HHS Public Access
Author manuscript
Asian Conf Pattern Recognit. Author manuscript; available in PMC 2019 January 07.

Published in final edited form as:
Asian Conf Pattern Recognit. 2018 November ; 2017: 917–922. doi:10.1109/ACPR.2017.147.

Learning Pairwise-Similarity Guided Sparse Functional Connectivity Network for MCI Classification

Xiaobo Chen[1,2], **Han Zhang**[1], **Yu Zhang**[1], **Jian Yang**[3], and **Dinggang Shen**[1]

[1]Department of Radiology and BRIC, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[2]Automotive Engineering Research Institute, Jiangsu University, Zhenjiang, China

[3]School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing, China

Abstract

Learning functional connectivity (FC) network from resting-state function magnetic resonance imaging (RS-fMRI) data via sparse representation (SR) or weighted SR (WSR) has been proved to be promising for the diagnosis of Alzheimer's disease and its prodromal stage, mild cognitive impairment (MCI). However, traditional SR/WSR based approaches learn the representation of each brain region independently, without fully taking into account the possible relationship between brain regions. To remedy this limitation, we propose a novel FC modeling approach by considering two types of possible relationship between different brain regions which are incorporated into SR/WSR approaches in the form of regularization. In this way, the representations of all brain regions can be jointly learned. Furthermore, an efficient alternating optimization algorithm is also developed to solve the resulting model. Experimental results show that our proposed method not only outperforms SR and WSR in the diagnosis of MCI subjects, but also leads to the brain FC network with better modularity structure.

Keywords

functional connectivity; sparse representation; resting-state fMRI; mild cognitive impairment

I. INTRODUCTION

Alzheimer's disease (AD) is the most prevalent form of dementia in elderly people over 70 years of age. It accounts for about 50% to 80% of age-related dementia [1]. Mild cognitive impairment (MCI), as an intermediate stage of brain cognitive decline between AD and normal aging, shows mild symptoms of cognitive impairment. MCI does not significantly impact daily functioning, but it is associated with increased risk of developing to AD. For example, individuals with MCI can progress to clinical AD at an annual conversion rate of 10–15%, compared with normal controls (NC) who develop to AD at much lower annual conversion rate of approximately 1–2% [2]. Therefore, early diagnosis of MCI is important for reducing the risk of developing AD by appropriate pharmacological treatments and behavioral interventions. However, diagnosing MCI from the undergoing normal aging is difficult, because of its mild symptoms of cognitive impairment.

Resting-state functional magnetic resonance imaging (RS-fMRI) [3] can measure the blood-oxygenation-level-dependent (BOLD) signals of the brain when a subject is not performing any explicit task. RS-fMRI technique provides an efficient and noninvasive way for the investigation of neurological disorders at a whole-brain level. It holds great promise and has been successfully applied to the early diagnosis of AD/MCI before development of clinical symptoms.

Learning brain functional connectivity (FC) network [4] from BOLD signals is a prevalent approach to infer brain functional organization. FC measures the temporal correlation of BOLD signals in structurally separated brain regions [5]. In most studies, FC network is characterized by graph, in which a node corresponds to a brain region and an edge connects a pair of brain regions. Previous studies have suggested altered FC network in MCI subjects [6] and such alterations occur many years before clinically detectable with cognitive tests [7, 8]. Pearson's correlation is probably the most widely used approach to capture the pairwise correlation between brain regions. This method is intuitive and computationally efficient because of its simplicity. However, it ignores the impact of other brain regions, thus insufficient to provide adequate and complete information on the interactions among multiple brain regions [9–11]. Moreover, the resulted FC network is often quite dense, containing possible spurious or insignificant connections.

To overcome the drawbacks of Pearson's correlation, some other FC modeling methods have been developed, such as partial correlation, high-order correlation [10, 12, 13, 28, 29], and another widely-accepted method such as sparse representation (SR). It is generally believed that a brain region interacts directly with only a few other brain regions [14]. Therefore, learning FC network from BOLD signals via SR has attracted increasing attention [15, 9, 14, 16]. In particular, l_1-norm regularization based SR [15, 16] is a typical approach to construct sparse FC network. Based on the BOLD signals of each brain region, SR represents one brain region using as few other brain regions as possible via linear combination. However, other brain regions are typically treated equally, without considering their strength of correlation with the target brain region. Intuitively, a brain region that highly correlates with the target brain region tends to have larger SR weight. To incorporate such information, weighted SR (WSR) was proposed [15] where the FC strengths (based on Pearson's correlation) between the target brain region and other brain regions were used as prior weights in l_1-norm regularization. Empirical results showed that the WSR achieved superior performance in comparison with the original SR. Nevertheless, we found that the representations of the brain regions in either SR or WSR are actually *independent* of each other, indicating that some important prior information (such as the inherent relationship among different SRs for different brain regions) is still ignored.

To this end, we propose a novel "pairwise-similarity guided sparse representation (PSGSR)" method for learning brain FC network from RS-fMRI data. The motivation comes from the following hypotheses. First, for the two brain regions with similar pattern of BOLD fluctuations, their representations over the other brain regions should be also similar. Second, the representations of any other brain region over the two brain regions with similar BOLD fluctuations should be also similar. Based on these intuitions, we propose two novel regularization terms and incorporate them into the framework of WSR. To solve the PSGSR model, we develop an efficient iterative algorithm by combing bound optimization and alternating optimization techniques. After constructing the brain FC network for each subject, a machine learning framework consisting of feature extraction, selection, and pattern classification is used to distinguish MCI from NC subjects.

II. METHODOLOGY

A. Sparse and weighted sparse representation

Suppose the whole brain is parcellated into N brain regions (or regions-of-interest, ROIs) based on a specific brain atlas. Let the regional mean BOLD signal of the ROI i be denoted by a column vector $x_i \in R^M$, where M is the total number of time points. In the SR-based approach, x_i is supposed to be approximated by a linear combination of BOLD signals of other ROIs. Mathematically, we have

$$x_i = \sum_{j=1, j \neq i}^{N} w_{ji} x_j + \varepsilon_i \quad (1)$$

where the weight w_{ji} reflects the contribution of x_j in representing x_i, and ε_i is the approximation error. The weight w_{ji} can be calculated by

$$\min_{w_{\cdot i}} \frac{1}{2} \left\| x_i - \sum_{j=1, j \neq i}^{N} w_{ji} x_j \right\|^2 + \lambda \| w_{\cdot i} \|_1 \quad (2)$$

where $w_{\cdot i}$ denotes the i-th column, $\|w_{\cdot i}\|_1 = \sum_{j=1, j \neq i}^{N} |w_{ji}|$ is the l1-norm, λ is a balance parameter used to control the degree of sparsity. In (2), each x_j is treated equally in the representation of x_i. On the other hand, the weighted SR (WSR) [15] is proposed recently as given below:

$$\min_{w_{\cdot i}} \frac{1}{2} \left\| x_i - \sum_{j=1, j \neq i}^{N} w_{ji} x_j \right\|^2 + \lambda \| c_{\cdot i} \odot w_{\cdot i} \|_1 \quad (3)$$

where $\odot$ denotes element-wise multiplication, $c_{ji} = \exp\left(-P_{ji}^2/\sigma\right)$, P_{ji} is the Pearson's correlation between ROIs j and i, σ is a positive parameter used to control the contribution of the different values of P_{ji}. From (3), we can observe that, if ROI j is highly correlated with ROI i, c_{ji} will have a small value such that the resulting w_{ji} tends to be large. For other ROIs, we also have similar formulation as (3). Therefore, the whole-brain weighted sparse FC network can be obtained by solving the following optimization problem:

$$\min_{W} \frac{1}{2} \| X - XW \|^2 + \lambda \| C \odot W \|_1 \quad s.t.\ diag(W) = 0 \quad (4)$$

where $X = [x_1, x_2, \cdots, x_N] \in R^{M \times N}$, $C = [c_{\cdot 1}, c\&_{\text{middot};2}, \cdots, c_{\cdot N}] \in R^{N \times N}$, $W = [w_{\cdot 1}, w_{\cdot 2}, \cdots, w_{\cdot N}] \in R^{N \times N}$, $\|W\|_1 = \sum_{i=1, j=1}^{i=N, j=N} |w_{ij}|$ denotes the l1-norm of W, and $diag(W)$ indicates the diagonal elements of W.

B. Pairwise-similarity guided sparse representation

In both SR (2) and WSR (3), the inherent relationship between the representations of different ROIs is largely ignored. To address this limitation, we propose to learn the representations of all ROIs *jointly* by introducing an ROI relationship induced penalty into the WSR model. An illustration of our method is shown in Figure 1.

First, if BOLD signals x_i and x_j are similar, their corresponding representation w_i and w_j are expected to be similar as well. By defining $d_{ij} = \| x_i - x_j \|$ as the distance between x_i and x_j, and also weight $u_{ij} = exp\left(-\frac{d_{ij}^2}{2\sigma^2}\right)$ as the similarity between x_i and x_j, we propose the following constraint:

$$\frac{1}{2}\sum_{i,j} u_{ij} \|w_{\cdot i} - w_{\cdot j}\|^2 = tr(WLW^T) \quad (5)$$

where $U = [u_{ij}]$, and $L = D - U$ is called Laplacian matrix in spectral graph theory [17] (D is a diagonal matrix whose entries are the row sums of U). As a result, if the BOLD signals associated with ROIs i and j have similar fluctuation patterns, we will obtain a larger weight u_{ij}, which further makes their representations $w_{\cdot i}$ and $w_{\cdot j}$ more similar.

From another view angle, suppose that the BOLD signals x_k associated with ROI k is linearly represented by the BOLD signals associated with other ROIs (including x_i and x_j). If x_i and x_j have similar fluctuation patterns, we expect that the resulting representation coefficients w_{ik} and w_{jk} also likely to be similar. Therefore, from a symmetrical perspective, as (5), we also define the following constraint:

$$\frac{1}{2}\sum_{i,j} u_{ij} \|w_{i\cdot} - w_{j\cdot}\|^2 = tr(W^T LW) \quad (6)$$

where w_i denotes the i-th row of W.

By integrating (4), (5) and (6) together, we finally obtain the following pairwise-similarity guided sparse representation (PSGSR) model for FC network construction:

$$\min_W \frac{1}{2}\|X - XW\|^2 + \lambda_1 \|C \odot W\|_1 + \frac{\lambda_2}{2} tr(WLW^T + W^T LW) \quad \text{s.t.}\ diag(W) = 0 \quad (7)$$

where λ_1, λ_2 are the parameters that control the importance of the respective constraint terms. Obviously, PSGSR degenerates to WSR when $\lambda_2 = 0$. After calculating W for each subject using (7), we follow [15] and take the average of W and its transposed one, e.g., $W \leftarrow \frac{1}{2}(W + W^T)$, in order to obtain a symmetric FC network.

C. Optimization algorithm

In order to solve the problem (7), we further develop an effective iterative algorithm by combing two optimization strategies, i.e., bound and alternating optimizations [18].

Let us define

$$f(W) = \frac{1}{2}\|X - XW\|^2 + \lambda_1\|C \odot W\|_1 + \frac{\lambda_2}{2}tr(WLW^T + W^TLW) \quad (8)$$

First, for any variable a, we have

$$|\ a\ | \le \frac{1}{2}\left(\frac{a^2}{\left|a^{(0)}\right|} + \left|a^{(0)}\right|\right) \quad (9)$$

where $a^{(0)}$ is an arbitrary nonzero value. The equality holds when $a = a^{(0)}$. Then, we have

$$\|C \odot W\|_1 = \sum_{ij} C_{ij}\left|w_{ij}\right| \le \frac{1}{2}\sum_{ij} C_{ij}\left(\frac{w_{ij}^2}{\left|w_{ij}^{(t)}\right|} + \left|w_{ij}^{(t)}\right|\right) = \frac{1}{2}\sum_j w_{\cdot j}^T D_j^{(t)} w_{\cdot j} + Const \quad (10)$$

where $\sum_i C_{ij}\frac{w_{ij}^2}{\left|w_{ij}^{(t)}\right|} = w_{\cdot j}^T D_j^{(t)} w_{\cdot j}$, $Const$ denotes a constant, and $D_j^{(t)}$ is a diagonal matrix with element $\frac{C_{ij}}{\left|w_{ij}^{(t)}\right|}$ as follows

$$D_j^{(t)} = \begin{bmatrix} \frac{C_{1j}}{\left|w_{1j}^{(t)}\right|} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{C_{Nj}}{\left|w_{Nj}^{(t)}\right|} \end{bmatrix} \in R^{N \times N} \quad (11)$$

Then, let us define

$$f(W, W^{(t)}) = \frac{1}{2}\|X - XW\|^2 + \frac{\lambda_1}{2}\sum_j w_{\cdot j}^T D_j^{(t)} w_{\cdot j} + \frac{\lambda_2}{2} tr(WLW^T + W^T LW) + Const \quad (12)$$

Thus, we have

$$\begin{aligned} f(W) &\le f(W, W^{(t)}), \\ f(W^{(t)}) &= f(W^{(t)}, W^{(t)}) \end{aligned} \quad (13)$$

It means that $f(W, W^{(t)})$ is a tight upper bound of original objective function (8). Therefore, following the principle of bound optimization, we can optimize $f(W, W^{(t)})$, instead of $f(W)$, so as to get the next approximate solution. Specifically, we have the following problem:

$$\min_W \frac{1}{2}\|X - XW\|^2 + \frac{\lambda_1}{2}\sum_j w_{\cdot j}^T D_j^{(t)} w_{\cdot j} + \frac{\lambda_2}{2} tr(WLW^T + W^T LW) \quad \text{s.t.}\ diag(W) = 0 \quad (14)$$

To solve (14), we further develop an effective alternating optimization algorithm where the column of W is alternatively optimized while fixing the other. For example, in order to optimize $w_{\cdot k}$, we can rewrite (14) as

$$\min_{w_{\cdot k}} \frac{1}{2}\|x_k - Xw_{\cdot k}\|^2 + \frac{\lambda_1}{2} w_{\cdot k}^T D_k^{(t)} w_{\cdot k} + \frac{\lambda_2}{4}\sum_{j, j \ne k} u_{kj}\|w_{\cdot k} - w_{\cdot j}\|^2 + \frac{\lambda_2}{2} w_{\cdot k}^T L w_{\cdot k} \quad \text{s.t.}\ w_{kk} = 0 \quad (15)$$

where the terms irrelevant to $w_{\cdot k}$ have been eliminated. In order to handle the constraint $w_{kk} = 0$, we remove the k-th element from $w_{\cdot k}$ and $w_{\cdot j}$, and denote the resulting vectors as $\overline{w_{\cdot k}}$ and $\overline{w_{\cdot j}}$. Similarly, let X_{-k}, $\overline{D_k^{(t)}}$, L_{-k} denote the original matrices with the k-th row and column removed. Then, (15) can be simplified as the following unconstrained problem

$$\min_{w_{\cdot k}} \frac{1}{2}\|x_k - X_{-k}\overline{w_{\cdot k}}\|^2 + \frac{\lambda_1}{2}\overline{w_{\cdot k}}^T \overline{D_k^{(t)}}\,\overline{w_{\cdot k}} + \frac{\lambda_2}{4}\sum_{j, j \ne k} u_{kj}\|\overline{w_{\cdot k}} - \overline{w_{\cdot j}}\|^2 + \frac{\lambda_2}{2}\overline{w_{\cdot k}}^T L_{-k}\overline{w_{\cdot k}} \quad (16)$$

To find the optimal solution for (16), we set the derivative *with respect to* $\overline{w_{\cdot k}}$ equal to zero, which leads to the closed-form solution as follows:

$$\overline{w_{\cdot k}} = \left(X_{-k}{}^{T}X_{-k} + \lambda_1 \overline{D_k^{(t)}} + \frac{\lambda_2}{2} I \sum_{j, j \neq k} u_{kj} + \lambda_2 L_{-k}\right)^{-1} \left(X_{-k}{}^{T} x_k + \frac{\lambda_2}{2} \sum_{j, j \neq k} u_{kj} \overline{w_{\cdot j}}\right)$$

(17)

where I is an identity matrix with appropriate dimension.

D. Feature extraction, selection and classification

After FC network has been constructed for each subject, local clustering coefficients for weighted graph (WLCC) [19] are used as features of all brain regions in the network. WLCC is computed for each node in a weighted graph to quantify the probability that the neighbors of this node are also connected to each other. Thus, each subject with N (N = 90 in this work) ROIs is represented by a feature vector with N elements. Many features might be redundant or irrelevant to classification. In order to alleviate overfitting and construct effective classification model with better generalization, reducing the number of features is a crucial issue. In this study, sparse regression LASSO [20] is used to select a small subset of features which are important for classification. After selecting a few number of features for each subject, a support vector machine (SVM) model with linear kernel [21, 22] is constructed for classification.

III. EXPERIMENTS AND RESULTS

Next, we evaluate the effectiveness of the proposed method in the application of MCI classification. Some closely related FC network construction methods, such as Pearson's correlation (PEC), SR, and WSR are also included for comparison. The proposed learning framework was implemented on MATLAB 2012b environment. Leave-one-out cross-validation (LOOCV) is applied to evaluate the generalization of different methods because of its popularity. Another nested LOOCV on the training samples is applied to select optimal hyperparameters in each method.

A. Data

The publicly available neuroimaging data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database are used. In this work, 97 age- and gender-matched subjects (49 MCI and 48 NC) are selected from the database. These subjects were scanned using 3.0T Philips scanner. To preprocess the acquired fMRI data, SPM8 software (http://www.fil.ion.ucl.uk/spm/software/spm8) was utilized. In particular, for magnetization equilibrium, the first three RS-fMRI volumes of each subject were discarded before preprocessing, leaving 137 volumes. To correct head motion, rigid body transformation was used where the subjects with head motion larger than 2 mm or 2 degrees were discarded. The fMRI images were normalized to the Montreal Neurological Institute (MNI) template.

The images were spatially smoothed using an isotropic 8-mm full-width-at-half-maximum (FWHM) Gaussian kernel. We did not conduct scrubbing to the data with frame-wise displacement larger than 0.5, as it would introduce additional artifacts. The subjects who had less than 4.5 min RS-fMRI data were excluded from further analysis. The images were parcellated into 90 ROIs, defined based on the automated anatomical labeling (AAL) atlas. Mean RS-fMRI time series of each ROI was calculated and band-pass filtered ($0.015 \leq f \leq 0.15$Hz). Head motion parameters (Friston24 model), together with the mean BOLD signals of white matter and cerebrospinal fluid, were regressed out from the regional mean BOLD signals.

B. Algorithm convergence

The optimization algorithm proposed to solve (7) is essentially iterative. Thus, we investigate how it converges. Figure 2 shows the convergence curve for the two randomly selected subjects. The y-axis denotes the value of the loss function, and x-axis denotes the iteration number. We can observe that the proposed method converges very fast. In most cases, the algorithm converges in less than 30 iterations.

C. Classification performance

To comprehensively measure the performance of the different methods, we use the following indices: accuracy (ACC), area under ROC curve (AUC), sensitivity (SEN), and specificity (SPE). As we can see from Table 1, SR achieves better performance than PEC. This may be because SR is able to measure complex interaction relationship among multiple brain regions and at the same time removes many spurious and redundant connections. Furthermore, we can see that WSR outperforms the original SR because the former focuses more on the brain regions with strong Pearson's correlation while the latter equally treats all brain regions during SR. After further introducing prior knowledge about brain regions, the proposed PSGSR further improved the performance of MCI classification. The result confirms that learning the representations of all brain regions *jointly* can effectively make use of the inherent high-level relationship among the brain regions.

D. FC networks

Figure 3 illustrates the FC networks constructed by PEC, SR, WSR, and the proposed PSGSR for a randomly selected subject. As can be seen, PEC leads to a dense network which contains too many potentially spurious or insignificant connections. In contrast, the other three methods all can generate sparse networks. Nevertheless, the result of SR is difficult to interpret because it lacks notable structure. WSR is slightly better than SR because WSR incorporates Pearson's correlation as prior. Among all the methods, PSGSR can generate a FC network with both sparsity and better modularity structure.

E. Discriminative ROIs

Note that the proposed approach is evaluated by nested LOOCV. In different folds of LOOCV, we have different training and testing subjects and the hyperparameters optimized could be different as well. In this study, LASSO-based feature selection was used to select a small number of features that are crucial for diagnosis. Only the features with sufficient discriminative ability were frequently selected. Therefore, those frequently selected features during the LOOCV are regarded as important for MCI diagnosis. Following this criterion, the most discriminative ROIs are shown in Figure 4.

In summary, these ROIs are generally believed to be related to AD. For example, parahippocampal gyrus is an important biomarker in distinguishing AD patients from NC because the volume of the parahippocampal gyrus is significantly reduced in AD [23]. Sofie *et al.* [24] proposed that FC of the occipital cortex is likely to be affected in early-onset AD patients, which indicates an increased vulnerability of this brain region. Latha *et al.* [25] found that gray matter volume in the right supramarginal gyrus was reduced in AD patients, thus leading to poor memory performance and visual recognition deficits. In [26], the authors found that the volume of thalamus was significantly reduced in the patients diagnosed with AD and this degenerative process may contribute to cognitive decline in AD. Some studies [27] found significant brain volume reduction in the caudate in AD patients, compared to NC subjects.

IV. CONCLUSION

In this paper, we have presented a novel sparse representation based FC modeling approach for MCI diagnosis. This work is motivated by the fact that previous SR/WSR-based approaches represent each brain region *independently* without considering their inherent interactions. Therefore, we propose that the brain regions with similar BOLD signals should have *similar representations* and play a mutually similar role in the representations of other brain regions. Based on this assumption, we step forward and propose a pairwise-similarity guided sparse representation for brain FC network modeling, and develop a machine learning framework for MCI diagnosis. The results show that our approach *not only* enhances the modularity structure of the brain FC network, *but also* leads to better diagnosis performance.

Acknowledgments

This work was supported in part by NIH grants (EB006733, EB008374, AG041721, AG049371, AG042599, AG053867, EB022880), the National Natural Science Foundation of China (Grant No. 61773184, 61203244, 61772254), Fujian Provincial Key Laboratory of Information Processing and Intelligent Control (Minjiang University) (MJUKF201724), and the Talent Foundation of Jiangsu University, China (No. 14JDG066).

References

1. A. s. Association. 2012 Alzheimer's disease facts and figures. Alzheimer's & Dementia. 8:131–168.2012;
2. Petersen RC, Doody R, Kurz A, Mohs RC, Morris JC, Rabins PV, et al. Current concepts in mild cognitive impairment. Archives of neurology. 58:1985–1992.2001; [PubMed: 11735772]
3. Smith SM, Miller KL, Salimi-Khorshidi G, Webster M, Beckmann CF, Nichols TE, et al. Network modelling methods for FMRI. Neuroimage. 54:875–891.2011; [PubMed: 20817103]
4. Sporns O. The human connectome: a complex network. Annals of the New York Academy of Sciences. 1224:109–125.2011; [PubMed: 21251014]
5. Friston K, Frith C, Liddle P, Frackowiak R. Functional connectivity: the principal-component analysis of large (PET) data sets. Journal of cerebral blood flow and metabolism. 13:5–5.1993; [PubMed: 8417010]

6. Rombouts SA, Barkhof F, Goekoop R, Stam CJ, Scheltens P. Altered resting state networks in mild cognitive impairment and mild Alzheimer's disease: an fMRI study. Human brain mapping. 26:231–239.2005; [PubMed: 15954139]
7. Johnson S, Schmitz T, Moritz C, Meyerand M, Rowley H, Alexander A, et al. Activation of brain regions vulnerable to Alzheimer's disease: the effect of mild cognitive impairment. Neurobiology of aging. 27:1604–1612.2006; [PubMed: 16226349]
8. Thompson P, Apostolova L. Computational anatomical methods as applied to ageing and dementia. The British journal of radiology. 80:S78–S91.2007; [PubMed: 18445748]
9. Huang S, Li J, Sun L, Ye J, Fleisher A, Wu T, et al. Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation. NeuroImage. 50:935–949.2010; [PubMed: 20079441]
10. Chen X, Zhang H, Gao Y, Wee CY, Li G, Shen D. High - order resting - state functional connectivity network for MCI classification. Human brain mapping. 37:3282–96.2016; [PubMed: 27144538]
11. Nie D, Wang L, Gao Y, Shen D. Fully convolutional networks for multi-modality isointense infant brain image segmentation. ISBI. :1342–1345.2016
12. Zhang H, Chen X, Shi F, Li G, Kim M, Giannakopoulos P, Haller S, Shen D. Topographic Information based High-Order Functional Connectivity and its Application in Abnormality Detection for Mild Cognitive Impairment. Journal of Alzheimer's Disease. 54(3):1095–1112.2016;
13. Chen, X; Zhang, H; Shen, D. Ensemble Hierarchical High-Order Functional Connectivity Networks for MCI Classification. International Conference on Medical Image Computing and Computer-Assisted Intervention; 2016; 18–25.
14. Wee, C-Y; Yap, P-T; Zhang, D; Wang, L; Shen, D. Constrained sparse functional connectivity networks for MCI classification. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012; Springer; 2012. 212–219.
15. Yu R, Zhang H, An L, Chen X, Wei Z, Shen D. Connectivity strength - weighted sparse group representation - based brain network construction for MCI classification. Human Brain Mapping. 38:2370–2383.2017; [PubMed: 28150897]
16. Wee C-Y, Yang S, Yap P-T, Shen D. A. s. D. N. Initiative. Sparse temporally dynamic resting-state functional connectivity networks for early MCI identification. Brain Imaging and Behavior. :1–15.2015 [PubMed: 25724689]
17. Chung, FR. Spectral graph theory. American Mathematical Soc; 1997.
18. Chen X, Yang J, Liang J. Optimal Locality Regularized Least Squares Support Vector Machine via Alternating Optimization. Neural processing letters. 33:301–315.2011;
19. Rubinov M, Sporns O. Complex network measures of brain connectivity: uses and interpretations. Neuroimage. 52:1059–1069.2010; [PubMed: 19819337]
20. Tibshirani R. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological). :267–288.1996
21. Cortes C, Vapnik V. Support-vector networks. Machine learning. 20:273–297.1995;
22. Chen X, Yang J, Ye Q, Liang J. Recursive projection twin support vector machine via within-class variance minimization. Pattern Recognition. 44:2643–2655.2011;
23. Echávarri C, Aalten P, Uylings H, Jacobs H, Visser P, Gronenschild E, et al. Atrophy in the parahippocampal gyrus as an early biomarker of Alzheimer's disease. Brain Structure and Function. 215:265–271.2011; [PubMed: 20957494]
24. Adriaanse SM, Binnewijzend MA, Ossenkoppele R, Tijms BM, van der Flier WM, Koene T, et al. Widespread disruption of functional brain organization in early-onset Alzheimer's disease. PloS one. 9:e102995.2014; [PubMed: 25080229]
25. Velayudhan L, Dury R, Susan F, Penny G, Kempton M, Sagnik B. REDUCED SUPRAMARGINAL GYRUS GRAY MATTER VOLUME ASSOCIATED WITH COGNITIVE IMPAIRMENT IN ALZHEIMER'S DISEASE: A 7-TESLA MRI STUDY. Alzheimer's & Dementia: The Journal of the Alzheimer's Association. 12:P943–P944.2016;
26. De Jong L, Van der Hiele K, Veer I, Houwing J, Westendorp R, Bollen E, et al. Strongly reduced volumes of putamen and thalamus in Alzheimer's disease: an MRI study. Brain. 131:3277–3285.2008; [PubMed: 19022861]

27. Jiji S, Smitha KA, Gupta AK, Pillai VPM, Jayasree RS. Segmentation and volumetric analysis of the caudate nucleus in Alzheimer's disease. European journal of radiology. 82:1525–1530.2013; [PubMed: 23664648]

28. Zhang H, Chen X, Zhang Y, Shen D. Test-retest reliability of "high-order" functional connectivity in young healthy adults. Frontiers in Neuroscience. Aug 2.2017 doi: 10.3389/fnins.2017.00439

29. Zhang Y, Zhang H, Chen X, Lee S-W, Shen D. Hybrid High-order Functional Connectivity Networks Using Resting-state Functional MRI for Mild Cognitive Impairment Diagnosis. Scientific Reports. 2017; doi: 10.1038/s41598-017-06509-0

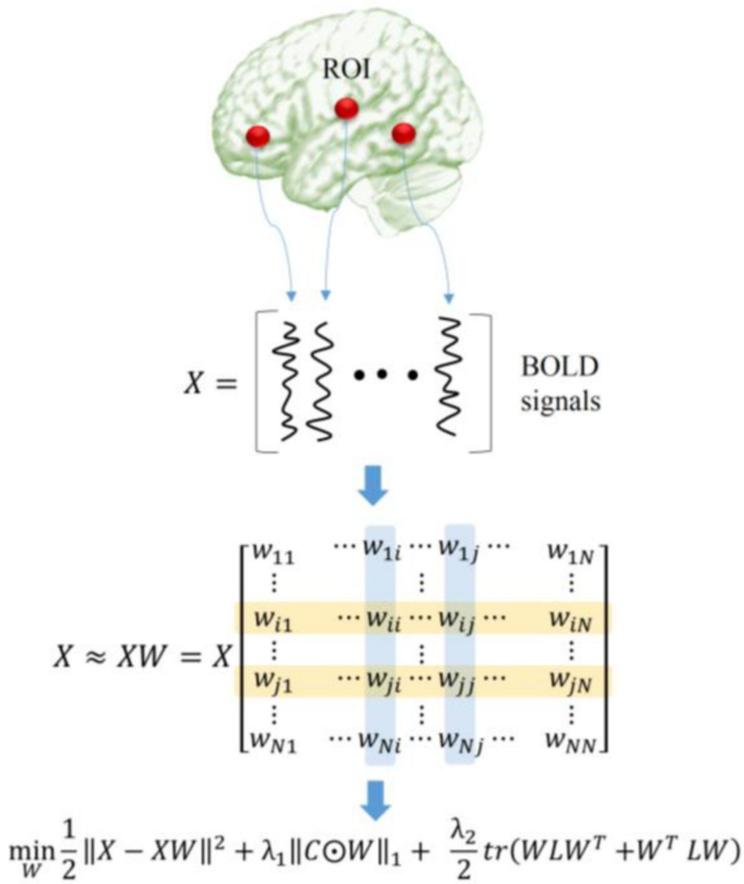

Figure 1.
Illustration of PSGSR.

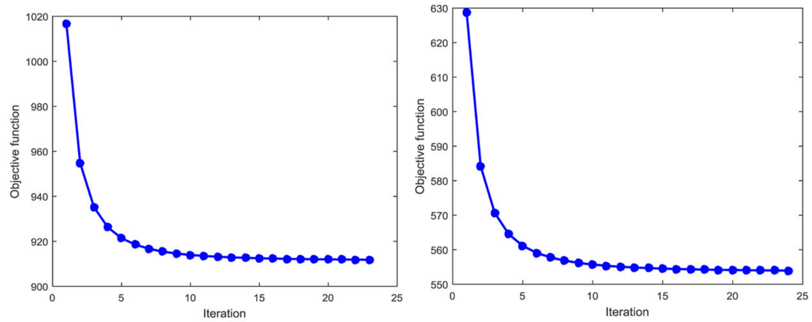

Figure 2.
Illustration of convergence of algorithm.

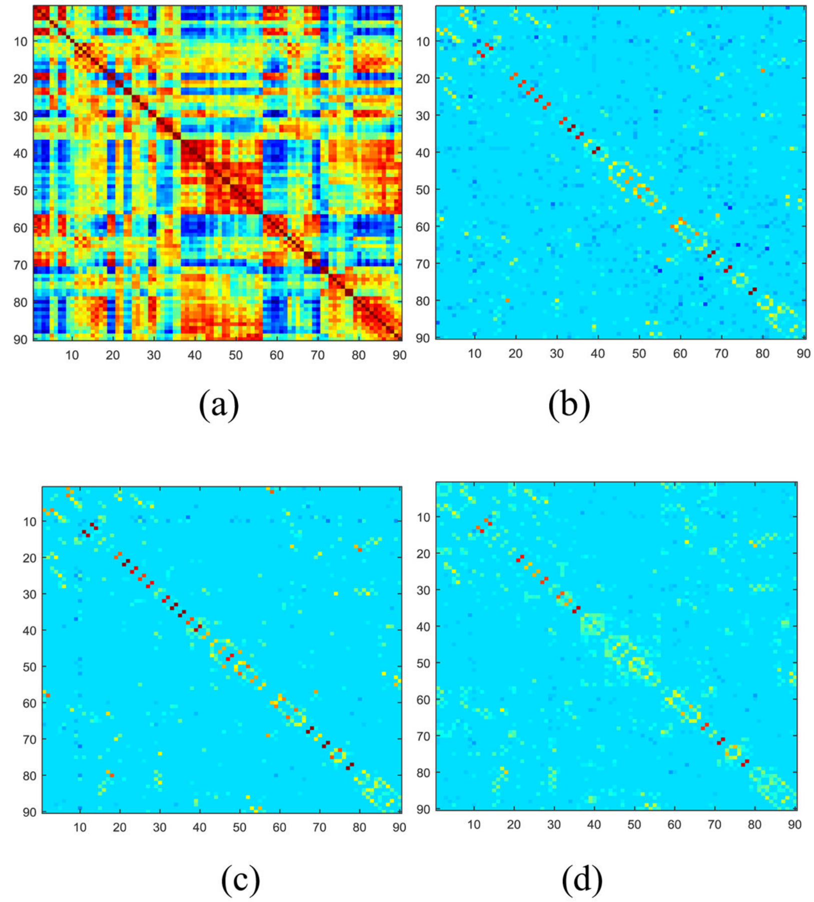

Figure 3.
FC networks derived from (a) PEC, (b) SR, (c) WSR, and (d) PSGSR.

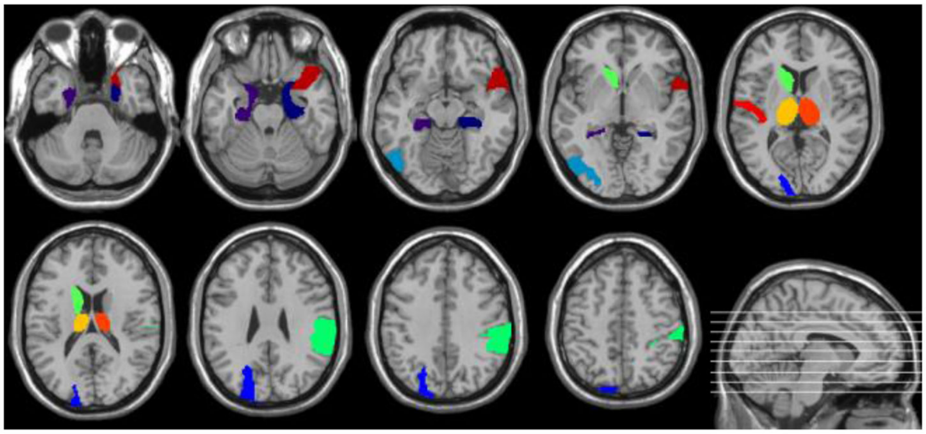

Figure 4.
The most discriminative brain regions.

TABLE I

Performance of different methods.

Method	ACC	AUC	SEN	SPE
PEC	63.92	0.6619	63.27	64.58
SR	68.04	0.7096	75.51	60.42
WSR	72.16	0.7300	75.51	68.75
PSGSR	81.44	0.8418	83.67	79.17