections, as well as the structural graph which includes more long-range connections. Our results indicated that functional graph generally performed better than spatial graph when using a localized small neighborhood in graph convolutions (Figure 3). The advantages of using the functional graph are mainly contributed by the long-range connections including within- and between-network as well as inter-hemisphere connectives. Meanwhile, the GCN architecture plays a complementary role to the transmission efficiency. Here, we used ChebNet convolution in the core GCN layers, which propagates brain activities and other attributes among the  $K$ th-order neighborhood, making it  $K$ -localized in space (Defferrard et al., 2016). With a relative large value of the order  $K$ , the graph convolutions can improve the efficiency of the chosen brain graph by expanding the range of information propagation per step. This characteristic is especially helpful for the spatial graph. In line with this, we found a jump in the state prediction accuracy by increasing to the 5th-order neighborhood (Figure 3 A). Meanwhile, it did not impact the other brain graphs which already included a sufficient portion of long-range connections. To conclude, in order to maintain a high performance in brain state annotation, we can either use a high globally efficient graph like the functional graph which was constructed from resting-state functional connectives, or use a high order in ChebNet convo-

lutions.

The embedded graph representations of fMRI data were validated in two ways. Firstly, the RSA matrix at both state- and trial-level showed a disassociation across different task conditions while maintaining high similarity within states. A 10-fold difference was observed among within- and between-state similarities. This indicated that the GCN models accurately captured some state-specific features. The pretrained GCN models can predict brain states in other individual brain mapping applications or the extracted graph representations of fMRI activities that are associated with behavioral, and cognitive measures. Secondly, we used t-SNE to map the high-level graph representations onto a 2d space, which indicated isolated clusters for different task conditions. Both RSA matrix and t-SNE mapping showed a distinct representation between hand, foot and tongue, which is in line with the somatotopic maps in the primary motor cortex. It is also interesting to notice that the representations between left and right foot were more largely overlapped than left-vs-right hands. Similar patterns have been shown in (Barch et al., 2013) that larger overlapping of activation patterns of left-right foot movements in the medial surface, than of left-right hand movements in the lateral surface of the primary motor cortex.

## Acknowledgments

This work was supported in part by the Courtois NeuroMod Project and IVADO Postdoctoral Scholarships Program.

## References

- Barch, D. M., Burgess, G. C., Harms, et al. (2013). Function in the human connectome: task-fMRI and individual differences in behavior. *Neuroimage*, *80*, 169–189.
- Bassett, D. S., & Bullmore, E. (2006). Small-world brain networks. *The neuroscientist*, *12*(6), 512–523.
- Defferrard, M., Bresson, X., & Vandergheynst, P. (2016). Convolutional neural networks on graphs with fast localized spectral filtering. In *Advances in neural information processing systems* (pp. 3844–3852).
- Glasser, M. F., Coalson, T. S., Robinson, et al. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, *536*(7615), 171.
- Glasser, M. F., Sotiropoulos, S. N., Wilson, et al. (2013). The minimal preprocessing pipelines for the human connectome project. *Neuroimage*, *80*, 105–124.
- Maaten, L. v. d., & Hinton, G. (2008). Visualizing data using t-sne. *Journal of machine learning research*, *9*(Nov), 2579–2605.
- Norman, K. A., Polyn, et al. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in cognitive sciences*, *10*(9), 424–430.
- Van Essen, D. C., Smith, et al. (2013). The wu-minn human connectome project: an overview. *Neuroimage*, *80*, 62–79.
- Van Essen, D. C., Ugurbil, et al. (2012). The human connectome project: a data acquisition perspective. *Neuroimage*, *62*(4), 2222–2231.