MULTISET FMRI DATA ANALYSIS WITH SUBJECT GROUP INFORMATION USING STRUCTURED DICTIONARY LEARNING

Shuai Xu [1], *Rui Jin* [1], *Seung-Jun Kim* [1*], *Vince D. Calhoun* [2], *and Tülay Adali* [1]

[1]Department of Computer Science and Electrical Engineering
University of Maryland, Baltimore County, Baltimore, MD, USA
[2]Tri-institutional Center for Translational Research in Neuroimaging and Data Science,
Georgia State University, Georgia Institute of Technology, and Emory University, GA, USA

ABSTRACT

A dictionary learning (DL)-based method for analyzing multiple functional magnetic resonance imaging (fMRI) data sets is developed. The algorithm can incorporate subject group information to extract neural activation maps indicative of the group differences and the maps shared across groups. Furthermore, multiple data sets are jointly analyzed to obtain the maps common across data sets and those unique to individual data sets. For this, a novel structured supervised DL problem is formulated. Numerical tests on synthetic and real fMRI data sets verify the effectiveness of the proposed method.

Index Terms— Dictionary learning, fMRI data analysis, group sparsity, multiset analysis.

1. INTRODUCTION

Functional magnetic resonance imaging (fMRI) is a technique for capturing brain activities noninvasively through blood oxygenation level dependent (BOLD) neurovascular coupling. To analyze the fMRI data, model-driven methods such as the general linear model (GLM) employ a postulated hemodynamic response function to regress spatial neural activation maps incurred by given temporal stimuli. On the contrary, data-driven approaches make minimal assumptions on the data to obtain neural activations.

Representative data-driven fMRI data analysis techniques include independent component analysis (ICA) and dictionary learning (DL). ICA postulates that the data is generated through a linear combination of statistically independent component maps. Various ICA-based algorithms have been developed to tackle multisubject fMRI data to estimate activation maps that are shared or statistically dependent among subjects [1]. The DL approach, on the other hand, assumes that component activation maps are sparse. The DL framework is flexible to incorporate additional prior information such as structural sparsity or label information [2].

The focus of this work is on jointly analyzing multiple fMRI data sets. For example, multiple task-based fMRI data sets acquired while performing various tasks can provide complementary views to brain activities. Identification of common and distinct components through multiset analysis can help understand the underlying relationship among the data sets [3]. Several ICA-based approaches to perform multiset fMRI data analysis have been developed, where common activations across multiple data sets were found in addition to the distinct activation maps unique to individual data sets [4, 5]. However, the methods did not incorporate subject group information at the map estimation stage.

In this work, we propose a DL-based multisubject multiset fMRI data analysis method that can incorporate subject group information as well as commonality across data sets. A structured supervised DL problem is formulated, where the group attributes are incorporated through Fisher's discriminant cost [6], while group-sparsity constraints are employed to capture shared traits across data sets [7].

Notations: The i-th row, the j-th column, and the (i,j)-entry of a matrix $\mathbf{A}$ are denoted by $\mathbf{A}(i,:)$, $\mathbf{A}(:,j)$, and $\mathbf{A}(i,j)$, respectively. The Frobenius norm, the ℓ_1-norm, and the $\ell_{1,2}$-norm of $\mathbf{A}$ is defined as $\|\mathbf{A}\|_F := \sqrt{\sum_{i,j}\mathbf{A}(i,j)^2}$, $\|\mathbf{A}\|_1 := \sum_i\sum_j|\mathbf{A}(i,j)|$, and $\|\mathbf{A}\|_{1,2} := \sum_i\sqrt{\sum_j\mathbf{A}(i,j)^2} = \sum_i\|\mathbf{A}(i,:)\|_2$, respectively. The transpose of $\mathbf{A}$ is denoted by $\mathbf{A}^\top$, and the trace of matrix $\mathbf{A}$ is denoted by $\mathrm{tr}\{\mathbf{A}\}$.

2. SUPERVISED DL FOR FMRI DATA ANALYSIS

Given a fMRI dataset $\mathbf{X} \in \mathbb{R}^{M\times V}$ involving M subjects and V voxels, DL finds sparse matrix $\mathbf{Z} \in \mathbb{R}^{K\times V}$ in which the rows capture K sparse neural activation maps and dictionary $\mathbf{D} \in \mathbb{R}^{M\times K}$ that includes the corresponding subject-wise weights for the individual spatial maps, such that $\mathbf{X} \approx \mathbf{DZ}$.

When subject group information is available, one can incorporate the information in the framework of supervised

This work was partly supported by NSF grants 1631838 and 2242412. This research study was conducted retrospectively using human subject data made available in open access. Ethical approval was not required with the open access data. *Corresponding author. E-mail: sjkim@umbc.edu.

DL. Let the M subjects be partitioned to C classes with $\mathcal{M}_c$ denoting the set of indices of the subjects that belong to class c, for $c = 1, \ldots, C$. Let M_c be the number of subjects in $\mathcal{M}_c$ with $\sum_{c=1}^{C} M_c = M$. Denote as $\mathbf{q}_m$ the feature vector of subject m for $m = 1, 2, \ldots, M$. Then, one can define the per-class centroid $\mathbf{m}_c := M_c^{-1} \sum_{m \in \mathcal{M}_c} \mathbf{q}_m$ and the overall centroid $\mathbf{m} := M^{-1} \sum_{m=1}^{M} \mathbf{q}_m$. With $\mathbf{Q} := [\mathbf{q}_1, \ldots, \mathbf{q}_M]$, we define also the within-class scatter matrix $\mathbf{S}_w$ and the between-class scatter matrix $\mathbf{S}_b$ as

$$\mathbf{S}_w(\mathbf{Q}) := \sum_{c=1}^{C} \sum_{m \in \mathcal{M}_c} (\mathbf{q}_m - \mathbf{m}_c)(\mathbf{q}_m - \mathbf{m}_c)^\top \quad (1)$$

$$\mathbf{S}_b(\mathbf{Q}) := \sum_{c=1}^{C} M_c(\mathbf{m}_c - \mathbf{m})(\mathbf{m}_c - \mathbf{m})^\top \quad (2)$$

respectively. Fisher's discriminant cost can then be defined as

$$f(\mathbf{Q}) := \operatorname{tr}\{\mathbf{S}_w(\mathbf{Q})\} - \operatorname{tr}\{\mathbf{S}_b(\mathbf{Q})\} + \|\mathbf{Q}\|_F^2. \quad (3)$$

Thus, by minimizing $f(\mathbf{Q})$, the features tend to cluster together within the same class and disperse apart across different classes. The last term in (3) ensures that $f(\mathbf{Q})$ is convex with respect to (w.r.t.) $\mathbf{Q}$.

To formulate a supervised DL problem for fMRI data, $\mathbf{Z}$ is split into group-discriminative maps $\widetilde{\mathbf{Z}} \in \mathbb{R}^{\widetilde{K} \times V}$ and non-discriminative maps $\bar{\mathbf{Z}} \in \mathbb{R}^{\bar{K} \times V}$ as $\mathbf{Z} := [\bar{\mathbf{Z}}^\top, \widetilde{\mathbf{Z}}^\top]^\top$, where $\widetilde{K} + \bar{K} = K$. One can also partition $\mathbf{D}$ compatibly as $\mathbf{D} = [\bar{\mathbf{D}}, \widetilde{\mathbf{D}}]$, where the rows of $\widetilde{\mathbf{D}}$ represent the discriminative features for the individual subjects. Thus, $\widetilde{\mathbf{D}}^\top$ can be used as the input to the Fisher's cost as $f(\widetilde{\mathbf{D}}^\top)$. Then, supervised DL is formulated as [2]

$$\min_{\mathbf{D}, \mathbf{Z}} \frac{1}{2} \|\mathbf{X} - \mathbf{D}\mathbf{Z}\|_F^2 + \lambda \|\mathbf{Z}\|_1 + \frac{\mu}{2} f(\widetilde{\mathbf{D}}^\top) \quad (4a)$$

$$\text{s.t. } \mathbf{D} \in \mathcal{D} := \{[\mathbf{d}_1, \ldots, \mathbf{d}_K] : \|\mathbf{d}_k\|_2 \leqslant 1, \forall k\} \quad (4b)$$

where the first term in (4a) fits the factorization $\mathbf{DZ}$ to data $\mathbf{X}$, and the second term encourages sparsity in $\mathbf{Z}$. The strengths of the regularizers are adjusted by nonnegative parameters λ and μ. The constraint (4b) ensures that the scaling ambiguity inherent in bi-factorization is resolved.

3. SUPERVISED DL FOR MULTISET ANALYSIS

When more than one fMRI dataset is available for the same set of subjects, performing a joint analysis of the multiset fMRI data is useful. Suppose that there are S fMRI datasets, with $\mathbf{X}^s \in \mathbb{R}^{M \times V}$ being the s-th dataset for $s = 1, \ldots, S$. Then the goal of joint DL analysis is again to factorize $\mathbf{X}^s$ into $\mathbf{X}^s \approx \mathbf{D}^s \mathbf{Z}^s$, with sparse maps in $\mathbf{Z}^s$ and corresponding weights in $\mathbf{D}^s$.

It is natural to expect that there are common neural activation maps across data sets and maps unique to individual data sets. Group sparsity priors can be employed to relate the common maps across data sets [8]. Thus, considering both

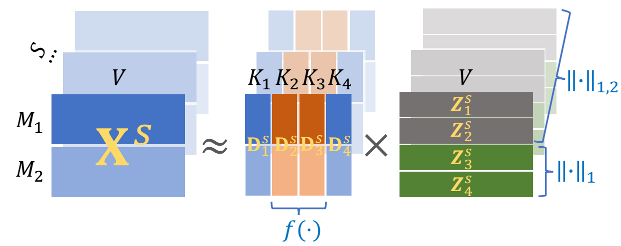

Fig. 1. Proposed structured DL model when there are two classes of M_1 and M_2 samples, respectively, and S data sets.

the subject groups and the multiple datasets, one can estimate four types of maps: **(m1)** maps that are common across datasets and shared across subject groups; **(m2)** maps that are common across datasets but discriminative across groups; **(m3)** maps that are distinct across datasets as well as group-discriminative; and **(m4)** maps that are distinct across datasets but shared across groups. Accordingly, we partition each map matrix $\mathbf{Z}^s$ as

$$\mathbf{Z}^s := [(\mathbf{Z}_1^s)^\top, (\mathbf{Z}_2^s)^\top, (\mathbf{Z}_3^s)^\top, (\mathbf{Z}_4^s)^\top]^\top, \; s = 1, \ldots, S \quad (5)$$

where $K = \sum_i K_i$, and $\mathbf{Z}_i^s \in \mathbb{R}^{K_i \times V}$, for $i \in \{1, \ldots, 4\}$, captures in its rows K_i maps of type **(mi)**. Similarly, a compatible partition of $\mathbf{D}^s$ is defined as

$$\mathbf{D}^s := [\mathbf{D}_1^s, \mathbf{D}_2^s, \mathbf{D}_3^s, \mathbf{D}_4^s], \; s = 1, \ldots, S \quad (6)$$

with $\mathbf{D}_i^s \in \mathbb{R}^{M \times K_i}$ representing the weights for maps of type **(mi)**. Collect also the group-discriminative features (corresponding to types **(m2)** and **(m3)**) in $\widetilde{\mathbf{D}}^s := [\mathbf{D}_2^s, \mathbf{D}_3^s] \in \mathbb{R}^{M \times K_{23}}$, where $K_{23} := K_2 + K_3$, upon which Fisher's cost is imposed. Next, aggregate the maps that are distinct across datasets (map types **(m3)** and **(m4)**) in $\widetilde{\mathbf{Z}}^s := [(\mathbf{Z}_3^s)^\top, (\mathbf{Z}_4^s)^\top]^\top \in \mathbb{R}^{K_{34} \times V}$ for each s, with $K_{34} := K_3 + K_4$, which simply needs to be (element-wise) sparse, as is expected for neural activations. Finally, with $K_{12} := K_1 + K_2$, define

$$\bar{\mathcal{Z}}_v := \begin{bmatrix} \mathbf{Z}_1^1(:, v), \mathbf{Z}_1^2(:, v), \ldots, \mathbf{Z}_1^S(:, v) \\ \mathbf{Z}_2^1(:, v), \mathbf{Z}_2^2(:, v), \ldots, \mathbf{Z}_2^S(:, v) \end{bmatrix} \in \mathbb{R}^{K_{12} \times S} \quad (7)$$

which collects the v-th voxels of all maps common across datasets (types **(m1)** and **(m2)**) for $s = 1, \ldots, S$. We impose row-wise group-sparsity on $\bar{\mathcal{Z}}_v$ to encode the prior that they have shared support across datasets, as they share similar activation patterns. Then the formulated structured supervised DL problem is given by

$$\min_{\{\mathbf{D}^s \in \mathcal{D}\}, \{\mathbf{Z}^s\}} \sum_{s=1}^{S} \left\{ \frac{1}{2} \|\mathbf{X}^s - \mathbf{D}^s \mathbf{Z}^s\|_F^2 + \frac{\lambda_1}{2} f\big((\widetilde{\mathbf{D}}^s)^\top\big) + \lambda_2 \|\widetilde{\mathbf{Z}}^s\|_1 \right\} + \sum_{v=1}^{V} \lambda_3 \|\bar{\mathcal{Z}}_v\|_{1,2} \quad (8)$$

where λ_1, λ_2, and λ_3 are pre-specified nonnegative parameters adjusting the strengths of the regularizers. Fig. 1 illustrates the structured DL model.

Step height / Method	0.7	0.9	1.1	1.3	1.5
Supervised DL [2]	66%	68%	70%	73%	77%
Proposed method	**82%**	**82%**	**90%**	**96%**	**96%**

Table 1. Average classification accuracies.

Step height / Method	0.7	0.9	1.1	1.3	1.5
Supervised DL [2]	0.79	0.81	0.82	0.85	0.87
DS-ICA [5]	0.90	0.91	0.93	0.93	0.94
MCCA-jICA [9]	0.90	0.91	0.91	0.91	0.91
Proposed method	**0.92**	**0.92**	**0.95**	**0.95**	**0.95**

Table 2. Average correlation coefficients between the estimated and the true maps.

Problem (8) is not convex w.r.t. $\{\mathbf{D}^s\}$ and $\{\mathbf{Z}^s\}$ jointly, but with one set of variables fixed, the problem is convex w.r.t. the other set of variables. Thus, the block coordinate descent (BCD) method can be applied to derive an efficient algorithm.

4. NUMERICAL EXPERIMENTS

The proposed formulation was tested using synthetic and real data sets. For comparison, the supervised DL formulation in (4) as well as two ICA-based algorithms were tested.

4.1. Test with Synthetic Data Sets

Three synthetic data sets were generated containing four types of maps and two groups of subjects with $K_1 = K_2 = K_3 = K_4 = 3$ and $M_1 = 138, M_2 = 109$. All entries in $\{\mathbf{D}^s\}$ and $\{\mathbf{Z}^s\}$ were initialized with standard Gaussian random variables, where each pair of the common maps had a correlation coefficient of 0.95, and the rest of the random variables were independently distributed. Then, 50% of the entries in the maps were randomly set to zero to simulate sparsity. To simulate the group difference, a step of height h was added to each element of the discriminative dictionaries corresponding to the second group. Table 1 lists the accuracies in predicting the subject groups from the supervised DL method in (4) and the proposed method. Our structured DL method achieves higher accuracies for various values of h. In Table 2, the correlation coefficients between the estimated and the true maps averaged over all maps from different algorithms are reported. It can be seen that our method provides the best map estimation performance.

Fig. 2 depicts the correlation coefficients between every pair of the true and the estimated maps for each data set for the proposed method. Since the maps $1-3$, $4-6$, $7-9$, and $10-12$ are of types (**m1**), (**m2**), (**m3**), and (**m4**), respectively, for both the true and the estimated $\{\mathbf{Z}^s\}$, the figure visualizes the true and the identified map types as well. It is observed that all but two map types have been correctly identified in each data set. The permutation and the sign ambiguities are inherent in bi-factorization methods. Notably, all the common

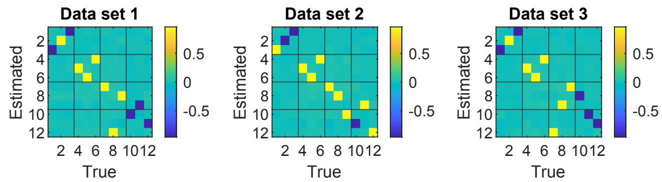

Fig. 2. Correlation coefficients between the true and the estimated maps. Maps 1-3, 4-6, 7-9, and 10-12 are type (**m1**), (**m2**), (**m3**), and (**m4**) maps, respectively.

maps (Maps $1-6$) have been correctly identified thanks to the group sparsity constraint.

4.2. Test with Real Data

The method was also validated on three real fMRI data sets from Mind Research Network Clinical Imaging Consortium Collection [10]. The data sets were collected when subjects were performing the auditory oddball (AOD), Sternberg item recognition paradigm (SIRP), and sensory-motor (SM) tasks, respectively. The data set contains $M_1 = 138$ healthy controls and $M_2 = 109$ schizophrenic subjects. The parameters were tuned based on the classification accuracy, which yielded $\lambda_1 = 0.1, \lambda_2 = 0.005, \lambda_3 = 0.006, K_1 = 8, K_2 = 25, K_3 = 3$, and $K_4 = 3$. Our method obtained an accuracy of 82%, with DS-ICA yielding 75%.

In Fig. 3, selected maps estimated from the proposed method are shown. First, it can be seen that the p-values for non-discriminative maps ((**m1**) and (**m4**)) are indeed larger than 0.05. The (**m2**) maps, which are common and discriminative, are often interpretable. For example, map $k = 20$ suggests a link between the activation in the auditory cortex and schizophrenia [11]. As the AOD and the SM tasks are auditory tasks, their maps associated with the auditory cortex show stronger group differences than the corresponding SIRP map. Similarly, maps 25, and 32 seem to indicate the linkage of the activations in the left and the right central executive networks (CENs) [12], respectively.

Due to the sparsity prior, our DL method extracts more localized activations than ICA-based methods. The last column in Fig. 3 depicts the motor region from DS-ICA, which is split into maps 28 and 29 in our algorithm. Interestingly, map 29 is discriminative in all tasks, while map 28 is discriminative only in the AOD task. Finally, our method could find map 24, which is a meaningful component and is highly significant in AOD and SM, which seems not to be obtained by DS-ICA.

5. CONCLUSION

A DL-based method for analyzing multiset fMRI data incorporating subject group information has been proposed. The group attributes were used to estimate the component neural activation maps that are discriminative across groups as well as those shared among groups in a supervised DL framework. Also, the maps that are common across data sets and unique to individual data sets were extracted employing a

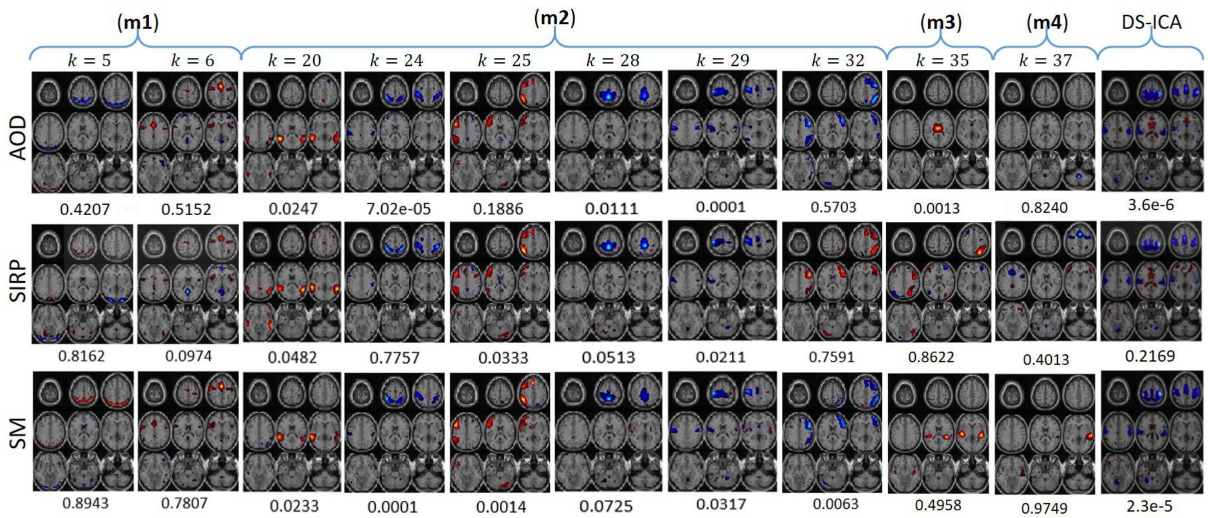

Fig. 3. Selected maps and their p-values from the proposed and the DS-ICA methods. Significant p-values are highlighted. The warm colors in the maps indicate higher activation in the controls and the cold colors the opposite. Some maps with types **(m1)** and **(m2)** have flipped signs, which is attributed to very small magnitudes of the t-statistics associated with the maps.

group sparsity-based structural prior. Numerical tests with synthetic and real fMRI data showed that the method can find four types of maps as designed, with promising performances in terms of group prediction accuracy, map estimation quality, and finding interpretable maps.

6. REFERENCES

[1] T. Adali, M. Anderson, and G. S. Fu, "Diversity in independent component and vector analyses: Identifiability, algorithms, and applications in medical imaging," *IEEE Signal Process. Mag.*, vol. 31, no. 3, pp. 18–33, May 2014.

[2] R. Jin, K. Dontaraju, S.-J. Kim, et al., "Dictionary learning-based fMRI data analysis for capturing common and individual neural activation maps," *IEEE J. Sel. Topics Signal Process.*, vol. 14, no. 6, pp. 12651279, May 2020.

[3] E. F. Lock, K. A. Hoadley, J. S. Marron, et al., "Joint and individual variation explained (JIVE) for integrated analysis of multiple data types," *Ann. Appl. Stat.*, vol. 7, no. 1, pp. 523, 2013.

[4] Y. Levin-Schwartz, V. D. Calhoun, and T. Adalı, "Quantifying the interaction and contribution of multiple datasets in fusion: Application to the detection of schizophrenia," *IEEE Trans. Med. Imaging*, vol. 36, no. 7, pp. 1385–1395, Jul. 2017.

[5] M. Akhonda, B. Gabrielson, S. Bhinge, et al., "Disjoint subspaces for common and distinct component analysis: Application to the fusion of multi-task fMRI data," *J. Neurosci. Methods*, vol. 358, pp. 109214, 2021.

[6] M. Yang, L. Zhang, X. Feng, and D. Zhang, "Fisher discrimination dictionary learning for sparse representation," in *Proc. Int. Conf. Computer Vis.*, Barcelona, Spain, Nov. 2011, p. 543550.

[7] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *J. R. Stat. Soc. B.*, vol. 68, no. 1, pp. 4967, 2006.

[8] S. Bahrampour, N. M. Nasrabadi, A. Ray, and W. K. Jenkins, "Multimodal task-driven dictionary learning for image classification," *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 24–38, Jan. 2016.

[9] J. Sui, T. Adali, G. Pearlson, et al., "A CCA+ICA based model for multi-task brain imaging data fusion and its application to schizophrenia," *Neuroimage*, vol. 51, no. 1, pp. 123–134, 2010.

[10] R. L. Gollub, J. M. Shoemaker, M. D. King, et al., "The MCIC collection: a shared repository of multi-modal, multi-site brain image data from a clinical investigation of schizophrenia," *Neuroinformatics*, vol. 11, pp. 367–388, 2013.

[11] C. Dondé, J. T. Kantrowitz, A. Medalia, et al., "Early auditory processing dysfunction in schizophrenia: Mechanisms and implications," *Neurosci. Biobehav. Rev.*, vol. 148, pp. 105098, 2023.

[12] K. Supekar, W. D. Cai, R. Krishnadas, et al., "Dysregulated brain dynamics in a triple-network saliency model of schizophrenia and its relation to psychosis," *Biol. Psychiatry.*, vol. 85, no. 1, pp. 60–69, 2019.