Medical Image Analysis 52 (2019) 80–96

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Multimodal hyper-connectivity of functional networks using functionally-weighted LASSO for MCI classification

Yang Li[a], Jingyu Liu[a], Xinqiang Gao[a], Biao Jie[b], Minjeong Kim[c], Pew-Thian Yap[d], Chong-Yaw Wee[e,*], Dinggang Shen[d,f,**]

[a] School of Automation Sciences and Electrical Engineering, Beijing Advanced Innovation Center for Big Data and Brain Computing, Beijing Advanced Innovation Center for Big Date-based Precision Medicine, Beihang University, Beijing, China
[b] Department of Computer Science and Technology, Anhui Normal University, Wuhu, China
[c] Department of Computer Science, University of North Carolina at Greensboro, NC, USA
[d] Department of Radiology and BRIC, University of North Carolina at Chapel Hill, NC, USA
[e] Department of Biomedical Engineering, National University of Singapore, Singapore
[f] Department of Brain and Cognitive Engineering, Korea University, Seoul 02841, Republic of Korea

ARTICLE INFO

Article history:
Received 23 April 2018
Revised 30 September 2018
Accepted 12 November 2018
Available online 13 November 2018

Keywords:
Hyper-connectivity network
Multimodality
Weighted LASSO
Mild cognitive impairment (MCI)
Arterial spin labeling (ASL)
Ultra-least squares (ULS)

ABSTRACT

Recent works have shown that hyper-networks derived from blood-oxygen-level-dependent (BOLD) fMRI, where an edge (called hyper-edge) can be connected to more than two nodes, are effective biomarkers for MCI classification. Although BOLD fMRI is a high temporal resolution fMRI approach to assess alterations in brain networks, it cannot pinpoint to a single correlation of neuronal activity since BOLD signals are composite. In contrast, arterial spin labeling (ASL) is a lower temporal resolution fMRI technique for measuring cerebral blood flow (CBF) that can provide quantitative, direct brain network physiology measurements. This paper proposes a novel sparse regression algorithm for inference of the integrated hyper-connectivity networks from BOLD fMRI and ASL fMRI. Specifically, a least absolution shrinkage and selection operator (LASSO) algorithm, which is constrained by the functional connectivity derived from ASL fMRI, is employed to estimate hyper-connectivity for characterizing BOLD-fMRI-based functional interaction among multiple regions. An ASL-derived functional connectivity is constructed by using an Ultra-GroupLASSO-UOLS algorithm, where the combination of ultra-least squares (ULS) criterion with a group LASSO (GroupLASSO) algorithm is applied to detect the topology of ASL-based functional connectivity networks, and then an ultra-orthogonal least squares (UOLS) algorithm is used to estimate the connectivity strength. By combining the complementary characterization conveyed by rs-fMRI and ASL fMRI, our multimodal hyper-networks demonstrated much better discriminative characteristics than either the conventional pairwise connectivity networks or the unimodal hyper-connectivity networks. Experimental results on publicly available ADNI dataset demonstrate that the proposed method outperforms the existing single modality based sparse functional connectivity inference methods.

1. Introduction

Alzheimer's disease (AD) is the most common cause of dementia in people over 65 years old (Barker et al., 2002; Wilson et al., 2012; Zhu et al., 2017). Recently, it has been reported that a new case of AD is expected to develop every 33 s, and by 2050 it will result in nearly a million new cases each year (Association, 2016; Li et al., 2018b). As a prodromal stage of AD, mild cognitive impairment (MCI) has gained a great deal of attention recently due to its high progression rate to AD. Existing studies show that people with MCI, especially MCI involving memory problems (i.e., amnestic MCI), are more likely to develop AD than people without MCI (Kantarci et al., 2009; Mitchell and Shiri-Feshki, 2009). Therefore, early diagnosis of MCI is of great importance for preventing, slowing or stopping its progression to AD (Li et al., 2017c; Ward et al., 2013). In the past years, the combination of neuroimaging-based techniques and graph theory has been shown powerful for exploring the pathological underpinnings of AD and MCI (Brier et al., 2014; Li et al., 2013; Park and Friston, 2013; Wang et al., 2013). For example, the functional connectivity networks derived from functional magnetic resonance imaging (fMRI) based on the blood-oxygenation-

* Corresponding author.
** Corresponding author at: Department of Radiology and BRIC, University of North Carolina at Chapel Hill, NC 27599, USA.
E-mail addresses: cywee2000@gmail.com (C.-Y. Wee), dgshen@med.unc.edu (D. Shen).

https://doi.org/10.1016/j.media.2018.11.006
1361-8415/

level-dependent (BOLD) signals, which can characterize the cognitive-related interaction patterns between brain regions at rest, have been used for the MCI classification (Chen et al., 2011; Feng et al., 2012; Li et al., 2014; Petrella et al., 2011).

Most of the existing functional connectivity network studies for MCI classification used the pairwise correlation-based approach to characterize the functional relationship between two brain regions (Hayasaka and Laurienti, 2010; Kaiser, 2011; Smith et al., 2013; Sporns, 2011; Wee et al., 2012a). However, such analysis may not be accurate in revealing the active cognitive activities of the brain since, based on neurological findings (Huang et al., 2010), a brain region predominantly interacts directly with a few of other brain regions in neurological processes. Moreover, recent neurological studies have demonstrated significant high-order interactions (i.e., interaction among more than two brain regions) in neuronal spiking, local field potentials, and cortical activities (Ganmor et al., 2011; Ohiorhenuan et al., 2010; Yu et al., 2011). Based on this finding, Jie et al. (2016) derived a functional connectivity hyper-network from BOLD fMRI data for characterizing the high-order interactions between three or more brain regions, and achieved a better AD classification performance compared to the conventional pairwise connectivity method. Davison et al. (2015) constructed hyper-networks and characterized hyper-connectivity on the basis of its topology structure, anatomy, and task-specificity. In their study, hyper-connectivity was demonstrated to be a more effective measure to characterize functional brain dynamics than dyadic connectivity. Gu et al. (2017) proposed a hypergraph representation method based on the BOLD fMRI data, which reveals three specific classes of hyper-edges: bridges, stars and clusters, representing bipartite, focal and spatially distributed architectures, respectively. Furthermore, a novel learning-based hyper-network was recently proposed to characterize complex connectivity patterns among multiple brain regions (Zu et al., 2018). A hypergraph learning method is first employed to construct the hyper-edges and then a sparse constraint based on the discrimination power and the intra-group consistency is applied to suppress the spurious hyper-connectivity (Zu et al., 2018). Nevertheless, one main deficiency of their hyper-networks method is that it only considers the interaction among multiple brain regions from a single modality (i.e., BOLD fMRI) data, ignoring the complementary information conveyed by other imaging modalities (Li et al., 2017b).

An ample number of studies have demonstrated the advantage of integrating complementary information from multiple modalities for improving representation capacity and classification performance (Gao et al., 2015; Liu et al., 2014; Wee et al., 2012b; Zhu et al., 2013). Multiple imaging modalities can reflect specific characteristics of the brain structure and function from different views. For instance, BOLD fMRI provides the regional interactions happened when the subject is in the absence of an explicit task. While arterial spin labeling (ASL) perfusion MRI, as a relatively new noninvasive functional MRI technique, measures cerebral blood flow (CBF) by using arterial water as an endogenous tracer (Borogovac and Asllani, 2012; Liang et al., 2014). Particularly, ASL perfusion MRI, which measures CBF directly and noninvasively, can provide a more direct quantitative correlate of neural activity than BOLD fMRI (Cavusoglu et al., 2012; Havlicek et al., 2015), and thus is applied to characterize the inter-regional interactions of brain activities via constructing the functional connectivity networks in a recent study (Liang et al., 2014). However, ASL perfusion MRI presents some drawbacks (i.e., lower signal-to-noise ratio (SNR), reduced temporal resolution, and increased power deposition) compared to BOLD fMRI (Alsop et al., 2015). To combine the benefits of both modalities, (Gao et al., 2015) integrated ASL fMRI information into BOLD fMRI sequences at the feature level in a unified MCI classification framework. Their results suggested that the fusion of BOLD and ASL fMRI can improve the performance of MCI classification. Additionally, Jann et al. (2015) presented a framework for independent and joint functional connectivity analyses of BOLD fMRI and ASL perfusion MRI to identify modality-specific brain networks. Their experimental results showed that the functional connectivity in brain networks was correlated with the regional CBF of the associated networks. These results suggest that joint functional connectivity network analyses by using BOLD and ASL can characterize the spatiotemporal and quantitative properties of brain networks in a relatively more comprehensive manner. However, to the best of our knowledge, there is no study that constructs hyper-network by integrating ASL perfusion MRI with BOLD fMRI time series for MCI classification.

It is well-known that the brain network modeling method has a great impact on the performance of the MCI classification (Jie et al., 2016). Recent studies have shown that the sparse representation based network modeling approaches can reduce spurious connections and improve performance for MCI classification (Rosa et al., 2015; Zanin et al., 2012). For example, (Wee et al., 2014) constructed the functional connectivity based on the Group LASSO with a $l_{2,1}$ regularizer for classifying patients with MCI from normal control (NC). Their experimental results demonstrated that the sparse functional connectivity approach yields markedly improved classification performance compared to the conventional Pearson correlation-based networks. Through the sparse representation model, all other regions will be considered simultaneously and those regions with genuine functional interactions to the target brain region would stand out. However, those sparse representation based network modeling methods inevitably face the overfitting problem due to the use of the classic least squares criterion (Li et al., 2017a) as the interruption by the highly noisy information in fMRI time series may not provide reliable estimations of the brain networks.

In this paper, we propose a novel ultra-least-squares-based functionally-weighted LASSO (FW-LASSO) method to construct multimodal hyper-network for MCI classification, which integrates information from BOLD fMRI and ASL perfusion MRI data. The key idea of our proposed method is to use the functional connectivity information derived from the ASL fMRI data as the constraint to guide the regression process of FW-LASSO for extracting interactions among multiple brain regions from the BOLD fMRI to construct hyper-connectivity networks. As BOLD fMRI has higher SNR and higher temporal resolution, hyper-connectivity analysis based on BOLD signals is more reliable than ASL fMRI (Jann et al., 2015). Moreover, since CBF is proposed as a close marker for metabolic activity (Jann et al., 2015), employing the ASL-derived functional connectivity information as a constraint can supplement metabolic-related information, which is inaccessible by BOLD because its effect originates from an intricate interplay between changes in CBF, cerebral blood volume (CBV) and oxygen consumption (Borogovac and Asllani, 2012), for the hyper-connectivity construction. Specifically, in order to obtain the metabolic-related functional connectivity information from ASL, an Ultra-GroupLASSO regression algorithm, which is composed of a group LASSO (GroupLASSO) algorithm and an ultra-least squares (ULS) criterion (Li et al., 2017a), is proposed to detect the topology of the ASL-based networks. Then, we apply an UOLS algorithm (Guo et al., 2016), which consists of the ULS criterion and an orthogonal least squares (OLS) algorithm (Li et al., 2018a), to estimate the strength of the identified ASL-based connections. In the Ultra-GroupLASSO algorithm, a group constraint item (i.e., $l_{2,1}$) is used to ensure an identical network topology across individuals while at the meantime preserving the individual-specific information via different connectivity values (Wee et al., 2014). In addition, the combination of the Ultra-Group-LASSO algorithm and the UOLS algorithm *not only* extracts the classical dependent relation between the ASL time series of a region pair, *but also* uses the dependent

relation of the associated weak derivatives to avoid the overfitting problem. As a generalization of the strong derivative (i.e., the derivative in the usual sense) that can only be calculated for the differentiable functions, the weak derivatives allow the computation of all integrable functions (Guo et al., 2016). The weak derivatives can solve nondifferentiable problems in the derivation process of the strong derivative. Considering both the strong derivative and the weak derivative helps the model to fit the ASL fMRI time series with the more precise relationship between different time points. Therefore, both the Ultra-GroupLASSO algorithm and the UOLS algorithm, which include the weak derivatives in the formulation, are able to avoid the overfitting problem, which is often encountered in the conventional least squares criterion (Li et al., 2017a). To validate the effectiveness of our proposed method, multimodal hyper-networks from ASL perfusion and BOLD fMRI data are constructed for MCI classification. We extract three types of hyper-network features constructed and use a manifold regularized multi-task feature selection method (M2TFS) (Jie et al., 2016) to jointly select the most discriminative region-related features. We finally use the selected features to train a multi-kernel support vector machine (SVM) (Li et al., 2018c) for MCI classification. Experimental results showed promising improvement in classification performance, demonstrating the superiority of our proposed method compared to the-state-of-the-art single modal hyper-connectivity-based methods.

In summary, the contribution of our proposed method is threefold. First, we proposed a novel MCI classification framework, which provides a new scheme for the fusion of multiple neuroimaging modalities. To the best of our knowledge, our method is among the first to use the multimodal hyper-network that fuses the complementary information from BOLD fMRI and ASL fMRI in neuroimaging studies. Second, we proposed a novel FW-LASSO method to construct the multimodal hyper-networks that combine BOLD fMRI and ASL fMRI data to characterize functional connectivities within the human brain. Compared with the conventional pairwise correlation approach, our proposed multimodal hyper-network modeling method is more comprehensive in characterizing the interactions among multiple brain regions that are working together. Third, we took into consideration the discrepancy between the weak derivatives of the observed signals and the model prediction function via the ULS criterion. The estimated functional connectivity networks are more reliable as the ULS criterion can avoid the overfitting problem.

The rest of the paper is organized as follows. Materials and Methodology section furnishes information on the data acquisition and post-processing, followed by the proposed framework for constructing multimodal hyper-networks using complementary information from BOLD fMRI and ASL fMRI. Then, we evaluate the performance of our proposed framework for MCI classification in Experiments and Results section. Further, findings, methodological issues, and limitations of our proposed framework are discussed extensively in Discussion section. Finally, we conclude this paper in Conclusion section.

2. Materials and methodology

2.1. Dataset

In this study, the BOLD fMRI and ASL perfusion data were collected from 28 MCI individuals and 33 normal controls (NCs). For BOLD fMRI, all the subjects were scanned using a standard echo-planar imaging (EPI) sequence on a 3 Tesla Siemens TRIO scanner with the following parameters: TR = 3000 ms, TE = 30 ms, acquisition matrix = 74×74, 45 slices, and voxel thickness = 3 mm. One-hundred and eighty resting-state fMRI volumes were acquired. Standard preprocessing pipeline of the fMRI images was performed using Statistical Parametric Mapping 8 (SPM8) software package which includes removal of first 10 fMRI volumes, slice timing correction, head-motion correction, regression of nuisance signals (ventricle, white matter, global signal, and head-motion with Friston's 24-parameter mode (Friston et al., 1996; Wee et al., 2016)), signal de-trending, and band-pass filtering (0.01–0.08 Hz). Next, the brain space was parcellated into 90 ROIs based on the automated anatomical labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002).

For ASL perfusion MRI, we used a pulsed ASL (pASL) sequence to record brain perfusion with the following parameters: acquisition matrix = 64×64, voxel size = $3.44 \times 3.44 \times 6$ mm^3, TE = 21 ms, TR = 5000 ms, inversion time (TI) = 700 ms and 1800 ms, saturation stop time = 1600 ms, flow limit = 5 cm/s, 71 images with alternation of labeled and non-labeled images, parallel imaging (iPAT) factor = 2, and PICORE Q2Tips labeling scheme (Luh et al., 1999). Preprocessing of ASL images was performed using a Matlab and SPM-based ASL perfusion fMRI data processing toolbox, ASLtbx (Wang et al., 2008). The preprocessing steps include raw image quality checking, integer-to-float format conversion, independent label and control motion correction, coregistration, spatial smoothing, cerebral blood flow (CBF) calculation, outlier cleaning, statistical analysis, and contrast definition (Wang et al., 2008). ASL images were registered to their structural images using affine registration algorithm provided in SPM.

2.2. Methods

Fig. 1 illustrates a framework of our proposed multimodal hyper-network for MCI classification, which includes three main steps: (1) employ an Ultra-GroupLASSO algorithm and an UOLS algorithm to detect functional network topology; (2) integrate the complementary information of ASL fMRI and BOLD fMRI by applying FW-LASSO algorithm to construct multimodal hyper-networks; (3) extract three types of network features from hyper-networks, followed by an M2TFS method to jointly select the most discriminative features, and use a multi-kernel SVM on the selected features for MCI classification. In addition, a detailed description of Step (2) is provided in Fig. 2. In the following sections, we give the detailed procedure of our proposed method.

2.2.1. Construction of multimodal hyper-connectivity networks

To integrate individual advantages of ASL fMRI and BOLD fMRI in the functional connectivity network construction, we propose a novel FW-LASSO algorithm that employs functional connectivity information derived from ASL fMRI data as a constraint to construct hyper-networks for characterizing BOLD-based interactions among multiple brain regions. Here, these hyper-networks combine the high-order interactions among brain regions and the mutual information between BOLD fMRI and ASL fMRI, called multimodal hyper-networks in this paper. Particularly, the major difference between our method and the adaptive LASSO (Chatterjee and Lahiri, 2013) is that our method determines the weight directly according to the functional connectivity analysis based on ASL data, while the latter uses l_2 initial estimation to reweight l_1 penalty. The LASSO is used to represent the functional interactions based on BOLD fMRI data and ASL-derived functional connectivity constraint will guide the regression process. The neuroscience basis of our method is that CBF changes are one of the primary contributors to the intrinsic BOLD fluctuations and the functional connectivity is positively correlated with the regional CBF (Jann et al., 2015; Tak et al., 2015). That is, if CBF signals of two ROIs are highly correlated, their BOLD-based functional connectivity should be penalized less. Otherwise, those weak ASL-derived functional connections should be penalized with larger weights. To measure the correlation between CBF signals over each pair of ROIs, a functional connectivity strength matrix will be derived from the ASL fMRI

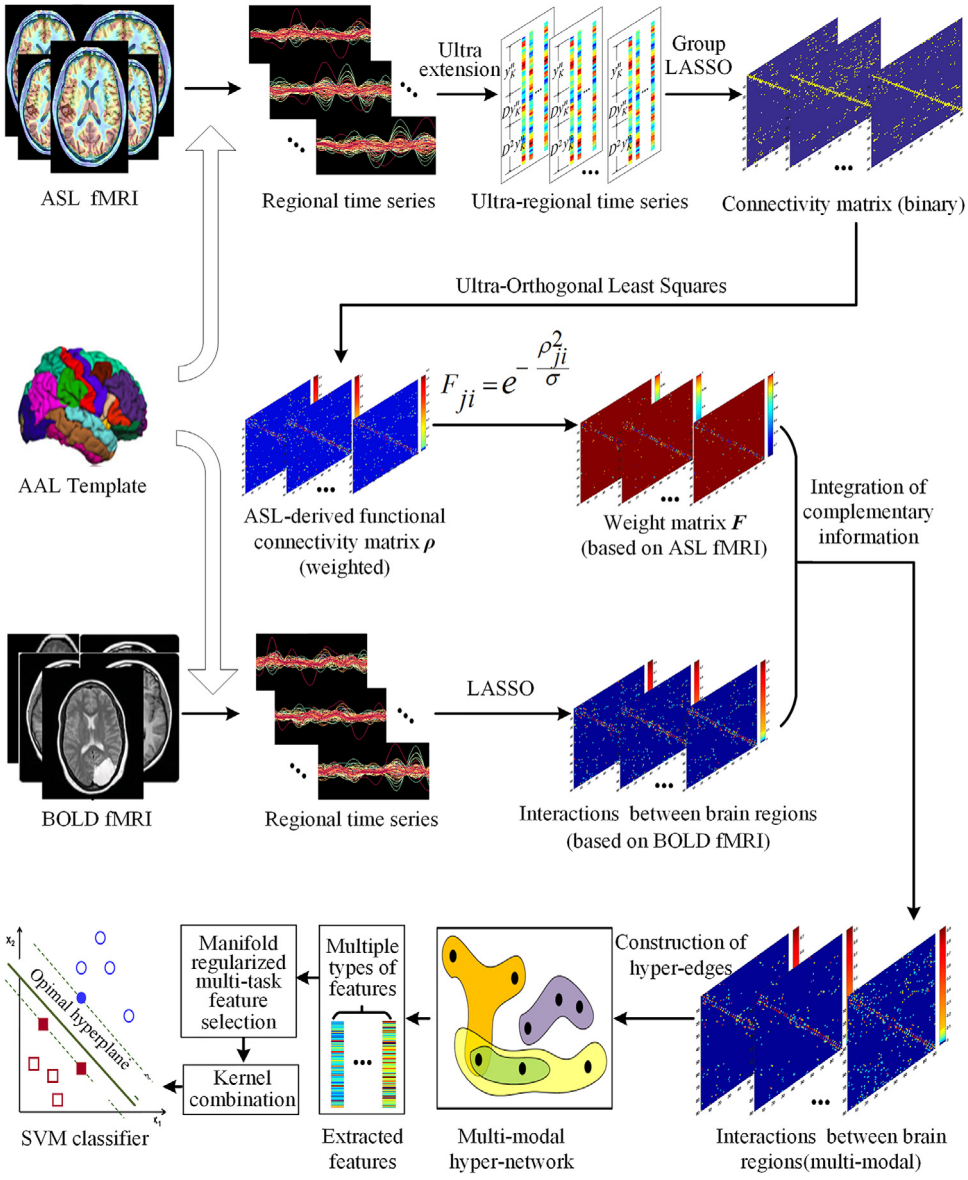

Fig. 1. The flowchart of our proposed method.

data. Then, the penalty weight F_{ji} for the connectivity between the i-th ROI and the j-th ROI is given as follows (Yu et al., 2016):

$$F_{ji}=e^{-\frac{\rho_{ji}^2}{\sigma}}. \tag{1}$$

where ρ_{ji} denotes the ASL-derived functional connectivity strength between the i-th ROI and the j-th ROI, and σ is the mean of all subjects' standard variances for the connectivity strength (i.e., ρ_{ji}). The ASL-derived functional connectivity matrix $\boldsymbol{\rho}$ is constructed from the ASL perfusion data by using the combination of an Ultra-GroupLASSO algorithm and an UOLS algorithm. Here, the Ultra-GroupLASSO algorithm is composed of a group LASSO algorithm (Wee et al., 2014) and the ULS criterion (Li et al., 2017a), while the UOLS algorithm consists of the ULS criterion and an OLS algorithm (Guo et al., 2016). The detailed description for the Ultra-GroupLASSO algorithm and the UOLS algorithm is given in the next section.

Accordingly, supposing that we have P ROIs and N subjects (i.e., $P=90,\ N=61$ in the experiments), our proposed FW-LASSO, which is applied to represent functional interactions among multiple brain regions based on the high SNR BOLD signals, can be given by:

$$\min_{\boldsymbol{\alpha}_i}\frac{1}{2}\left\|\boldsymbol{x}_i-\sum_{j\neq i}^{P}\boldsymbol{x}_j\alpha_{ji}\right\|_2^2+\lambda\sum_{j\neq i}^{P}F_{ji}\left|\alpha_{ji}\right|_1. \tag{2}$$

where $\boldsymbol{F}\in\mathrm{R}^{P\times P}$ is the weight matrix, which is constructed from the ASL data by using the functional connectivity information, with each element F_{ji} being inversely proportional to the functional connectivity strength between the j-th ROI and the i-th ROI, $\boldsymbol{\alpha}_i=[\alpha_{1i},\cdots,\alpha_{ji},\cdots,\alpha_{Pi}]$ denotes the coefficient vector that quantifies the extent of the influence from other ROIs to the i-th ROI, $\boldsymbol{x}_i$ is the BOLD fMRI time series reflecting integrated information of changes in CBF, CBV and oxygen consumption for the i-th ROI, and λ is a regularization parameter used to controls the sparsity of the connectivity matrix. In particular, different λ values correspond to different sparse model solutions, with a larger λ value indicating a sparser connectivity network. Additionally, Eq. (2) is similar in concept, but not identical, to the adaptive LASSO, since the weight F_{ji} in Eq. (2) is from the ASL-derived functional connectivity, instead of learning from the dataset. Hence, the FW-LASSO estimates

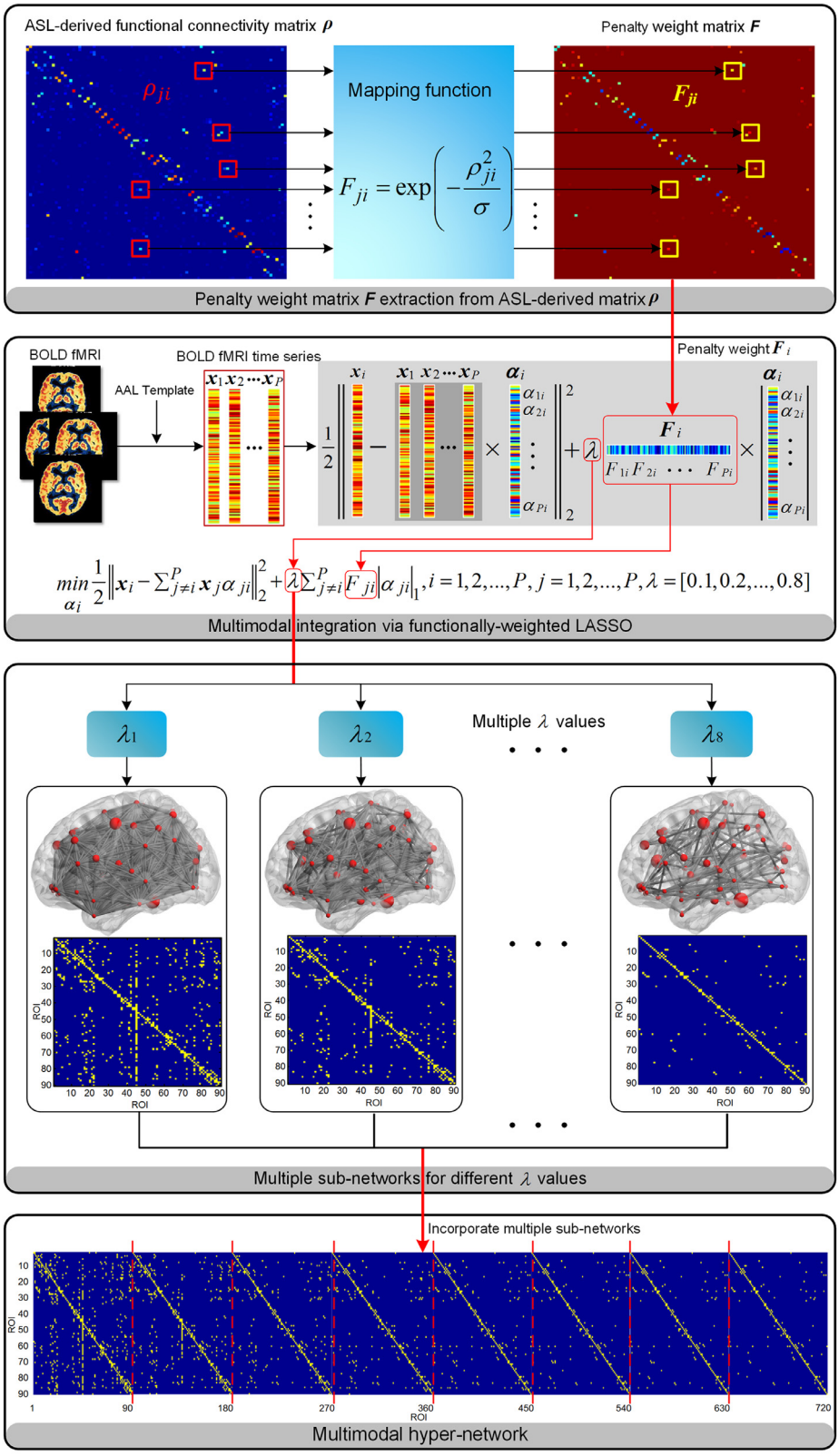

Fig. 2. Illustration of the procedure for constructing multimodal hyper-networks.

in Eq. (2) can also be solved by the least angle regress (LARS) algorithm (Efron et al., 2004) as follows:

$$\min_{\alpha'_i} \frac{1}{2}\left\|x_i - \sum_{j\neq i} x_j \alpha'_{ji}\right\|_2^2 + \lambda \sum_{j\neq i} |\alpha'_{ji}|_1 \quad (3)$$

where $\alpha'_{ji} = F_{ji}\alpha_{ji}$ $(F_{ji} > 0)$, $x'_{ji} = x_{ji}/F_{ji}$ and $\boldsymbol{\alpha}'_i = [\alpha'_{1i}, \cdots, \alpha'_{ji}, \cdots, \alpha'_{Pi}]$. Eq. (3) is a standard l_1-norm regularized optimization problem.

For investigating the interactions among different brain regions, we use the FW-LASSO to construct a multimodal hyper-network for

each subject, with each ROI as a node. Let V and E be the vertex set and the hyper-edge set of a hyper-network G, respecrtively. Then, for each subject with P ROIs, a hyper-network $G=(V,E)$ with P vertices can be constructed, where each of its vertices represents an ROI. Specifically, a hype-edge e_i is constructed by linking the centroid ROI (i.e., i-th ROI) to other ROIs with the corresponding non-zero elements in the coefficient vector $\boldsymbol{\alpha}_i$ in Eq. (2). Furthermore, following the work in (Jie et al., 2016), to capture multi-level interactions among ROIs, we employ multiple λ values in the FW-LASSO method for generating a group of sub-networks for each subject.

2.2.2. Construction of the ASL-derived connectivity strength matrix ρ

During the construction of multimodal hyper-networks, the ASL-derived connectivity strength matrix $\boldsymbol{\rho}$ plays an important role in the fusion of complementary information from ASL fMRI and BOLD fMRI. In order to obtain the ASL-derived connectivity strength matrix $\boldsymbol{\rho}$, two necessary steps are implemented: 1) the topology detection of the connectivity networks via an Ultra-GroupLASSO algorithm, by incorporating the ultra-least squares (ULS) algorithm with the group LASSO (GroupLASSO) algorithm; 2) estimation of connectivity strength based on the detected network topology via an ultra-orthogonal least squares (UOLS) algorithm. In particular, we further extract the dependent relation of the associated weak derivatives between the fMRI time series of a region pair, using the ULS criterion in both the Ultra-GroupLASSO algorithm and the UOLS algorithm to avoid the overfitting problem that is commonly encountered in the conventional least squares criterion (Li et al., 2017a). Additionally, it is worth noting that the first step is crucial since the Ultra-GroupLASSO algorithm can filter out insignificant or spurious connections for the subsequent estimation of the connectivity strength in the second step. In the following subsection, we provided a detailed description to demonstrate how the ASL-derived connectivity matrix $\boldsymbol{\rho}$ is obtained from the ASL fMRI data.

An Ultra-GroupLASSO algorithm is proposed to detect the topology structure of ASL-derived functional networks as below:

$$f(\boldsymbol{\theta}_i)=\sum_{n=1}^{N}\left\|\boldsymbol{\psi}_i^n-\boldsymbol{A}_i^n\boldsymbol{\theta}_i^n\right\|_2^2+\sum_{n=1}^{N}\sum_{l=1}^{L}\left\|D^l\boldsymbol{\psi}_i^n-D^l\boldsymbol{A}_i^n\boldsymbol{\theta}_i^n\right\|_2^2+\varpi\left\|\boldsymbol{\theta}_i\right\|_{2,1}. \tag{4}$$

where $\boldsymbol{\psi}_i^n$ denotes the ASL fMRI time series of the i-th ROI for the n-th subject, $\boldsymbol{A}_i^n=[\boldsymbol{\psi}_1^n,\ldots,\boldsymbol{\psi}_{i-1}^n,\boldsymbol{\psi}_{i+1}^n,\ldots,\boldsymbol{\psi}_P^n]$ denotes a matrix including all ROIs time series except the i-th ROI, $\boldsymbol{\theta}_i=[\boldsymbol{\theta}_i^1,\ldots,\boldsymbol{\theta}_i^{n-1},\boldsymbol{\theta}_i^{n+1},\ldots,\boldsymbol{\theta}_i^N]$ indicates the coefficient matrix, $\boldsymbol{\theta}_i^n=[\theta_{1i}^n;\ \ldots;\ \theta_{(i-1)i}^n;\ \theta_{(i+1)i}^n;\ \ldots;\theta_{Pi}^n]$ is a coefficient vector roughly showing the impact of other brain regions on the i-th ROI since the Ultra-GroupLASSO is a biased estimation, $\varpi>0$ is the regularization parameter used to control the sparsity of the model, and $\|\boldsymbol{\theta}_i\|_{2,1}$ is the summation of l_2-norms of $\boldsymbol{\theta}_i^n$, i.e., $\sum_i\|\boldsymbol{\theta}_i^n\|_2$. Moreover, $D^l\boldsymbol{\psi}_i^n(t)$, where D^l is the l-th order weak derivative, is defined by (Li et al., 2017a):

$$\int_{[0,\ T]}\boldsymbol{\psi}_i^n(t)D^l\varphi(t)dt=(-1)^l\int_{[0,\ T]}\varphi(t)D^l\boldsymbol{\psi}_i^n(t)dt. \tag{5}$$

for all test function $\varphi(t)\in C_0^\infty(0,\ T)$ which is smooth and compact support on $[0,\ T]$.

Although Ultra-GroupLASSO can accurately filter out insignificant or spurious connections, the non-zero estimated coefficients cannot be directly regarded as the connectivity strengths between brain regions since they are biased due to group-constrained sparse penalization. Hence, to obtain an unbiased estimate of the connectivity strength, the UOLS algorithm is used to estimate the connectivity strength between the ASL fMRI time series of two brain regions (Guo et al., 2016). ROIs with the corresponding non-zero elements in the coefficient vector $\boldsymbol{\theta}_i^n$ in Eq. (4) are deemed to be correlated to the i-th ROI. Supposing that Q ROIs have been found correlated to the i-th ROI, the ASL-derived functional connectivity between the i-th ROI and the other Q ROIs can be estimated by:

$$f\left(\boldsymbol{\rho}_i^n\right)=\left\|\boldsymbol{\psi}_i^n-\boldsymbol{\Phi}_i^n\boldsymbol{\rho}_i^n\right\|_2^2+\sum_{l=1}^{L}\left\|D^l\boldsymbol{\psi}_i^n-D^l\boldsymbol{\Phi}_i^n\boldsymbol{\rho}_i^n\right\|_2^2. \tag{6}$$

where $\boldsymbol{\Phi}_i^n=[\psi_{j_1}^n,\psi_{j_2}^n,\ldots,\psi_{j_Q}^n]\,(i.e.,\ j_q\neq i,q=1,2,\cdots,Q)$ is the matrix including ASL fMRI time series of Q ROIs, and $\boldsymbol{\rho}_i^n=[\rho_{j_1,i}^n;\rho_{j_2,i}^n;\ldots;\rho_{j_Q,i}^n]$ $(i.e.,\ j_q\neq i,q=1,2,\cdots,Q)$ is the coefficient vector (i.e., the connectivity strength). Then, for the n-th subject, the ASL-derived connectivity strength matrix is $\boldsymbol{\rho}=[\boldsymbol{\rho}_1^n,\ \cdots,\boldsymbol{\rho}_i^n,\cdots,\ \boldsymbol{\rho}_P^n]$. Detailed computational process for the UOLS algorithm is provided in Appendix A.

2.2.3. Feature extraction and selection

Topological properties derived from a hyper-network provide quantitative measures to effectively study the differences in terms of brain functional organization between MCI and NC (Jie et al., 2016). Thus, we extract three types of clustering coefficients from the constructed multimodal hyper-networks as features for MCI classification. Given a multimodal hyper-network $G=(V,E)$, let $M(v)$ be the hyper-edges adjacent to the vertex v, i.e., $M(v)=\{e\in E:v\in e\}$, and let $N(v)$ be the neighboring vertices to v, i.e., $N(v)=\{u\in V:\exists e\in E,\ u,v\in e\}$. Accordingly, three different types of clustering coefficients on vertex v can be computed as follows (Jie et al., 2016):

$$\mathrm{HCC}^1(v)=\frac{2\sum_{u,q\in N(v)}I(u,q,\neg v)}{|N(v)|(|N(v)|-1)}. \tag{7}$$

$$\mathrm{HCC}^2(v)=\frac{2\sum_{u,q\in N(v)}I'(u,q,v)}{|N(v)|(|N(v)|-1)}. \tag{8}$$

$$\mathrm{HCC}^3(v)=\frac{2\sum_{e\in M(v)}(|e|-1)-|N(v)|}{|N(v)|(|M(v)|-1)}. \tag{9}$$

where $u,\ q,\ v\in V$ and $e\in E$, $I(u,q,\neg v)=1$ if there exists $e\in E$ such that $u,\ q\in e$ but $v\notin e$, and 0 otherwise. $I'(u,q,v)=1$ if there exists $e\in E$ such that $u,\ q,\ v\in e$, and 0 otherwise. Three types of clustering coefficient features represent the topological properties of the multimodal hyper-connectivity network from three different perspectives. Specifically, the HCC^1 denotes the number of neighboring vertices that have connections not facilitated by vertex v. In contrast, the HCC^2 denotes the number of neighboring vertices with connections facilitated by vertex v, showing that these nodes may share some brain functions with each other and vertex v. The HCC^3 denotes the amount of overlap among adjacent hyper-edges of vertex v.

Additionally, note that not all features extracted from the multimodal hyper-networks are relevant to MCI pathology. Thus, in order to remove those irrelevant or redundant features, following the work in (Jie et al., 2016), we perform a manifold regularized multi-task feature selection (M2TFS) method to jointly select the most discriminative features. Let $\boldsymbol{S}^c=[\boldsymbol{s}_1^c,\cdots,\boldsymbol{s}_n^c,\cdots,\boldsymbol{s}_{N_{tr}}^c]^T\in R^{N_{tr}\times P}$ be three sets of features from a total of N_{tr} training subjects, and $\boldsymbol{Y}=[y_1,\cdots,y_n,\cdots,y_{N_{tr}}]^T\in R^P$ be the response vector for those N_{tr} training subjects, where $\boldsymbol{s}_n^c=[\mathrm{HCC}^c(v_i)]_{i=1:P}\in R^P$ is the vector of clustering coefficient features from the n-th training subject on task c (in our case, each task represents feature learning on one type of clustering coefficients), and y_n be the class label for the n-th training subject. Then, the M2TFS method is given by

Jie et al. (2016):

$$\min_{\boldsymbol{Z}} \frac{1}{2}\sum_{c=1}^{C}\|\boldsymbol{Y}-\boldsymbol{S}^c\boldsymbol{z}^c\|_2^2+\beta\sum_{c=1}^{C}(\boldsymbol{S}^c\boldsymbol{z}^c)^T\boldsymbol{D}^c(\boldsymbol{S}^c\boldsymbol{z}^c)+\gamma\|\boldsymbol{Z}\|_{2,1}\ . \quad (10)$$

where $\boldsymbol{D}^c=\boldsymbol{L}^c-\boldsymbol{W}^c$ is the combinatorial Laplacian matrix on task c. $\boldsymbol{W}^c$ is a matrix, which describes the similarity of subjects on the c-th task, with each element defined as $W^c_{nm}=\exp(-\|\boldsymbol{s}^c_n-\boldsymbol{s}^c_m\|^2/h)$. Here, h is a free parameter to be tuned empirically. $\boldsymbol{L}^c$ is a diagonal matrix with $L^c_{nn}=\sum_{m=1}^{N_{tr}}W^c_{nm}$. Moreover, $\boldsymbol{Z}=[\boldsymbol{z}^1,\boldsymbol{z}^2,\cdots,\boldsymbol{z}^C]\in R^{P\times C}$ is a coefficient matrix with C as the total number of tasks (i.e., $C=3$), and $\|\boldsymbol{Z}\|_{2,1}=\sum_{i=1}^{P}\|\boldsymbol{z}_i\|_2$ is the group sparsity regularizer that encourages features from different tasks to be jointly selected, where $\boldsymbol{z}_i$ is the i-th row vector of $\boldsymbol{Z}$. In Eq. (10), β and γ are the associated regularization coefficients. The values of h,β and γ can be determined via an inner cross-validation on the training subjects.

2.2.4. Classification

A multi-kernel SVM (Li et al., 2015) technique is employed to fuse three types of clustering coefficient features selected by the M2TFS method for classification, which is given as follows:

$$k(\boldsymbol{f}_n,\boldsymbol{f}_m)=\sum_{c=1}^{C}\mu^c k^c(\boldsymbol{f}^c_n,\boldsymbol{f}^c_m). \quad (11)$$

where $\boldsymbol{f}^c_n$ is the selected features of the n-th subject from the c-th task, $k^c(\boldsymbol{f}^c_n,\boldsymbol{f}^c_m)$ denotes the linear kernel function on the c-th task between the n-th and m-th subjects, and μ^c is a non-negative weight coefficient with $\sum_{c=1}^{C}\mu^c=1$. In addition, we adopt a coarse-grid search strategy via a cross-validation on the training subjects to optimize μ^c. Once we obtain the optimal μ^c, the standard SVM can be performed for MCI classification.

2.2.5. Implementation details

In this work, a nested leave-one-out cross-validation (LOOCV) scheme is implemented to evaluate the performance and generalization power of our proposed method. Specifically, in each cross-validation, one subject is selected as a testing subject, the remaining subjects are used to train the classifier. Repeat the entire process with each subject being left out once as a testing subject to estimate the classification performance. We employ LIBSVM toolbox to implement the multi-kernel SVM classifier, and a grid search is applied to determine the optimal weights μ^c for integrating multiple kernels based on another LOOCV on the training subjects within the range [0, 1] at a step size of 0.1. Besides, we normalize each extracted feature from all training subjects, and then the associated feature of testing data is normalized based on the training data. It should be noted that the inner LOOCV on the training data is used to optimize the parameters h, β and γ, which identify a set of the most discriminative features for MCI classification, while the outer LOOCV is used to evaluate the generalizability of learning model. In addition, the optimal λ value in Eq. (2) and the optimal ϖ value in Eq. (4) are determined via a grid search.

3. Experiments and results

3.1. Classification performance

As a common practice, the classification performance of our proposed method is evaluated by measuring the classification accuracy (ACC), sensitivity (SEN), specificity (SPE), and area under receiver operating characteristic (ROC) curve (AUC). Besides, the balanced accuracy (BAC), which is defined as the arithmetic mean of sensitivity and specificity, is computed to avoid the inflated performance on imbalanced datasets. To construct the optimal multimodal hyper-networks for MCI classification, the proposed method employs a grid search to find the optimal values of parameters λ and ϖ, where λ varies in 9 group values (i.e., [0.1], [0.1, 0.2], [0.1, 0.2, 0.3], ..., [0.1, 0.2, ..., 0.9]) and ϖ varies in a specified range (i.e., [0.001, 0.002, ..., 0.009]). Fig. 3 gives the corresponding classification results for partial parameters. The results show the best classification accuracy of 86.9% when using $\lambda=$[0.1, 0.2, ..., 0.8] and $\varpi=0.006$.

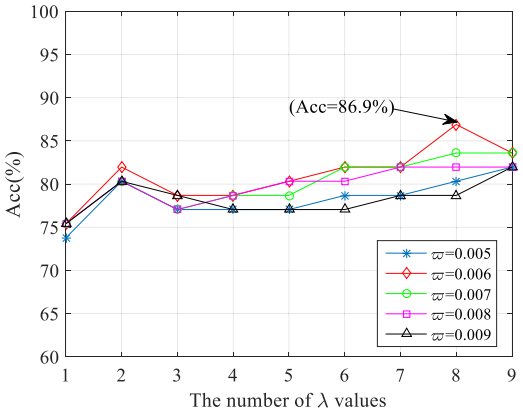

Fig. 3. The classification performance of different regularization parameters λ and ϖ. The abscissa denotes nine groups of different regularization parameter λ, where 1 represents that λ value is [0.1], 2 denotes that λ value is [0.1, 0.2], and 3 denotes that λ value is [0.1, 0.2, 0.3], ..., 9 represents that [0.1, 0.2, ..., 0.9].

Our proposed method is first compared with the single-modality-based methods and a multimodal-based comparing approach. The single-modal-based methods include the approach that employs the conventional LASSO to construct hyper-networks only from a single imaging modality (i.e., BOLD fMRI and ASL fMRI separately) (called LASSO-BOLD and LASSO-ASL, respectively), and the Ultra-GroupLASSO-ASL method that constructs hyper-networks only from ASL fMRI via the Ultra-GroupLASSO regression algorithm. The multimodal comparison methods include Ultra-GroupLASSO-ASL-constrained-BOLD, UOLS-ASL-constrained-BOLD, and BOLD-constrained-ASL. To demonstrate the superiority of using the combination of Ultra-GroupLASSO and UOLS algorithms for the hyper-network construction and MCI classification, we compare our proposed method with the Ultra-GroupLASSO-ASL-constrained-BOLD and UOLS-ASL-constrained-BOLD methods. In the Ultra-GroupLASSO-ASL-constrained-BOLD method, only the Ultra-GroupLASSO algorithm is used alone to estimate the matrix $\boldsymbol{\rho}$, while in the UOLS-ASL-constrained-BOLD method only the UOLS algorithm is used alone to estimate the matrix $\boldsymbol{\rho}$. In order to investigate the difference between the ASL fMRI constrained hyper-network and the BOLD fMRI constrained hyper-network, we compare the proposed method with the BOLD-constrained-ASL method. The only difference between the two multimodal methods is that our proposed method uses the ASL fMRI functional connectivity to constrain the l_1 penalty in the BOLD fMRI-based hyper-networks, while the comparison method uses the BOLD fMRI functional connectivity to constrain the l_1 penalty in the ASL fMRI-based hyper-networks. Furthermore, we also compare the performance of our proposed method with a recently proposed method called Weighted Sparse Group Representation (WSGR) (Yu et al., 2017). This method integrates the functional connectivity strength, group structure, and sparsity in a unified framework for MCI classification. In all above comparison methods, we extract three types of clustering coefficients (i.e., HCC^1, HCC^2 and HCC^3) from hyper-networks, and perform feature selection using the M2TFS method,

Table 1
Classification performances for all the comparison methods.

Method	ACC (%)	AUC	SEN (%)	SPE (%)	BAC (%)
LASSO-ASL	65.6	0.68	53.6	75.8	64.7
Ultra-GroupLASSO-ASL	73.8	0.76	71.4	75.8	73.6
LASSO-BOLD	75.4	0.82	64.3	84.9	74.6
UOLS-ASL-constrained-BOLD	70.5	0.78	64.3	75.8	70.0
Ultra-GroupLASSO-ASL-constrained-BOLD	78.7	0.83	71.4	84.9	78.1
BOLD-constrained-ASL	72.1	0.77	67.9	75.8	71.8
WSGR	78.7	0.89	71.4	84.9	78.1
Proposed	**86.9**	**0.90**	**82.1**	**90.9**	**86.5**

LASSO-ASL: This single-modal-based method employs the conventional LASSO to construct hyper-networks from ASL fMRI; Ultra-GroupLASSO-ASL: This approach uses the Ultra-GroupLASSO algorithm to construct hyper-networks from ASL fMRI; LASSO-BOLD: This single-modal-based method employs the conventional LASSO to construct hyper-networks from BOLD fMRI; UOLS-ASL-constrained-BOLD: This multimodal-based method uses a weighted LASSO to construct the hyper-network based on BOLD fMRI data, l_1 penalty in the weighted LASSO is reweighted based on the ASL-derived functional connectivity information obtained by the UOLS algorithm; Ultra-GroupLASSO-ASL-constrained-BOLD: On the basis of the UOLS-ASL-constrained-BOLD method, replace the UOLS algorithm with the Ultra-GroupLASSO algorithm; BOLD-constrained-ASL: The only difference with the proposed method is that this method uses the BOLD fMRI functional connectivity to constrain the l_1 penalty in the ASL fMRI-based hyper-networks.

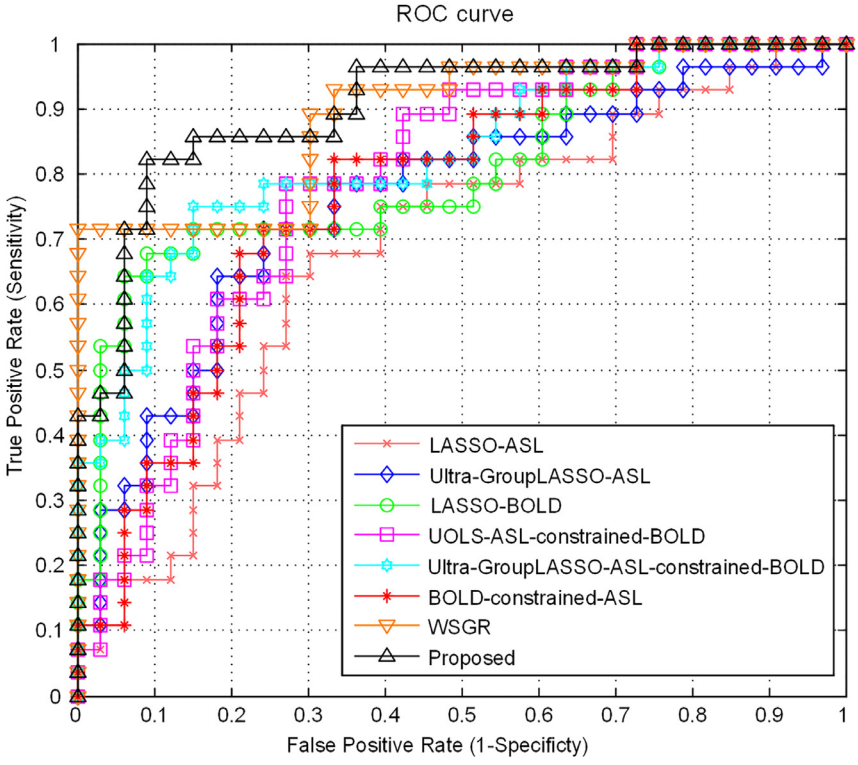

Fig. 4. ROC curves for all comparison methods.

followed by a multi-kernel SVM for classification. Table 1 gives the classification performance of all competing methods, while Fig. 4 graphically illustrates their ROC curves.

As shown in Table 1, our proposed method outperforms all competing methods with a relatively large margin. Specifically, our proposed method yields an accuracy of 86.9% while the second best method yields only 78.7%, an 8.2% of improvement. Our proposed method obtains an AUC of 0.90, indicating excellent diagnostic power. Furthermore, our proposed method performs significantly better than all competing methods in all other performance measures. Fig. 4 shows the ROC curves of all comparison methods. It can be observed that the proposed method achieves the largest area under ROC curves, indicating its excellent diagnostic ability with respect to the changes of discrimination threshold.

3.2. The most discriminative regions

Besides achieving high classification performance, it is also important to selected biologically meaningful features for classification. We thus investigate the most discriminative features (corresponding to ROIs) that are selected by our method. Since the features selected in each LOOCV fold are different, we define the features selected in all folds that contribute to the best classification performance as the most discriminative features, and then refer the corresponding regions as the most discriminative regions. In order to evaluate the discriminative power of each discriminative feature, we perform the standard t-test between two groups, i.e., MCI group and NC group. Table 2 gives the most discriminative regions and their corresponding p-values, where regions with their p-values smaller than 0.05 are shown in bold text. Those

Table 2
List of the most discriminative regions selected during MCI classification.

No.	Regions	*p*-value(HCC1)	*p*-value(HCC2)	*p*-value(HCC3)
1	SFGdor.L	0.459	**0.014**	0.060
2	SFGdor.R	**0.008**	0.839	0.971
3	MFG.L	0.539	**0.003**	**0.033**
4	IFGoperc.L	0.670	**0.025**	**0.004**
5	ORBinf.R	0.972	0.308	0.508
6	SMA.R	0.366	0.090	0.448
7	OLF.R	0.269	0.654	0.073
8	ORBsupmed.R	0.571	0.239	**0.048**
9	REC.L	0.790	0.384	0.609
10	INS.L	0.451	0.825	0.841
11	HIP.R	**0.019**	**0.027**	0.294
12	CAL.L	0.509	0.209	**0.021**
13	SOG.R	0.114	0.192	0.164
14	MOG.L	0.504	**0.025**	**0.017**
15	MOG.R	0.575	0.056	**0.048**
16	PoCG.R	0.965	0.100	**0.029**
17	SPG.L	0.106	**0.001**	**0.012**
18	SPG.R	0.963	0.756	0.209
19	IPL.L	0.715	0.123	0.247
20	ANG.L	**0.049**	0.465	0.685
21	PCUN.L	0.288	0.099	0.068
22	PCL.L	0.325	0.130	0.098
23	CAU.L	**0.031**	0.103	**0.007**
24	CAU.R	0.858	0.283	0.081
25	PAL.L	0.343	0.113	**0.024**
26	STG.R	0.306	0.100	0.137
27	ITG.R	0.241	0.257	0.121

identified most discriminative regions include orbitofrontal cortex, frontal gyri, rectus gyrus, temporal gyri, insula, hippocampus, parietal gyrus, angular gyrus, precuneus, occipital gyri, caudate nucleus, and paracentral lobule, which have been found to be associated with MCI pathology in the previous studies (Bokde et al., 2006; Feng et al., 2012; Kim et al., 2010; Mascalchi et al., 2014; Wee et al., 2016; Zhu et al., 2016). On the other hand, as shown in Table 2, most of the *p*-values of those selected most discriminative features are smaller than 0.05, indicating that those selected features have good discriminative power. This also partly explains why our proposed method shows an excellent classification performance.

The bold part indicates that the corresponding *p*-values are less than 0.05.

3.3. Analysis of connectivity on the most discriminative regions

We further analyze the connectivity among brain regions to further understand the pathology of MCI by computing the average hyper-edges based on the selected regions for each group (i.e., MCI group and NC group), as disruption of functional connectivity in the brain network may suggest the aberrant of healthy brain functions and activities. Specifically, we repeat the following steps to construct hyper-edges of each group for each selected region listed in Table 2. First, we construct 90 hyper-edges (for the case of totally 90 brain regions used in this paper) for each subject by using Eq. (2), in which λ is a fixed value and the penalty weight F_{ij} is decided by setting a fixed parameter ϖ in Eq. (4). Then, for each group, we compute the number of occurrence of each region in the hyper-edges that share the same centroid node. Next, we calculate the average degree of the hyper-edges for all subjects, which is denoted as d. Finally, for each group, in the hyper-edges whose centroid node belongs to the selected regions, top d regions with the highest occurrence number are selected to construct the corresponding average hyper-edge. Here, we round d to the nearest integer greater than or equal to d, since the degree of the average hyper-edges should be an integer. Fig. 5 graphically illustrates the average hyper-edges constructed on 8 selected regions with p-values smaller than 0.05 (as the centroid node of the average hyper-edges) between the MCI group and the NC group for the fixed $\lambda = 0.5$ and $\varpi = 0.006$. Here, $\lambda = 0.5$ and $\varpi = 0.006$ are the optimal parameters determined by a grid search for constructing hyper-networks. Each average hyper-edge is shown in a subfigure with the red nodes denoting the centroid nodes and linked to other nodes (i.e., the blue nodes). As shown in Fig. 5, there are obvious differences in terms of the hyper-network structure between the MCI group and the NC group.

Additionally, a connectivity graph (CG) is induced from the constructed hyper-networks to graphically analyze the connectivity among brain regions. Here, nodes in the CG represent brain regions and the connectivity weight between each pair of nodes in the CG equal to the number of adjacent hyper-edges between each pair of nodes in the constructed hyper-network of each subject. Therefore, larger connectivity weight between a pair of nodes indicates that this pair of nodes involves more in hyper-edges. For each subject, we construct a CG by the constructed hyper-network. Then, we perform the standard two-sample *t*-test to evaluate the difference of each connection in CG between NC and MCI groups. Connections with their *p*-value smaller than 0.05 are considered as the most discriminative connections and are shown in Fig. 6. Fig. 6(a) graphically shows the *p*-values of all connections in the CG, and Fig. 6(b) shows the connections with *p*-value smaller than 0.05 (i.e., the *p*-values larger than 0.05 are set to 1). As shown in Fig. 6(a) and (b), most of the connections are concentrated in some specific brain regions, including frontal gyrus, amygdala, occipital gyrus, inferior parietal gyri, precuneus, paracentral lobule, and temporal gyri. Table 3 provides the top 10 regions with the highest occurrence frequency. Note that the regions listed in Table 3 were identified to be associated with MCI pathology (Bozzali et al., 2006; Nobili et al., 2010; Wee et al., 2012b; Xu et al., 2007) and are partly overlapped with the regions selected in the classification (i.e., brain regions in Table 2), indicating the efficacy of our proposed method.

3.4. Subject consistency

Since subjects within the same population should have similar brain connection structure, the subject consistency of the constructed networks can be used as an indicator to estimate the efficacy of the brain network construction method (Varoquaux et al., 2010). Hence, the higher the subject consistency of the constructed networks, the better the network construction method is. To quantify how similar the functional (hyper-)connectivity of subjects from the same group, we employ a commonly used metric, i.e., Dice Similarity Coefficient (DSC), as defined as follows:

$$DSC = \frac{2TP}{(2TP + FP + FN)}. \tag{12}$$

where $TP, FP,$ and FN are the true positive, false positive, and false negative, respectively. High DSC indicates high subject consistency in the brain connection structure. A high DSC indicates a high consistency of network topology among subjects from the same population. We calculate the DSC value of the hyper-networks constructed by our proposed method and the DSC values of six comparison methods (i.e., LASSO-ASL, LASSO-BOLD, UOLS-ASL-constrained-BOLD, Ultra-GroupLASSO-ASL-constrained-BOLD, BOLD-constrained-ASL, and WSGR). Fig. 7 shows the DSC values of NC and MCI populations for the proposed method and the comparison methods.

There is no significant difference in terms of DSC values between MCI and NC populations for all methods. The DSC value of our proposed method is higher than the DSC values of all the competing methods, demonstrating that integrating ASL-fMRI and BOLD-fMRI data based on the Ultra-GroupLASSO-UOLS algorithm and FW-LASSO algorithm increases subject consistency in

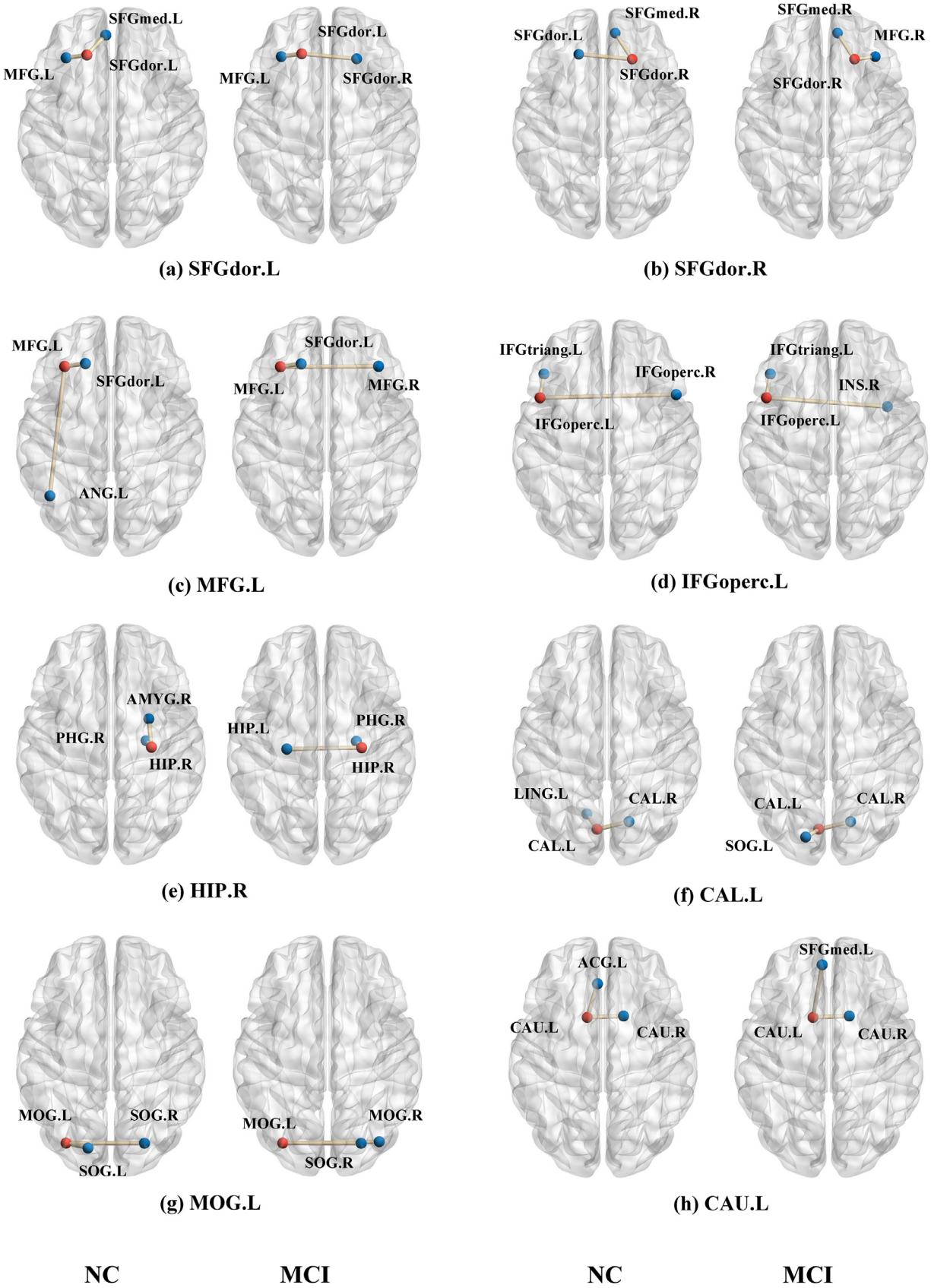

Fig. 5. The average degree of hyper-edges for NC and MCI for 8 brain regions listed in Table 2. Each sub-figure represents a hyper-edge between the corresponding brain region (indicated by the red node) and other nodes. The average degree of hyper-edges for a node is computed from the top d ROIs with the highest occurrence number among all subjects.

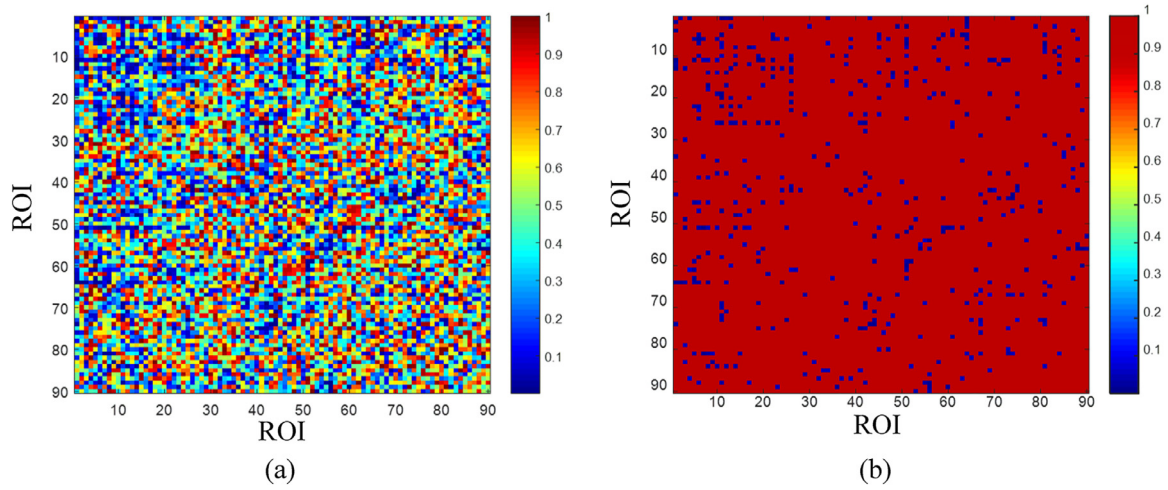

Fig. 6. Visualization of *p*-values on connections between regions. Colors denote the corresponding *p*-value. (a) The *p*-values on connections between regions. (b) The connections with the *p*-value less than 0.05.

Table 3
The top 10 regions with the highest occurrence frequency in the discriminative connections and their frequency of occurrence.

No	Regions	Frequency of occurrence	No	Regions	Frequency of occurrence
1	IFGoperc.L	16	6	STG.L	11
2	MOG.L	15	7	AMY.R	10
3	ORBsup.R	13	8	CUN.L	10
4	MFG.R	11	9	SMG.R	10
5	ORBsupmed.R	11	10	PCUN.L	9

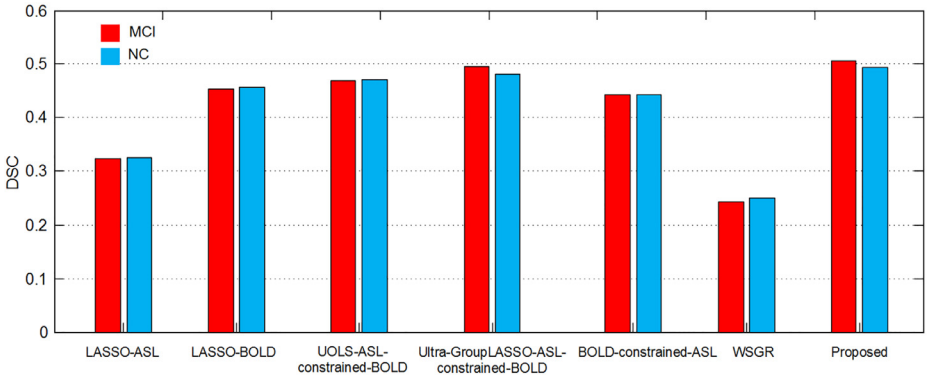

Fig. 7. DSC values for the proposed method and the comparison methods.

the learned connection structure. Furthermore, the DSC values of the ASL fMRI constrained multimodal methods (i.e. UOLS-ASL-constrained-BOLD, Ultra-GroupLASSO-ASL-constrained-BOLD, and the proposed method) is higher than the DSC value of BOLD fMRI constrained multimodal method (i.e. BOLD-constrained-ASL), indicating that the intra-group consistency of the hyper-network structure generated by using ASL fMRI as a constraint in the multimodal integration process is higher than the case of using BOLD fMRI. This is reasonable since the ASL fMRI, which measures cerebral blood flow (CBF) by using arterial water as an endogenous tracer, enjoys marked advantages in inter-subject consistency and reliability relative to BOLD fMRI (Weber et al., 2013).

4. Discussion

4.1. Significance of results

As the basic step of brain network analysis, accurate construction of brain network enables a deeper understanding of the structural and functional organizations of the brain. Since different modality characterizes the brain from different perspectives, some studies have been resorted to constructing multimodal brain networks that can provide a more comprehensive characterization of the brain. However, most of the existing multimodal network modeling methods are solely based on the conventional pair-wise correlation networks, which only characterize the low-order interactions between paired brain regions. We therefore propose a new method to construct the multimodal hyper-connectivity network for the brain disease diagnosis. In our method, the multimodal hyper-network is constructed based on BOLD fMRI and ASL perfusion MRI using a newly derived FW-LASSO regression algorithm, where the weights of the sparse penalty are obtained based on the functional connectivity derived from ASL time series. Different from the existing methods, our proposed method can *not only* characterize high-order interactions among multiple brain regions, *but also* integrate the complementary information from different modalities. The experimental results on a MCI dataset sug-

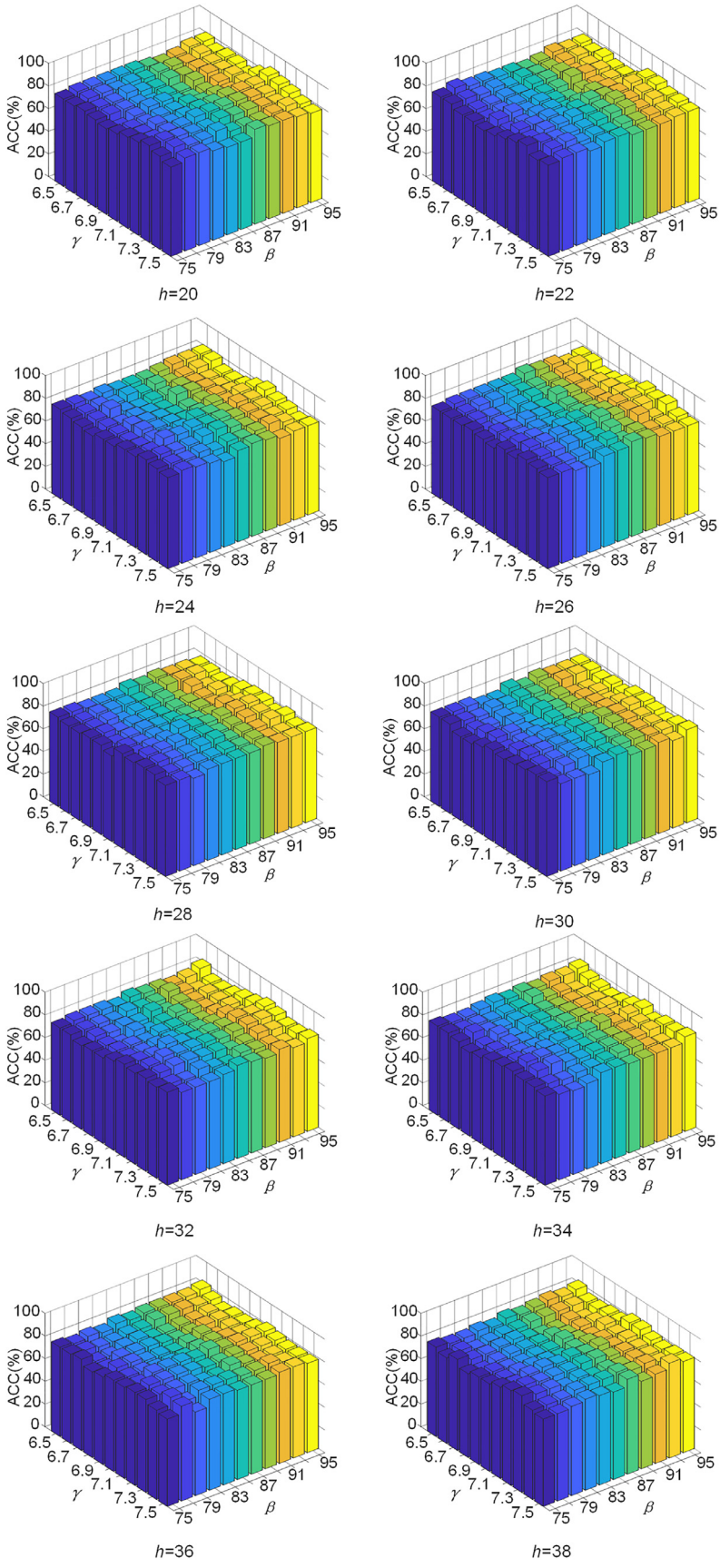

Fig. 8. The classification performance with different combinations of β, γ and h values.

gest that, compared to the existing pairwise connectivity-network-based methods, our proposed method can significantly improve the performance of brain disease classification.

Through these results shown in Table 1, we demonstrate the superiority of our proposed method from three aspects. First, the Ultra-GroupLASSO-ASL method outperforms the LASSO-ASL method, implying the advantages of considering both the group structure in the network and the derivative information of the fMRI time series during the ASL-derived functional connectivity construction. Second, all the classification performance of the Ultra-GroupLASSO-ASL-constrained-BOLD method and the UOLS-ASL-constrained-BOLD method are inferior to our proposed method, demonstrating that the combination of the Ultra-GroupLASSO and the UOLS algorithm can provide a more accurate representation that eventually facilitates the MCI classification. This is reasonable since the estimated coefficients based on Ultra-GroupLASSO algorithm alone are biased as the result of group-constrained sparse penalization, while only using the UOLS algorithm inevitably generates different network topologies for different subjects, which may degrade the generalization performance of the classifiers. Third, our proposed method significantly outperforms the BOLD-constrained-ASL method in all performance measures, indicating that the ASL fMRI constrained BOLD fMRI hyper-network performs better than the BOLD fMRI constrained ASL fMRI hyper-network. This can be interpreted as the ASL fMRI, which has higher inter-subject consistency, can provide a more reliable constraint for the hyper-network construction than BOLD fMRI. Moreover, the LASSO-BOLD method outperforms the LASSO-ASL method, suggesting that constructing hyper-networks based on BOLD fMRI can represent functional interactions among multiple brain regions more precise than ASL fMRI in this paper. To further verify the effect of the ASL-derived functional connectivity constraint on the hyper-edges construction, we compare DSC of hyper-networks modeled by our proposed method with that of the BOLD-fMRI method. DSC of our proposed method is higher than that of the method without the ASL-derived functional connectivity constraint (i.e., the LASSO-BOLD method). This result suggests that the method with the ASL-derived functional connectivity information can model the brain functional connectivity with larger subject consistency. This result demonstrates the benefits of multimodal integration for hyper-connectivity estimation and the effectiveness of employing the ASL-derived functional connectivity information to constrain the BOLD-fMRI-based hyper-network construction. Overall, these results complement each other and further indicate the efficacy of our proposed method.

In addition, our findings also suggest that the brain regions identified by our proposed method are highly relevant to MCI pathology as reported in previous studies, including orbitofrontal cortex (Khazaee et al., 2014), frontal gyri (Bell-McGinty et al., 2005; Liu et al., 2013; Zhang et al., 2015), rectus gyrus (Fleisher et al., 2009), temporal gyri (Bokde et al., 2006), insula (Fei et al., 2014), hippocampus (Bell-McGinty et al., 2005), parietal gyrus (Bell-McGinty et al., 2005; Bokde et al., 2006), angular gyrus (Bokde et al., 2006; Khazaee et al., 2014), precuneus (Bozzali et al., 2006), occipital gyri (Nobili et al., 2010), caudate nucleus (Khazaee et al., 2014), and paracentral lobule (Wang et al., 2012). On the other hand, through analyzing the interaction of the selected brain regions, we found that the hyper-network structure of the MCI group is obviously different from that of the NC group. Alterations of the hyper-network structure may indicate a disruption in the brain functional organization in MCI patients. In particular, the pattern of alteration in functional connectivity involving the hippocampus may provide clues on the underpinnings of cognitive deficit in MCI (Schapiro et al., 2016). Further analysis on the connectivity of hyper-networks demonstrates that the significantly altered connections observed in MCI patients are primarily located on the specific brain regions that are usually linked to AD and MCI pathology (Yao et al., 2013). Moreover, we also find the increased functional connectivity within each brain lobe and the decreased functional connectivity between brain lobes in MCI patients, in line with other resting-state fMRI based AD study (Wang et al., 2007).

4.2. Regularization parameters λ and ϖ in network construction

In our proposed method, a group of multimodal hyper-networks is constructed using the FW-LASSO defined in Eq. (2), with multiple λ values (or sparsity level) to capture multi-level interactions among brain regions. Besides, in the Ultra-GroupLASSO algorithm defined by Eq. (4), the regularization parameter ϖ can control the sparsity of the ASL-derived connectivity network. The network will be dense and susceptible to noise when a small ϖ is used, while a large ϖ will make the network have the relatively sparse connectivity. Selecting the suitable value of ϖ can help to filter out spurious connections in the ASL-derived functional connectivity networks, thereby providing a reliable prior knowledge for the construction of hyper-network using FW-LASSO regression algorithm on BOLD fMRI data. As the parameters λ and ϖ have a great impact on the construction of multimodal hyper-network, the identification of suitable network model parameters is crucial for achieving a good classification performance. Thus, a grid search is employed to find the optimal values of parameters λ and ϖ. As shown in Fig. 3, the classification accuracy can be improved with the increase of the number of λ values with fixed ϖ. This result suggests that the hyper-networks, which contain more multi-level interaction information, reflect greater structural differences between the MCI group and the NC group, indicating the advantage of using hyper-networks with multiple sparsity levels. Also, Fig. 3 shows that, with a fixed λ value, the classification performance is largely affected by the change of ϖ value, suggesting the importance of selecting the optimal ϖ value for final classification. Actually, this is reasonable since the parameter ϖ controls the sparsity of the ASL-derived functional networks and hence the structure of the finally constructed hyper-networks.

4.3. Parameters h, β and γ in feature selection

In this study, we perform a M2TFS method to select the most discriminative features, which includes two regularization items, i.e. a Laplacian regularization term and a group sparsity regularizer. The regularization parameters β and γ balance the relative contributions of the Laplacian regularization term and the group sparsity regularizer. The Laplacian regularization term preserves the discriminative information of the data from each type of network topological features (i.e., HCC^1, HCC^2 and HCC^3), and thus can induce more discriminative features for classification (Jie et al., 2015). The group sparsity regularizer is used to eliminate the irrelevant and redundant features while encourage different types of features of the same vertex to be jointly selected. A larger value of γ leads to a less number of the selected features. In the M2TFS method, h is a parameter controlling the magnitude of similarity of subjects. A larger h value results in a larger similarity value of subjects.

We perform a series of experiments to investigate how the parameters of M2TFS (i.e. β, γ and h) jointly affect the classification performance of our proposed method. Specifically, we calculate the classification accuracies with varied parameters (i.e., $\beta = [75 : 2 : 95]$, $\gamma = [6.5 : 0.1 : 7.5]$, $h = [20 : 2 : 38]$). Fig. 8 provides the ACCs with respect to different combinations of β, γ and h values. The classification accuracy changes smoothly with varied β, γ and h, demonstrating the robustness of the proposed framework. The optimal classification performance (ACC = 86.9%) is achieved at multiple combinations of β, γ and h, including (79, 7.2, 24), (83,

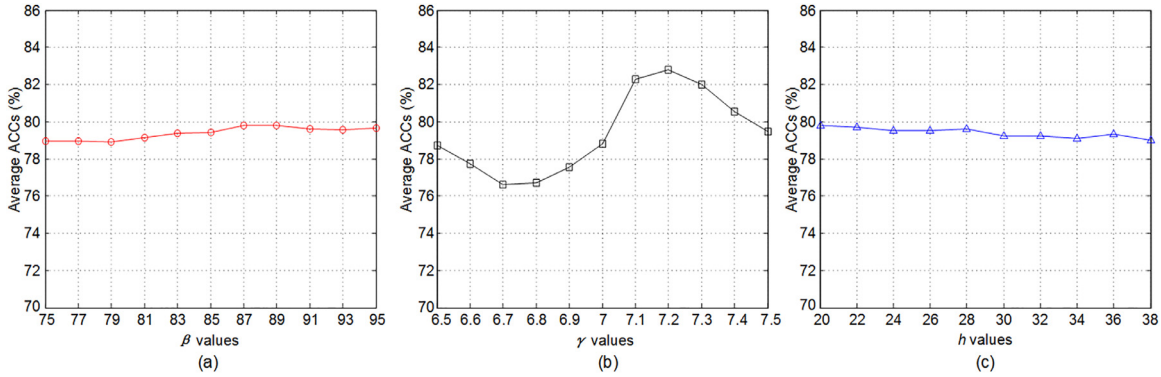

Fig. 9. The average ACCs for different parameters values, (a) β, (b) γ and (c) h.

7.1, 28), (83, 7.2, 24), (85, 7.3, 24), (87, 7.2, 26), (89, 7.2, 22), (95, 7.1, 20), indicating that the optimal performance is not achieved by chance. We further calculate the average ACCs of each β, γ and h value and the classification results are illustrated in Fig. 9. The average ACCs of each β and h value are in the relatively narrow range of [78.9%, 79.9%] and [79.0%, 79.8%], respectively. This indicates that the classification performance is relatively stable with respect to the change of β and h. Furthermore, the average ACC of more than 82% can be achieved over a range of $0.71 \leq \gamma \leq 0.73$. When γ becomes too small or too large, the classification accuracy decreases. This is reasonable since γ controls the sparsity of the feature selection process and determines the scale of the selected feature subset for classification. When γ is too small, the irrelevant and redundant features may be reserved, thus deteriorating the generalization performance of the classifier. Meanwhile, when γ is too large, some discriminative features may be excluded, causing the loss of the critical classification information.

4.4. Test-retest reliability of features

In this study, we assess the test-retest reliability of three types of network topological features (i.e., HCC^1, HCC^2 and HCC^3) separately using the mean IntraClass Correlation Coefficient (ICC) (Koo and Li, 2016). The ICC of each type of features is calculated when using the first half and second half of the BOLD fMRI time series. The experimental results demonstrate that all the features used in our proposed method are able to achieve large ICC values (i.e., ICC of $HCC^1 = 0.73 \pm 0.12$, ICC of $HCC^2 = 0.81 \pm 0.13$, and ICC of $HCC^3 = 0.83 \pm 0.04$, respectively), indicating excellent reliability of these features. Furthermore, we also calculate the ICC of these features from nine multimodal hyper-networks constructed with nine sets of λ values (i.e., $\{0.1\}$, $\{0.1, 0.2\}$, $\ldots$, $\{0.1, 0.2, \ldots, 0.9\}$), respectively. The experimental results shown in Fig. 10 indicate that (1) ICC value of HCC^1 increases with the number of λ values included and becomes relatively stable when the number of λ values included is larger than 7, (2) ICC values of HCC^2 and HCC^3 are relatively robust to the number of λ values.

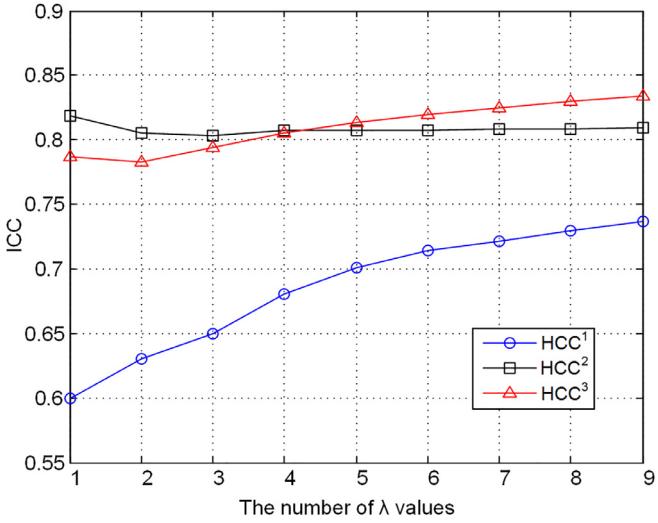

Fig. 10. ICC values of three types of features with the different number of λ values.

4.5. Hyper-connectivity analysis

The hyper-connectivity inferred using ASL fMRI data and BOLD fMRI data simultaneously measures the correlation of the cerebral blood flow (CBF) and the blood-oxygenation-level-dependent (BOLD) signals among multiple brain regions, providing a high-order characterization of the interactions within the whole-brain network. As shown in Fig. 5, there are some obvious differences in terms of the hyper-connectivity between the MCI group and the NC group. For example, in Fig. 5(e), the right hippocampus (HIP.R) is connected to the left hippocampus (HIP.L) and right parahippocampal gyrus (PHG.R) in MCI, while it is connected to the right parahippocampal gyrus (PHG.R) and right amygdala (AMYG.L) in NC. This can be interpreted as the sub-network of HIP.R, PHG.R, and AMYG.L for the MCI patients is disrupted by the pathological attacks. The connection between the HIP.R and AMYG.L in this sub-network is broken, while HIP.R is connected to the HIP.L for compensating the loss of network efficiency (Jie et al., 2016). It has been reported that AMYG plays a primary role in the processing of memory, emotional responses and decision making. The volume reduction of AMYG is considered as a very early manifestation of AD/MCI (Striepens et al., 2010; Yao et al., 2014). Previous works also reported that the HIP is a key brain region for memory (Yue et al., 2018). The atrophy in the HIP is believed to be an early biomarker of AD/MCI, implying that our findings are in line with the previous studies (Pereira et al., 2016; McDonald and Mott, 2017).

Noted that all brain regions in this hyper-connectivity (i.e. HIP, PHG, and AMYG) belong to the limbic system, which supports some important functions including long-term memory, cognition, behavior and emotion (Nishijo et al., 2018). The damage to the structure of the limbic system may be a clinical symptom of AD/MCI (Li et al., 2016). Furthermore, the left and right superior frontal gyri (dorsolateral) (SFGdor) are both centroid brain regions in the abnormal hyper-connectivity displayed in Fig. 5 (SFGdor.L in Fig. 5(a) and SFGdor.R in Fig. 5(b)). Therefore, the SFGdor is probably an important brain area associated with AD/MCI pathology.

However, further exploration is required in the future to better understand the role of the SFGdor in human cognitive function.

4.6. Limitations of the proposed method

Although the classification results of this study using the multimodal hyper-networks are encouraging, there are four major limitations. (1) Our proposed method ignores the influence of structural connectivity information on the functional brain network construction. A variety of neuroscience research studies have reported that axonal fiber connections are the structural substrates of functional interactions, and stronger structural connections among ROIs indicate higher functional interactions (Honey et al., 2009; Li et al., 2010; Zhu et al., 2012). As a future work, we will introduce anatomical constraint during the construction of multimodal hyper-network. (2) To obtain a reliable constraint for the hyper-network construction, a time-consuming complex method is employed to derive the functional connectivity information from ASL fMRI in this paper. In the future work, we will employ a more efficient functional connectivity estimation algorithm to construct the ASL-derived functional connectivity constraint. (3) In our study, when the FW-LASSO is used for integrating different neuroimaging modalities, there should be a certain prior relationship between these neuroimaging modalities similar to the relationship between BOLD fMRI and ASL fMRI. But this relationship doesn't exist between most neuroimaging modalities. Hence, it will be interesting to explore the integration of neuroimaging modalities other than ASL fMRI and BOLD fMRI and derive appropriate integration method that able to fully utilize the complementary information of multiple modalities. (4) We used a relatively small number of subjects in the current study. Nevertheless, the obtained results do provide evidence on the efficacy of the proposed method for distinguishing MCI individuals from healthy elderly subjects. In the future, we will validate the effectiveness of the proposed framework using a larger dataset.

5. Conclusion

In this paper, we propose a novel multimodal hyper-network modeling method for brain disease diagnosis. The proposed multimodal hyper-connectivity networks encode complementary information conveyed by multiple modalities, and thus provide a more comprehensive representation of the brain functional organization. We demonstrate the superiority of our proposed method by applying it for MCI classification. For the purpose of integrating complementary information from multiple modalities and high-order interaction information among different brain regions, we employ a novel FW-LASSO regression algorithm to construct multimodal hyper-networks from BOLD fMRI and ASL fMRI. Our major methodological contribution is that, by introducing the ASL-derived functional connectivity constraint, we can effectively examine the BOLD-fMRI-based functional dependences of brain regions within the brain hyper-network. In addition, in order to acquire the ALS-derived functional connectivity constraint for the FW-LASSO regression algorithm, the combination of Ultra-GroupLASSO and UOLS algorithm is employed to extract the connectivity strength between brain regions from the ALS fMRI data. Specifically, the GroupLASSO regression algorithm achieves a better generalization performance in the brain disease classification, since the group constraint ensures that the network topology tends to be identical for all individuals while still preserving individual information via different connectivity strength values. Moreover, both the Ultra-GroupLASSO algorithm and the UOLS algorithm with the ULS criterion overcome the overfitting problem, which is common in the conventional least squares criterion, by further extracting the dependent relation of the associated weak derivatives. The experimental results on a MCI dataset suggest that our proposed method can improve the classification performance of brain disease.

Acknowledgments

This work was supported by the National Natural Science Foundation of China [61671042, 61403016, 61573023], Beijing Natural Science Foundation [4172037], Open Fund Project of Fujian Provincial Key Laboratory in Minjiang University [MJUKF201702], and Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/I011056/1 and Platform Grant EP/H00453X/1, U.K.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This study was approved by the local ethical committee.

Informed consent

Informed consent was obtained from all individual participants included in the study.

Conflict of interest

The authors declare that they have no conflicts of interest.

Appendix A

The strength estimation of ASL-derived functional connectivity networks using UOLS

Inputs: ASL fMRI time series of the i-th ROI ψ_i and its correlated Q ROIs $\psi_{j_q} = \{1, 2, \cdots, Q\}$;
Output: $\boldsymbol{Corr}$ ($\boldsymbol{Corr}$ is the i-th row of the ASL-derived functional connectivity networks)
$index = \{1, 2, \cdots, Q\}, \boldsymbol{Corr} = zero(P, 1)$
$T^l_{\psi_i}(\tau) = \int_{\tau}^{T_0+\tau} \psi_i(t) D^l \varphi(t-\tau) dt$
$\boldsymbol{\psi} = \psi_i \cup \{T^l_{\psi_i}(\tau) | l = 1, \cdots, L, \tau = 0, 1, \cdots, T - T_0\}$
for all $q \in index$ **do**
    $T^l_{\psi_{j_q}}(\tau) = \int_{\tau}^{T_0+\tau} \psi_{j_q}(t) D^l \varphi(t-\tau) dt$
    $\boldsymbol{\omega}_q = \psi_{j_q} \cup \{T^l_{\psi_{j_q}}(\tau) | l = 1, \cdots, L, \tau = 0, 1, \cdots, T - T_0\}$
end for
while $index \neq []$ **do**
    for all $q \in index$ **do**
        $err_q = \frac{\langle \omega_q, \boldsymbol{\psi} \rangle^2}{\langle \omega_q, \omega_q \rangle \langle \boldsymbol{\psi}, \boldsymbol{\psi} \rangle}$, where $\langle \boldsymbol{\omega}_q, \boldsymbol{\psi} \rangle$ is the inner product of vectors $\boldsymbol{\omega}_q$ and $\boldsymbol{\psi}$
    end for
    $maxerr = 0, k = 0$
    for all $q \in index$ **do**
        if $err_q > maxerr$ then
            $maxerr = err_q, \ k = q$
        end if
    end for
    $Corr_{n_k} = maxerr, index = index \backslash \{k\}$
    for all $q \in index$ **do**
        $\boldsymbol{\omega}_q = \boldsymbol{\omega}_q - \frac{\langle \omega_q, \omega_k \rangle}{\langle \omega_k, \omega_k \rangle} \boldsymbol{\omega}_k$
    end for
end while

References

Alsop, D.C., Detre, J.A., Golay, X., Günther, M., Hendrikse, J., Hernandez-Garcia, L., Lu, H., MacIntosh, B.J., Parkes, L.M., Smits, M., 2015. Recommended implementation of arterial spin-labeled perfusion MRI for clinical applications: A consensus of the ISMRM perfusion study group and the European consortium for ASL in dementia. Magn. Reson. Med. 73, 102–116.

Association, A.s., 2016. 2016 Alzheimer's disease facts and figures. Alzheimers Dement. 12, 459–509.
Barker, W.W., Luis, C.A., Kashuba, A., Luis, M., Harwood, D.G., Loewenstein, D., Waters, C., Jimison, P., Shepherd, E., Sevush, S., 2002. Relative frequencies of Alzheimer disease, Lewy body, vascular and frontotemporal dementia, and hippocampal sclerosis in the State of Florida Brain Bank. Alzheimer Dis. Assoc. Disord. 16, 203–212.
Bell-McGinty, S., Lopez, O.L., Meltzer, C.C., Scanlon, J.M., Whyte, E.M., DeKosky, S.T., Becker, J.T., 2005. Differential cortical atrophy in subgroups of mild cognitive impairment. Arch. Neurol. 62, 1393–1397.
Bokde, A., Lopez-Bayo, P., Meindl, T., Pechler, S., Born, C., Faltraco, F., Teipel, S., Möller, H.-J., Hampel, H., 2006. Functional connectivity of the fusiform gyrus during a face-matching task in subjects with mild cognitive impairment. Brain 129, 1113–1124.
Borogovac, A., Asllani, I., 2012. Arterial spin labeling (ASL) fMRI: advantages, theoretical constrains and experimental challenges in neurosciences. Int. J. Biomed. Imaging 2012, 818456.
Bozzali, M., Filippi, M., Magnani, G., Cercignani, M., Franceschi, M., Schiatti, E., Castiglioni, S., Mossini, R., Falautano, M., Scotti, G., 2006. The contribution of voxel-based morphometry in staging patients with mild cognitive impairment. Neurology 67, 453–460.
Brier, M.R., Thomas, J.B., Fagan, A.M., Hassenstab, J., Holtzman, D.M., Benzinger, T.L., Morris, J.C., Ances, B.M., 2014. Functional connectivity and graph theory in preclinical Alzheimer's disease. Neurobiol. Aging 35, 757–768.
Cavusoglu, M., Bartels, A., Yesilyurt, B., Uludağ, K., 2012. Retinotopic maps and hemodynamic delays in the human visual cortex measured using arterial spin labeling. Neuroimage 59, 4044–4054.
Chatterjee, A., Lahiri, S., 2013. Rates of convergence of the adaptive LASSO estimators to the oracle distribution and higher order refinements by the bootstrap. Ann. Stat. 41, 1232–1259.
Chen, G., Ward, B.D., Xie, C., Li, W., Wu, Z., Jones, J.L., Franczak, M., Antuono, P., Li, S.-J., 2011. Classification of Alzheimer disease, mild cognitive impairment, and normal cognitive status with large-scale network analysis based on resting-state functional MR imaging. Radiology 259, 213–221.
Davison, E.N., Schlesinger, K.J., Bassett, D.S., Lynall, M.E., Miller, M.B., Grafton, S.T., Carlson, J.M., 2015. Brain network adaptability across task states. PLoS Comput. Biol. 11, 14.
Efron, B., Hastie, T., Johnstone, I., Tibshirani, R., 2004. Least angle regression. Ann. Stat. 32, 407–499.
Fei, F., Jie, B., Zhang, D., 2014. Frequent and discriminative subnetwork mining for mild cognitive impairment classification. Brain Connect. 4, 347–360.
Feng, Y., Bai, L., Ren, Y., Chen, S., Wang, H., Zhang, W., Tian, J., 2012. FMRI connectivity analysis of acupuncture effects on the whole brain network in mild cognitive impairment patients. Magn. Reson. Imaging 30, 672–682.
Fleisher, A.S., Sherzai, A., Taylor, C., Langbaum, J.B., Chen, K., Buxton, R.B., 2009. Resting-state BOLD networks versus task-associated functional MRI for distinguishing Alzheimer's disease risk groups. Neuroimage 47, 1678–1690.
Friston, K.J., Williams, S., Howard, R., Frackowiak, R.S., Turner, R., 1996. Movement-related effects in fMRI time-series. Magn. Reson. Med. 35, 346–355.
Ganmor, E., Segev, R., Schneidman, E., 2011. Sparse low-order interaction network underlies a highly correlated and learnable neural population code. Proc. Natl. Acad. Sci. 108, 9679–9684.
Gao, Y., Wee, C.-Y., Kim, M., Giannakopoulos, P., Montandon, M.-L., Haller, S., Shen, D., 2015. MCI identification by joint learning on multiple MRI data. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 78–85.
Gu, S., Yang, M.Z., Medaglia, J.D., Gur, R.C., Gur, R.E., Satterthwaite, T.D., Bassett, D.S., 2017. Functional hypergraph uncovers novel covariant structures over neurodevelopment. Hum. Brain Mapp. 38, 3823–3835.
Guo, Y., Guo, L.Z., Billings, S.A., Wei, H.L., 2016. Ultra-orthogonal forward regression algorithms for the identification of non-linear dynamic systems. Neurocomputing 173, 715–723.
Havlicek, M., Roebroeck, A., Friston, K., Gardumi, A., Ivanov, D., Uludag, K., 2015. Physiologically informed dynamic causal modeling of fMRI data. NeuroImage 122, 355–372.
Hayasaka, S., Laurienti, P.J., 2010. Comparison of characteristics between region-and voxel-based network analyses in resting-state fMRI data. Neuroimage 50, 499–508.
Honey, C., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.-P., Meuli, R., Hagmann, P., 2009. Predicting human resting-state functional connectivity from structural connectivity. Proc. Natl. Acad. Sci. 106, 2035–2040.
Huang, S., Li, J., Sun, L., Ye, J., Fleisher, A., Wu, T., Chen, K., Reiman, E., Initiative, A.s.D.N., 2010. Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation. NeuroImage 50, 935–949.
Jann, K., Gee, D.G., Kilroy, E., Schwab, S., Smith, R.X., Cannon, T.D., Wang, D.J., 2015. Functional connectivity in BOLD and CBF data: similarity and reliability of resting brain networks. Neuroimage 106, 111–122.
Jie, B., Wee, C.-Y., Shen, D., Zhang, D., 2016. Hyper-connectivity of functional networks for brain disease diagnosis. Med. Image Anal. 32, 84–100.
Jie, B., Zhang, D.Q., Cheng, B., Shen, D.G., 2015. Manifold regularized multitask feature learning for multimodality disease classification. Hum. Brain Mapp. 36, 489–507.
Kaiser, M., 2011. A tutorial in connectome analysis: topological and spatial features of brain networks. Neuroimage 57, 892–907.
Kantarci, K., Weigand, S., Przybelski, S., Shiung, M., Whitwell, J.L., Negash, S., Knopman, D.S., Boeve, B.F., O'Brien, P., Petersen, R.C., 2009. Risk of dementia in MCI Combined effect of cerebrovascular disease, volumetric MRI, and 1H MRS. Neurology 72, 1519–1525.
Khazaee, A., Ebrahimzadeh, A., Babajani-Feremi, A., 2014. Automatic classification of Alzheimer's disease with resting-state fMRI and graph theory. In: Biomedical Engineering (ICBME), 2014 21th Iranian Conference on. IEEE, pp. 252–257.
Kim, S.H., Seo, S.W., Yoon, D.S., Chin, J., Lee, B.H., Cheong, H.-K., Han, S.-H., Na, D.L., 2010. Comparison of neuropsychological and FDG-PET findings between early-versus late-onset mild cognitive impairment: a five-year longitudinal study. Dement. Geriatr. Cognit. Disord. 29, 213–223.
Koo, T.K., Li, M.Y., 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J. Chiropr. Med. 15, 155–163.
Li, K., Guo, L., Faraco, C., Zhu, D., Deng, F., Zhang, T., Jiang, X., Zhang, D., Chen, H., Hu, X., 2010. Individualized ROI optimization via maximization of group-wise consistency of structural and functional profiles, Advances in. Neural Information Inf. Processing Process. SystemsSyst., pp. 1369–1377.
Li, K., Liu, Y., Wang, Q., Wu, Y., Song, S., Sun, Y., Liu, T., Wang, J., Li, Y., Du, S., 2015. A spacecraft electrical characteristics multi-label classification method based on off-line FCM clustering and on-line WPSVM. PLoS One 10, e0140395.
Li, X.S., Wang, H.B., Tian, Y.H., Zhou, S.S., Li, X.H., Wang, K., Yu, Y.Q., 2016. Impaired white matter connections of the limbic system networks associated with impaired emotional memory in Alzheimer's Disease. Front. Aging Neurosci. 8, 14.
Li, Y., Cui, W.-G., Guo, Y.-Z., Huang, T., Yang, X.-F., Wei, H.-L., 2017a. Time-varying system identification using an ultra-orthogonal forward regression and multiwavelet basis functions with applications to EEG. IEEE Trans. Neural Netw. Learn. Syst., 1–13.
Li, Y., Cui, W.-G., Huang, H., Guo, Y.-Z., Li, K., Tan, T., 2018a. Epileptic seizure detection in EEG signals using sparse multiscale radial basis function networks and the Fisher vector approach. Knowledge-Based Syst. doi:10.1016/j.knosys.2018.10.029.
Li, Y., Gao, X., Jie, B., Yap, P.-T., Kim, M.-j., Wee, C.-Y., Shen, D., 2017b. Multimodal Hyper-connectivity Networks For MCI Classification, International Conference On Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 433–441.
Li, Y., Jewells, V., Kim, M., Chen, Y., Moon, A., Armao, D., Troiani, L., Markovic–Plese, S., Lin, W., Shen, D., 2013. Diffusion tensor imaging based network analysis detects alterations of neuroconnectivity in patients with clinically early relapsing-remitting multiple sclerosis. Hum. Brain Mapp. 34, 3376–3391.
Li, Y., Liu, J., Huang, J., Li, Z., Liang, P., 2018b. Learning brain connectivity sub-networks by group- constrained sparse inverse covariance estimation for Alzheimer's Disease classification. Front. Neuroinform. 12.
Li, Y., Wang, X.-D., Luo, M.-L., Li, K., Yang, X.-F., Guo, Q., 2018c. Epileptic seizure classification of EEGs using time – frequency analysis based multiscale radial basis functions. IEEE J. Biomed. Health Inform. 22, 386–397.
Li, Y., Wee, C.-Y., Jie, B., Peng, Z., Shen, D., 2014. Sparse multivariate autoregressive modeling for mild cognitive impairment classification. Neuroinformatics 12, 455–469.
Li, Y., Yang, H., Li, K., Yap, P.-T., Kim, M., Wee, C.-Y., Shen, D., 2017c. Novel effective connectivity network inference for MCI identification. In: International Workshop on Machine Learning in Medical Imaging. Springer, pp. 316–324.
Liang, X., Connelly, A., Calamante, F., 2014. Graph analysis of resting-state ASL perfusion MRI data: nonlinear correlations among CBF and network metrics. Neuroimage 87, 265–275.
Liu, F., Wee, C.-Y., Chen, H., Shen, D., 2014. Inter-modality relationship constrained multi-modality multi-task feature selection for Alzheimer's Disease and mild cognitive impairment identification. NeuroImage 84, 466–475.
Liu, R., Hu, B., Yao, Z., Ratcliffe, M., Wang, W., Liang, C., Cai, Q., Yang, J., Zhao, Q., 2013. Abnormal neural activity and functional connectivity in amnestic mild cognitive impairmet: a resting state fMRI study, Neural Engineering (NER). In: 2013 6th International IEEE/EMBS Conference on. IEEE, pp. 765–769.
Luh, W.-M., Wong, E.C., Bandettini, P.A., Hyde, J.S., 1999. QUIPSS II with thin-slice TI1 periodic saturation: a method for improving accuracy of quantitative perfusion imaging using pulsed arterial spin labeling. Magn. Reson. Med. 41, 1246–1254.
Mascalchi, M., Ginestroni, A., Toschi, N., Poggesi, A., Cecchi, P., Salvadori, E., Tessa, C., Cosottini, M., De Stefano, N., Pracucci, G., 2014. The burden of microstructural damage modulates cortical activation in elderly subjects with MCI and leuko-araiosis. A DTI and fMRI study. Hum. Brain Mapp. 35, 819–830.
McDonald, A.J., Mott, D.D., 2017. Functional neuroanatomy of amygdalohippocampal interconnections and their role in learning and memory. J. Neurosci. Res. 95, 797–820.
Mitchell, A.J., Shiri-Feshki, M., 2009. Rate of progression of mild cognitive impairment to dementia-meta-analysis of 41 robust inception cohort studies. Acta Psychiatr. Scand. 119, 252–265.
Nishijo, H., Rafal, R., Tamietto, M., 2018. Editorial: limbic-brainstem roles in perception, cognition, emotion, and behavior. Front. Neurosci. 12, 3.
Nobili, F., Mazzei, D., Dessi, B., Morbelli, S., Brugnolo, A., Barbieri, P., Girtler, N., Sambuceti, G., Rodriguez, G., Pagani, M., 2010. Unawareness of memory deficit in amnestic MCI: FDG-PET findings. J. Alzheimer's Dis. 22, 993–1003.
Ohiorhenuan, I.E., Mechler, F., Purpura, K.P., Schmid, A.M., Hu, Q., Victor, J.D., 2010. Sparse coding and high-order correlations in fine-scale cortical networks. Nature 466, 617.
Park, H.-J., Friston, K., 2013. Structural and functional brain networks: from connections to cognition. Science 342, 1238411.

Pereira, J.B., Mijalkov, M., Kakaei, E., Mecocci, P., Vellas, B., Tsolaki, M., Kloszewska, I., Soininen, H., Spenger, C., Lovestone, S., Simmons, A., Wahlund, L.O., Volpe, G., Westman, E., AddNeuroMed, C., 2016. Disrupted network topology in patients with stable and progressive mild cognitive impairment and Alzheimer's Disease. Cereb. Cortex 26, 3476–3493.

Petrella, J., Sheldon, F., Prince, S., Calhoun, V., Doraiswamy, P., 2011. Default mode network connectivity in stable vs progressive mild cognitive impairment. Neurology 76, 511–517.

Rosa, M.J., Portugal, L., Hahn, T., Fallgatter, A.J., Garrido, M.I., Shawe-Taylor, J., Mourao-Miranda, J., 2015. Sparse network-based models for patient classification using fMRI. Neuroimage 105, 493–506.

Schapiro, A.C., Turk-Browne, N.B., Norman, K.A., Botvinick, M.M., 2016. Statistical learning of temporal community structure in the hippocampus. Hippocampus 26, 3–8.

Smith, S.M., Vidaurre, D., Beckmann, C.F., Glasser, M.F., Jenkinson, M., Miller, K.L., Nichols, T.E., Robinson, E.C., Salimi-Khorshidi, G., Woolrich, M.W., 2013. Functional connectomics from resting-state fMRI. Trends Cognit. Sci. 17, 666–682.

Sporns, O., 2011. The human connectome: a complex network. Ann. N. Y. Acad. Sci. 1224, 109–125.

Striepens, N., Scheef, L., Wind, A., Popp, J., Spottke, A., Cooper-Mahkorn, D., Suliman, H., Wagner, M., Schild, H.H., Jessen, F., 2010. Volume loss of the medial temporal lobe structures in subjective memory impairment. Dement. Geriatr. Cognit. Disord. 29, 75–81.

Tak, S., Polimeni, J.R., Wang, D.J., Yan, L., Chen, J.J., 2015. Associations of resting-state fMRI functional connectivity with flow-BOLD coupling and regional vasculature. Brain Connect. 5, 137–146.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273–289.

Varoquaux, G., Gramfort, A., Poline, J.-B., Thirion, B., 2010. Brain covariance selection: better individual functional connectivity models using population prior, Advances in. Neural Information Inf. Processing Process. SystemsSyst., pp. 2334–2342.

Wang, J., Zuo, X., Dai, Z., Xia, M., Zhao, Z., Zhao, X., Jia, J., Han, Y., He, Y., 2013. Disrupted functional brain connectome in individuals at risk for Alzheimer's disease. Biol. Psychiatry 73, 472–481.

Wang, K., Liang, M., Wang, L., Tian, L., Zhang, X., Li, K., Jiang, T., 2007. Altered functional connectivity in early Alzheimer's disease: a resting-state fMRI study. Hum. Brain Mapp. 28, 967–978.

Wang, Z., Aguirre, G.K., Rao, H., Wang, J., Fernández-Seara, M.A., Childress, A.R., Detre, J.A., 2008. Empirical optimization of ASL data analysis using an ASL data processing toolbox: ASLtbx. Magn. Reson. Imaging 26, 261–269.

Wang, Z., Jia, X., Liang, P., Qi, Z., Yang, Y., Zhou, W., Li, K., 2012. Changes in thalamus connectivity in mild cognitive impairment: evidence from resting state fMRI. Eur. J. Radiol. 81, 277–285.

Ward, A., Tardiff, S., Dye, C., Arrighi, H.M., 2013. Rate of conversion from prodromal Alzheimer's disease to Alzheimer's dementia: a systematic review of the literature. Dement. Geriatr. Cognit. Disord. Extra 3, 320–332.

Weber, M.J., Detre, J.A., Thompson-Schill, S.L., Avants, B.B., 2013. Reproducibility of functional network metrics and network structure: a comparison of task-related BOLD, resting ASL with BOLD contrast, and resting cerebral blood flow. Cognit. Affect. Behav. Neurosci. 13, 627–640.

Wee, C.-Y., Yang, S., Yap, P.-T., Shen, D., Initiative, A.s.D.N., 2016. Sparse temporally dynamic resting-state functional connectivity networks for early MCI identification. Brain Imaging Behav. 10, 342–356.

Wee, C.-Y., Yap, P.-T., Denny, K., Browndyke, J.N., Potter, G.G., Welsh-Bohmer, K.A., Wang, L., Shen, D., 2012a. Resting-state multi-spectrum functional connectivity networks for identification of MCI patients. PLoS One 7, e37828.

Wee, C.-Y., Yap, P.-T., Zhang, D., Denny, K., Browndyke, J.N., Potter, G.G., Welsh-Bohmer, K.A., Wang, L., Shen, D., 2012b. Identification of MCI individuals using structural and functional connectivity networks. Neuroimage 59, 2045–2056.

Wee, C.-Y., Yap, P.-T., Zhang, D., Wang, L., Shen, D., 2014. Group-constrained sparse fMRI connectivity modeling for mild cognitive impairment identification. Brain Struct. Funct. 219, 641–656.

Wilson, R.S., Segawa, E., Boyle, P.A., Anagnos, S.E., Hizel, L.P., Bennett, D.A., 2012. The natural history of cognitive decline in Alzheimer's disease. Psychol. Aging 27, 1008.

Xu, G., Antuono, P., Jones, J., Xu, Y., Wu, G., Ward, D., Li, S.-J., 2007. Perfusion fMRI detects deficits in regional CBF during memory-encoding tasks in MCI subjects. Neurology 69, 1650–1656.

Yao, H., Liu, Y., Zhou, B., Zhang, Z., An, N., Wang, P., Wang, L., Zhang, X., Jiang, T., 2013. Decreased functional connectivity of the amygdala in Alzheimer's disease revealed by resting-state fMRI. Eur. J. Radiol. 82, 1531–1538.

Yao, H., Zhou, B., Zhang, Z., Wang, P., Guo, Y.e., Shang, Y., Wang, L., Zhang, X., An, N., Liu, Y., 2014. Longitudinal alteration of amygdalar functional connectivity in mild cognitive impairment subjects revealed by resting-state FMRI. Brain Connect. 4, 361–370.

Yu, R., Zhang, H., An, L., Chen, X., Wei, Z., Shen, D., 2016. Correlation-weighted sparse group representation for brain network construction in MCI classification. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 37–45.

Yu, R., Zhang, H., An, L., Chen, X., Wei, Z., Shen, D., 2017. Connectivity strength-weighted sparse group representation-based brain network construction for MCI classification. Hum. Brain Mapp. 38, 2370–2383.

Yu, S., Yang, H., Nakahara, H., Santos, G.S., Nikolić, D., Plenz, D., 2011. Higher-order interactions characterized in cortical activity. J. Neurosci. 31, 17514–17526.

Yue, L., Wang, T., Wang, J.Y., Li, G.J., Wang, J.H., Li, X., Li, W., Hu, M.X., Xiao, S.F., 2018. Asymmetry of hippocampus and amygdala defect in subjective cognitive decline among the community dwelling Chinese. Front. Psychiatry 9.

Zanin, M., Sousa, P., Papo, D., Bajo, R., García-Prieto, J., Del Pozo, F., Menasalvas, E., Boccaletti, S., 2012. Optimizing functional network representation of multivariate time series. Sci. Rep. 2, 630.

Zhang, X., Hu, B., Ma, X., Xu, L., 2015. Resting-state whole-brain functional connectivity networks for mci classification using l2-regularized logistic regression. IEEE Trans. Nanobiosci. 14, 237–247.

Zhu, D., Li, K., Faraco, C.C., Deng, F., Zhang, D., Guo, L., Miller, L.S., Liu, T., 2012. Optimization of functional brain ROIs via maximization of consistency of structural connectivity profiles. NeuroImage 59, 1382–1393.

Zhu, D., Li, X., Jiang, X., Chen, H., Shen, D., Liu, T., 2013. Exploring high-order functional interactions via structurally-weighted LASSO models, International Conference on Information Processing in Medical Imaging. NIH Public Access 13–24.

Zhu, X., Suk, H.-I., Lee, S.-W., Shen, D., 2016. Subspace Regularized Sparse Multitask Learning for Multiclass Neurodegenerative Disease Identification. IEEE Trans. Biomed. Eng. 63, 607–618.

Zhu, X., Suk, H.-I., Wang, L., Lee, S.-W., Shen, D., 2017. A novel relational regularization feature selection method for joint regression and classification in AD diagnosis. Med. Image Anal. 38, 205–214.

Zu, C., Gao, Y., Munsell, B., Kim, M., Peng, Z., Cohen, J.R., Zhang, D., Wu, G., 2018. Identifying disease-related subnetwork connectome biomarkers by sparse hypergraph learning. Brain Imaging Behav. doi:10.1007/s11682-018-9899-8.