CHARACTERIZING AND DIFFERENTIATING TASK-BASED AND RESTING STATE FMRI SIGNALS VIA TWO-STAGE DICTIONARY LEARNING

Shu Zhang[1], Xiang Li[1*], Jinglei Lv[1, 2], Xi Jiang[1], Bao Ge[1, 3], Lei Guo[2], Tianming Liu[1]*

[1]Cortical Architecture Imaging and Discovery Lab, Department of Computer Science, The University of Georgia, Athens, GA [2]School of Automation, Northwestern Polytechnical University, Xi'an, China. [3]School of Physics & Information Technology, Shaanxi Normal University, Xi'an, China. *Joint first authors.

ABSTRACT

A relatively underexplored question in fMRI is whether there are intrinsic differences in terms of signal composition patterns that can effectively characterize and differentiate task-based and resting state fMRI (tfMRI or rsfMRI) signals. In this paper, we propose a novel two-stage sparse representation framework to examine the fundamental difference between tfMRI and rsfMRI signals. In the first stage, subject-wise whole-brain tfMRI and rsfMRI signals are factorized into dictionary matrix and the corresponding loading coefficients via dictionary learning method. In the second stage, dictionaries learned at the first stage across multiple subjects are aggregated into the matrix which serve as the input for another round of dictionary learning, obtaining groupwise common dictionaries and their loading coefficients. This framework had been applied on the recently publicly released Human Connectome Project (HCP) data, and experimental results revealed that there exist distinctive and descriptive atoms in the groupwise common dictionary that can effectively differentiate tfMRI and rsfMRI signals, achieving 100% classification accuracy. Moreover, certain common dictionaries learned by our framework have a clear neuroscientific interpretation. For example, the well-known default mode network (DMN) activities can be recovered from the heterogeneous and noisy large-scale groupwise whole-brain signals.

Index Terms— tfMRI, rsfMRI, sparse representation, classification, big data

1. INTRODUCTION

Functional magnetic resonance imaging (fMRI) based on blood-oxygen-level dependent (BOLD) techniques has been widely used to study the functional activities and cognitive behaviors of the brain based on the induced stimulus by tasks, i.e., task fMRI (tfMRI) [1-3] or during task-free resting-state, i.e., resting state fMRI (rsfMRI) [4-5]. To infer meaningful neuroscientific patterns within fMRI data, various computational/statistical methods have been proposed, including the widely-used general linear model (GLM) for tfMRI [2, 6], independent component analysis (ICA) for rsfMRI [7], as well as many others methods including wavelet algorithms [8], Markov random field (MRF) models [9], Bayesian approaches [10] and etc.

However, a relatively underexplored question in tfMRI and rsfMRI is whether there are intrinsic, fundamental differences in signal composition patterns that can effectively characterize and differentiate these two types of fMRI signals. As far as we know, there are at least three challenges in addressing the above question. Firstly, the variability of fMRI signals across brain scans and across individual subjects could be remarkable. Secondly, the amount of whole-brain, voxel-wise fMRI signals from multiple subjects could be immense. For example, high-resolution tfMRI scan from the recently publicly released Human Connectome Project (HCP) has around 150,000-200,000 time series signals for one subject during a single task/resting-state scan [15]. In total, for Q1 release of HCP data, there are around 10,200,000-13,600,000 time series signals for all 68 subjects of a single task. Thirdly, there are a variety of noise sources in fMRI signals. During fMRI scans, several factors including scanner instability, experiment design deficits, and effects of susceptibility of high fields may all lead to noises. For an individual subject, head motion, lack of attention and other factors that are not related to the experiment design could also introduce noises. Inspired by the successes of using sparse representation in pattern recognition [11, 12] and in functional brain imaging analysis [13, 14], we propose a novel two-stage sparse representation framework to obtain a groupwise characterization of fMRI signals obtained during both task and resting-state, which could address the abovementioned three challenges and have the capability of effectively classifying datasets from task and resting-state.

2. METHOD

2.1. Data acquisition and preprocessing

The dataset used in this work was obtained from the Human Connectome Project Q1 release [15]. The acquisition parameters of tfMRI data are: 90×104 matrix, 220mm FOV, 72 slices, TR=0.72s, TE=33.1ms, flip angle = 52°, BW=2290 Hz/Px, in-plane FOV = 208×180 mm, 2.0mm isotropic voxels. For tfMRI images, the preprocessing pipelines included motion correction, spatial smoothing,

temporal pre-whitening, slice time correction, global drift removal. More detailed data acquisition and preprocessing could be referred to [15], rsfMRI data were acquired with the same EPI pulse sequence parameters as tfMRI [16].

In this work, we have split the total 68 subjects in the Q1 release into two equal subsets. One set would be used for the training process, which includes two-stage dictionary learning and training of the classifier. The other set (testing) would only go through the first-stage dictionary learning, then we applied common dictionaries and the inferred features obtained from training set on the testing set for the classification and verification.

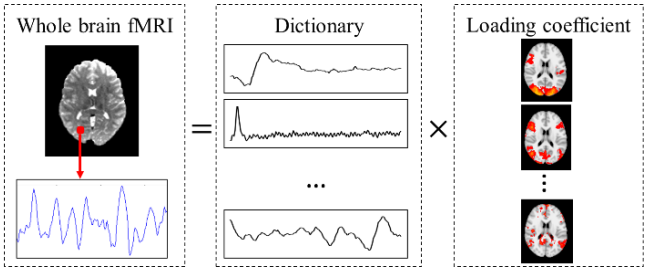

Figure 1. Illustration of the first-stage dictionary learning process. Left panel: individual whole brain voxel-wise fMRI data; Middle panel: dictionary matrix D as time series learned from the input; Right panel: corresponding loading coefficient α learned simultaneously with D as spatial maps.

2.2. Two-stage dictionary learning

In the first stage, the online dictionary learning algorithm [11] is adopted to learn a dictionary with sparsity constraint from the whole-brain fMRI signals (with time length t and voxel number n) from fMRI datasets during each task (and resting-state) of each subject. The algorithm would learn a meaningful and over-complete dictionary D consisting of k_1 atoms ($k_1>t$, $k_1<<n$) to represent input data with the corresponding sparse loading coefficient matrix α. Specifically, for the fMRI signal set $\mathbf{S}=[s_1,s_2,\ldots s_n]\epsilon R^{t\times n}$, the loss function for the dictionary learning algorithm to minimize is defined in Eq. (1) with a l_1 regularization that yields to a sparse constraint to the loading coefficient α, where λ is a regularization parameter to trade-off the regression residual and sparsity level:

$$\min_{\boldsymbol{D}\epsilon\mathbb{R}^{t\times k},\boldsymbol{\alpha}\epsilon\mathbb{R}^{k1\times n}} \frac{1}{2}||\mathbf{S}-\boldsymbol{D\alpha}||_F+\boldsymbol{\lambda}||\boldsymbol{\alpha}||_{1,1} \quad (1)$$

To prevent D from being arbitrarily large which leads to trivial solution of the optimization, its columns d_1, d_2 ... d_k are constrained by Eq. (2):

$$\boldsymbol{C}\triangleq\{\boldsymbol{D}\epsilon\mathbb{R}^{t\times k1}\ \ s.t.\ \ \forall j=1,\ldots k_1,\boldsymbol{d}_j^T\boldsymbol{d}_j\leq 1\} \quad (2)$$

The optimization procedure is done by iteratively updating D and α in Eq. (1). It should be noted that we employ the same assumption as in previous studies [17] that the atomic components (which are dictionary atoms in D) involved in each voxel's fMRI signal are a few major ones and the neural integration of those components is linear. The value of λ and dictionary size k_1 were determined experimentally in our previous studies (λ=0.1, k=400) [13, 14]. After the dictionary learning, the resulting D matrix contains temporal variation of each atomic basis functional component of fMRI data, while the corresponding sparse loading coefficient matrix α contains the spatial distribution of each component. A sample visualization showing the result of first-stage dictionary learning is shown in Fig. 1.

Based on the first-stage dictionary learning results, our next goal is to obtain the groupwise characterization of the dictionaries that could reveal the distinctive organization patterns between fMRI data under different conditions (in this work, task vs. resting-state). Thus we combine the dictionaries obtained from each tfMRI with the dictionaries obtained from the one rsfMRI data across all subjects in the training set, forming one multi-subjects, across task/rest, combined data matrix $\mathbf{S}^*$ of dimension $t\times(2k_1p)$, p is the number of subjects in the dataset. Note that in HCP dataset, rsfMRI data has a longer time length than tfMRI data for all tasks and so does the learned dictionaries. So the dictionaries learned during resting-state are truncate to the same length with their task counterparts. $\mathbf{S}^*$ would then be used as the input for the second-stage dictionary learning based on the same algorithm as used in the first-stage, where the parameters are λ=0.1, k_2=50, aiming at obtaining a group-wise common dictionary D^* and the corresponding loading coefficients α^*, which could reflect the groupwise temporal and spatial organization patterns of the given dataset (in this case, the combined tfMRI/rsfMRI data). Compared with the original fMRI data which are defined on the whole brain voxels, our two-stage framework achieves a huge size reduction while still maintaining the major functional characterization for each individual. More importantly, noises and undesired voxel-wise signal fluctuations are largely removed in $\mathbf{S}^*$, thus we can ensure that most of the common dictionaries can represent the groupwise consistent functional activities, and their differences are more likely to be originated from the intrinsic features of functional brain activity patterns. As the common dictionaries are learned from the dictionary space, rather than the original voxel space, we need to estimate their spatial maps over the brain by first aligning all the brains into the same template using linear registration, thus transforming each α into α'. Then the spatial map of the i-th common dictionary could be obtained by:

$$\sum_{x}^{p}\sum_{y}^{k1}\boldsymbol{\alpha}'_{\boldsymbol{x,y}}\cdot\boldsymbol{\alpha}^*_{\boldsymbol{i,(x-1)m+y}} \quad (3)$$

where $\alpha'_{x,y}$ is the loading coefficient matrix of the y-th dictionary (over the total of k_1) of the x-th subject (over the total of p) obtained from the first stage dictionary learning registered to the template, and $\alpha^*_{i,,xm+y}$ is the loading coefficient value of the y-th dictionary of the x-th subject for the i-th common dictionary from the second stage dictionary learning. In other words, the spatial maps of the common dictionaries are the weighted average of the loading

coefficient each individual dictionary of each subject. The spatial map of one sample common dictionary, along with its time series is visualized in Fig. 2.

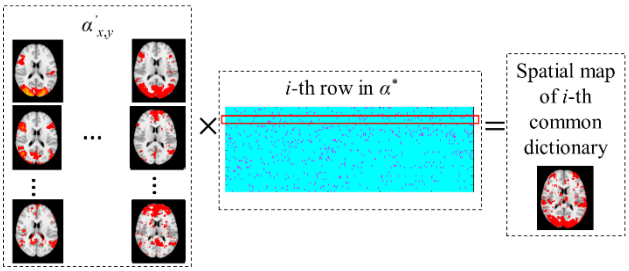

Figure 2. Visualization of the procedure to obtain the spatial map of one sample common dictionary. The corresponding row in α^* of the i-th common dictionary is highlighted by the red block.

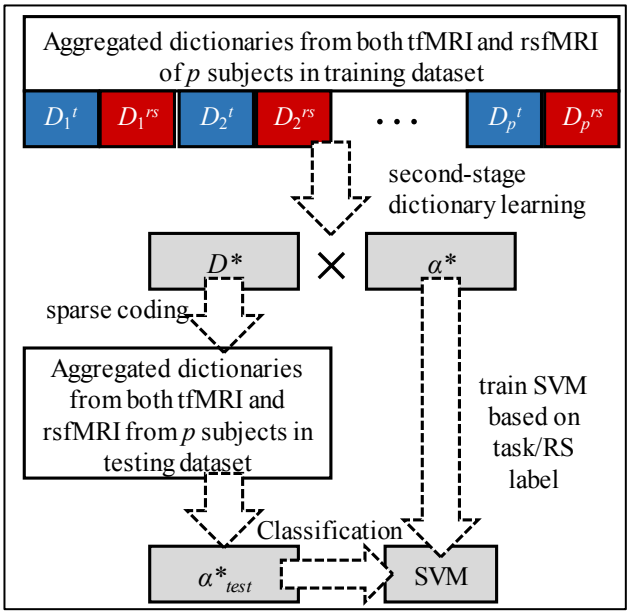

Figure 3. Illustration of the classification pipeline.

2.3. Sparse coding of the testing data and classification

For the purpose of verification of the proposed framework, we performed the classification analysis on the testing dataset. Firstly, we use loading coefficient of common dictionary α^* to train an SVM, establishing the relationship between the common dictionary composition of the dictionary learned from the first-stage with the type of dataset it is coming from (tfMRI or rsfMRI). Then the common dictionary D^* is used to sparse code (solving an l_1-regularized LASSSO problem) the dictionaries learned at the first-stage from the testing dataset, obtaining the loading coefficients of common dictionaries in the testing set α^*_{test}. Thus the relationship between training and testing dataset is established by the fact that both of their individual dictionaries learned during the first stage are sparsely coded by the same common dictionary D^*, making the rows in α^* and α^*_{test} corresponding to the same common dictionary. The rationale that we can project the dictionaries learned from the training dataset to the testing dataset of the same task is because they are from the same population and supposed to share the common underlying basis functional activation pattern. α^*_{test} would be classified by the trained SVM, an illustrative diagram of the whole classification scheme is shown in Fig. 3. After obtaining the classification results of the dictionaries in each fMRI dataset from each subject (i.e. dictionary-wise classification), we would use a simple majority voting strategy to determine the label of that dataset (i.e. dataset-wise classification), as the dataset constituted by those dictionaries shall have only one label.

3. RESULTS

By using the HCP dataset described in section 2.2, we combined each of the tfMRI data from seven different tasks with the same rsfMRI data, forming seven combined datasets including emotion/rsfMRI, gambling/rsfMRI, language/rsfMRI, motor/rsfMRI, social/rsfMRI, relational/rsfMRI and working memory/rsfMRI for each subject. Then we applied the proposed framework on the seven combined datasets. In all the datasets, tfMRI and rsfMRI can be effectively differentiated with dataset-wise classification accuracy of 100%. The results demonstrate that there exist fundamental differences between the component composition of tfMRI and rsfMRI, while the intrinsic spatial/temporal pattern underlying such difference could be characterized by the learned common dictionaries from the large and noisy group-wise data. There are two types of common dictionaries that have been learned from the propose two-stage dictionary learning framework: task-evoked components and resting-state components. Quite naturally they also played the key role in differentiating the tfMRI/rsfMRI dataset, which would be illustrated below.

3.1. Task-evoked common functional components

The most prominent and intuitive common dictionaries obtained by our framework are the task-evoked components. Here we will use the results from two tasks as examples to showcase their characteristics. In both the working memory/rsfMRI and emotion/rsfMRI combined datasets, there exists task-evoked component with very high percentage of presence (89.5% and 79.97%, respectively) in tfMRI, and similar spatial distributions were obtained from group-wise GLM activation detection applied on the tfMRI of Emotion or WM task from the 34 subjects in training dataset, as shown in Fig. 4(a) and (b). Their time series, plotted in Fig. 4(c), are correspondent with the task design contrast curves (Pearson correlation value: 0.6653 and 0.6679). Similar results were also obtained from the other 5 tasks in the HCP dataset. Based on the spatial and temporal characteristic and its sole presence in tfMRI, we can be assured that our framework could effectively identify task-evoked functional component from the combined fMRI data.

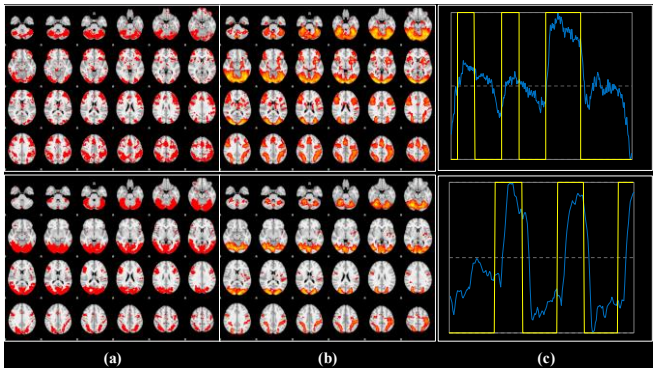

Figure 4. Example task-evoked common dictionaries from WM/RS (top panel) and Emotion/RS (bottom panel) dataset. (a): volume map of the common dictionaries; (b) volume map of the contrast maps obtained by group-wise GLM; (c): time series of the common dictionaries (blue), task design contrast curve (yellow);

3.2. Resting-state components

Opposite to the task-evoked components, there exist resting-state components with high presence in rsfMRI (75%, 82%) both in the emotion/RS and WM/RS dataset. As visualized in Fig. 5(a), their spatial map largely resembles the widely-reported default mode network (DMN) [4]. We had applied the group-wise independent component analysis (ICA) on the same dataset, certain components obtained by ICA were found to have the spatial pattern, as shown in Fig. 5(b) (spatial overlapping rate with ICA resting-state map: 83% and 79.97%, respectively).

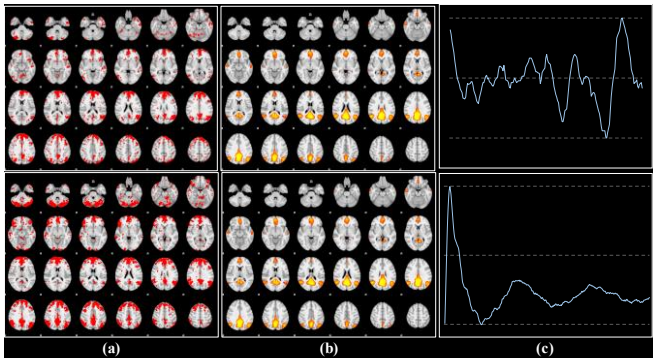

Figure 5. Example resting-state common dictionaries from Emotion/RS (top panel) and WM/RS (bottom panel) dataset. (a): volume map of the common dictionaries; (b) volume map of the contrast maps obtained by group-wise GLM; (c): time series of the common dictionaries (blue), task design contrast curve (yellow).

4. CONCLUSION

We have presented a novel two-stage sparse representation framework to examine the intrinsic differences in tfMRI/rsfMRI signals. The major methodological novelty of the two-stage sparse representation is that the framework can effectively remove the noises and undesired voxel-wise signal fluctuations, efficiently deal with the big-data (a matrix of millions times hundreds data points), and infer distinctive and descriptive common dictionary atoms that can well characterize and differentiate tfMRI/rsfMRI signals in task performance and resting state. The applications of this framework on seven HCP tfMRI datasets and one rsfMRI dataset have demonstrated promising results. In the future, we plan to better interpret other dictionary atoms and apply this framework to clinical to elucidate possible alterations of functional activities in brain disorders.

5. REFERENCES

[1] Worsley, KJ., Friston, KJ., "Analysis of fMRI time series revisited Again," *NeuroImage* 2, 173–181, 1995

[2] Worsley KJ., "An overview and some new developments in the statistical analysis of PET and fMRI data," *Hum Brain Mapp*, 5(4):254-8, 1997

[3] Linden, D.E, Prvulovic, D., et al., "The Functional Neuroanatomy of Target Detection: An fMRI Study of Visual and Auditory Oddball Tasks," *Cerebral Cortex*, 9(8): 815-823, 1999.

[4] Raichle M.E., MacLeod A.M., Snyder A.Z., Powers W.J., Gusnard D.A., Shulman G.L., "A default mode of brain function," *Proc Natl Acad Sci USA*, 98(2): 676-682, 2001.

[5] Fox, M.D., and Raichle, M.E., "Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging," *Nat Rev Neurosci*, 8: 700-711, 2007.

[6] Friston KJ., Holmes AP., Worsley KJ., "Statistical parametric maps in functional imaging: a general linear approach," *Human Brain Mapping*, V2-I4:189-210, 1994

[7] McKeown, M.J., et al., "Spatially independent activity patterns in functional MRI data during the Stroop color-naming task," *PNAS*, 1998. 95(3): p. 803, 1998.

[8] Bullmore E., Fadili J., et al., "Wavelets and statistical analysis of functional magnetic resonance images of the human brain," *Stat Methods Med Res*, Oct 12(5):375–399, 2003.

[9] Descombes X., Kruggel F., von Cramon DY., "fMRI signal restoration using a spatio-temporal markov random field preserving transitions." *NeuroImage*. Nov;8(4):340–349, 1998.

[10] Luo H., Puthusserypady S., "fMRI data analysis with nonstationary noise models: a Bayesian approach," *IEEE Trans Biomed Eng.* Sep;54:1621–1630, 2007.

[11] Mairal J., Bach Francis., Ponce J., and Sapiro G., "Online dictionary learning for sparse coding," *In Proceedings of the International Conference on Machine Learning (ICML)*, 2009.

[12] Kreutz-Delgado K., Murray JF., Rao BD., Engan K., Lee TW., Sejnowski TJ., "Dictionary learning algorithms for sparse representation," *Neural Comput*, Feb;15(2):349-96, 2003.

[13] Lv J., Jiang X., Li X., et al., "Sparse Representation of Whole-brain FMRI Signals for Identification of Functional Networks," in press, *Medical Image Analysis*, 2014a.

[14] Lv J., Jiang X., Li X., et al, "Holistic Atlases of Functional Networks and Interactions Reveal Reciprocal Organizational Architecture of Cortical Function," in press, *IEEE Transactions on Biomedical Engineering*, 2014b.

[15] Barch DM., Burgess GC., Harms MP., et al,. "Function in the human connectome: task-fMRI and individual differences in behavior," *Neuroimage*, 2013.

[16] Smith S., Johansen-Berg H., Snyder AZ., et al., "Function in the human connectome: task-fMRI and individual differences in behavior," *Neuroimage*, 2013.

[17] Li Y., Long J., He L., Lu H., Gu Z., et al., "A Sparse Representation-Based Algorithm for Pattern Localization in Brain Imaging Data Analysis," *PLoS ONE* 7(12): e50332, 2012.