

# Predicting Spatio-Temporal Human Brain Response Using fMRI

Chongyue Zhao<sup>1</sup>, Liang Zhan<sup>1</sup>, Paul M. Thompson<sup>2</sup>, and Heng Huang<sup>1\*</sup>

<sup>1</sup> Department of Electrical and Computer Engineering, University of Pittsburgh,  
Pittsburgh, PA, USA

<sup>2</sup> Imaging Genetics Center, University of Southern California, Los Angeles, CA, USA  
cyzxidian@gmail.com, liang.zhan@pitt.edu, pthomp@usc.edu,  
heng.huang@pitt.edu

**Abstract.** The transformation and transmission of brain stimuli reflect the dynamical brain activity in space and time. Compared with functional magnetic resonance imaging (fMRI), magneto- or electroencephalography (M/EEG) fast couples to the neural activity through generated magnetic fields. However, the MEG signal is inhomogeneous throughout the whole brain, which is affected by the signal-to-noise ratio, the sensors' location and distance. Current non-invasive neuroimaging modalities such as fMRI and M/EEG excel high resolution in space or time but not in both. To solve the main limitations of current technique for brain activity recording, we propose a novel recurrent memory optimization approach to predict the internal behavioral states in space and time. The proposed method uses Optimal Polynomial Projections to capture the long temporal history with robust online compression. The training process takes the pairs of fMRI and MEG data as inputs and predicts the recurrent brain states through the Siamese network. In the testing process, the framework only uses fMRI data to generate the corresponding neural response in space and time. The experimental results with Human connectome project (HCP) show that the predicted signal could reflect the neural activity with high spatial resolution as fMRI and high temporal resolution as MEG signal. The experimental results demonstrate for the first time that the proposed method is able to predict the brain response in both milliseconds and millimeters using only fMRI signal.

**Keywords:** Brain dynamics · Recurrent neural network · fMRI.

## 1 Introduction

In computational neuroscience, brain state often refers to wakefulness, sleep, and anesthesia. However, the precise and dynamical brain complexity is still missing. Dynamical neural representation is thought to arise from neural firing, which do not fire in isolation[11]. The development of modern brain measuring

---

\* corresponding author

techniques makes it possible to infer the large-scale structural and functional connectivity and characterize the anatomical and functional patterns in human cortex. Electrophysiological methods provide a direct and non-invasive way to record brain with milliseconds temporal resolution which is not affected by the problems commonly caused by intermediate processes. fMRI uses brain activity-related blood-oxygen-level-dependent (BOLD) to understand the neural representations in millimeter level, however it is too sluggish in the time domain. None of the current non-invasive brain recording techniques could measure the high resolution dynamics in space and time. This is a major challenge in neuroscience which draws attention of researchers to develop simultaneous simulation models of brain dynamics.

Since each brain region interacts with other regions, it is hard to measure the intraregional brain dynamics with single technique such as fMRI or MEG. With the increase of the strength of brain region interactions, the measured brain dynamical signals decrease. Recent study on complex biological system provides the powerful mathematical functions to model the coupling structural and functional brain activity using multimodal brain measurements[19]. Previous study use principle component analysis (PCA) [3, 12, 13] or linear models [9, 15] to quantify the inherent neural representations in either high dimensional or low dimensional. The high dimensionality is used for information encoding. The low dimensionality is used to encode information in complex cognitive or motor tasks[8]. Modern methods such as representational similarity analysis (RSA) [7] and multivoxel pattern analysis (MVPA) [10] is used to quantify the brain activity across task conditions. Further developments utilize the states defined as the transmission between different task activity to function higher cognitive dynamics [14, 16, 2, 3, 18].

The question remains: how to model the change and transmission between neurons with high spatial and temporal resolution? In this paper, we propose a novel method that use fMRI to predict the corresponding neural activity with high resolution in space and time. Our main contributions are summarized as follows:

- 1) The current non-invasive brain measuring technique could not identify the neural activity with high resolution in both spatial and temporal domain. MEG/EEG could resolve the brain activity in milliseconds but the spatial resolution is low. In contrast, fMRI could represent the neural dynamics in millimeters, but it is too slow for the rapid transmission between cortical sites. To link the mapping between brain regions and time points, we propose a general framework to project the time points onto polynomial basis. In the training step, the proposed method takes the ROI parcellated brain signals (fMRI and MEG) to learn the dynamical representation. In the testing step, the proposed method is the first to use original 4D fMRI signal to predict the internal brain response in space and time.

- 2) Most of the existing recurrent neural network (RNN) suffers from the vanishing gradient problem when dealing with long-term time series. It is challenging to represent the brain response at milliseconds with current RNN model.

The proposed method could discretize and project the time points onto the polynomial basis, which is able to deal with any time length data using the memory compression.

3) The two experimental results demonstrate that the proposed method is able to resolve the neural activity in space and time. Its feasibility is validated when compared with fMRI in structural domain using brain networks of independent component analysis (ICA). The predicted results also show the similar pattern to MEG signal in temporal domain.

The proposed method is related to the fMRI and MEG data fusion research. The difference between the proposed method and existing data fusion method is summarized as follows. 1) The proposed method does not use the MEG data in the testing step. It is able to use 4D fMRI data to predict the brain internal states in space and time. The MEG data is used in the training step as a reference for the proposed method to learn the temporal pattern of brain states. The research problem could be considered as the “super-resolution” in the temporal domain of fMRI data and keeps the structural details. However, as far as we know, most of the previous methods require both fMRI and MEG data to predict the internal states. It is hard to obtain simultaneous fMRI and MEG data at the same time. 2) Secondly, the proposed method could predict the high resolution temporal signals for each voxel of fMRI data. In the testing step, the proposed method does not require the preprocessing steps like ROI average or beamforming which could introduce noise in the preprocessed data. As far as we know, the proposed method is the first to predict the spatio-temporal brain internal states for each voxel using only fMRI data.

## 2 Spatio-Temporal Dynamical Modeling of Brain Response

Our approach benefits from the framework of Legendre Memory Unit (LMU) and online function approximation to learn the memory representation [17, 6]. The dynamical brain network with  $N$  neurons could be modeled as

$$\begin{aligned}\dot{z}(t) &= f(z, l, t) + e_1(t), \\ y(t) &= O(z(t)) + e_2(t),\end{aligned}\tag{1}$$

where  $z(t) = [z_1(t), z_2(t), \dots, z_N(t)]^T$  represents the internal states of  $N$  neuron nodes at time  $t$ .  $f(\cdot)$  denotes the nonlinear dynamical function of each node. And  $l(t) = [l_1(t), l_2(t), \dots, l_S(t)]^T$  represents the external stimuli for  $S$  neurons.  $y(t) = [y_1(t), y_2(t), \dots, y_N(t)]^T \in R^N$  is the observed time series of brain measurements (i.e. MEG or fMRI signals).  $e_1(t)$  and  $e_2(t)$  are time series noise of the internal states and measurements.  $O(\cdot)$  is the output function which controls the output of the brain measurements at time  $t$ . The nonlinear brain function  $f(\cdot)$  is modeled using recurrent neural network (RNN).

The whole framework of the proposed method is shown in Fig. 1. In the training stage, the proposed method consists of two parts, the RNN model with

Polynomial Projection and the Siamese network to score the agreement between the network prediction and the original time series of brain activity. Both Region-of-interest (ROI) extracted fMRI and MEG signals are used to train the network. However, in the testing stage, only fMRI signal is included to predict the internal states in space and time. The experiments demonstrate that the proposed method could deal with both the ROI extract fMRI signal and the original 4D fMRI image. The predicted internal behavioral states show high resolution in spatial and temporal domain.

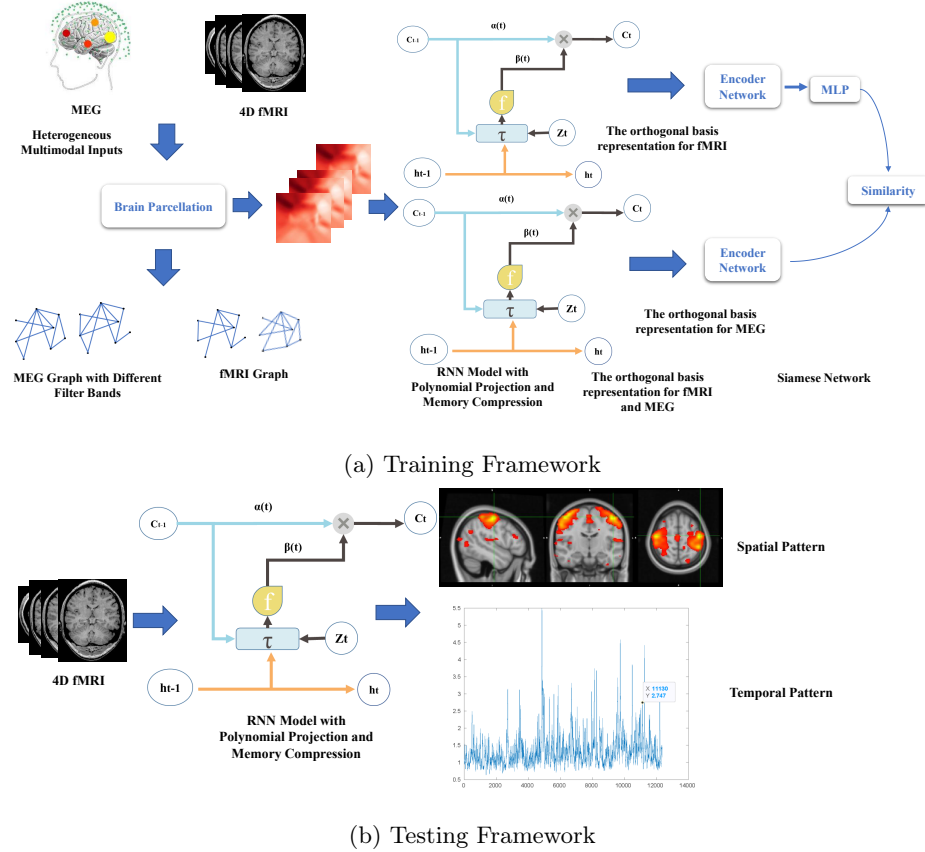


Fig. 1: Illustration of the proposed spatio-temporal brain dynamical network (a) Training framework. Both ROI extracted fMRI and MEG signals are used in the training process. (b) Testing framework for generating the internal states which uses 4D fMRI as input.

We propose a Polynomial Projection Operators together with Recurrent memory to solve the dynamical internal state  $z(t)$ . Given the input  $z(t)$  or

$f(z, l, t)$ , the proposed method aims to solve the future prediction  $z(t+1)$  based on the cumulative history  $f_{\leq t}$ . However, as temporal resolution for MEG signal is intractably high, there exists vanishing gradient problem when the model evolves over all time states. To solve the vanishing gradient problem, we project the input state onto the subspace and maintain the compressed historical representation. So there are two problems to be solved: how to quantify the approximation, and the way to learn the subspace.

**Function approximation.** We introduce the probability measure  $\mu \in [0, \infty)$  to define the space of square function  $\langle f, p \rangle_\mu = \int_0^\infty f(x)p(x)d\mu(x)$ .  $p(\cdot)$  is defined on a  $M$  subspace  $\mathcal{P}$ . The function  $p_t \in \mathcal{P}$  is utilized to minimize the approximation of  $f_{\leq t}$  with  $\|f_{\leq t} - p_t\|_{L_2(\mu_t)}$ , where  $\mu_t$  is the measure ranges in  $(-\infty, t)$ .

**Polynomial basis expansion for subspace learning.** To learn the suitable subspace, we define the polynomial basis  $\mathcal{P}$  with parameter  $M$  to represent the projected history, which means the historical dynamics could be represented using  $M$  coefficients with the basis  $\mathcal{P}$ .  $M$  represents the size of the compression. However, it is challenging to maintain the parameter  $\mu_t$  when  $t \rightarrow \infty$ . We show more details of the suitable polynomial basis in Supplementary material.

The first step is to choose the appropriate basis  $p_m \in \mathcal{P}$  in the projection operator. The projection operator takes the historical memory of  $t$  and minimize the approximation of  $f_{\leq t}$  using  $\|f_{\leq t} - p_t\|_{L_2(\mu_t)}$ . According to approximation theory, we use the orthogonal polynomials of  $\mu_t$  as the orthogonal basis and represent the coefficients  $c_m(t) := \langle f_{\leq t}, p_m \rangle_{\mu_t}$ .

The second key step is to differentiate the projection using the inner product  $\langle f, p \rangle_\mu = \int_0^\infty f(x)p(x)d\mu(x)$ , which will lead to the similar result that  $\frac{d}{dt}c_m(t)$  is expressed using  $(c_k(t))_{k \in M}$  and  $f(t)$ . Thus,  $c(t) \in R^M$  satisfies the ODE form  $\frac{d}{dt}c(t) = \alpha(t)c(t) + \beta(t)f(t)$ , where  $\alpha(t) \in R^{M \times M}$  and  $\beta(t) \in R^{M \times 1}$ . For each time  $t$ , We could change  $c(t)$  to the ODE form,

$$\frac{d}{dt}c(t) = -\frac{1}{t}\alpha(t)c(t) + \frac{1}{t}\beta(t)f(t), \quad (2)$$

The whole derivation of  $c(t)$  is shown in Supplementary material. Next we want to solve the ODE problem and obtain the coefficient  $c(t)$  by discretizing the continuous function  $c(t)$  and  $f(t)$ , which yields the recurrence  $c_{k+1} = -\alpha_k c_k + \beta_k f_k$  using  $(f_k)_{k \in M}$ . The orthogonal basis is defined as follows,

$$c_{k+1} = (1 - \frac{\alpha}{k})c_k + \frac{1}{k}\beta f_k, \quad (3)$$

## 2.1 The Siamese Network for Behavioral Prediction in Space and Time

To score the agreement between the time series prediction and original brain measurements, we introduce the Siamese network [1] which is the weight sharing network for comparing two views. It should be noted that the Siamese network is only used in the training process to help learn the structural and functional

patterns of fMRI and MEG. We give detailed description of training and testing procedure in Supplementary material. The Siamese network is used to measure the similarity between fMRI and MEG signals. The Siamese network consists of two views, the predicted MEG signal using fMRI and the corresponding ground truth MEG signal. The two input views are preprocessed using the encoder network  $\acute{O}(\cdot)$ . The encoder network shares the same weights between two views. The encoder network  $\acute{O}(\cdot)$  of the predicted MEG-like signal is followed by an MLP head  $\Omega(\cdot)$  to match the output of the other view.

**Encoder network with continuous convolutions** To predict the multi-modal brain measurements using the orthogonal basis in Equation 3, We could use the continuous convolutions to represent the output signals. Following the previous work of bilinear method, we converts the state  $\alpha$  and  $\beta$  in Equation 3 into an approximation  $\tilde{\alpha} = (I - \Delta/2 \cdot \alpha)^{-1}(I + \Delta/2 \cdot \alpha)$  and  $\tilde{\beta} = (I - \Delta/2 \cdot \alpha)^{-1} \Delta \beta$ . Based on Equation 3, the continuous convolutions is defined as,

$$\begin{aligned} y_k &= \widetilde{O\alpha}^k \tilde{\beta} f_0 + \widetilde{O\alpha}^{k-1} \tilde{\beta} f_1 + \cdots + \widetilde{O\alpha} \tilde{\beta} f_{k-1} + \widetilde{O\beta} f_k \\ y &= \acute{O} * f \quad \text{and} \quad \acute{O} \in R^K = (\widetilde{O\beta}, \widetilde{O\alpha\beta}, \dots, \widetilde{O\alpha}^k \beta), \end{aligned} \quad (4)$$

**Similarity measurement** Given two kinds of inputs  $y_1$  and  $y_2$  at time  $t$  with  $y_1(t) = O(z_1(t))$ .  $y_1(t)$  represents the predicted brain measurement of the output model  $O(\cdot)$  in Equation 1.  $y_2(t)$  is the original time series of brain activity, such as fMRI or MEG signals. The similarity of two view is defined as

$$D(\rho_1(t), \eta_2(t)) = -\frac{\rho_1(t)}{\|\rho_1(t)\|_2} * \frac{\eta_2(t)}{\|\eta_2(t)\|_2}, \quad (5)$$

where  $\rho_1(t) = \Omega(\acute{O}(y_1(t)))$  and  $\eta_2(t) = \acute{O}(y_2(t))$ .  $\|\cdot\|_2$  is  $l_2$ -norm.

In the experiment, we adopt the stop-gradient operation. The stop-gradient treat the second input  $\eta_2(t)$  as constant. The symmetrized loss is denoted by

$$\mathcal{L}_s = \frac{1}{2} D(\rho_1(t), \text{stopgrad}(\eta_2(t))) + \frac{1}{2} D(\rho_2(t), \text{stopgrad}(\eta_1(t))), \quad (6)$$

### 3 Experimental Results

We tested the fidelity of the proposed method based on the resting state fMRI and MEG from HCP dataset. The resting-state fMRI was pre-processed following the minimal preprocessing pipeline [5]. Then the pre-processed data was registered into a standard cortical surface using MSMAll [5]. The artefacts were removed using ICA-FIX. The cortical surface was parcellated into N=360 major ROI[4]. In addition, the averaged time course of each ROI was normalized using z-score. The resting-state MEG was pre-processed using ICA to remove out artefacts related to head and eye movement. Sensor-space data were down-sampled to 300Hz using anti-aliasing filter. Next the MEG data were source-reconstructed with a scalar beamformer and registered into the standard space of the Montreal Neuroimaging Institute (MNI). MEG signals were then filtered into 1-30Hz

and beamformed onto 6 mm grid. We used theta (4-8 Hz), alpha (8-13 Hz) and beta (13-30 Hz) bands to filter the source-space data. The parcellation atlas and z-score normalization method of MEG were similar to the resting-state fMRI.

### 3.1 Spatio-Temporal Patterns of the Predicted Results Using 4D fMRI Image

We next applied the proposed method to acquire the dynamics of the behavioral representation for each voxel in fMRI signal. We used the whole 4D fMRI image to generate the corresponding behavioral states in the spatio-temporal domain. We show the spatial map of independent temporal signal using original MEG and the predicted results in Fig. 2. The spatial map was generated using the ICA. With the 25 generated ICA temporal components of MEG and predicted results, we paired the 8 resting state brain networks (RSNs) spatial map with that derived using resting state fMRI. The DMN pattern is shown in Fig. 2a. The nodes are highlighted in the medial frontal cortex and inferior parietal lobules. The patterns of left lateralized frontoparietal and sensorimotor network are shown in Fig. 2b and Fig. 2c. From Fig. 2, we could see that the spatial pattern of predicted results could match that of MEG in all the network. However, the spatial resolution of predicted result is much higher than MEG signal.

The results have shown that the temporal ICA components originate from the brain regions correlated with the RSNs spatial maps. We show more results in the Supplementary material. The proposed method could generate the high resolution behavioral states that are the perfect matches to the spatial pattern of the observed fMRI. Thus, the proposed method could provide fundamental role of brain dynamics related to behavioral measurements.

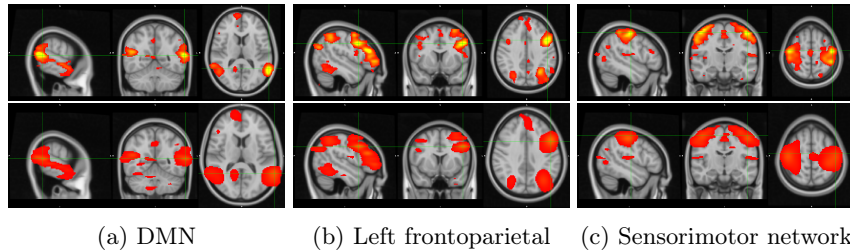


Fig. 2: Brain networks acquired with ICA shown in the order predicted results (top) and MEG (bottom). (a) DMN, (b) left frontoparietal network, (c) sensorimotor network.

### 3.2 Temporal Pattern of Predicted Results

We finally evaluated temporal pattern of the predicted internal states compared with original 4D fMRI and MEG image. Fig. 3 shows the averaged neural time

series of the predicted results, fMRI and MEG signal in visual network. From Fig. 3, we could see that the proposed method inherits the high temporal resolution of MEG signal in a dynamical system. The proposed method predicts the hidden observation of the dynamical neural transmission locally and globally. In Table 1, we use the mean squared error (MSE) to measure the similarity between the ROI averaged predicted results and ROI averaged MEG signals. In addition, we also compare the proposed method with three baselines LSTM, GRU-D and proposed method without Siamese network. Polynomial Projection Operators could be combined with The introduction of the spatio-temporal constraints into the dynamical system is consistent with the fundamental role in biological system. The proposed method provides chance to understand how behavioral representations evolve over time.

Table 1: Temporal pattern prediction results with different baseline models

| Methods | LSTM   | GRU-D  | Without Siamese Network | Proposed Method |
|---------|--------|--------|-------------------------|-----------------|
| MSE     | 0.3199 | 0.5805 | 0.1134                  | 0.0577          |

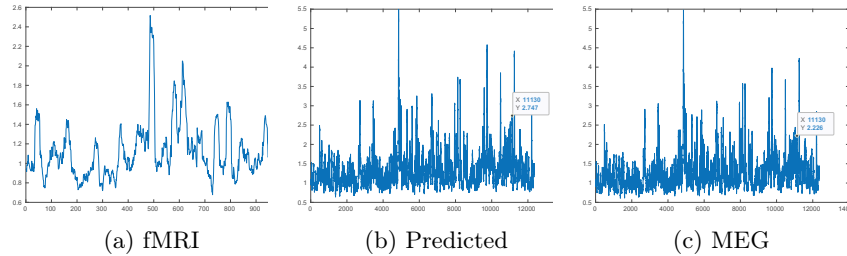


Fig. 3: Averaged neural time series in visual network (a) fMRI, (b) predicted result, (c) MEG.

## 4 Conclusions

To understand the complex brain dynamics, we need to record the activity in space and time, which could not be solved using current noninvasive techniques of brain measurements. We propose a novel computational model that could combine the information from several techniques and predict the internal brain activity with both high spatial and high temporal resolution. The proposed framework address the vanishing gradient problem by abstracting the long term temporal relationship with functional approximation. In the training step, we use both fMRI and MEG data to learn the brain dynamical representation with Siamese network. While in the testing step, for the first time, the proposed method solves



the problem of predicting the internal behavioral states with high resolution in spatial and temporal domain using only fMRI data. The potential of the proposed method to represent the spatio-temporal dynamics has been demonstrated using two experiments with HCP data.

## Acknowledgement

This work was partially supported by NIH R01AG071243, R01MH125928, R01AG049371, U01AG068057, and NSF IIS 2045848, 1845666, 1852606, 1838627, 1837956, 1956002, IIA 2040588.

## References

1. Chen, X., He, K.: Exploring simple siamese representation learning. arXiv preprint arXiv:2011.10566 (2020)
2. Cornblath, E.J., Ashourvan, A., Kim, J.Z., Betzel, R.F., Ciric, R., Adebimpe, A., Baum, G.L., He, X., Ruparel, K., Moore, T.M., et al.: Temporal sequences of brain activity at rest are constrained by white matter structure and modulated by cognitive demands. *Communications biology* **3**(1), 1–12 (2020)
3. Gallego, J.A., Perich, M.G., Naufel, S.N., Ethier, C., Solla, S.A., Miller, L.E.: Cortical population activity within a preserved neural manifold underlies multiple motor behaviors. *Nature communications* **9**(1), 1–13 (2018)
4. Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., et al.: A multi-modal parcellation of human cerebral cortex. *Nature* **536**(7615), 171–178 (2016)
5. Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., et al.: The minimal preprocessing pipelines for the human connectome project. *Neuroimage* **80**, 105–124 (2013)
6. Gu, A., Dao, T., Ermon, S., Rudra, A., Ré, C.: Hippo: Recurrent memory with optimal polynomial projections. arXiv preprint arXiv:2008.07669 (2020)
7. Kriegeskorte, N., Mur, M., Bandettini, P.A.: Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in systems neuroscience* **2**, 4 (2008)
8. McIntosh, A.R., Mišić, B.: Multivariate statistical analyses for neuroimaging data. *Annual review of psychology* **64**, 499–525 (2013)
9. Musall, S., Kaufman, M.T., Juavinett, A.L., Gluf, S., Churchland, A.K.: Single-trial neural dynamics are dominated by richly varied movements. *Nature neuroscience* **22**(10), 1677–1686 (2019)
10. Norman, K.A., Polyn, S.M., Detre, G.J., Haxby, J.V.: Beyond mind-reading: multi-voxel pattern analysis of fmri data. *Trends in cognitive sciences* **10**(9), 424–430 (2006)
11. Saxena, S., Cunningham, J.P.: Towards the neural population doctrine. *Current opinion in neurobiology* **55**, 103–111 (2019)
12. Shine, J.M., Breakspear, M., Bell, P.T., Martens, K.A.E., Shine, R., Koyejo, O., Sporns, O., Poldrack, R.A.: Human cognition involves the dynamic integration of neural activity and neuromodulatory systems. *Nature neuroscience* **22**(2), 289–296 (2019)

13. Stringer, C., Pachitariu, M., Steinmetz, N., Carandini, M., Harris, K.D.: High-dimensional geometry of population responses in visual cortex. *Nature* **571**(7765), 361–365 (2019)
14. Taghia, J., Cai, W., Ryali, S., Kochalka, J., Nicholas, J., Chen, T., Menon, V.: Uncovering hidden brain state dynamics that regulate performance and decision-making during cognition. *Nature communications* **9**(1), 1–19 (2018)
15. Tang, E., Mattar, M.G., Giusti, C., Lydon-Staley, D.M., Thompson-Schill, S.L., Bassett, D.S.: Effective learning is accompanied by high-dimensional and efficient representations of neural activity. *Nature neuroscience* **22**(6), 1000–1009 (2019)
16. Tavares, R.M., Mendelsohn, A., Grossman, Y., Williams, C.H., Shapiro, M., Trope, Y., Schiller, D.: A map for social navigation in the human brain. *Neuron* **87**(1), 231–243 (2015)
17. Voelker, A.R., Kajić, I., Eliasmith, C.: Legendre memory units: Continuous-time representation in recurrent neural networks. *Proceedings of the 33rd International Conference on Neural Information Processing Systems* (2019)
18. Zhao, C., Gao, X., Emery, W.J., Wang, Y., Li, J.: An integrated spatio-spectral-temporal sparse representation method for fusing remote-sensing images with different resolutions. *IEEE Transactions on Geoscience and Remote Sensing* **56**(6), 3358–3370 (2018)
19. Zhao, C., Li, H., Jiao, Z., Du, T., Fan, Y.: A 3d convolutional encapsulated long short-term memory (3dconv-lstm) model for denoising fmri data. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 479–488. Springer (2020)