

# SCIENTIFIC REPORTS



OPEN

## Hybrid High-order Functional Connectivity Networks Using Resting-state Functional MRI for Mild Cognitive Impairment Diagnosis

Yu Zhang<sup>1</sup>, Han Zhang<sup>1</sup>, Xiaobo Chen<sup>1</sup>, Seong-Whan Lee<sup>1,2</sup> & Dinggang Shen<sup>1,2</sup>

Conventional functional connectivity (FC), referred to as low-order FC, estimates temporal correlation of the resting-state functional magnetic resonance imaging (rs-fMRI) time series between any pair of brain regions, simply ignoring the potentially high-level relationship among these brain regions. A high-order FC based on “correlation’s correlation” has emerged as a new approach for abnormality detection of brain disease. However, separate construction of the low- and high-order FC networks overlooks information exchange between the two FC levels. Such a *higher-level* relationship could be more important for brain diseases study. In this paper, we propose a novel framework, namely “hybrid high-order FC networks” by exploiting the *higher-level* dynamic interaction among brain regions for early mild cognitive impairment (eMCI) diagnosis. For each sliding window-based rs-fMRI sub-series, we construct a whole-brain associated high-order network, by estimating the correlations between the topographical information of the high-order FC sub-network from one brain region and that of the low-order FC sub-network from another brain region. With multi-kernel learning, complementary features from multiple time-varying FC networks constructed at different levels are fused for eMCI classification. Compared with other state-of-the-art methods, the proposed framework achieves superior diagnosis accuracy, and hence could be promising for understanding pathological changes of brain connectome.

Alzheimer’s disease (AD) is an irreversible serious neurological disease in the elderly population, characterized by progressive perceptive and cognitive deficits<sup>1</sup>. The incidence of AD doubles every five years after the age of 65<sup>2</sup>. AD symptoms, such as impaired memory function, get worse over time due to the neurodegenerative processes<sup>3</sup>. Mild cognitive impairment (MCI) is an intermediate stage of cognitive decline between AD and normal aging<sup>4</sup>. Recent researches have reported that individuals with MCI tend to progress to AD at a rate of about 10–15% per year<sup>5,6</sup>. Such a high conversion rate may possibly be reduced if early interventions could be applied to the early stage of MCI (eMCI)<sup>7</sup>. Therefore, timely diagnosis of eMCI is of great clinical significance<sup>8</sup>. Accurate brain imaging-based eMCI diagnosis is still challenging since brain anatomical and functional changes in this stage are considerably subtle<sup>9,10</sup>. By far, compared with numerous computer-aided diagnosis studies on AD and MCI with various neuroimaging modalities<sup>11–20</sup>, those on eMCI diagnosis are still quite few<sup>8,21</sup>. Although the accuracy is still not so satisfactory for clinical application, these preliminary studies have already indicated that resting-state functional magnetic resonance imaging (rs-fMRI) can serve as a promising imaging technique for eMCI diagnosis.

Rs-fMRI is an *in vivo* brain functional imaging modality, measuring blood oxygen level-dependent (BOLD) signals<sup>22</sup> when subjects are in natural rest. With rs-fMRI, temporal synchronization of the spontaneous brain activity among different brain regions can be adopted to measure brain functional connectivity (FC), a metrics reflecting brain intrinsic functional organization<sup>23</sup>. Based on FC between each pair of brain regions, a whole-brain functional network can be constructed, which opens a new avenue for brain disease study by leveraging brain

<sup>1</sup>Department of Radiology and BRIC, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599, USA.

<sup>2</sup>Department of Brain and Cognitive Engineering, Korea University, Seoul, 02841, Republic of Korea. Correspondence and requests for materials should be addressed to D.S. (email: [dgshen@med.unc.edu](mailto:dgshen@med.unc.edu))

connectomics and complex network analysis<sup>24–30</sup>. For early AD diagnosis, it is usually hypothesized that the FC network-based biomarkers show up earlier than macroscopic anatomical changes<sup>8, 31, 32</sup>. However, previous studies on rs-fMRI based AD early diagnosis often utilized simply calculated FC by measuring inter-regional BOLD signal temporal synchronization with Pearson's correlation or, more generally, with sparse representation<sup>8, 33–35</sup>. This type of networks is low-order by definition because they characterize BOLD signal synchronizations and are insufficient to characterize high-level inter-regional interactions. In a recent study<sup>36</sup>, we proposed a high-order FC network construction method by measuring the similarity between two regions' FC topographical profiles (i.e., correlation's correlation). Preliminary group comparison between MCI and health controls has suggested a great potential of using this metrics to provide complementary information to the low-order FC metrics in the context of early AD biomarker detection.

The above two types of FC studies separately calculate low-order and high-order FCs. However, there could be an intriguing relationship and functional association between the two FC levels. Such an inter-level interaction exists in many biological networks, reflecting hierarchical organization and self-resemblance across multiple spatial scales<sup>37</sup>. Supposing that in human brain the low-level connections collect information and the high-level connections abstract information via the hierarchy, the functions of the inter-level connections could be 1) to facilitate two-level information “talking” to each other, 2) to let the low-level information guide high-level abstraction and, 3) to change the way of low-level information collection for achieving a better high-level integration. Moreover, from a robust system point of view, a network or a biological system could make itself less fragile and more resistant to targeted pathological attacks through the inter-level connections. Taking brain as an example, via psychophysiological and physiophysiological interactions, high-level preset of a psychological status (e.g., attention level) may change both sensory information collection and synthesis; their co-varying status may indicate such inter-level functional associations. In early AD neuropathological model, one may hypothesize that subtle pathological changes in the stage of eMCI may not only alter high-order FC while leaving the low-order FC largely intact<sup>21, 36</sup>, but also affect the functional association between the high- and low-order FCs. Collectively, the three types of FC networks (i.e., low-order FC, high-order FC and such an inter-level associated FC) complement each other, characterizing brain functional organization from different aspects. By integrating the three types of FCs, eMCI classification may be more accurate compared with that using only a single type of FCs.

To this end, we propose a novel approach called “hybrid high-order FC networks” to comprehensively explore the brain's complicated functional associations and search for subtle early imaging biomarkers which could be related to pathological changes, for better eMCI diagnosis. Besides the three different types of FC, we also need to take advantage of dynamic FC by integrating brain dynamics into the study and to calculate three types of time-varying FC networks for diagnosis. This is because that the pattern of brain FC networks may change along time while brain is switching among different status<sup>22, 38–44</sup>, and such a dynamic information may also provide sensitive features for revealing early brain functional abnormalities and for even better improved eMCI diagnosis.

Three main contributions of our study can thus be summarized: 1) A new FC metrics, namely associated high-order FC network, is proposed to characterize previously untouched inter-level interaction between the low- and the high-order FC networks; 2) For the first time, we investigate dynamics of the three types of FC networks and utilize them for disease diagnosis; 3) We propose a novel applicable machine learning framework to effectively fuse various types of dynamics FC (thus, it is called hybrid high-order) networks with a multi-kernel learning strategy for computer-aided eMCI diagnosis. Experiments are carried out to compare the accuracy of eMCI diagnosis between our proposed approach and other state-of-the-art methods. Experimental results indicate that our method achieves superior performance than those using only the static FC or only the traditional low- and high-order FC networks.

## Results

**Data acquisition.** To demonstrate the effectiveness of our method, we apply it to real rs-fMRI data from a putative public-accessible dataset, consisting of eMCI and normal aging subjects. One of our hypotheses is that, with features extracted from our newly developed associated high-order FC networks, eMCI classification could be more accurate, compared with those extracted using either the traditional low-order or high-order FC. Another hypothesis is that our computational framework of the hybrid high-order FC networks could effectively conduct multi-kernel fusion of the three types of brain dynamic networks and further boost classification performance. In addition, higher diagnosis performance would be achieved using our method compared with those using other state-of-the-art methods<sup>8, 21</sup>.

The rs-fMRI data are obtained from Alzheimer's Disease Neuroimaging Initiative (ADNI) project (<http://adni.loni.usc.edu>). ADNI was launched in 2003 by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, the Food and Drug Administration, private pharmaceutical companies and non-profit organizations. The original goal was to define biomarkers for use in clinical trials to determine the most appropriate way to measure treatment effects of AD. The current goal has been extended to discover more effective methods to early detect AD at its pre-dementia stage. We use ADNI phase-2 dataset since it includes eMCI subjects.

Data from twenty-nine eMCI subjects (13 F/16 M, aged  $73.7 \pm 4.8$  years) and 30 age-matched ( $p = 0.6174$ ) normal controls (NCs) subjects (17 F/13 M, aged  $74.4 \pm 5.7$  years) are used. All subjects were scanned with the same protocol using 3.0 T Philips Achieva scanners. The following parameters were used: repetition time (TR) = 3000 ms, echo time (TE) = 30 ms, flip angle = 80°, imaging matrix = 64 × 64, 48 slices, 140 volumes, and slice thickness = 3.3 mm. The rs-fMRI data are preprocessed using SPM8 software (<http://www.fil.ion.ucl.ac.uk/spm/software/spm8>) according to the previous studies<sup>8, 21</sup>. Briefly, the first 10 volumes of each subject are discarded to ensure magnetization equilibrium. After head motion correction, spatial registration to the standard space, spatial smoothing and temporal filtering (0.01–0.08 Hz), averaged signals from brain ventricle and white matter as well as head-motion parameters are regressed out from rs-fMRI data to reduce the nuisance effect on

Method	ACC (%)	AUC	SEN (%)	SPE (%)
SN <sub>L</sub>	66.1	0.614	55.2	76.7
SN <sub>H</sub>	71.2	0.741	70.0	72.4
SN <sub>A</sub>	<b>72.9</b>	<b>0.816</b>	<b>70.0</b>	<b>75.9</b>
SN <sub>L</sub> + SN <sub>H</sub> + SN <sub>A</sub>	<b>83.1</b>	<b>0.875</b>	<b>79.3</b>	<b>86.7</b>
DN <sub>Wee</sub> <sup>8</sup>	79.7	0.792	75.9	83.3
DN <sub>Chen</sub> <sup>18</sup>	86.4	0.900	86.2	86.7
DN <sub>L</sub>	72.9	0.743	69.0	76.7
DN <sub>H</sub>	81.4	0.875	79.3	83.3
DN <sub>A</sub>	<b>83.1</b>	<b>0.907</b>	<b>82.8</b>	<b>83.3</b>
DN <sub>L</sub> + DN <sub>H</sub> + DN <sub>A</sub>	<b>91.5</b>	<b>0.940</b>	<b>89.7</b>	<b>93.3</b>

**Table 1.** Performance comparison among the methods using different FC networks for eMCI diagnosis.

**FC estimation.** According to the Automated Anatomical Labeling (AAL) brain atlas, the mean regional rs-fMRI time series are extracted from each of the 116 brain regions.

**Performance evaluation.** A nested leave-one-out cross validation (LOOCV) scheme is adopted for performance evaluation of the proposed approach. Specifically,  $N$  subjects are involved in our study,  $N - 1$  of them are used for training the classifier while the left-out one is used for evaluating the classification performance. The procedure is repeated  $N$  times until each subject serves once for testing. In each repeat of the above procedure, an additional inner LOOCV is carried out on the  $N - 1$  training samples to determine optimal parameters, which include regularization parameters in the LASSO feature selection and also the weighting factors in the multiple-kernel learning. The parameter values leading to the best performance on the  $N - 1$  tests are selected and used for learning the optimal classification model. The soft-margin parameter in SVM is set as  $C = 1$ . For the dynamic FC network construction, the length of sliding window  $L$  and the step size  $S$  are set to be 70 and 1 (consistent with literature<sup>21</sup>), respectively, which results in the number of time sub-series as  $K = 61$ .

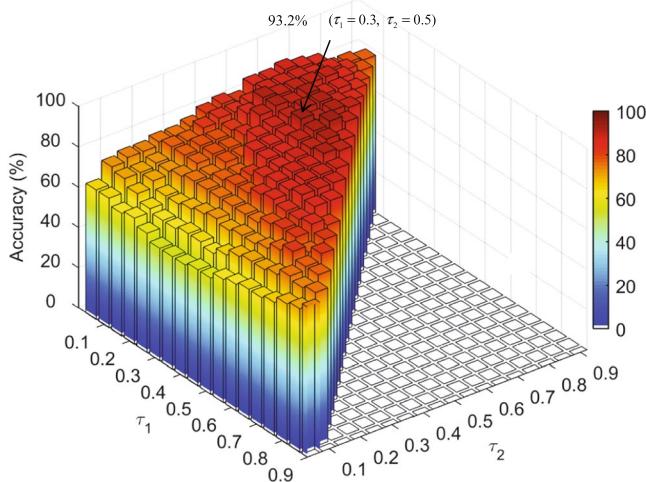
Extensive experiments are carried out to validate the effectiveness of our proposed method in comparison with other state-of-the-art methods. We investigate the diagnosis performance of various methods based on either static FC networks or dynamic FC networks. These compared methods include: (1) Static low-order Network (SN<sub>L</sub>, where the subscript “L” indicates conventional “low-order” FC); (2) Static high-order Network (SN<sub>H</sub>, where “H” indicates conventional “high-order” FC used in literature<sup>36</sup>); (3) Static associated high-order Network (SN<sub>A</sub>, where “A” denotes “associated”, i.e., our newly proposed *higher-level* type of FC); (4) Hybrid Static Networks (SN<sub>L</sub> + SN<sub>H</sub> + SN<sub>A</sub>) which fuse the three aforementioned networks; (5) Dynamic low-order Network (DN<sub>L</sub>); (6) Dynamic high-order Network (DN<sub>H</sub>); (7) Dynamic associated high-order Network (DN<sub>A</sub>); (8) Hybrid Dynamic Networks (DN<sub>L</sub> + DN<sub>H</sub> + DN<sub>A</sub>), which is our proposed framework combining three types of networks in a dynamic way; (9) Sparse temporally dynamic networks (DN<sub>Wee</sub>) that are constructed using group graphical LASSO<sup>8</sup>; (10) High-order network<sup>21</sup> that is constructed by estimating the correlation between two regions’ low-order FC dynamics and that of another two regions (DN<sub>Chen</sub>). Among them, SN<sub>A</sub> and DN<sub>A</sub> are the novel network modeling methods, while SN<sub>L</sub> + SN<sub>H</sub> + SN<sub>A</sub> and DN<sub>L</sub> + DN<sub>H</sub> + DN<sub>A</sub> are the novel network fusion frameworks for classification.

We evaluate the classification performance based on classification accuracy (ACC), area under ROC curve (AUC), Sensitivity (SEN), and Specificity (SPE). ACC is defined as the ratio of the number of correctly predicted labels to the number of whole samples. AUC measures the probability that a classifier will rank a randomly chosen positive sample higher than a randomly chosen negative one. SEN and SPE are defined as true positive rate and one minus false positive rate, respectively:

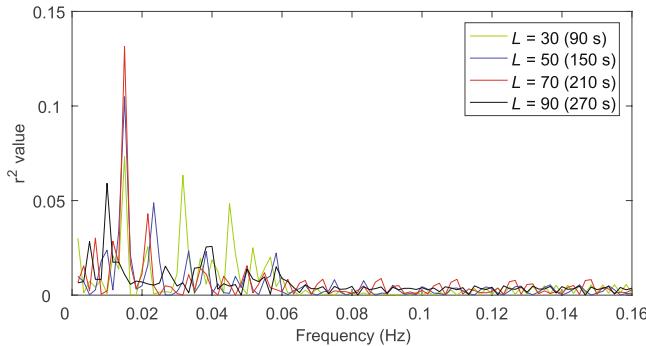
$$\text{SEN} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad \text{SPE} = \frac{\text{TN}}{\text{TN} + \text{FP}}. \quad (1)$$

**Experimental results.** Table 1 summarizes the performance on eMCI classification for all of the ten aforementioned methods. Consistent with our hypotheses, the main results are: 1) The high-order FC networks enhanced the classification performance and our associated high-order FC networks gained the highest one if only single type of FC network was used; 2) Integrating all the three types of networks with multi-kernel learning, eMCI classification yielded better performance compared to that using only single type of networks (even using only static networks, we reached an ACC of 83.1%; while using dynamic networks, we obtained the best ACC of 91.5% amongst all competing methods); and 3) The classifications based on the dynamic FC networks consistently outperformed those based on the static FC networks, indicating the necessity of integrating dynamic FC into classification. Of note, our method always achieved better ACC, AUC, SEN and SPE compared with the most recently developed, state-of-the-art methods (DN<sub>Wee</sub> and DN<sub>Chen</sub>).

It should be noted that the classification performance of our proposed framework depends on the selection of some parameters. For example, three weighting factors  $\tau_1 \in [0, 1]$ ,  $\tau_2 \in [0, 1]$  and  $\tau_3 \in [0, 1]$  ( $\tau_1 + \tau_2 + \tau_3 = 1$ ) in the multi-kernel SVM need to be estimated to fuse the kernel matrices that are derived from the DN<sub>L</sub>, DN<sub>H</sub> and DN<sub>A</sub>, respectively. A larger value of a weighting factor indicates the larger importance of the corresponding kernel



**Figure 1.** Classification accuracy of the hybrid high-order FC networks by integrating  $DN_L$ ,  $DN_H$  and  $DN_A$  ( $DN_L + DN_H + DN_A$ ) with different kernel weighting factors  $\tau_1$  and  $\tau_2$  for  $DN_L$  and  $DN_H$ , respectively (the weight factors for  $DN_A$  are determined by  $\tau_3 = 1 - (\tau_1 + \tau_2)$ ). The best accuracy was achieved with  $\tau_1 = 0.3$  and  $\tau_2 = 0.5$ , which is 93.2% as indicated by an arrow.

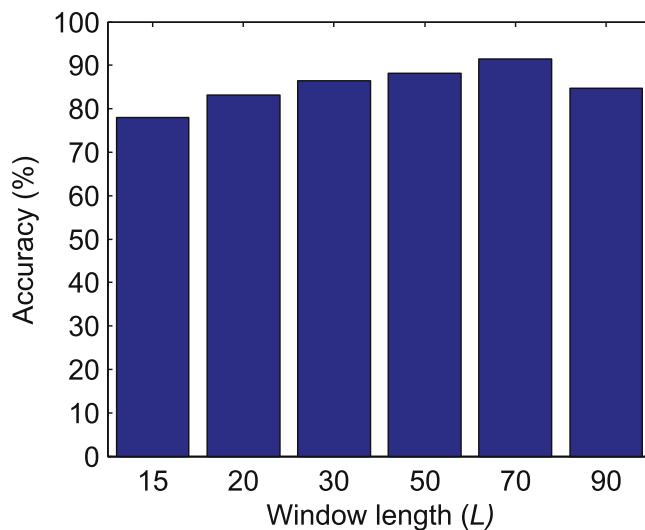


**Figure 2.** Separability ( $r^2$  value) between NC and eMCI for different frequency dynamic FC spectrum. The FC is measured for the discriminative link connecting the left inferior frontal gyrus and the left angular gyrus. Results from the use of various window lengths are shown.

matrices for classification. Although we have used inner LOOCV to estimate the best parameters for above classifiers, we can use all of the  $N$  subjects to better evaluate the possible dependency on parameter selection. Figure 1 shows accuracy of the classification model based on  $DN_L + DN_H + DN_A$  estimated using LOOCV on all of the  $N$  subjects, with different combinations of the three weighting factors  $\tau_1$ ,  $\tau_2$  and  $\tau_3$  (where is  $\tau_3 = 1 - (\tau_1 + \tau_2)$ ). After the exhaustive searching, the best accuracy of 93.2% is achieved with  $\tau_1 = 0.3$  (for  $DN_L$ ),  $\tau_2 = 0.5$  (for  $DN_H$ ), and  $\tau_3 = 1 - (\tau_1 + \tau_2) = 0.2$  (for  $DN_A$ ). The results indicate that  $DN_L$ ,  $DN_H$  and  $DN_A$  indeed provide complementary information to each other and all of them are necessary for classification. On the other hand, our method with the parameters estimated by inner LOOCV yielded a 91.5% accuracy, which is close to the best accuracy 93.2% with specific parameters estimated on all subjects.

In addition, the window length  $L$  used in sliding window strategy is an important factor for dynamic FC network analysis. The window length should be large enough to permit a reliable estimation of FC and resolve the lowest frequencies of interest in BOLD signals, while small enough to capture the dynamics of FC<sup>40</sup>. Leonardi and Van De Ville<sup>42</sup> have recommended using a window length that exceeds the longest wavelength composing the BOLD signals in order to suppress spurious fluctuations of dynamic FC. According to this criterion, the widow length in our study should be set to be larger than  $1/0.01 = 100$  s since our high-pass cutoff frequency in bandpass filtering is 0.01 Hz. On the contrary, Zalesky and Breakspear<sup>43</sup> suggested that non-stationary fluctuations in dynamic FC could be fairly robustly detected with a shorter window length (40–60 s). Although these two studies as well as the former review paper<sup>40</sup> are trying to set up a guideline to decide the window length, it is still far away from consensus. The above two studies are mainly based on simulated data and simulated SNR condition, which might not always be the case in real fMRI data because there is still no “ground truth” of the FC dynamics.

Note that, as a trade-off, Leonardi and Van De Ville<sup>42</sup> also mentioned to focus on the frequency interval  $[0-1/w]$  where  $w$  denotes the window length in second when interpreting the dynamic FC spectrum. To show the dynamic FC spectrum with varied window length, we calculated the separability (as measured by  $r^2$ ) (see Fig. 2)



**Figure 3.** Effects of varying the sliding window length  $L$  on eMCI classification accuracy, derived by the hybrid dynamic networks ( $DN_L + DN_H + DN_A$ ).

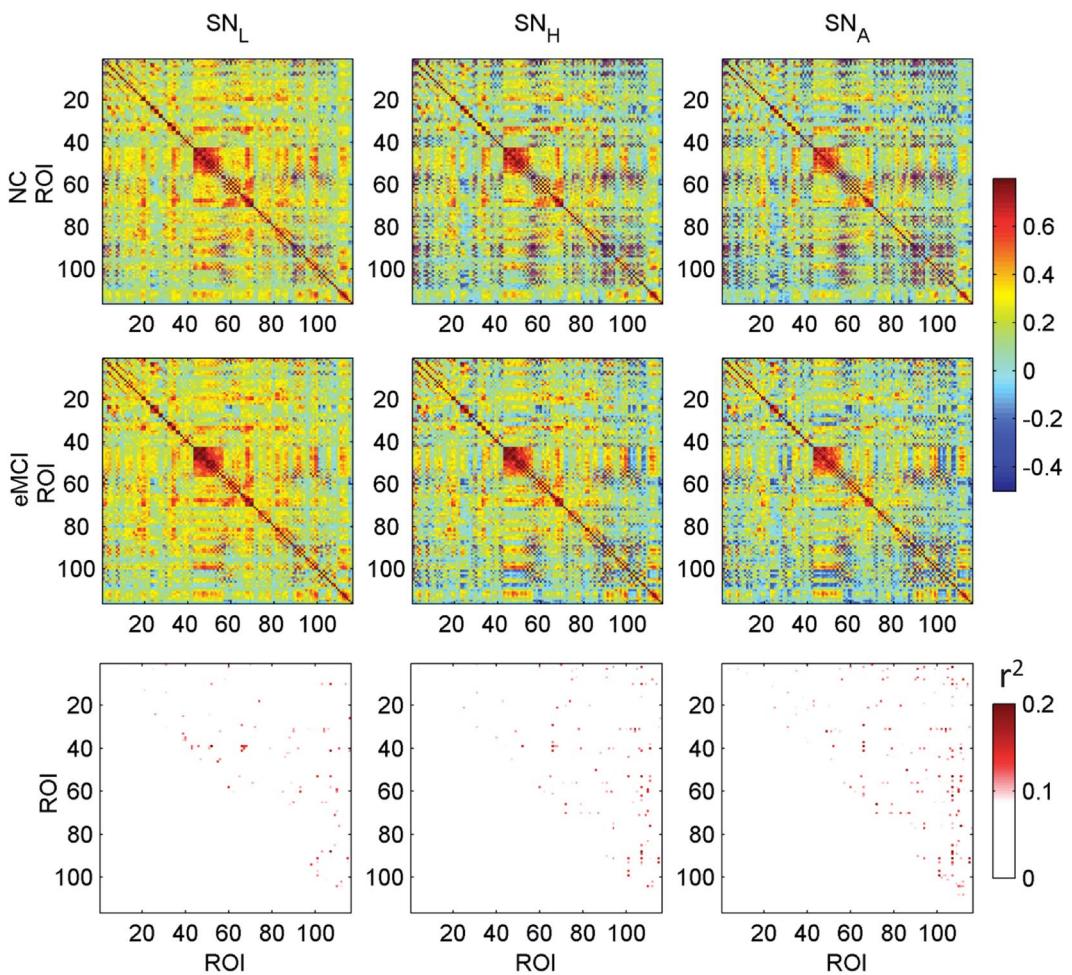
between NC and eMCI subjects across different frequencies for the discriminative link connecting the left inferior frontal gyrus and left angular gyrus. Consistent with the result in literature<sup>42</sup>, a longer window length presented a low-pass filtering effect with a lower cutoff frequency. However, we found that useful discriminative information was located at the frequencies higher than the suggested upper limit frequency  $1/w$ , which could contribute to good diagnosis performance of eMCI.

Therefore, we still consider using a window length larger than 100 s ( $L = 70$  volumes, i.e., 210 s). Such a long window length was chosen based on the performance of the eMCI classification using cross-validation (LOOCV) with training data. Specifically, we compared the eMCI classification accuracies derived by  $DN_L + DN_H + DN_A$  using various  $L$  ( $L = 15, 20, 30, 50, 70$  and  $90$ ) (see Fig. 3). Consistent with the observation in previous eMCI classification studies<sup>8,21</sup>, the selection of  $L = 70$  yielded the best classification accuracy, which was thus adopted for the subsequent analysis in our experiment.

## Discussion

To investigate the contribution of the associated high-order FC network to the diagnosis, we calculate the group-level separability (defined by the differences in the group averaged associated high-order FC networks of the two groups) between the NC and the eMCI subjects in both static and dynamic cases, and compare them with those obtained from the traditional low-order and high-order FC networks. Figure 4 shows the group-level  $SN_L$ ,  $SN_H$  and  $SN_A$  for NC (first row) and eMCI groups (second row), respectively. The discriminability index, calculated by squared pointwise biserial correlation coefficients ( $r^2$  values)<sup>44,45</sup> for all connections in each type of the FC networks is shown in the third row. Larger  $r^2$  value indicates higher separability of the feature distribution patterns between two classes. From Fig. 4, we can see the separability using the static low-order and high-order FC networks are smaller and involve fewer FC connections. However, the static associated high-order networks reveal more discriminative nodes and higher separability. This explains why  $SN_A$  yielded better diagnosis performance than  $SN_L$  and  $SN_H$ . On the other hand, we found that the three types of FC networks identified several different discriminative FC connections that may serve as complementary features for eMCI diagnosis. This indicates that it is suboptimal to utilize only the new FC network modeling method for disease diagnosis since a better performance can be achieved by combining different types of FC networks. In this sense, our newly developed associated high-order FC metrics does not intend to replace the conventional ones but may provide unique, essential and meaningful information to conventional FC metrics for comprehensive brain connectome research. As a result, further improvement of diagnosis performance can be achieved by integrating these complementary features under our proposed framework of hybrid high-order FC networks.

One of the highlights of our study is that we estimated various types of dynamic FC networks and demonstrated the feasibility of using these dynamic networks to improve classification accuracy. An increasing number of studies<sup>22,39–41</sup> have suggested that FC network is not stationary but spontaneously changes over time. We, for the first time, use a frequency power spectrum method to effectively take advantage of such discriminative spontaneous changes and demonstrate that such information can be adopted to further improve classification. Figure 5 shows the time-varying FC matrices of  $DN_L$ ,  $DN_H$  and  $DN_A$  estimated from different sliding windows, for one randomly selected NC subject and one eMCI subject. We found that all the dynamic FC networks could capture the temporal variation of FC patterns. This assists to explore rich features from time-evolving FC networks, resulting in better (with accuracy improved by about 7–11% if using dynamic networks compared to static ones) diagnosis performance. Most importantly, compared with the  $DN_L$ , both the  $DN_H$  and the  $DN_A$  may enlarge network topology differences along time, which could be used for better differentiation between NC and eMCI. This has actually been proved with profound improvement in diagnosis accuracy (with the increment of



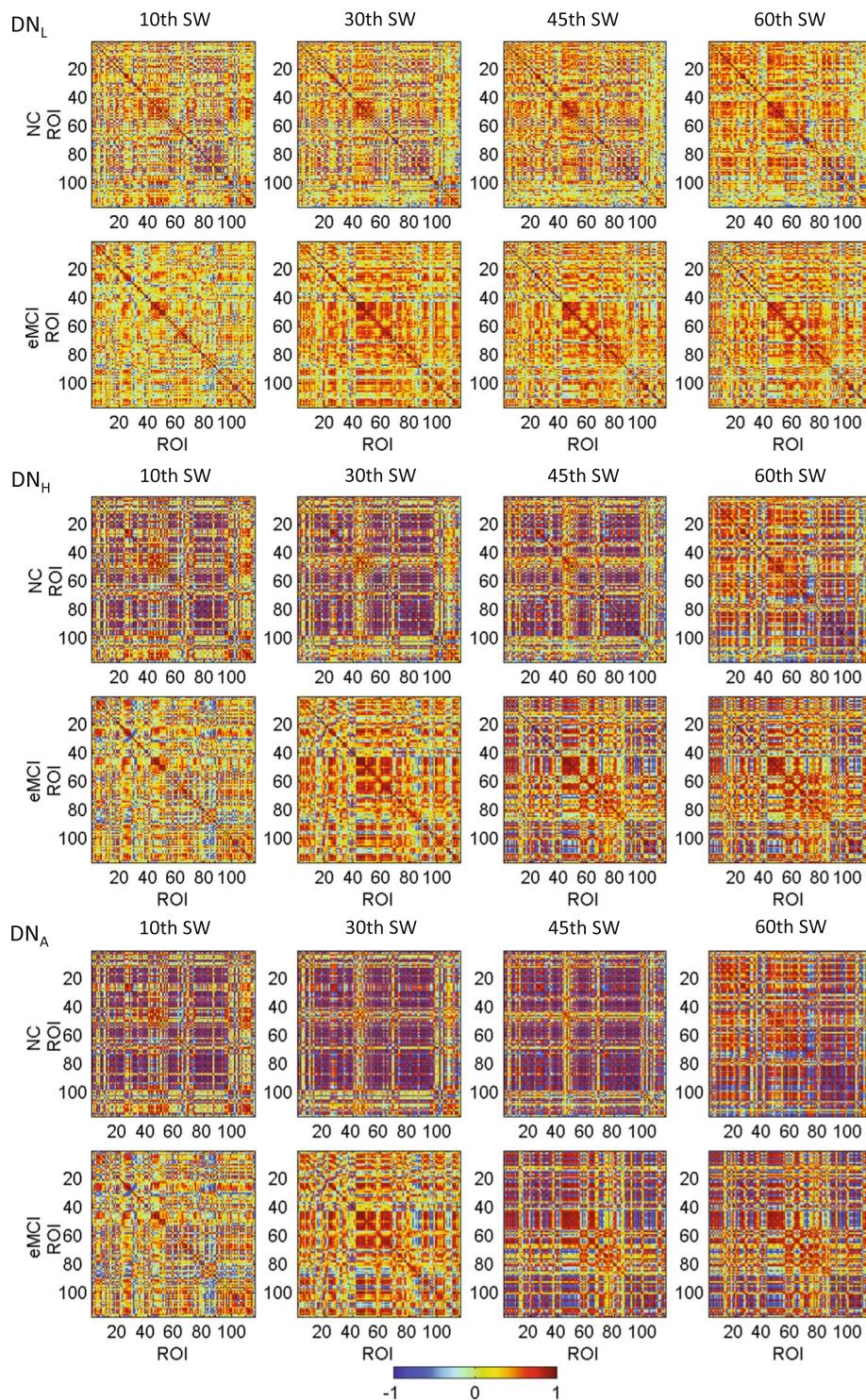
**Figure 4.** Group-averaged FC matrices of SN<sub>L</sub>, SN<sub>H</sub>, and SN<sub>A</sub> for NC and eMCI groups, and the separability matrices between the two groups for each type of the FC networks. The separability between the NC and eMCI subjects is calculated for each link in the network by computing  $r^2$  value, and is color-coded from light to dark red (from less to better separability). Only the  $r^2$  values at the upper triangle are shown since the FC matrices are symmetric. SN<sub>L</sub>: static low-order FC network; SN<sub>H</sub>: static high-order FC network; SN<sub>A</sub>: static associated high-order FC network.

about 9–11%) by using the features from dynamic high-order FC networks compared with that using dynamic low-order FC network-based features.

We also investigate the potential biological meaning of the machine learning algorithm selected brain regions as biomarkers for early AD detection and compare the results among different types of dynamic networks. Figure 6 presents the principal component coefficients corresponding to the most discriminative features selected by LASSO for DN<sub>L</sub>, DN<sub>H</sub> and DN<sub>A</sub>, respectively. Most of the important information is mainly concentrated in the range of very low frequencies (<0.033 Hz). This is reasonable since the biologically meaningful fluctuations of the dynamic FC time series are believed to be relatively slow<sup>22</sup>. From the zoomed-in areas of Fig. 6, we also found that the spatial-frequency locations of the most discriminative features from different types of FC networks are quite different.

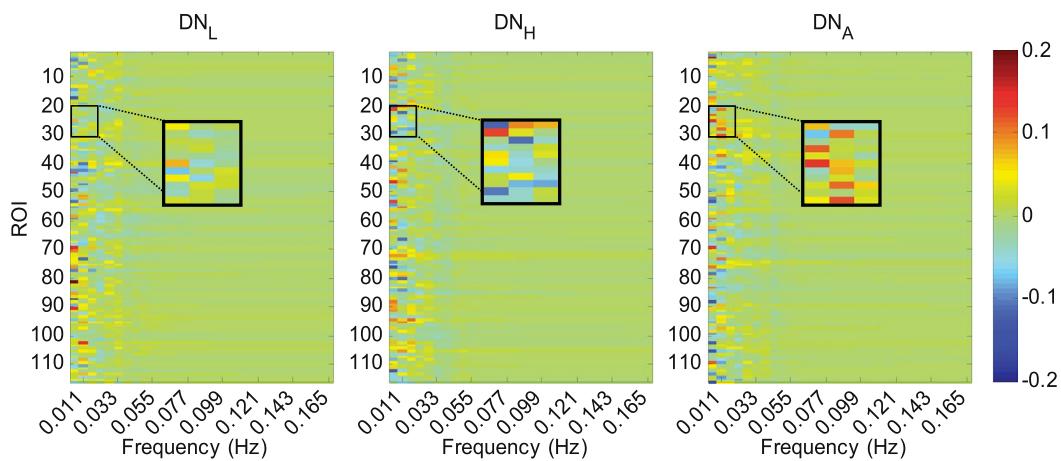
Figure 7 shows spatial locations of the most discriminative ROIs included in the top ten principal component coefficients selected by LASSO for DN<sub>L</sub>, DN<sub>H</sub> and DN<sub>A</sub>, respectively. A total of 22 ROIs are selected from the different FC networks and they have been consistently reported by previous studies on biomarker detection for AD and MCI, including the left superior temporal gyrus<sup>46</sup>, right inferior temporal gyrus<sup>47</sup>, right lobule VIIb of the cerebellar hemisphere<sup>48</sup>, left rolandic operculum<sup>48</sup>, right middle temporal pole<sup>49</sup>, left paracentral lobule<sup>44</sup>, left putamen<sup>50</sup>, right paracentral lobule<sup>51</sup>, right cuneus<sup>52</sup>, right globus pallidus<sup>53</sup>, left parahippocampal gyrus<sup>54, 55</sup>, left inferior frontal gyrus<sup>56</sup>, left olfactory cortex<sup>57</sup>, right supplementary motor area<sup>48</sup>, left transverse temporal gyrus<sup>58</sup>, right olfactory cortex<sup>57</sup>, left fusiform gyrus<sup>59, 60</sup>, left medial orbitofrontal cortex<sup>56</sup>, left lobule IX of cerebellar hemisphere<sup>48</sup>, left superior frontal gyrus<sup>48, 59</sup>, right insula<sup>52, 55</sup> and the right putamen<sup>50</sup>.

From Fig. 7, we also found that many selected ROIs were different if using different FC networks for classification. Consistent with our observation of the discriminative static FC connections in Fig. 4, different types of dynamic networks also produced complementary features, which are integrated using the framework of our hybrid high-order FC networks to further improve eMCI identification. This is supported by the superior

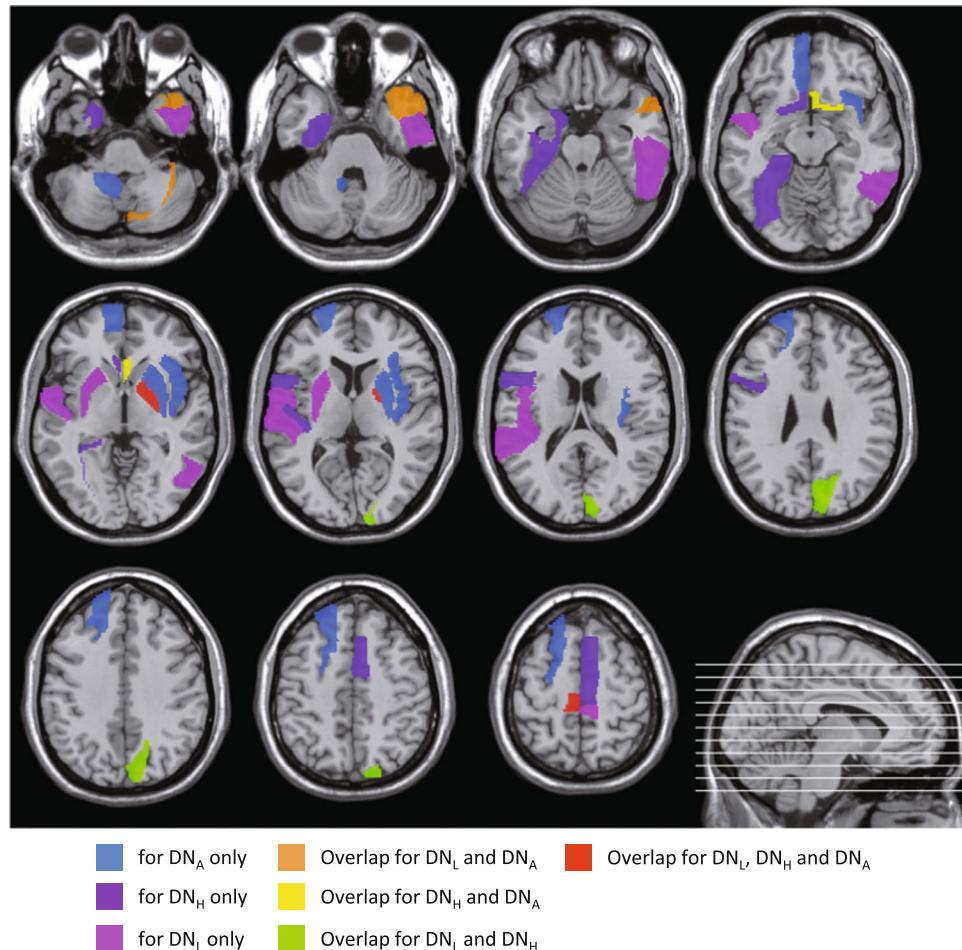


**Figure 5.** Respective FC matrices of  $\text{DN}_L$ ,  $\text{DN}_H$ , and  $\text{DN}_A$  from one NC subject and one eMCI subject at the 10-th, 30-th, 45-th and 60-th sliding windows (SW). The NC and eMCI subjects are randomly selected from each group.  $\text{DN}_L$ : dynamic low-order FC network;  $\text{DN}_H$ : dynamic high-order FC network;  $\text{DN}_A$ : dynamic associated high-order FC network. Please note that we intend to show the network pattern changes along time, but not mean to compare the network pattern at a specific sliding window between two subjects.

diagnosis performance (91.5% when using all three networks vs. 72.9–83.1% when using them separately) obtained by the proposed approach of hybrid high-order FC networks. Further investigation on the similarities and the differences among the features selected from different network types will be needed using more data sets and via extensive applications.



**Figure 6.** Principal component coefficients corresponding to the most discriminative features selected by LASSO for  $\text{DN}_L$ ,  $\text{DN}_H$  and  $\text{DN}_A$ , respectively. For better visualization and comparison among the different networks for the different spatial-frequency characteristic of the selected features, a small patch at the same location of the three matrices was zoomed in.



**Figure 7.** ROIs with the top ten principal component coefficients corresponding to the most discriminative features selected by LASSO for  $\text{DN}_L$ ,  $\text{DN}_H$  and  $\text{DN}_A$ , respectively.

To more clearly explain the characteristics of these two new FC metrics beyond the low-order FC, we provide two intuitive examples here. First, let us consider the low-level visual area V1 (primal visual cortex) and the high-level visual processing areas, i.e., posterior parietal cortex (PPC). The well-established model for dorsal

visual stream begins with V1, goes through secondary and associated visual areas, and finally to the PPC<sup>61</sup>. Therefore, the BOLD signal synchronization between the V1 and PPC is supposed to be not strong since they are responsible for processing visual information at different levels (instead, the bilateral V1 could have strong BOLD synchronization since they are at the same level). However, due to tight feedforward and feedback between them (e.g., they are all modulated by attention<sup>62</sup>), their inter-region high-order functional association could be strong. That is, there could exist some potentially indirect relationships between these two brain regions, which may not be effectively revealed by the low-order FC. Consistent with the above hypothesized model, based on the data from a randomly selected subject, we found that the low-order FC (direct BOLD signal synchronization) between the left lingual gyrus (encompassing V1) and the left inferior parietal lobule (covering most of the PPC) is low (0.36). However, both high-order FC (FC topographical profile-based similarity) and associated high-order FC are strong (0.61 and 0.74, respectively). This indicates that the lifted functional association, when measured from a high level, could reflect the close relationship between the two regions in the visual pathway.

Another example is from our previous study<sup>36</sup>, where the three FC metrics are calculated between the left posterior cingulate cortex (PCC) and the anterior cingulate cortex (ACC). Since the PCC is within the default mode networks while the ACC is included in several other attention-related functional networks, their low-order FC is observed as weak (0.36). However, enhanced strengths of the high-order FC (0.59) and the associated high-order FC (0.70) between the two regions indicate that the two new high-order FC metrics could be able to capture a close relationship among these high-level cognition-related functional networks<sup>63</sup>.

Interestingly, both above examples show higher associated high-order FC, compared with the low-order FC, which further indicates that the former could be able to capture more complicated functional interaction between two regions. The associated high-order FC measures the modulatory interaction between the low-order FC and the high-order FC, i.e., a cross-level functional association. Since our present study aims to demonstrate the feasibility of using high-order FC metrics for disease diagnosis, the biological meaning of these metrics requires more dedicated studies with the aid from existing neurocognitive models in future.

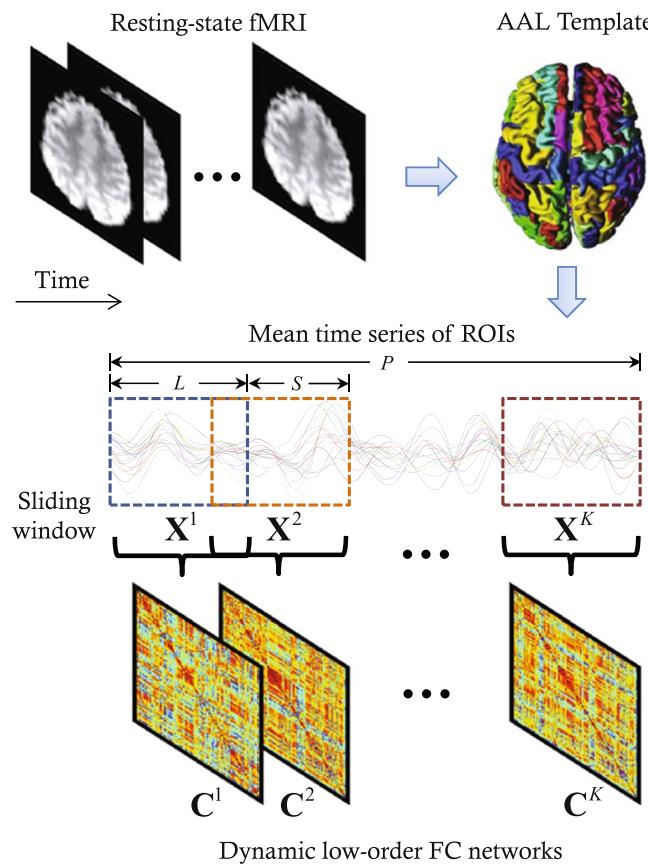
It should be noted that our introduced high-order FC and associated high-order FC reveal higher-level functional interactions between FC profiles between any pair of brain regions, yet ignore potential complex relationships among multiple brain regions. A recent method<sup>64</sup>, called hyper network, has been proposed to reveal more complex FC among multiple regions, and hence provide new approaches to investigate FC, which, however, is fundamentally different from our high-order FC and associated high-order FC metrics. Specifically, this method was developed based on hyper-graph theory for exploring the complex interactions among multiple brain regions, where an edge in the hyper network connects with more than two brain regions. A combination of our method and the dynamic hyper network method could further improve the diagnosis performance, which will be investigated in our future work.

In summary, we propose a novel approach, namely hybrid high-order FC networks, to effectively integrate multiple types of FC networks by using multi-kernel learning strategy for eMCI diagnosis. Associated high-order network, characterizing *higher-level* functional interactions between high- and low-level FC networks, is newly proposed to reveal the previously untouched relationship among brain regions. Three types of dynamic whole-brain FC networks are systematically defined and jointly used to provide complementary discriminative features for early MCI identification. Our method achieves superior performance (accuracy = 91.5%) in this challenging problem, which is even racing ahead of the most recently developed state-of-the-art solutions. This study reveals the complexity of our brain connectome, and the feasibility of using it as an effective computer-aided individual diagnosis tool for future clinical applications towards precise medicine.

## Methods

In our hybrid high-order FC network approach, we combine three types of FC networks and dynamics FC analysis for comprehensive feature extraction. To achieve dynamic FC networks, sliding window strategy is adopted to segment the entire rs-fMRI time series into multiple sub-series, from each of which the three types of FC networks are constructed. We first construct a traditional low-order FC network and a topographical FC profile-based high-order FC network. This produces two different FC fingerprints for each brain region: the one is the topographical low-order FC profiles between this region and other regions; the other is the high-order FC profiles between the sub-network centering at this region and those centering at other regions. We then calculate a *higher-level* associated FC between the two types of the FC fingerprints for each pair of brain regions, which consequently forms an associated high-order FC network. Finally, all the dynamic FC networks are integrated into a unified model with multi-kernel learning strategy to make features from one FC networks support those from others, for better classification performance. Each step is detailed in the following sub-sections.

**Low-order FC network construction.** With sliding window approach, an rs-fMRI time series can be segmented into multiple sub-series, each generating one FC matrix (Fig. 8). In particular,  $K = [(P - L)/S] + 1$  sub-series can be generated from an rs-fMRI time series with  $P$  time points, where  $L$  and  $S$  are the window length and step size, respectively. Suppose that  $\mathbf{X}^k = [\mathbf{x}_1^k, \mathbf{x}_2^k, \dots, \mathbf{x}_R^k] \in \mathbb{R}^{L \times R}$  ( $k = 1, 2, \dots, K$ ) denotes the  $k$ -th time sub-series for a total of  $R = 116$  ROIs, and  $\mathbf{x}_i^k = [x_{1,i}^k, x_{2,i}^k, \dots, x_{L,i}^k]^T \in \mathbb{R}^L$  is the  $k$ -th time sub-series corresponding to the  $i$ -th ROI. The correlation strength  $C_{ij}^k$  between the  $i$ -th and  $j$ -th ROIs of the  $k$ -th time sub-series can be typically computed using Pearson's correlation. Such correlation strength, in graph theoretic analysis of complex brain FC networks, is called the edge weight. By computing such correlation strength of the  $k$ -th time sub-series between each pair of ROIs (or nodes), an FC network can be constructed as a symmetric correlation matrix  $\mathbf{C}^k = [C_{ij}^k] \in \mathbb{R}^{R \times R}$ . Without loss of generality, we assume that  $\mathbf{x}_i^k$  has been centralized by  $\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k$  and further



**Figure 8.** Illustration of the construction of dynamic low-order FC networks using sliding window-based Pearson's correlation on rs-fMRI data.

normalized by  $\sqrt{(\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k)^T (\mathbf{x}_i^k - \bar{\mathbf{x}}_i^k)}$  for  $i = 1, 2, \dots, R$ . The computation of FC network on the  $k$ -th time sub-series can then be equivalently written as:

$$\mathbf{C}^k = (\mathbf{X}^k)^T \mathbf{X}^k. \quad (2)$$

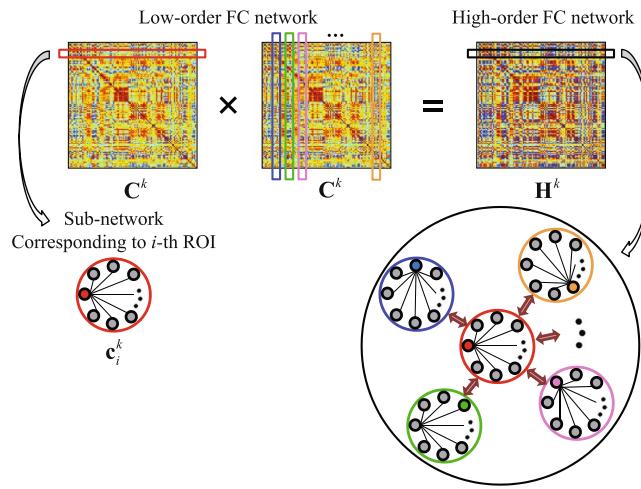
The dynamic FC networks can be derived by estimating the correlation matrices for all the  $k = 1, 2, \dots, K$  time sub-series. Note that Eq. (2) defines dynamic low-order FC networks, while a static network is an extreme case where window length is maximized to the entire time scale ( $L = P$ ). Thus, multiple FC matrices in Fig. 8 merge into one.

**Hybrid high-order FC networks for eMCI diagnosis.** The “hybrid high-order FC networks” refer to as a framework which fuses three types of FC networks for improving diagnosis performance. In this framework, a key step is to construct the associated high-order network. Of note, this type of network can be regarded to as a *higher-level* FC network as it characterizes the interaction between the conventional low-order and the high-order networks. The construction of associated high-order network will be followed by feature extraction and selection, and a multi-kernel learning strategy for multi-type feature fusion.

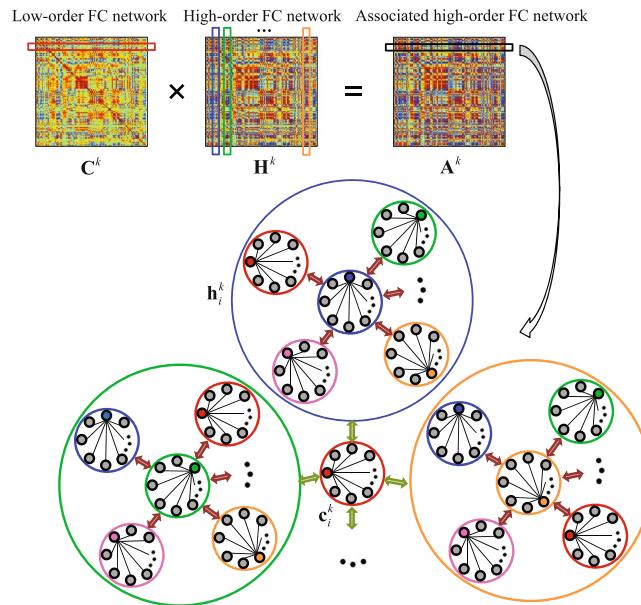
**Associated high-order FC network construction.** Low-order FC network on the  $k$ -th time sub-series can be rewritten as  $\mathbf{C}^k = [\mathbf{c}_1^k, \mathbf{c}_2^k, \dots, \mathbf{c}_R^k] \in \mathbb{R}^{R \times R}$ , where the  $i$ -th column  $\mathbf{c}_i^k$  (or the  $i$ -th row due to the symmetry of  $\mathbf{C}^k$ ) defines the connectivity pattern between the  $i$ -th ROI and all other ROIs. Therefore, we regard  $\mathbf{c}_i^k$  as a low-order “sub-network” between node  $i$  and other regions. Of note, it is quite important to have this “sub-network” definition based on the low-order FC network, since, similarly, we can define a high-order FC by the topographical similarity between any pair of these low-order sub-networks. Then, a high-order FC network can be constructed by calculating the FC between every pair of the low-order sub-networks. Figure 9 illustrates the construction of the high-order FC network.

Assuming  $\mathbf{c}_i^k$  ( $i = 1, 2, \dots, R$ ) has been centralized and normalized, a high-order network construction on the  $k$ -th time sub-series can be similarly written as:

$$\mathbf{H}^k = (\mathbf{C}^k)^T \mathbf{C}^k, \quad (3)$$



**Figure 9.** Illustration of high-order FC network construction based on the topographical similarity between each pair of the low-order sub-networks. The high-order FC network is built in a dynamic way using the  $k$ -th rs-fMRI time subseries ( $k = 1, 2, \dots, K$ ).



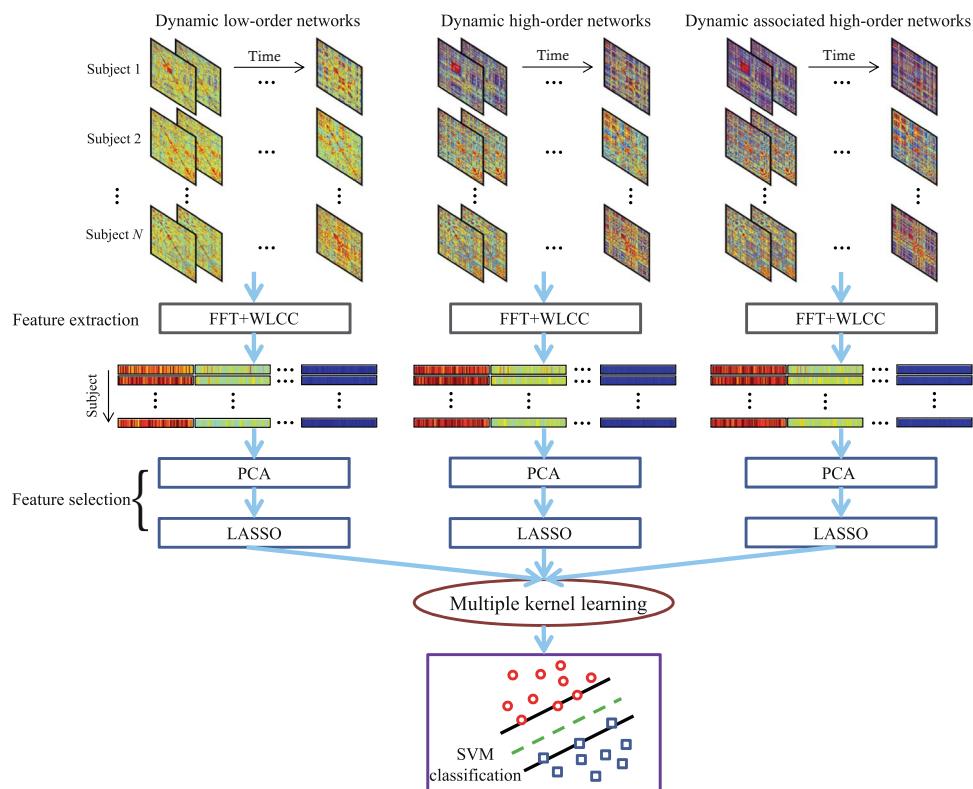
**Figure 10.** Illustration of an associated high-order FC network construction using topographical similarity between each high-order sub-network and each low-order sub-network. The associated high-order FC network is built in a dynamic way using the  $k$ -th rs-fMRI time sub-series.

where a certain element of  $H^k$ ,  $H_{ij}^k$ , denotes the topographical similarity (measured by Pearson's correlation) between the  $i$ - and  $j$ -th low-order sub-networks, and  $C^k = (\mathbf{X}^k)^T \mathbf{X}^k$ . Based on the same form of Eqs (2) and (3), both low- and high-order FC calculations can be mathematically unified. By using the whole length of rs-fMRI time series, a static high-order network can similarly be calculated.

To construct associated high-order FC network, we further refer the high-order FCs corresponding to the same “node”  $i$  (here the “node” is actually a low-order sub-network centering at region  $i$ ) as a high-order sub-network  $\mathbf{h}_i^k$ , and  $\mathbf{H}_i^k = [\mathbf{h}_1^k, \mathbf{h}_2^k, \dots, \mathbf{h}_R^k] \in \mathbb{R}^{R \times R}$ . Supposing that the  $\mathbf{h}_i^k$  ( $i = 1, 2, \dots, R$ ) is centralized and normalized, we can characterize the inter-level interactions between the low-order sub-networks  $\mathbf{c}_i^k$  ( $i = 1, 2, \dots, R$ ) and high-order sub-networks  $\mathbf{h}_i^k$  ( $i = 1, 2, \dots, R$ ), which can be written in the following form:

$$\mathbf{A}^k = (\mathbf{C}^k)^T \mathbf{H}^k = (\mathbf{C}^k)^T ((\mathbf{C}^k)^T \mathbf{C}^k), \quad (4)$$

where  $\mathbf{A}^k$  defines the associated high-order FC network, a *higher-level* FC network, using the  $k$ -th time sub-series. The element  $A_{ij}^k$  in  $\mathbf{A}^k$  denotes the interaction between  $i$ -th low-order sub-network and  $j$ -th high-order sub-network (Fig. 10). From Eq. (4), we can see that the associated correlation matrix  $\mathbf{A}^k \in \mathbb{R}^{R \times R}$  is an asymmet-



**Figure 11.** Our proposed eMCI diagnosis framework based on hybrid high-order FC networks. It has the three types of dynamic FC networks as input; after feature extraction and selection, a multiple kernel learning strategy is used to achieve effective feature fusion for classification. (FFT: Fast Fourier transform, WLCC: Weighted local clustering coefficient, PCA: Principal component analysis, LASSO: Least absolute shrinkage and selection operation, SVM: Support vector machine).

rical matrix. To improve interpretation, we further transform the asymmetrical  $\mathbf{A}^k$  to symmetrical by using  $\mathbf{A}^k = (\mathbf{A}^k + (\mathbf{A}^k)^T)/2$ , similarly to that also used in a previous study<sup>32</sup>. Of note, our experiment has shown that this additional symmetry operation does not significantly affect final classification accuracy. Similar to the other two types of dynamic FC networks, a static associated high-order FC network can be estimated using the whole length of rs-fMRI time series.

**Feature extraction and selection.** In this section, we introduce the procedure of feature extraction and selection for the dynamic low-order FC networks. Please note that the same procedure is also carried out to extract and select features from the dynamic high-order and the dynamic associated high-order FC networks.

Because of the unconstrained mental activity during resting state, features directly extracted from each sliding window-based FC network for a subject do not have temporal correspondence with those extracted from the same sliding window for other subjects. Therefore, for one subject, the different features extracted from different sliding windows cannot be concatenated along time. To allow feature concatenation, one must have a hypothesis that features extracted from the same temporal window for different subjects belong to the same instantaneous brain network, which cannot be guaranteed<sup>65</sup>. To enforce the feature correspondence across subjects, we transform the temporally dynamic FC networks for each subject into the frequency domain, obtaining multiple frequency-specific FC networks. Specifically, for the dynamic low-order networks  $\mathbf{C}^1, \mathbf{C}^2, \dots, \mathbf{C}^K$ , where  $\mathbf{C}^k = [C_{ij}^k] \in \mathbb{R}^{R \times R}$ , dynamic FC time series between regions  $i$  and  $j$  can be obtained by concatenating the elements  $C_{ij}^k$  across  $K$  temporal windows as  $\mathbf{g}_{ij} = [C_{ij}^1, C_{ij}^2, \dots, C_{ij}^K] \in \mathbb{R}^K$ , representing how the FC fluctuates along time. Fast Fourier transform (FFT) is then applied to transform this dynamic FC time series into power spectrums  $[Z_{ij}^1, Z_{ij}^2, \dots, Z_{ij}^Q]$ , where  $Q$  is the number of effective frequency bins. Thus, we can construct time-invariant FC networks for all spectrums as  $\mathbf{Z}^1, \mathbf{Z}^2, \dots, \mathbf{Z}^Q$ , where  $\mathbf{Z}^q = [Z_{ij}^q] \in \mathbb{R}^{R \times R}$ . These time-invariant FC networks characterize the frequency characteristics of the temporally dynamic FC networks. Note that the FFT is not implemented for the static network.

Graph theory-based feature extraction and selection are implemented based on the FC spectrum networks. In this study, we adopt weighted local clustering coefficient (WLCC)<sup>66</sup> as a nodal feature for each brain region and each frequency. The WLCC quantifies the “cliqueness” of each node in a weighted network. The cliqueness is originally a graph theoretic concept, which characterizes a network’s local topology for every node. This metric has been widely used as a sensitive feature in eMCI diagnostic studies<sup>8,21</sup>. For each network  $\mathbf{Z}^q$  ( $q = 1, 2, \dots, Q$ ), the WLCC for the  $i$ -th node can be defined as:

$$f_i^q = \frac{2\sum_{j:j \in \Omega_i} (Z_{ij}^q)^{\frac{1}{3}}}{v_i(v_i - 1)}, \quad (5)$$

where  $\Omega_i$  is a set of nodes directly connected to the  $i$ -th node and  $v_i$  denotes the number of elements in  $\Omega_i$ . After extracting the features from all the nodes at all  $Q$  frequency bins, we concatenate them to form a feature vector according to:

$$\mathbf{f} = [f_1^1, \dots, f_R^1, f_1^2, \dots, f_R^2, \dots, f_1^Q, \dots, f_R^Q]^T. \quad (6)$$

This feature vector is of a relatively high dimension and may contain irrelevant or redundant features which need to remove. To do this, we construct a feature vector for each subject according to Eq. (6), thereby obtaining a feature vector set as  $\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N]^T$ , where  $N$  is the number of subjects. Principal component analysis (PCA)<sup>67</sup> is implemented on  $\mathbf{F}$  to reduce feature dimension. The original features with the dimensionality of  $Q \times R$  are transformed into a new feature space defined by all  $N - 1$  principal components with non-zero eigenvalues. Subsequently, a supervised feature selection strategy based on the least absolute shrinkage and selection operation (LASSO)<sup>68,69</sup> is adopted to select discriminative features from the  $N - 1$  principal components. The features corresponding to non-zero LASSO regression coefficients are retained as crucial features for classification.

*Multi-kernel SVM for classification.* With the above-mentioned feature extraction and selection, we obtain three feature vector sets for the three FC network types, respectively. Since one of our hypotheses is that these FC networks could provide complementary information to each other for classification, we can fuse all features to generate better classification performance. The simplest way for this is to concatenate all features from different types of FC networks into a longer feature vector. However, such simple concatenation may not be optimal for achieving effective feature combination<sup>13</sup>. On the other hand, a kernel-based feature combination using multi-kernel learning offers more flexibility for feature fusion by estimating different weights on the features from different modalities<sup>70–72</sup>, which could provide a better way to integrate the features derived from different types of FC networks.

Therefore, we adopt multi-kernel learning to fuse the features by a linear combination of kernels that are estimated from the low-order, the conventional high-order and the novel associated high-order FC networks, respectively. An SVM classifier with a linear kernel  $K(\mathbf{a}, \mathbf{b}) = \mathbf{a}^T \mathbf{b}$  based on LIBSVM<sup>73</sup> is used for the multi-kernel learning based classification. Specifically, we first perform normalization on each feature vector to make sure that all the features from different types of FC networks are comparable. Based on the normalized features, a linear kernel is calculated across subjects for each type of the FC networks. Effective feature fusion is then achieved by computing a composite kernel through an optimal linear combination of the multiple kernels. Finally, classification is carried out using SVM with the composite kernel. Figure 11 summarizes the overall framework of our proposed eMCI diagnosis method based on hybrid high-order FC networks.

## References

1. Association, A. Alzheimer's disease facts and figures. *Alzheimer's and Dementia* **9**, 208–245 (2013).
2. Bain, L. *et al.* Healthy brain aging: A meeting report from the Sylvan M. Cohen Annual Retreat of the University of Pennsylvania Institute on Aging. *Alzheimer's and Dementia* **4**, 443–446 (2008).
3. McKhann, G. *et al.* Clinical diagnosis of Alzheimer's disease: Report of the NINCDS–ADRDA Work Group\* under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* **34**, 939–939 (1984).
4. Davatzikos, C., Bhatt, P., Shaw, L., Batmanghelich, K. & Trojanowski, J. Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. *Neurobiology of Aging* **32**, 2322.e2319–2322.e2327 (2011).
5. Gauthier, S. *et al.* Mild cognitive impairment. *The Lancet* **367**, 1262–1270 (2006).
6. Grundman, M. *et al.* Mild cognitive impairment can be distinguished from Alzheimer disease and normal aging for clinical trials. *Archives of Neurology* **61**, 59–66 (2004).
7. DeCarli, C. Mild cognitive impairment: prevalence, prognosis, aetiology, and treatment. *The Lancet Neurology* **2**, 15–21 (2003).
8. Wee, C., Yang, S., Yap, P. & Shen, D. Sparse temporally dynamic resting-state functional connectivity networks for early MCI identification. *Brain Imaging and Behavior* **10**, 342–356 (2016).
9. Petersen, R. Challenges of epidemiological studies of mild cognitive impairment. *Alzheimer Disease and Associated Disorders* **18**, 1–2 (2004).
10. Chong-Yaw Wee, Pew-Thian Yap, Dinggang Shen. Prediction of Alzheimer's disease and mild cognitive impairment using cortical morphological patterns. *Human Brain Mapping* **34**(12), 3411–3425 (2013).
11. Amlien, I. & Fjell, A. Diffusion tensor imaging of white matter degeneration in Alzheimer's disease and mild cognitive impairment. *Neuroscience* **276**, 206–215 (2014).
12. Liu, H., Zhou, X., Jiang, H., He, H. & Liu, X. A semi-mechanism approach based on MRI and proteomics for prediction of conversion from mild cognitive impairment to Alzheimer's disease. *Scientific Reports* **6**, 26712 (2016).
13. Zhang, D., Wang, Y., Zhou, L., Yuan, H. & Shen, D. Multimodal classification of Alzheimer's disease and mild cognitive impairment. *NeuroImage* **55**, 856–867 (2011).
14. Zhu, X., Suk, H.-I., Wang, L., Lee, S.-W. & Shen, D. A novel relational regularization feature selection method for joint regression and classification in AD diagnosis. *Medical Image Analysis* (2015).
15. Liu, M., Zhang, D. & Shen, D. Relationship induced multi-template learning for diagnosis of Alzheimer's disease and mild cognitive impairment. *IEEE Transactions on Medical Imaging* **35**, 1463–1474 (2016).
16. Willette, A., Calhoun, V., Egan, J. & Kapogiannis, D. Prognostic classification of mild cognitive impairment and Alzheimer's disease: MRI independent component analysis. *Psychiatry Research: Neuroimaging* **224**, 81–88 (2014).
17. Thung, K., Wee, C., Yap, P. & Shen, D. Identification of progressive mild cognitive impairment patients using incomplete longitudinal MRI scans. *Brain Structure and Function* **221**, 3979–3995 (2015).
18. Zhu, X., Suk, H.-I., Lee, S.-W. & Shen, D. Subspace regularized sparse multitask learning for multiclass neurodegenerative disease identification. *IEEE Transactions on Biomedical Engineering* **63**, 607–618 (2016).

19. Yong Fan, Hengyi Rao, Hallam Hurt, Joan Giannetta, Marc Korczykowski, David Shera, Brian B. Avants, James C. Gee, Jiongjiong Wang, Dinggang Shen. Multivariate examination of brain abnormality using both structural and functional MRI. *NeuroImage* **34**(4), 1189–1199 (2007).
20. Xiaofeng Zhu, Heung-Il Suk, Dinggang Shen. A novel matrix-similarity based loss function for joint regression and classification in AD diagnosis. *NeuroImage* **100**, 91–105 (2014).
21. Chen, X. *et al.* High-order resting-state functional connectivity network for MCI classification. *Human Brain Mapping* **37**, 3282–3296 (2016).
22. Allen, E. *et al.* Tracking whole-brain connectivity dynamics in the resting state. *Cerebral Cortex* **24**, 663–676 (2012).
23. Chen, X., Zhang, H., Zhang, L., Shen, C., Lee, S.W. and Shen, D. Extraction of dynamic functional connectivity from brain grey matter and white matter for MCI classification. *Human Brain Mapping* **2**, doi:[10.1002/hbm.23711](https://doi.org/10.1002/hbm.23711) (2017).
24. Smith, S. *et al.* Network modelling methods for fMRI. *NeuroImage* **54**, 875–891 (2011).
25. Smith, S. *et al.* Functional connectomics from resting-state fMRI. *Trends in Cognitive Sciences* **17**, 666–682 (2013).
26. Stam, C. *et al.* Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer's disease. *Brain* **132**, 213–224 (2009).
27. Honey, C. *et al.* Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences* **106**, 2035–2040 (2009).
28. Mohan, A., De Ridder, D. & Vanneste, S. Graph theoretical analysis of brain connectivity in phantom sound perception. *Scientific Reports* **6**, 19683 (2016).
29. Wang, J. *et al.* Multi-task diagnosis for autism spectrum disorders using multi-modality features: A multi-center study. *Human Brain Mapping* (2017).
30. Wang, P. *et al.* Aberrant intra-and inter-network connectivity architectures in Alzheimer's disease and mild cognitive impairment. *Scientific Reports* **5**, 14824 (2015).
31. Huang, S. *et al.* Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation. *NeuroImage* **50**, 935–949 (2010).
32. Qiao, L. *et al.* Estimating functional brain networks by incorporating a modularity prior. *NeuroImage* **141**, 399–407 (2016).
33. Lv, J. *et al.* Sparse representation of whole-brain fMRI signals for identification of functional networks. *Medical Image Analysis* **20**, 112–134 (2015).
34. Wright, J., Yang, A., Ganesh, A., Sastry, S. & Ma, Y. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **31**, 210–227 (2009).
35. Zhang, Y., Jin, J., Qing, X., Wang, B. & Wang, X. Lasso based stimulus frequency recognition model for ssvep bcis. *Biomedical Signal Processing and Control* **7**, 104–111 (2012).
36. Zhang, H. *et al.* Topographical information-based high-order functional connectivity and its application in abnormality detection for mild cognitive impairment. *Journal of Alzheimer's Disease* **54**, 1095–1112 (2016).
37. Guimera, R., Danon, L., Diaz-Guilera, A., Giralt, F. & Arenas, A. Self-similar community structure in a network of human interactions. *Physical review E* **68**, 065103 (2003).
38. Deng, L., Sun, J., Cheng, L. & Tong, S. Characterizing dynamic local functional connectivity in the human brain. *Scientific Reports* **6**, 26976 (2016).
39. Damaraju, E. *et al.* Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia. *NeuroImage: Clinical* **5**, 298–308 (2014).
40. Hutchison, R. M. *et al.* Dynamic functional connectivity: promise, issues, and interpretations. *NeuroImage* **80**, 360–378 (2013).
41. Thompson, W. H. & Fransson, P. Bursty properties revealed in large-scale brain networks with a point-based method for dynamic functional connectivity. *Scientific Reports* **6**, 39156 (2016).
42. Leonardi, N. & Van De Ville, D. On spurious and real fluctuations of dynamic functional connectivity during rest. *NeuroImage* **104**, 430–436 (2015).
43. Zalesky, A. & Breakspear, M. Towards a statistical test for functional connectivity dynamics. *NeuroImage* **114**, 466–470 (2015).
44. Zhang, Y. *et al.* Sparse Bayesian classification of EEG for brain-computer interface. *IEEE Transactions on Neural Networks and Learning Systems* **27**, 2256–2267 (2016).
45. Zhang, Y., Wang, Y., Jin, J. & Wang, X. Sparse Bayesian learning for obtaining sparsity of EEG frequency bands based feature vectors in motor imagery classification. *International Journal of Neural Systems* **27**, 1650032 (2017).
46. Kosicek, M. & Hecimovic, S. Phospholipids and alzheimer's disease: alterations, mechanisms and potential biomarkers. *International Journal of Molecular Sciences* **14**, 1310–1322 (2013).
47. Jacobs, H. *et al.* Functional integration of parietal lobe activity in early alzheimer disease. *Neurology* **78**, 352–360 (2012).
48. Wang, K. *et al.* Altered functional connectivity in early alzheimer's disease: a resting-state fmri study. *Human Brain Mapping* **28**, 967–978 (2007).
49. Arnold, S. E., Hyman, B. T. & Van Hoesen, G. W. Neuropathologic changes of the temporal pole in alzheimer's disease and pick's disease. *Archives of Neurology* **51**, 145–150 (1994).
50. Ding, B. *et al.* Correlation of iron in the hippocampus with mmse in patients with alzheimer's disease. *Journal of Magnetic Resonance Imaging* **29**, 793–798 (2009).
51. Yao, Z. *et al.* Abnormal cortical networks in mild cognitive impairment and alzheimer's disease. *PLoS Comput Biol* **6**, e1001006 (2010).
52. Kogure, D. *et al.* Longitudinal evaluation of early alzheimer's disease using brain perfusion spect. *Journal of Nuclear Medicine* **41**, 1155–1162 (2000).
53. Baloyannis, S. J. Mitochondrial alterations in alzheimer's disease. *Journal of Alzheimer's Disease* **9**, 119–126 (2006).
54. Echávarri, C. *et al.* Atrophy in the parahippocampal gyrus as an early biomarker of alzheimer's disease. *Brain Structure and Function* **215**, 265–271 (2011).
55. Magnin, B. *et al.* Support vector machine-based classification of alzheimer's disease from whole-brain anatomical mri. *Neuroradiology* **51**, 73–83 (2009).
56. Salvatore, C. *et al.* Magnetic resonance imaging biomarkers for the early diagnosis of alzheimer's disease: a machine learning approach. *Frontiers in Neuroscience* **9**, 1–13 (2015).
57. Li, Y. *et al.* Discriminant analysis of longitudinal cortical thickness changes in alzheimer's disease using dynamic and network features. *Neurobiology of Aging* **33**, 427.e415–427.e430 (2012).
58. Peters, F. *et al.* The neural correlates of verbal short-term memory in alzheimer's disease: an fmri study. *Brain* awp075 (2009).
59. Golby, A. *et al.* Memory encoding in alzheimer's disease: an fmri study of explicit and implicit memory. *Brain* **128**, 773–787 (2005).
60. He, Y. *et al.* Regional coherence changes in the early stages of alzheimer's disease: a combined structural and resting-state functional mri study. *NeuroImage* **35**, 488–500 (2007).
61. Ungerleider, L. G. Two cortical visual systems. *Analysis of visual behavior* 549–586 (1982).
62. Buffalo, E. A., Fries, P., Landman, R., Liang, H. & Desimone, R. A backward progression of attentional effects in the ventral stream. *Proceedings of the National Academy of Sciences* **107**, 361–365 (2010).
63. Menon, V. Large-scale brain networks and psychopathology: a unifying triple network model. *Trends in Cognitive Sciences* **15**, 483–506 (2011).

64. Jie, B., Wee, C.-Y., Shen, D. & Zhang, D. Hyper-connectivity of functional networks for brain disease diagnosis. *Medical Image Analysis* **32**, 84–100 (2016).
65. Damoiseaux, J. *et al.* Consistent resting-state networks across healthy subjects. *Proceedings of the National Academy of Sciences* **103**, 13848–13853 (2006).
66. Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: uses and interpretations. *NeuroImage* **52**, 1059–1069 (2010).
67. Abdi, H. & Williams, L. Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics* **2**, 433–459 (2010).
68. Tibshirani, R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)* **58**, 267–288 (1996).
69. Zhang, Y. *et al.* Aggregation of sparse linear discriminant analyses for event-related potential classification in brain-computer interface. *International Journal of Neural Systems* **24**, 1450003 (2014).
70. De Bie, T., Tranchevent, L.-C., Van Oeffelen, L. M. & Moreau, Y. Kernel-based data fusion for gene prioritization. *Bioinformatics* **23**, i125–i132 (2007).
71. Jie, B., Zhang, D., Cheng, B. & Shen, D. Manifold regularized multitask feature learning for multimodality disease classification. *Human Brain Mapping* **36**, 489–507 (2015).
72. Yu, S. *et al.* L 2-norm multiple kernel learning and its application to biomedical data fusion. *BMC Bioinformatics* **11**, 1 (2010).
73. Chang, C.-C. & Lin, C.-J. Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* **2**, 1–27 (2011).

## Acknowledgements

This study is partially supported by NIH grants (EB006733, EB008374, EB009634, MH107815, AG041721, and AG042599). This study is also partially supported by Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korea government (No. 2017-0-00451).

## Author Contributions

Y.Z. implemented the code and experimental study. Y.Z. and H.Z. drafted the manuscript. X.C., S.W.L. and D.S. participated in idea discussion and revised the manuscript.

## Additional Information

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017