Correlation-Weighted Sparse Group Representation for Brain Network Construction in MCI Classification

Renping Yu[1,2], Han Zhang[2], Le An[2], Xiaobo Chen[2], Zhihui Wei[1], and Dinggang Shen[2(✉)]

[1] School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China
[2] Department of Radiology and BRIC, UNC at Chapel Hill, Chapel Hill, NC, USA
dgshen@med.unc.edu

Abstract. Analysis of brain functional connectivity network (BFCN) has shown great potential in understanding brain functions and identifying biomarkers for neurological and psychiatric disorders, such as Alzheimer's disease and its early stage, mild cognitive impairment (MCI). In all these applications, the accurate construction of biologically meaningful brain network is critical. Due to the sparse nature of the brain network, sparse learning has been widely used for complex BFCN construction. However, the conventional l_1-norm penalty in the sparse learning equally penalizes each edge (or link) of the brain network, which ignores the link strength and could remove strong links in the brain network. Besides, the conventional sparse regularization often overlooks group structure in the brain network, *i.e.*, a set of links (or connections) sharing similar attribute. To address these issues, we propose to construct BFCN by integrating both *link strength* and *group structure information*. Specifically, a novel correlation-weighted sparse group constraint is devised to account for and balance among (1) sparsity, (2) link strength, and (3) group structure, in a unified framework. The proposed method is applied to MCI classification using the resting-state fMRI from ADNI-2 dataset. Experimental results show that our method is effective in modeling human brain connectomics, as demonstrated by superior MCI classification accuracy of 81.8 %. Moreover, our method is promising for its capability in modeling more biologically meaningful sparse brain networks, which will benefit both basic and clinical neuroscience studies.

1 Introduction

Study of brain functional connectivity network (BFCN), based on resting-state fMRI (rs-fMRI), has shown great potentials in understanding brain functions

R. Yu was supported by the Research Fund for the Doctoral Program of Higher Education of China (RFDP) (No. 20133219110029), the Key Research Foundation of Henan Province (15A520056) and NFSC (No. 61171165, No. 11431015).

S. Ourselin et al. (Eds.): MICCAI 2016, Part I, LNCS 9900, pp. 37–45, 2016.
DOI: 10.1007/978-3-319-46720-7_5

and identifying biomarkers for neurological disorders [1]. Many BFCN modeling approaches have been proposed and most of them represent the brain network as a graph by treating brain regions as nodes, and the connectivity between a pair of region as an edge (or link) [2]. Specifically, the brain can be first parcellated into different regions-of-interest (ROIs) and then the connectivity in a pair of ROIs can be estimated by the correlation between the mean blood-oxygen-level dependent (BOLD) time series of these ROIs.

The most common BFCN modeling approach is based on pairwise Pearson's correlation (PC). However, PC is insufficient to account for the interaction among multiple brain regions [3], since it only captures pairwise relationship. Another common modeling approach is based on sparse representation (SR). For example, the sparse estimation of partial correlation with l_1-regularization can measure the relationship among certain ROIs while factoring out the effects of other ROIs [4]. This technique has been applied to construct brain network in the studies of Alzheimer's disease (AD), mild cognitive impairment (MCI) [3], and autism spectrum disorder [5]. However, human brain inherently contains *not only* sparse connections *but also* group structure [6], with the latter considered more in the recent BFCN modeling methods. A pioneer work in [7] proposed non-overlapping group sparse representation by considering group structures and supporting group selections. The group structure has been utilized in various ways. For example, Varoquaux *et al.* [8] used group sparsity prior to constrain all subjects to share the same network topology. Wee *et al.* [9] used group constrained sparsity to overcome inter-subject variability in the brain network construction. To introduce the sparsity within each group, sparse group representation (SGR) has also been developed by combining l_1-norm and $l_{q,1}$-norm constraints. For example, a recent work [10] defined "group" based on the anatomical connectivity, and then applied SGR to construct BFCN from the whole-brain fMRI signals.

Note that, in all these existing methods, the l_1-norm constraint in both SR and SGR penalizes each edge equally. That is, when learning the sparse representation for a certain ROI, BOLD signals in all other ROIs are treated equally. This process ignores the similarity between BOLD signals of the considered ROI and the other ROIs during the network reconstruction. Actually, if BOLD signals of two ROIs are highly similar, their strong connectivity should be kept or enhanced during the BFCN construction, while the weak connectivity shall be restrained. In light of this, we introduce a *link-strength related penalty* in sparse representation. Moreover, to further make the penalty consistent across all similar links in the whole brain network, we propose a *group structure based constraint* on the similar links, allowing them to share the same penalty during the network construction. In this way, we can *jointly* model the whole brain network, instead of separately modeling a sub-network for each ROI. This is implemented by a novel weighted sparse group regularization that considers *sparsity*, *link strength*, and *group structure* in a unified framework.

To validate the effectiveness of our proposed method in constructing brain functional network, we conduct experiments on a real fMRI dataset for the BFCN construction and also for BFCN-based brain disorder diagnosis. The experimental results in distinguishing MCI subjects from normal controls (NCs) confirm that our proposed method, with simple t-test for feature selection and linear SVM for classification, can achieve superior classification performance compared to the competing methods. The selected feature (*i.e.*, network connections) by our method can be utilized as potential biomarkers in future studies on early intervention of such a progressive and incurable disease.

2 Brain Network Construction and MCI Classification

Suppose that each brain has been parcellated into N ROIs according to a certain brain atlas. The regional mean time series of the i^{th} ROI can be denoted by a column vector $x_i = [x_{1i}; x_{2i}; ...; x_{Ti}] \in \mathbb{R}^T$, where T is the number of time points in the time series, and thus $X = [x_1, ..., x_i, ..., x_N] \in \mathbb{R}^{T \times N}$ denotes the data matrix of a subject. Then the key step of constructing the BFCN for this subject is to estimate the connectivity matrix $W \in \mathbb{R}^{N \times N}$, given the N nodes (*i.e.*, $x_i, i = 1, 2, ..., N$), each of which represents signals in a ROI.

Many studies model the connectivity of brain regions by a sparse network [4]. The optimization of the BFCN construction based on SR can be formulated as

$$\min_W \sum_{i=1}^{N} \frac{1}{2} ||x_i - \sum_{j \neq i} x_j W_{ji}||_2^2 + \lambda \sum_{i=1}^{N} \sum_{j \neq i} |W_{ji}|. \tag{1}$$

The l_1-norm penalty involved in Eq. (1) penalizes each representation coefficient with the same weight. In other words, it treats each ROI equally when reconstructing a target ROI (x_i). As a result, sparse modeling methods based on this formulation tend to reconstruct the target ROI by some ROIs that have very different signals as the target ROI. Furthermore, the reconstruction of each ROI is independent from the reconstructions of other ROIs; thus, the estimated reconstruction coefficients for the similar ROIs could vary a lot, and this could lead to an unstable BFCN construction. Hence, the link strength that indicates signal similarity of two ROIs should be considered in the BFCN construction.

2.1 Correlation-Weighted Sparse Group Representation for BFCN Construction

To take into account the *link strength*, we introduce a correlation-weighted sparse penalty in Eq. (1). Specifically, if BOLD signals of the two ROIs have high similarity, *i.e.*, their link is strong, then this link should be penalized less. On the other hand, weak link should be penalized more with larger weight. To measure the link strength between signals of two ROIs, PC coefficient can be calculated. Then the penalty weight for W_{ji}, *i.e.*, the link between the i^{th} ROI x_i and the j^{th} ROI x_j, can be defined as:

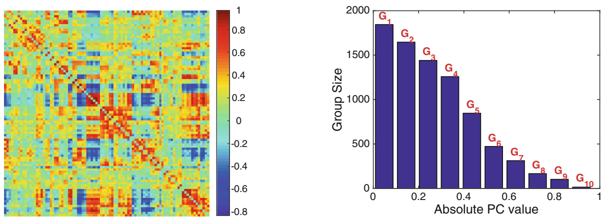
$$C_{ji} = e^{-\frac{P_{ji}^2}{\sigma}}, \tag{2}$$

where P_{ji} is the PC coefficient between the i^{th} ROI x_i and the j^{th} ROI x_j, and σ is a parameter used to adjust the weight decay speed for the *link strength* adaptor. Accordingly, the correlation-weighted sparse representation (WSR) can be formulated as

$$\min_{W} \sum_{i=1}^{N} \frac{1}{2}||x_i - \sum_{j \neq i} x_j W_{ji}||_2^2 + \lambda \sum_{i=1}^{N} \sum_{j \neq i} C_{ji}|W_{ji}|, \tag{3}$$

where $C \in \mathbb{R}^{N \times N}$ is the link strength adaptor matrix with each element C_{ji} being inversely proportional to the similarity (*i.e.*, PC coefficient) between the signals in ROI x_j and the signals in the target ROI x_i.

Note that the reconstruction of x_i, *i.e.*, the i^{th} sub-network construction, is still independent from the reconstructions of sub-networks for other ROIs. In order to further make this *link-strength* related penalty consistent across all links with similar *strength* in the whole network, we propose a group structure constraint on the similar links, allowing them to share the same penalty during the whole BFCN construction. In this way, we can model the whole brain network jointly, instead of separately modeling sub-networks of all ROIs.



(a) PC Matrix (b) Group partition

Fig. 1. Illustration of group partition for a typical subject in our data. (a) Pearson correlation coefficient matrix P. (b) The corresponding group partition (K = 10) of (a).

To identify the group structure, we partition all links, *i.e.*, the pairwise connections among ROIs, into K groups based on the PC coefficients. Specifically, K non-overlapping groups of links are pre-specified by their corresponding PC coefficients. Assuming the numerical range of the absolute value of the PC coefficient $|P_{ij}|$ is $[P_{min}, P_{max}]$ with $P_{min} \geq 0$ and $P_{max} \leq 1$, we partition $[P_{min}, P_{max}]$ into K uniform and non-overlapping partitions with the same interval $\Delta = (P_{max} - P_{min})/K$. The k^{th} group is defined as $G_k = \{(i,j) \mid |P_{ij}| \in [P_{min} + (k-1)\Delta, P_{min} + k\Delta]\}$. Figure 1 shows the grouping results by setting $K = 10$ for illustration purpose. Most link's strength in the network is weak, while the strong connectivity accounts for a small number of links.

To integrate constraints on *link strength*, *group structure*, as well as the sparsity in a unified framework, we propose a novel weighted sparse group regularization formulated as:

$$\min_{W} \sum_{i=1}^{N} \frac{1}{2}||x_i - \sum_{j\neq i} x_j W_{ji}||_2^2 + \lambda_1 \sum_{i=1}^{N} \sum_{j\neq i} C_{ji}|W_{ji}| + \lambda_2 \sum_{k=1}^{K} d_k ||W_{G_k}||_q, \quad (4)$$

where $||W_{G_k}||_q = \sqrt[q]{\sum_{(i,j)\in G_k}(W_{ij})^q}$ is the l_q-norm (with q=2 in this work). $d_k = e^{-\frac{E_k^2}{\sigma}}$ is a pre-defined weight for the k^{th} group and $E_k = \frac{1}{|G_k|}\sum_{(i,j)\in G_k} P_{ij}$. σ is the same parameter as in Eq. (2), set as the mean of all subjects' standard variances of absolute PC coefficients. In Eq. (4) the first regularizer (l_1-norm penalty) controls the overall sparsity of the reconstruction model, and the second regularizer ($l_{q,1}$-norm penalty) contributes the sparsity at the group level.

2.2 MCI Classification

The estimated BFCN are applied to classification of MCI and NC. Note that the learned connectivity matrix W could be asymmetric. Therefore, we simply make a symmetric matrix by $W^* = (W + W^T)/2$, and use W^* to represent the final network that contains $N(N-1)/2$ effective connectivity measures due to symmetry. These connectivity measures are used as the imaging features, with the feature dimensionality of 4005 for the case of $N = 90$. For feature selection, we use two-sample t-test with the significance level of $p < 0.05$ to select features that significantly differentiate between MCI and NC classes. After feature selection, we employ a linear SVM [11] with $c = 1$ for classification.

3 Experiments

The Alzheimers Disease Neuroimaging Initiative (ADNI) dataset is used in this study. Specifically, 50 MCI patients and 49 NCs are selected from the ADNI-2 dataset in our experiments. Subjects from both groups were scanned using 3.0T Philips scanners. SPM8 toolbox (http://www.fil.ion.ucl.ac.uk/spm/) was used to preprocess the rs-fMRI data according to the well-accepted pipeline [6].

3.1 Brain Functional Network Construction

Automated Anatomical Labeling (AAL) template is used to define 90 brain ROIs, and the mean rs-fMRI signals are extracted from each ROI to model BFCN. For comparison, we also construct the brain networks using other methods, including PC, SR, WSR, and SGR (corresponding to $C_{ji} = 1$ in Eq. (4)). The SLEP toolbox [12] is used to solve the sparse related models in this paper.

Figure 2 shows the visualization of the constructed BFCNs of one typical subject based on 5 different methods. As can be seen from Fig. 2(a), the intrinsic grouping in brain connectivity can be indirectly observed, although the PC only measures pairwise ROI interaction. Comparing Fig. 2(b) and (d), we can observe that there are relatively fewer non-zero elements in the SGR-based model due to the use of group sparse regularization. Similarly, the group structure is more obvious in Fig. 2(e) by our WSGR method than that in Fig. 2(c) by WSR.

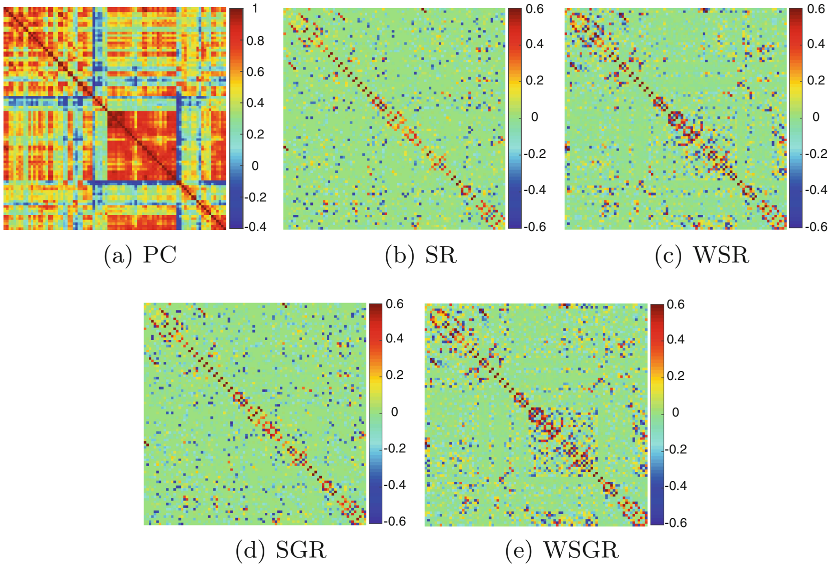


Fig. 2. Comparison of BFCNs of the same subject reconstructed by 5 different methods.

Regarding the effectiveness of using the link-strength related weights, we can see from Fig. 2(c) and (e) that the sparse constraint with the link-strength related weights is more reasonable for modeling the BFCN than their counterparts without weights (shown in Fig. 2(b) and (d)).

3.2 Classification Results

A leave-one-out cross-validation (LOOCV) strategy is adopted in our experiments. To set the values of the regularization parameter (*i.e.*, λ in SR, WSR, and λ_1, λ_2 in SGR, WSGR), we employed a nested LOOCV strategy on the training set. The parameters are grid-searched in the range of $[2^{-5}, 2^{-4}, ..., 2^{0}, ..., 2^{4}, 2^{5}]$.

To evaluate the classification performance, we use seven evaluation measures: accuracy (ACC), sensitivity (SEN), specificity (SPE), area under curve (AUC), Youden's index (YI), F-Score and balanced accuracy (BAC).

As shown in Fig. 3, the proposed WSGR model achieves the best classification performance with an accuracy of 81.8 %, followed by WSR (78.8 %). By comparison, we can verify the effectiveness of the *link strength* related weights from two aspects. First, it can be observed that the WSR model with *link strength* related weights from PC performs much better than both the PC and SR models. Second, the classification result our model outperforms the SGR model (72.73 %). Similarly, by comparing the results by the SR and WSR model with those by the SGR and WSGR models, the effectiveness of our introduced group structure based penalty can be well justified. With the DeLong's non-parametric statistical significance test [13], our proposed method

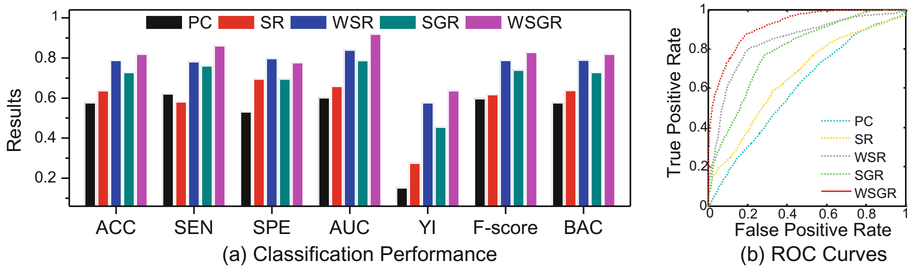


Fig. 3. Comparison of classification results by five methods, using both seven classification performance metrics and ROC curve.

significantly outperforms PC, SR, WSR and SGR under 95 % confidence interval with $p-value = 1.7 \times 10^{-7}, 3.6 \times 10^{-6}$, 0.048 and 0.0017, respectively. The superior performance of our method suggests the weighted group sparsity is beneficial in constructing brain networks and also able to improve the classification performance.

As the selected features by t-test in each validation might be different, we record all selected features during the training process. The 76 most frequently selected features are visualized in Fig. 4, where the thickness of an arc indicating the discriminative power of an edge, which is inversely proportional to the estimated p-values. The colors of arcs are randomly generated to differentiate ROIs

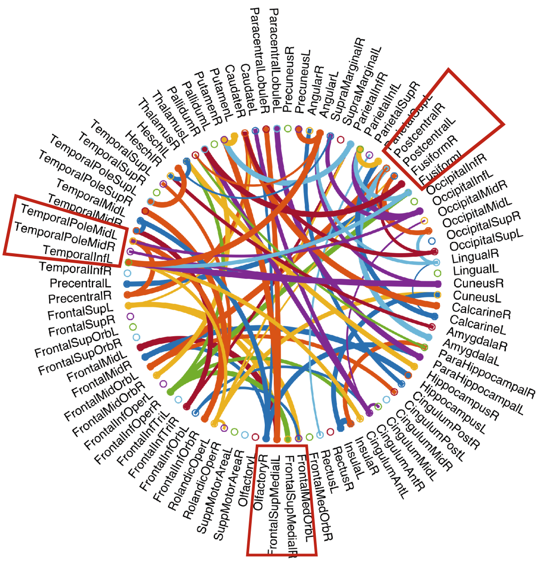


Fig. 4. The most frequently selected connections for the 90 ROIs of AAL template. The thickness of an arc indicates the discriminative power of an edge for MCI classification.

and connectivity for clear visualization. We can see that several brain regions (as highlighted in the figure) are jointly selected as important features for MCI classification. For example, a set of brain regions in the temporal pole, olfactory areas and medial orbitofrontal cortex, as well as bilateral fusiform, are found to have dense connections which are pivotal to MCI classification [14].

4 Conclusion

In this paper, we have proposed a novel weighted sparse group representation method for brain network modeling, which integrates *link strength*, *group structure*, as well as *sparsity* in a unified framework. In this way, the complex brain network can be more accurately modeled as compared to other commonly used methods. Our proposed method was validated in the task of MCI and NC classification, and superior results were obtained compared to the classification performance of other brain network modeling approaches. In future work, we plan to work on more effective grouping strategy in order to further improve the modeling accuracy and the MCI diagnosis performance.

References

1. Fornito, A., Zalesky, A., Breakspear, M.: The connectomics of brain disorders. Nat. Rev. Neurosci. **16**, 159–172 (2015)
2. Smith, S.M., Miller, K.L., et al.: Network modelling methods for FMRI. NeuroImage **54**, 875–891 (2011)
3. Huang, S., Li, J., Sun, L., Ye, J., Fleisher, A., Wu, T.: Alzheimer's Disease NeuroImaging Initiative: learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation. NeuroImage **50**, 935–949 (2010)
4. Meinshausen, N., Bühlmann, P.: High-dimensional graphs and variable selection with the lasso. Ann. Stat., 1436–1462 (2006)
5. Lee, H., Lee, D.S., et al.: Sparse brain network recovery under compressed sensing. IEEE Trans. Med. Imaging **30**, 1154–1165 (2011)
6. Rubinov, M., Sporns, O.: Complex network measures of brain connectivity: uses and interpretations. NeuroImage **52**, 1059–1069 (2010)
7. Yuan, M., Lin, Y.: Model selection and estimation in regression with grouped variables. J. R. Stat. Soc. Series. B. Stat. Methodol **68**, 49–67 (2006)
8. Varoquaux, G., Gramfort, A., Poline, J.B., Thirion, B.: Brain covariance selection: better individual functional connectivity models using population prior. In: Advances in Neural Information Processing Systems, pp. 2334–2342 (2010)
9. Wee, C.Y., et al.: Group-constrained sparse fMRI connectivity modeling for mild cognitive impairment identification. Brain Struct. Funct. **219**, 641–656 (2014)
10. Jiang, X., Zhang, T., Zhao, Q., Lu, J., Guo, L., Liu, T.: Fiber connection pattern-guided structured sparse representation of whole-brain fMRI signals for functional network inference. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A. (eds.) MICCAI 2015. LNCS, vol. 9349, pp. 133–141. Springer, Heidelberg (2015)
11. Chang, C.C., Lin, C.J.: LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. **2**, 27 (2011)

12. Liu, J., Ji, S., Ye, J.: SLEP: sparse learning with efficient projections. Arizona State Univ. **6**, 491 (2009)
13. DeLong, E.R., DeLong, D.M., Clarke-Pearson, D.L.: Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics, 837–845 (1988)
14. Albert, M.S., DeKosky, S.T., Dickson, D., et al.: The diagnosis of mild cognitive impairment due to Alzheimers disease: recommendations from the National Institute on Aging-Alzheimers Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's Dement. **7**, 270–279 (2011)