Check for updates

A Tensor-Based Framework for rs-fMRI Classification and Functional Connectivity Construction

Ali Noroozi and Mansoor Rezghi*

Department of Computer Science, Tarbiat Modares University, Tehran, Iran

Recently, machine learning methods have gained lots of attention from researchers seeking to analyze brain images such as Resting-State Functional Magnetic Resonance Imaging (rs-fMRI) to obtain a deeper understanding of the brain and such related diseases, for example, Alzheimer's disease. Finding the common patterns caused by a brain disorder through analysis of the functional connectivity (FC) network along with discriminating brain diseases from normal controls have long been the two principal goals in studying rs-fMRI data. The majority of FC extraction methods calculate the FC matrix for each subject and then use simple techniques to combine them and obtain a general FC matrix. In addition, the state-of-the-art classification techniques for finding subjects with brain disorders also rely on calculating an FC for each subject, vectorizing, and feeding them to the classifier. Considering these problems and based on multi-dimensional nature of the data, we have come up with a novel tensor framework in which a general FC matrix is obtained without the need to construct an FC matrix for each sample. This framework also allows us to reduce the dimensionality and create a novel discriminant function that rather than using FCs works directly with each sample, avoids vectorization in any step, and uses the test data in the training process without forcing any prior knowledge of its label into the classifier. Extensive experiments using the ADNI dataset demonstrate that our proposed framework effectively boosts the fMRI classification performance and reveals novel connectivity patterns in Alzheimer's disease at its early stages.

Keywords: Alzheimer's disease (AD) classification, functional connectivity, tensor, high order singular value decomposition, dimension reduction

1. INTRODUCTION

Alzheimer's disease (AD) is a progressive neurodegenerative disorder with a long pre-morbid asymptomatic period, which affects millions of elderly individuals worldwide (Caselli et al., 2004). It is predicted that the number of affected people will double in the next 20 years, and 1 in 85 people will be affected by 2050 (Brookmeyer et al., 2007). The predominant clinical symptoms of AD include a decline in some important brain cognitive and intellectual abilities such as memory, thinking, and reasoning. Early detection is important for possible delay of the progression of mild MCI to moderate and severe stages (Folch et al., 2016). However, diagnosis of MCI is difficult due to its mild symptoms of cognitive impairment, causing most computer-aided diagnosis to achieve lower-than-desired performance (Musha et al., 2013; Li R. et al., 2018). Precise diagnosis of AD, especially in its early warning stage, that is, early Mild Cognitive Impairment (eMCI), enables treatments to delay or even avoid such disorders.

In recent years, medical imaging techniques such as positron emission tomography (PET) (Chandra et al., 2019), electroencephalography (EEG) (Bi and Wang, 2019), computed tomography (CT) scan (Ozdemir et al., 2019; van de Leemput et al., 2019), intracoronary imaging (Gao et al., 2019), and functional magnetic resonance imaging, which is a non-invasive brain imaging technique (fMRI) (Golby et al., 2005), have been used in order to analyze and detect disorders within body and brain (Zhang et al., 2012; Han et al., 2013). Due to high spatial resolution, fMRI is vastly used among researchers in order to monitor brain activities, especially in AD and all its stages in which detecting abnormalities within small brain regions is essential (Dennis and Thompson, 2014). An fMRI sample is naturally a 4D tensor consisting of 3D time-varying voxels, and each voxel contains an intensity value that is proportional to the strength of the **B**lood **O**xygenation **L**evel **D**ependent (BOLD) signal, which is a measure of the changes in blood flow to estimate the activity of different brain regions. Resting-state fMRI (rs-fMRI) is an fMRI technique in which the patient is asked to rest during the whole scan and it focuses on the low-frequency ($< 0.1Hz$) oscillations of BOLD signal presenting the underlying neuronal activation patterns of brain regions. rs-fMRI is usually used in order to analyze brain diseases like AD or Autism (Leonardi et al., 2013; Kazeminejad and Sotero, 2019; Nguyen et al., 2019). Different toolboxes such as GraphVar (Waller et al., 2018), Graph CNN (Zhang et al., 2019), and BrainNetClass (Zhou et al., 2020) are also developed to aid this cause.

Since each fMRI series consists of hundreds of thousands of voxels, which are often highly correlated with the surrounding voxels in the brain volume, parcellation of the brain for further analysis has moved toward the use of anatomical atlases. These atlases are strictly defined using anatomical features of the brain like locations of common gyri and do not rely on any functional information. To generate data using an atlas-based approach, the BOLD signal from all voxels is averaged within each brain region called region of interest (ROI) (Stanley et al., 2013). By putting together the average time series for all the ROIs, the ith series becomes $X_i \in \mathbb{R}^{T \times R}, i = \{1, 2, \cdots, S\}$, in which R, T, and S are the number of ROIs, time points, and samples, respectively. This process is illustrated in **Figure 1**

There are two major studies associated with rs-fMRI data: finding common brain disorders caused by diseases such as AD, autism, schizophrenia, and so on, and more recently detecting patients with brain disorders using classification techniques (de Vos et al., 2018; Du et al., 2018). Due to the high dimensionality of data along with the nature of diseases such as eMCI, which does not show any reliable clinical symptoms, researchers have moved toward advanced machine learning techniques in order to achieve more reliable analysis (Cuingnet et al., 2011).

A powerful tool that is commonly used in order to achieve aforementioned goals is the functional connectivity (FC) network. Let X_i be the ith sample, its corresponding FC, and $\bar{X}$ is a *region × region* matrix in which $\bar{x}_{ij}$ represents the FC between the ith and jth ROI. Functional connectivity is an observable phenomenon quantifiable with measures of statistical dependencies such as correlations, coherence, or transfer entropy
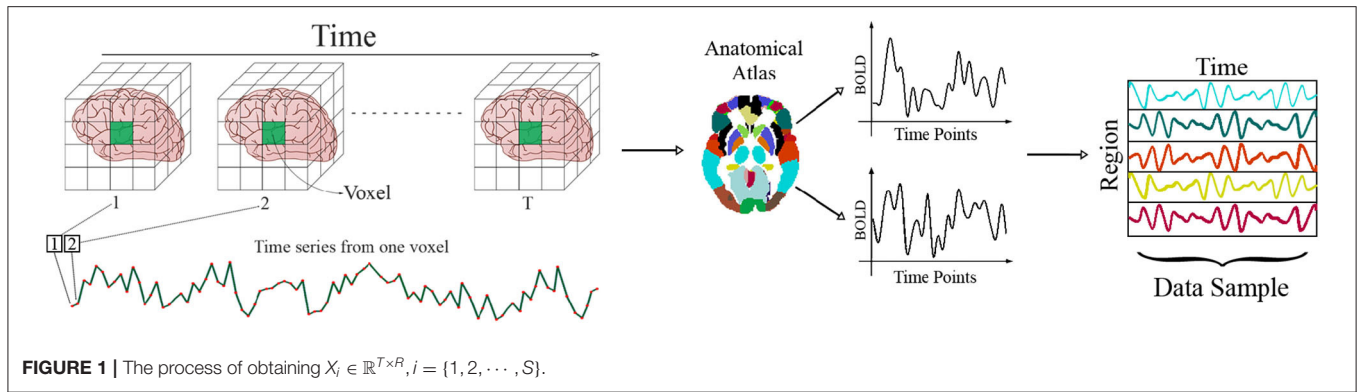
(Friston, 2011). Recent studies have shown that some brain disorders such as AD could alter the way some brain regions interact with each other. For example, compared with healthy subjects, AD patients have been found with decreased FC between the hippocampus and other brain regions, and MCI patients have been observed with increased FC between the frontal lobe and other brain regions (Dennis and Thompson, 2014).

FCs are also used as features in classification. So, instead of using X_i as the ith sample, its corresponding FC, that is, $\bar{X}_i$, is used as a feature. Common techniques for calculating FC, that is, simple statistical measures such as coherence and Pearson correlation, allow for different ambiguities (Smith, 2012; Reid et al., 2019). And since brain alterations in early MCI are tiny, more sophisticated and computationally expensive methods such as partial correlation (Li Y. et al., 2018; Pervaiz et al., 2020), high-order networks (Chen et al., 2016), and spectral clustering (Liu et al., 2018) are required in order to achieve a better FC. The computational cost of a sophisticated FC is usually high and also its quality affects the performance of the learning process massively. Also, since the conventional classifiers like **S**upport **V**ector **M**achine (SVM) or k-NN works on data in vector format, these matrix features should be vectorized in order be fed to these classifiers. This vectorization leads to high-dimensional vectors that produce poor performance due to the phenomenon known as the curse of dimensionality. Alongside the curse of dimensionality, vectorization also destroys potential information embedded in the structure of data. This problem has been studied especially in image data in which vectorization destroys the spatial relations within an image (Ahmadi and Rezghi, 2020).

In this paper, based on high-order tensor decomposition, we have created a framework in which the aforementioned goals, that is finding a general FC and detecting a disorder through classification, could be achieved via a single **H**igh-**O**rder **S**ingular **V**alue **D**ecomposition (HOSVD) of each class. Here based on latent variables obtained by HOSVD, a general representative pattern of FC for eMCI and normal controls are obtained. As it was mentioned before, finding a proper FC is a challenging task. Obtaining an FC via the proposed method is not only fast and straightforward, but also very accurate. The majority of connectivity patterns detected by this method have been observed and studied in several separated types of research (cited in the experimental studies section), which show the reliability and power of the proposed method. Along with these connections, we have also detected novel connectivities especially regarding the cerebellum, which is usually discarded in the analysis of AD.

The proposed classifier is also much faster than the state-of-the art classifiers, and also uses the obtained rs-fMRI data directly in the classification process rather than calculating its corresponding FC matrix. Being able to feed X_is directly to the classifier saves us from a lot of problems related to calculating the proper FC. It also shows that the proposed classifier is strong enough to detect tiny alterations, which other state-of-the-art methods rely on finding FC to highlight.

To verify our approach, we conducted an extensive experimental study on rs-fMRI data from the benchmark

FIGURE 1 | The process of obtaining $X_i \in \mathbb{R}^{T \times R}, i = \{1, 2, \cdots, S\}$.

dataset ADNI. As will be seen, the results demonstrate the effectiveness and advantages of our method. Specifically, the proposed framework not only grants us superior classification accuracy to that from other methods, but it is also much faster and more stable against different data selection schemes. We have also confirmed our achieved general FC matrix using empirical data on the eMCI and normal FC patterns.

2. RELATED WORKS

As it was mentioned previously, obtaining and classifying FC matrices have become the dominant approach toward eMCI analysis. Variety of methods such as pairwise Pearson's correlation coefficient, sparse representation (Jie et al., 2013), and **S**parse **I**nverse **C**ovariance **E**stimation (SICE) (Huang et al., 2010) exist to obtain an FC. While the first two are easy to understand and can capture pairwise functional relationship based on a pair of ROIs, the latter can account for more complex interactions among multiple ROIs, but the estimation of partial correlation involves an inversion of a covariance matrix, which may be ill posed due to the singularity of the covariance matrix. These methods result in vastly different networks (Du et al., 2018). On the other hand, computing the correlations, based on the entire time series of fMRI data simply measures the FC between ROIs with a scalar value, which is fixed across time. This actually implicitly hypothesizes the Stationary interaction patterns among ROIs, which will result in a static functional connectivity (sFC). As a result, this method may overlook the complex and dynamic interaction patterns among ROIs, which are essentially time-varying (since the phase is not locked for every subject). In order to overcome this issue, Non-stationary methods have been proposed, which result in more complex networks and also known as dynamic functional connectivity (dFC) (Leonardi and Van De Ville, 2015; Kam et al., 2019). The most common and straightforward way to investigate dFC is using windowed FC, which consists of calculating a given FC measure, for example, the Pearson correlation coefficient, over consecutive windowed segments of the data (Zalesky et al., 2014). Although such an analysis seems straightforward, there are also pitfalls associated with it, which may cause in a non-accurate FC network (Hindriks et al., 2016).

In the following, we briefly discuss two state-of-the-art eMCI classification techniques belonging to these two paradigms:

Kernel compact SICE (K−SIEC): SICE matrix have proven itself to be one of the best sFC models (Huang et al., 2010; Ng et al., 2013; Colclough et al., 2018; Foti and Fox, 2019), which is extracted via the following optimization:

$$S^* = \arg\max_{S \succ 0} \quad \log\left(\det(S)\right) - \mathrm{tr}(CS) - \lambda \|S\|_1 \qquad (1)$$

where C is the sample-based covariance matrix; $\det(\Delta)$, $\mathrm{tr}(\Delta)$, and $\|.\|_1$ denote the determinant, trace, and the sum of the absolute values of the entries of a matrix, respectively. In classification with FC features, the vectorized SICE of each sample is used (Leonardi et al., 2013). The occurrence of the curse of dimensionality and losing useful information contained in the SICE matrices [like symmetric positive definite (SPD) property] are two main drawbacks of this vectorization approach. As an inverse covariance matrix, an SICE matrix is SPD. This inherent property restricts SICE matrices to a lower dimensional Riemannian manifold rather than the full dimensional Euclidean space. This property allows some SPD manifold-based distances, like log-Euclidean distance (Arsigny et al., 2006) and Root Stein divergence (Sra, 2012) to be employed in kernel-based PCA to extract a compact representation of brain network (Zhang et al., 2015). The power of this method resides in a massive dimension reduction of SICE using its SPD property. The performance of this method also heavily relies on the choice of sparsity parameter λ for SICE calculations and the number of top eigenvectors m.

High−order networks (HON): This method which is proposed in Chen et al. (2016) belongs to non-stationary paradigm and uses the so-called high-order networks as features for classification purposes. It uses the sliding-window technique in order to split the time series into smaller pieces and then find the relation between them (Chang and Glover, 2010; Handwerker et al., 2012; Allen et al., 2014). Let $x_i^{(l)}(k) \in \mathbb{R}^N$ denotes the kth segment of the ith region in the lth sample. For each sample, a network with nodes $x_i^{(l)}(k)$ could be constructed, in which its edge weights are obtained as

$$C_{ij}^{(l)}(k) = \mathrm{corr}\left(x_i^{(l)}(k), x_j^{(l)}(k).\right)$$

Here, the weight $C_{ij}^{(l)}(k)$ represents the pairwise Pearson's correlation coefficients between the ith and the jth ROIs of the lth subject using the kth segment of subseries. Now

$$y_{ij}^{(l)} = \left[C_{ij}^{(l)}(1), C_{ij}^{(l)}(2), \cdots, C_{ij}^{(K)}(1) \right] \in \mathbb{R}^K$$

represents the similarity of the ith and jth ROIs of the lth sample in all segments. For each l by considering $y_{ij}^{(l)}$ as nodes of a networks with weights

$$H_{ij,pq}^{(l)} = \mathrm{corr}\left(y_{ij}^{(l)}, y_{pq}^{(l)} \right)$$

a higher-order network is obtained for each sample. Here for each pair of correlation time series y_{ij} and y_{pq}, $H_{ij,pq}^{(l)}$ indicates how the correlation between the ith and the jth ROIs influences the correlation between the pth and the qth ROIs. So for each sample its higher-order networks $\{H_{ij,pq}^{(l)}\}$ will be a matrix with size $R^4 \times R^4$ (R is the number of regions), which will lead to a large-scale high-order FC network, containing at least thousands of vertices and millions of edges. In order to overcome this issue, the correlation time series within each subject are grouped into different clusters. Then, the correlation computations are carried out between the means of clusters. After reducing the network size, the weighted-graph local clustering coefficients is used to select the key features for each network and then an SVM classifier is trained in order to classify the obtained features. As a result of constructing a high-order network, the notion of a physical ROI become vague and thus such networks are not preferable choices in order to analyze functional connectivities.

Our method overcomes the dynamic-stationary problem of FC construction by working in HOSVD-based domain, which considers the dynamic nature of data and is much more sophisticated than using a windowed FC. The obtained FC also considers all subjects within a class simultaneously, rather than calculating FC for each subject separately that highlights common patterns in a class and eliminates possible outliers within data. The proposed framework also does not require any FC calculations for classification, which is a major advantage since finding a proper FC for each subject might be a very challenging task.

Multilinear approaches have been used before in order to analyze fMRI data. For example, Park (2011) uses multilinear PCA to classify fMRI data by Subject and Motor Task. Ozdemir et al. (2017) and Al-sharoa et al. (2018) use tensor decomposition and clustering techniques for analyzing brain connectivity networks and proves the dynamic nature of rs-fMRI. Recently, Ma et al. (2016) and He et al. (2017) proposed a multilinear method for voxel-wise analysis of rs-fMRI, which is used in order to detect late AD and some other diseases. Leonardi and Van De Ville (2013) considers dynamic whole-brain FC estimated from fMRI data acquired during alternating epochs of resting and watching of movie excerpts, and uses HOSVD in order to retrieve connectivity maps with associated time courses and subject loadings. This method uses the sliding-window technique in order to estimate the dynamic connectivity matrix for each

subject, and then it constructs a 3-way tensor $\mathcal{R} \in \mathrm{R}^{C \times T \times T}$, by stacking the dynamic correlation matrices R of all subjects. Considering the HOSVD of R, this method obtains a matrix columns of which could be interpreted as group connectivity maps. There are similarities between this method and ours since they both take advantage of HOSVD. But our framework introduces major advantages such as (1) our framework does not require any FC calculations for its classifier. And (2), it is able to work with rs-fMRI, which is harder due to less constraint status of subjects.

3. NOTATION AND PRELIMINARIES

Tensors can be considered as a generalization of vectors and matrices of high dimensions. We use calligraphic letters to denote the tensors, for example, $(\mathcal{A}, \mathcal{B})$. Let $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ denote an order-3 tensor. Different "dimensions" of tensors are referred to as *modes*. We will use both standard subscripts and "MATLAB-like" notation to show tensor elements as follows:

$$\mathcal{A}(i,j,k) = a_{ijk}.$$

A *fiber* is a subtensor, where all indices but one are fixed. For example, mode-2 fibers of $\mathcal{A}$ have the following form:

$$\mathcal{A}(i,:,j) \in \mathbb{R}^{I_2}.$$

The mode-n product of an order-M tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times \cdots \times I_M}$ by a matrix $X \in \mathbb{R}^{K \times I_n}$ is defined as:

$$\mathbb{R}^{I_1 \times \cdots \times I_{n-1} \times K \times I_{n+1} \times \cdots \times I_M} \ni \mathcal{B} = (X)_n \cdot \mathcal{A}, \quad (2)$$

where,

$$b_{i_1, \cdots, i_M} = \sum_{l=1}^{I_n} x_{i_n,l}, a_{i_1, \ldots, i_{n-1}, l, i_{n+1}, \ldots, i_M}.$$

This means that all mode-n fibers of $\mathcal{A}$ are multiplied by the matrix X. The notation (2) was suggested by De Silva and Lim (2008). An alternative notation was earlier given in De Lathauwer et al. (2000). $(X)_n \cdot \mathcal{A}$ is the same as $\mathcal{A} \times_n X$ in that system. The Frobenius norm of the order-M tensor $\mathcal{A}$ can be defined as $\|\mathcal{A}\| = \sum_{i_1, \cdots, i_M} a_{i_1, \cdots, i_M}^2$.

3.1. Higher-Order Singular Value Decomposition

HOSVD is one common extension of singular value decomposition to the tensors (De Lathauwer et al., 2000). Using HOSVD, every order-M tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times \cdots \times I_M}$ can be decomposed as:

$$\mathcal{A} = \left(U^{(1)}, \cdots, U^{(M)} \right) \cdot \mathcal{S} \quad (3)$$

where orthogonal matrices $U^{(i)}$ are singular matrices of tensor $\mathcal{A}$. Here, $U^{(i)}$ is the left singular matrix of $A^{(i)}$, in which its column

vectors are the mode-n fibers of $\mathcal{A}$. The core tensor $\mathcal{S}$ is a real tensor of the same dimensions as $\mathcal{A}$ and

$$\mathcal{S} = \left(U^{(1)^\mathsf{T}}, \cdots, U^{(M)^\mathsf{T}} \right) \cdot \mathcal{A}$$

Although this core tensor is not diagonal as in the case of SVD of matrices, it satisfies the following conditions:

- **All orthogonality property**: Any two different slices along the same mode are orthogonal. This property of core tensor S is named as all orthogonality.
- **The ordering property**: The values $s_j^k = \left\| \mathcal{S}(:, \cdots, :, j, :, \cdots, :) \right\|$, where j is in the kth mode of $\mathcal{S}$, are named mode-k singular values of $\mathcal{A}$. It can be shown that for every k

$$s_1^k \geq s_2^k \geq \cdots \geq s_n^k \geq 0, \qquad k = 0, \cdots, M, \tag{4}$$

  are equal to the singular values of the matrix $A^{(k)}$. This means that the norms of the slices along every mode are ordered.
- **Oscillation**: It can be shown that as the indices increase, the singular vectors of each mode shows more oscillation. Based on this property, it can be shown that noises and outliers within the data are transferred into these high oscillation parts (Rezghi, 2017). Based on this fact and also the ordering property, the truncated version of HOSVD can be deployed as a noise reduction and compression tool (Lv and Wang, 2019).

The ordering property (4) demonstrates that, in the same way as matrices, singular values measure the *energy* of the tensor. So, it is easy to see that the energy of core tensor $\mathcal{S}$ focused on the elements of $\mathcal{S}$ with small indices. This property of HOSVD (similar to SVD) is very useful in the applications that encounter denoising problems. So, if $U_{k_l}^l$ contains the first k_l singular matrix and $\hat{\mathcal{S}} = \mathcal{S}(1 : k_1, \cdots, 1 : k_M)$, the following truncated HOSVD:

$$\hat{\mathcal{A}} = \left(U_{k_1}^{(1)}, \cdots, U_{k_M}^{(M)} \right)_{1:M} \cdot \hat{\mathcal{S}},$$

is a rank-$(k_1, \cdots .k_M)$ approximation of $\mathcal{A}$. Although this is not an optimal rank-$(k_1, \cdots .k_M)$ approximation of $\mathcal{A}$, it is still a good approximation and we have:

$$\left\| \mathcal{A} - \hat{\mathcal{A}} \right\| = \sum_{i=1}^{M} \sum_{j=k_i+1}^{r_i} s_j^{(i)^2},$$

where r_i is the rank of A^i (De Lathauwer et al., 2000).

4. PROPOSED FMRI ANALYSIS FRAMEWORK BASED ON HOSVD

In this section, which is divided into three subsections we first tackle the problem of classification, that is, designing a discriminant function that could predict the label of an unknown test subject. The second part describes a technique, which would enhance the designed classifier and the third part is allocated to find a general connectivity network for each class (e.g., eMCI

subjects). All three aforementioned goals, that is, classification and its enhancement and finding a general FC for each class, evolve around a single HOSVD of each class, which provides us with basis for each mode (*time*, *region*, and *sample*) and enables us to capture the essence of each feature in a few low dimensional slices. We will use the obtained low-dimensional bases along the sample and region mode in order to design our discriminant function and obtain the general FC. The enhancement technique also comes from HOSVD characteristics, which enables us to involve test samples in the training process without forcing any a priori knowledge into the classifier.

4.1. eMCI Classification

Let tensors $\mathcal{X}^{(i)} \in \mathbb{R}^{T \times R \times S_i}$ consists of normal and eMCI data for $i = 1, 2$, respectively. Here S_1, S_2 are the number of normal and eMCI samples. For tensor $\mathcal{X}^{(i)}$, the decomposition

$$\mathcal{X}^{(i)} = \left(U^{(i)}, V^{(i)}, W^{(i)} \right) \cdot \mathcal{S}^{(i)}, \tag{5}$$

is known as HOSVD, where orthogonal matrices $U^{(i)} \in \mathbb{R}^{T \times T}$, $V^{(i)} \in \mathbb{R}^{R \times R}$, and $W^{(i)} \in \mathbb{R}^{S_i \times S_i}$ are known as modes-1,2,3 singular matrices of $\mathcal{X}^{(i)}$, and $\mathcal{S}^{(i)}$ is the corresponding core tensor (Rezghi, 2017). Here, $U^{(i)}$ is a base of all mode-1 fibers $\mathcal{X}^{(i)}(:, l, k)$, which indicates the behavior of lth region of the kth sample of the ith class in all time points. Also $V^{(i)}$ is a base of all mode-2 fibers $\mathcal{X}(l, :, k)$, which indicates the behavior of all regions of lth sample of the ith class in the kth time. Due to the properties of HOSVD inherited from svd, the first columns of the kth singular matrix ($k = 1, 2, 3$) have more ability in construction of main parts of kth fibers (Rezghi, 2017). Therefore, a suitable dimension reduction would be to project the mode-1 and mode-2 fibers into space spanned by the first k_1^i and k_2^i singular vectors of modes-1,2, which will be denoted by $U_{k_1^i}^{(i)}$ and $V_{k_2^i}^{(i)}$, respectively. This dimension reduction could be done as:

$$\mathbb{R}^{k_1 \times k_2 \times S_i} \ni \bar{\mathcal{X}}^{(i)} = \left(U_{k_1^i}^{(i)\mathsf{T}}, V_{k_2^i}^{\mathsf{T}} \right)_{1,2} \cdot \mathcal{X}^{(i)} \tag{6}$$

It is clear that this reduction could be done separately on each mode without the need to fold any of them. This means that the structural integrity of data is preserved during the dimension reduction process, which is a key aspect in our work. It has been shown that even choosing relatively small values for k_1^1 and k_2^i would result in a very good reconstruction error (Ahmadi and Rezghi, 2020).

Inspired by the structure of this reduction, in the following we present a tensor-based discriminant function. By HOSVD decomposition of $\mathcal{X}^{(i)}$, the projected data $\overline{\mathcal{X}}^{(i)}$ in Equation (6) becomes

$$\bar{\mathcal{X}}^{(i)} = \left(\left[I_{k_1^i} \ 0 \right], \left[I_{k_2^i} \ 0 \right], W^{(i)} \right) \cdot \mathcal{S}^{(i)}$$
$$= \left(W^{(i)} \right)_3 \cdot \mathcal{S}^{(i)}(1 : k_1, 1 : k_2, :)$$

So, each sample of the ith class in the reduced space has the following form:

$$\bar{\mathcal{X}}^{(i)}(:,:,k) = \left(W^{(i)}(k,:) \right)_3 \cdot \mathcal{S}^{(i)}(1:k_1^i, 1:k_2^i, :)$$

$$= \sum_{k'=1}^{S_i} W^{(i)}(k,k') \cdot \mathcal{S}^{(i)}(1:k_1^i, 1:k_2^i, k').$$

This means that each sample in the ith class could be represented as linear combination of the slices of the tensor $\bar{\mathcal{S}}^{(i)} = \mathcal{S}^{(i)}(1:k_1^i, 1:k_2^i, :)$. So if a test data like $X \in \mathbb{R}^{T \times R}$ belongs to the ith class, it is natural to expect that its projected version into principle region and times spaces, spanned by $U_{k_1^i}, V_{k_2^i}$, that is,

$$Z^{(i)} = \left(U_{k_1^i}^{(i)\mathsf{T}}, V_{k_2^i}^{(i)\mathsf{T}} \right)_{1,2} \cdot X$$

could be approximated well as a linear combination of the slices of the tensor $\bar{\mathcal{S}}^{(i)}$ as follows:

$$Z^{(i)} \approx \sum_{k=1}^{S_i} \lambda_k^i \bar{\mathcal{S}}^{(i)}(:,:,k). \tag{7}$$

Based on this viewpoint, each test data X could be assigned to a class that its projected version has the best approximation in the form (7). Due to the importance of core tensor elements with small indices in the reconstruction of the signal part of data in comparison with its last parts, the small number $k_3^i < S_i$ of slices $\bar{\mathcal{S}}^{(i)}(:,:,k)$ could be used in (7). In this viewpoint, each test data X would be assigned to the lth class, if $r_l = \min_{i=1,2} r_i$, where

$$r_i = \min_{\lambda^i} \| Z^{(i)} - \sum_{k=1}^{k_3^i} \lambda_k^i \bar{\mathcal{S}}^{(i)}(:,:,k) \|, \quad \lambda^i = \begin{pmatrix} \lambda_1^i \\ \vdots \\ \lambda_{s_i}^i \end{pmatrix} \tag{8}$$

r_i shows the reconstruction error of the projected version of X in the ith class.

4.2. Enhancing the Classifier

Consider that the test data X is added to dataset $\mathcal{X}^{(i)}$ of the ith class. So the new dataset will be $\widetilde{X} \in \mathbb{R}^{T \times R \times (S_i+1)}$:

$$\widetilde{\mathcal{X}}^{(i)}(:,:,1:S_i) = \mathcal{X}^{(i)},$$
$$\widetilde{\mathcal{X}}^{(i)}(:,:,S_i+1) = X.$$

If X belongs to the ith class, then in the decomposition of $\widetilde{\mathcal{X}}^{(i)}$, X would be able to reinforce all slices of the core tensor and singular matrices. And thus enhances the reconstruction ability of (8) that would lead into a lower reconstruction error for the test subject X. On the other hand, if X does not belong to the ith class, HOSVD would naturally consider it as noise [based on ordering property (4)], since X is not similar to other samples and thus does not play a key role in reconstructing them, so its effect would be on the last slices of the core tensor and singular matrices, that is, slices with higher indices that are ignored in reconstruction (8).

In order to better demonstrate this effect, we conducted the following experiment: we randomly chose a test subject X^n from the class of normal subjects in ADNI dataset (this dataset is explained in detail in the experimental study section). The remaining normal samples are then gathered in a tensor $\mathcal{X}^{(1)} \in \mathbb{R}^{130 \times 116 \times 37}$. By adding X^n to this tensor, we obtained the incremented tensor $\widetilde{\mathcal{X}}^{(1)} \in \mathbb{R}^{130 \times 116 \times 38}$. We compute the HOSVD of these two tensors and plot the absolute mode-3 differences in **Figure 2**. As can be seen in this figure, since X^n belongs to the normal class, it effectively changes almost all singular values and so could improve the approximation in Equation (8). Then we randomly select an eMCI sample X^e and add it to $\mathcal{X}^{(1)}$ to construct another incremented version of it. The orange line in equation (**Figure 2**) shows the absolute mode-3 differences between these two tensors. It can be observed that adding an eMCI subject to the class of normal subjects only affected the last singular values and have a very low impact on the first singular values.

It can be concluded that adding unknown labeled test data to all classes before the basis extraction process would heavily impact the true class bases, and it has a rather negligible or in ideal case zero impact on the bases of other classes. As a result, after extracting the basis for each class in this manner, the reconstruction error (Equation 8) would be lower for the true class. Note that in the training process, the test data are added to all classes and they are uninformed of its label. Thus, no a priori knowledge is sneaked into the decision-making process. Algorithm (1) summarizes the proposed classification method.

Algorithm 1: Tensor-based classification method

1) **Input**: Normal train data $\mathcal{X}^{(1)}$, eMCI train data, $\mathcal{X}^{(2)}$
   $k_j^i, i, j = 1, 2.$
   Test data X
2) Construct $\widetilde{\mathcal{X}}^{(i)}$ for $i = 1, 2$, by adding X to both tensors.
3) Compute $U_{k_1^i}, V_{k_2^i}$ and $\mathcal{S}(1:k_1^i, 1:k_2^i, :)$ of $\widetilde{\mathcal{X}}^{(i)}$, for i = 1,2.
4) Compute $Z^{(i)} = \left(U_{k_1^i}^{(i)\mathsf{T}}, V_{k_2^i}^{(i)\mathsf{T}} \right)_{1,2} \cdot X, i = 1, 2.$
5) Comput r_1, r_2 from (8)
6) Assign X to class l, if $l = \arg\min_i \{r_i\}$

4.3. General Functional Connectivity

In the ith class, which is represented by $\mathcal{X}^{(i)}$, the slice $\mathcal{X}^{(i)}(:,l,:)$ denotes the behavior of lth region of all samples in all times. This slice could be considered as a feature for the lth region of the ith class, so each region is represented as a Times-sample feature matrix. By the properties of singular matrices in modes-1,3, and for appropriate values k_1^i, k_3^i, each region $\mathcal{X}(:,l,:)$ could be reduced in both time and sample features separately based on mode-1 and mode-3 truncated singular matrices $U_{k_1^i}^{(i)}$ and $W_{k_3}^{(i)}$ as follows:

$$\mathcal{Y}^{(i)}(:,l,:) = \left(U_{k_1^i}^{(i)\mathsf{T}}, W_{k_3}^{(i)\mathsf{T}} \right)_{1,3} \cdot \mathcal{X}^{(i)}(:,l,:). \tag{9}$$
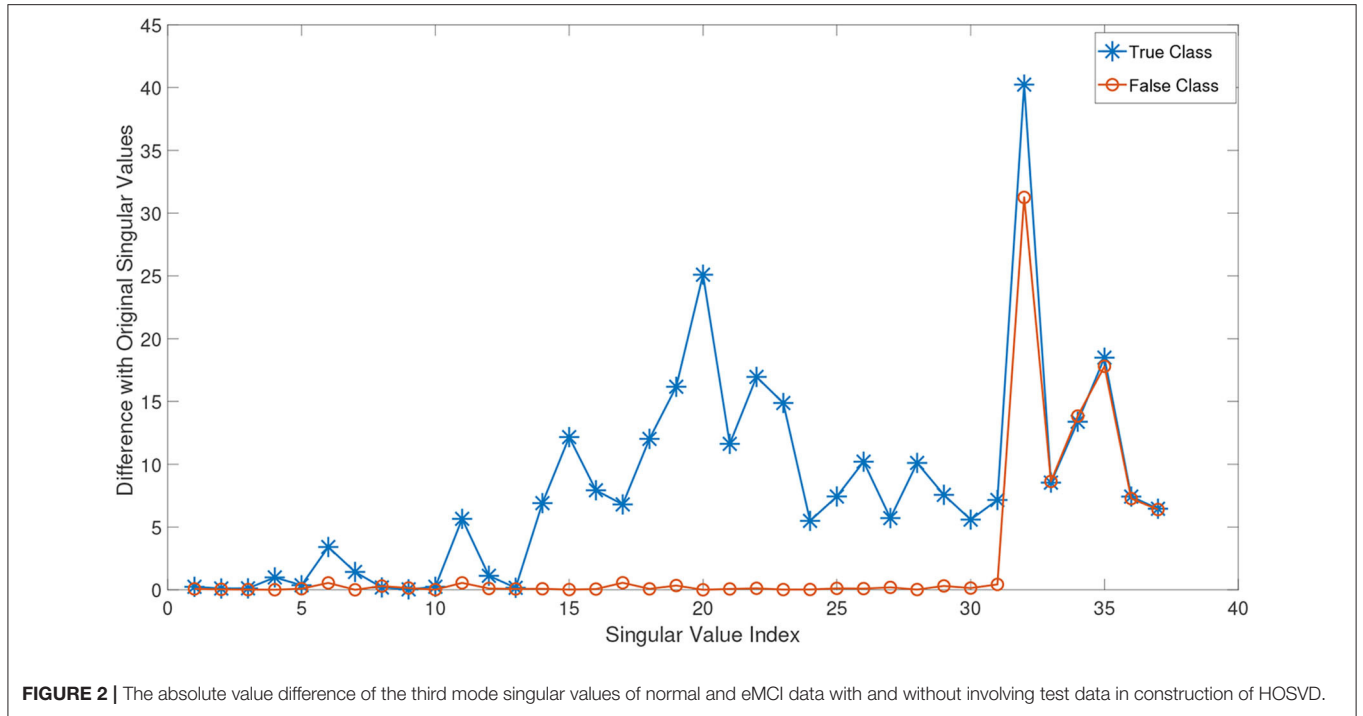
FIGURE 2 | The absolute value difference of the third mode singular values of normal and eMCI data with and without involving test data in construction of HOSVD.

Here $\mathcal{Y}^{(i)}(:, l, :)$ denotes a reduced version of $\mathcal{X}^{(i)}(:, l, :)$ into space spanned by $U_{k_1^i}^{(i)}$ and $W_{k_3^i}^{(i)}$ in modes-1,3. So,

$$\mathbb{R}^{k_1^i \times R \times k_3^i} \ni \mathcal{Y}^{(i)} = \left(U_{k_1^i}^{(i)\mathsf{T}}, W_{k_3^i}^{(i)\mathsf{T}} \right)_{1,3} \cdot \mathcal{X}^{(i)} \quad (10)$$

denotes all reduced regions of the ith class. By this structure and substituting the HOSVD decomposition of $\mathcal{X}^{(i)}$ in (10), we obtain

$$\mathcal{Y}^{(i)} = \left(\begin{bmatrix} I_{k_1^i} & 0 \end{bmatrix}, V, \begin{bmatrix} I_{k_3^i} & 0 \end{bmatrix} \right) \cdot \mathcal{S}^{(i)}$$
$$= (V)_2 \cdot \mathcal{S}^{(i)}(1 : k_1^i, :, 1 : k_3^i)$$

thus

$$\mathcal{Y}^{(i)}(:, k, :) = \sum_{k'}^{R} V^{(i)}(k, k') \bar{\mathcal{C}}^{(i)}(:, k', :)$$
$$= \left(V^{(i)}(k, :) \right)_2 \cdot \mathcal{C}^{(i)} \quad (11)$$

in which

$$\mathbb{R}^{k_1^i \times R \times k_3^i} \ni \mathcal{C}^{(i)} = \mathcal{S}(1 : k_1^i, :, 1 : k_3^i).$$

Equation (11) shows that the reduced version of each region in the ith class could be written as the linear combinations of mode-2 slices of $\mathcal{C}^{(i)}$. So the coefficients of slices in this linear combination could be considered as a new feature for the lth region of the ith class. Also as we mentioned before, the first slices are better than the last ones to reflect the principle properties of the data. So for appropriate k_3^i we could select only

the first coefficients in (11) as new features for the lth region. Mathematically, this means each region in the ith class could be represented by a new feature vector $V(l, 1; k_3^i) \in \mathbb{R}^{k_3^i}$.

This approach has two main benefits: (1) each region could be represented only by a vector with size k_3^i instead of a large time-sample matrix, and (2) the bases for each region is obtained in an HOSVD-based domain that is similar to Fourier frequency domain; but unlike Fourier, this transformation to HOSVD-domain is data dependent and hence the time-varying nature of rs-fMRI signals (2) would be taken into consideration (Rövid et al., 2013; Ozdemir et al., 2017). After representing each region with a single low dimensional vector, variety of methods such as SICE and other mentioned similarity measures could be deployed in order to construct a general FC for each class.

5. EXPERIMENTAL STUDY

5.1. Data Acquisition and Experimental Settings

RS-fMRI data of early MCI and NC patients were downloaded from ADNI website[1]. After removing subjects that had problems in the preprocessing steps, 44 eMCI and 38 NC subjects remained. The IDs of the 82 (38 NC and 44 early MCI) subjects are provided in the **Supplementary Material**.

The data are acquired on a 3-T (Philips) scanner with TR/TE set as 3,000/30 ms and flip angle of 80. Each series has 140 volumes, and each volume consists of 48 slices of image matrices with dimensions 64×64 with voxel size of $3.31 \times 3.31 \times 3.31 \ mm^3$. The preprocessing is carried out using

---

[1]http://adni.loni.usc.edu

SPM12 and DPARSFA (Chao-Gan and Yu-Feng, 2010). The first 10 acquired rs-fMRI volumes of each subject were initially discarded before any further processing to ensure magnetization equilibrium. The remaining 130 volumes were then corrected for the staggered order of slice acquisition that was used during echoplanar scanning. The correction ensures the data on each slice correspond to the same point in time. To further reduce the effects of nuisance signals, regression of ventricle and WM signals as well as six head-motion profiles was performed. rs-fMRI images were then normalized to the MNI space with resolution of $3.31 \times 3.31 \times 3.31 \ mm^3$ (Wee et al., 2016). Participants with too much head motion are excluded. The normalized brain images are warped into automatic anatomical labeling (AAL) (Tzourio-Mazoyer et al., 2002) atlas to obtain 116 ROIs as nodes. By following common practice (Park, 2011; Leonardi and Van De Ville, 2013; Al-sharoa et al., 2018), the ROI mean time series are extracted by averaging the time series from all voxels within each ROI and then bandpass filtered to obtain multiple sub-bands as in Al-sharoa et al. (2018). After the preprocessing steps, we obtained the normal samples $\mathcal{X}^{(1)} \in \mathbb{R}^{130 \times 116 \times 38}$ and eMCI samples $\mathcal{X}^{(2)} \in \mathbb{R}^{130 \times 116 \times 44}$.

5.2. Classification

Almost every subject in ADNI dataset has several ($\approx$ 6) individual rs-fMRI data series, that is, a patient might be scanned several times during a period of time. Usually, a random rs-fMRI data are selected and enters the processing step (Zhang et al., 2015). This random selection may cause several problems. Since the number of train data is very low, a small alteration in the samples could drastically change the set of input parameters in order to achieve the highest accuracy. Also achieving high-quality results with a classifier does not guarantee its effectiveness on other datasets even with fine-tuning the parameters, since the training set may contain outliers and unidentified corrupted data. In order to show that the proposed framework is less sensitive against the choice of different permutations of data, we have selected 18 different random permutations (i.e., each permutation contains a different rs-fMRI series, for each subject) and tested two state of the art eMCI classification methods on them: **HON** (Chen et al., 2016) and **k−SICE** (Zhang et al., 2015). We have used five evaluation measures: accuracy (ACC), sensitivity (SEN), Youden's index (YI), F-score, and balanced accuracy (BAC). The detailed definitions of these five statistical measures are provided in equation (**Table 1**), where TP, TN, FP, and FN denote the true positive, true negative, false positive, and false negative, respectively, and precision $= \frac{TP}{TP+FP}$ and recall $= \frac{TP}{TP+FN}$. In this article, we treat the eMCI samples as positive class and the NC samples as negative class.

5.2.1. Classification Performance

After fine-tuning the input parameter set for each method, the classification accuracy measure (ACC) shows that for 16 out of 18 different *random* selected datasets, our approach performs better than k-SICE the same also holds for 15 datasets comparing to HON, that is, in 88.8% of datasets the proposed method works better than k-SICE, and in 83.3% of datasets, it works better than FON. The highest classification accuracy (86.59%) is achieved

TABLE 1 | Definitions of five statistical measurement indices.

Measurement	Definition
Acc	$\frac{TP + TN}{TP + FP + TN + FN}$
SEN	$\frac{TP}{TP + FN}$
YI	$SEN + SPE - 1$
F-Score	$2 \times \frac{precession \times recall}{precession + recall}$
BAC	$\frac{1}{2}(SEN + SPE)$

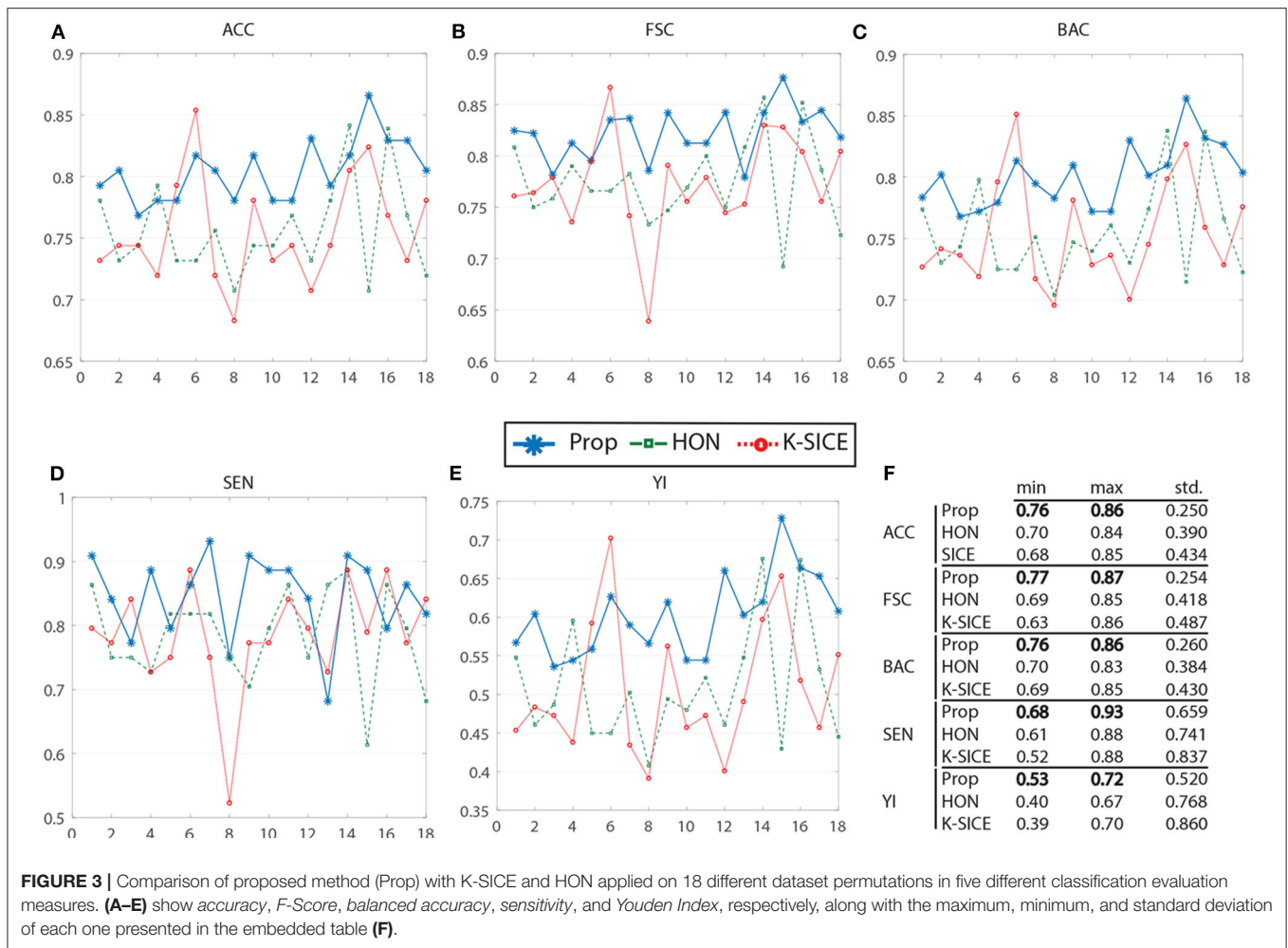**TABLE 2 |** The average of different classification measurements in all dataset permutations in percent.

Method	ACC	F-Score	SEN	YI	BAC
k-SICE	75.57	77.36	78.50	50.69	75.34
HON	75.66	77.44	78.40	50.89	75.44
Proposed	**80.43**	**82.20**	**84.60**	**60.20**	**80.09**

Higher values are indicated by bold numbers.

with the proposed method in the 15th sample data. The highest accuracy for the HON (84.15%) is achieved in the 14th, and the highest accuracy for the SICE method (85.37%) is achieved in the 6th sample data. As it was mentioned before, being stable when the input dataset changes is a very important aspect for a classifier, in order to measure the stability, the standard deviation of accuracy along with other measures is calculated. The standard of accuracy for the proposed method is 0.64 times less than HON and 1.73 times less than k-SICE method. Similar results hold for other classification measures as well.

Figure 3 shows the performance of these three methods in all five measurements. For a better demonstration, **Table 2** provides the average of several classification measurements scores for all dataset permutations. As it can be seen in this table, the average accuracy of proposed method, which is 80.43%, is 4.77% higher than the next method HON and 4.86% better than k-SICE. It is noteworthy that the other two methods, that is, HON and SICE, show similar results in average. The average F-score of the proposed method is also higher than other two, which shows a balanced prediction for both classes. Having a higher sensitivity (SEN) score, which measures the proportion of actual positives that are correctly identified as such, shows that the proposed method works better in detecting eMCI subjects. The YI is a measure for evaluating the biomarker effectiveness and having a higher YI yields a more informative decision (Youden, 1950). Our YI score is roughly 1.2 times better that other two methods. Similar to F-score, having a higher Balanced Accuracy Score (BAC) yields more balanced predictions. It is also noteworthy that the proposed method have much less standard deviation in all five measurements, which indicates its effectiveness and robustness toward different datasets.

One other key aspect of the proposed classifier is that it works significantly faster that the other two, especially in the training process. Our method is more than 600 times faster than HON and 20 times faster than SICE. Having a huge execution

FIGURE 3 | Comparison of proposed method (Prop) with K-SICE and HON applied on 18 different dataset permutations in five different classification evaluation measures. **(A–E)** show *accuracy*, *F-Score*, *balanced accuracy*, *sensitivity*, and *Youden Index*, respectively, along with the maximum, minimum, and standard deviation of each one presented in the embedded table **(F)**.

time especially affects the parameter selection scheme since all these methods use cross-validation procedure in order to find the optimal parameters, which itself requires several runs of the algorithm.

5.3. Functional Connectivity Network

The vector features for both normal and eMCI classes were obtained via the proposed method as it is described in section 4.3. Due to the aforementioned qualities of partial correlation, SICE is deployed in order to obtain the final FC. In order to better highlight the differences between normal and eMCI subjects, a difference graph D is constructed by subtracting the normal FC from the eMCI FC. This graph could be seen in **Figure 4**. The nodes of D show the ROIs according to the AAL atlas. The size of each node is proportional to its graph clustering coefficient, that is, the bigger node demonstrates higher activity in eMCI subjects in the corresponding ROI. Similar to nodes, the size of each edge is also proportional to the correlation between two ROIs. In addition, the edges are also color coded in a way that the green edges show the positive edges in D and the red edges show the negative edges in D. In this manner, the green edges demonstrate a decrease in activity between

the corresponding nodes in eMCI subjects, and the red edges show increasing activity between corresponding ROIs in the eMCI subjects.

As it can be seen in the difference graph, the big nodes, that is, ROIs with higher activities, do not necessarily establish strong connections with the other nodes. As an obvious example, higher activities in lingual gyrus (ROI index: 47, 48) (He et al., 2007), calcarine sulcus (ROI index: 43, 44) (Bakkour et al., 2013; Brewer and Barton, 2014), supplementary motor area (ROI index: 19, 20) (Brewer and Barton, 2014; Jacobsen et al., 2015), and temporal_mid_L (ROI index: 85) (Kosicek and Hecimovic, 2013) are easily detectable. The majority of ROIs located in frontal lobe also show rather high activities compared to normal subjects (Dennis and Thompson, 2014; Salvatore et al., 2015).

Similar to the nodes, the strong edge between two ROIs does not necessarily require the nodes to be highly active in eMCI, although a strong edge does indicate high activities and FC between the two corresponding ROIs. The difference in graph shows a significant increase in connectivity between Rectus (ROI index: 28, 27 in frontal lobe) and Parietal_Sup_R (ROI index: 60 in parietal lobe) (Brickman et al., 2015; De Reuck et al., 2015), Frontal_Inf_Orb_R (ROI index: 16 in frontal lobe) and
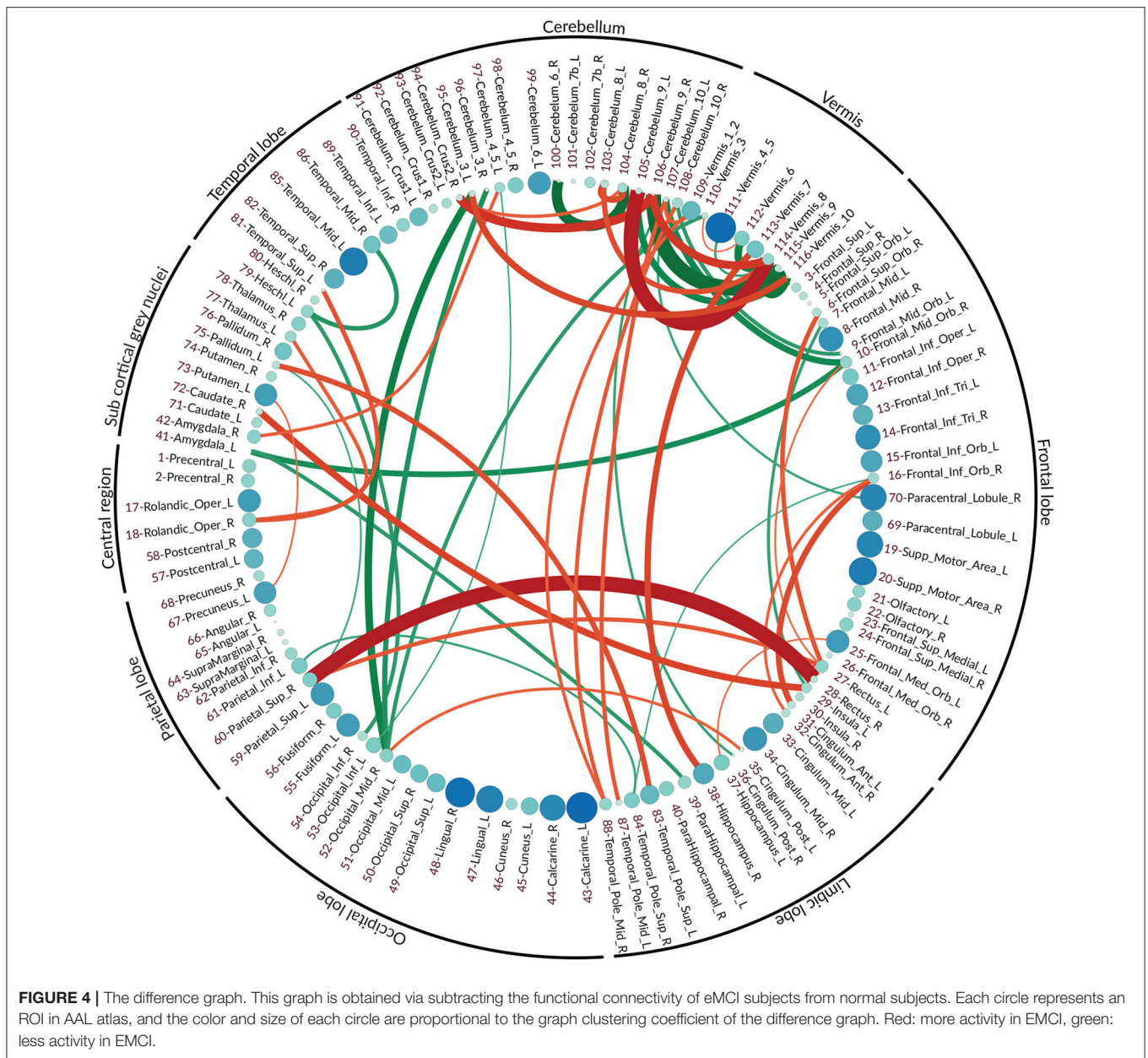
FIGURE 4 | The difference graph. This graph is obtained via subtracting the functional connectivity of eMCI subjects from normal subjects. Each circle represents an ROI in AAL atlas, and the color and size of each circle are proportional to the graph clustering coefficient of the difference graph. Red: more activity in EMCI, green: less activity in EMCI.

Cingulum_Ant (ROI index: 31, 32 in limbic lobe) (Perani et al., 2017), Insula_L, Temporal_Pole_Sup_L (ROI index: 29, 83 in limbic lobe) and Pallidum_R, Caudate_R (ROI index: 29, 83 in sub-cortical gray nuclei) (Watson et al., 2016). It can also be seen that within activities, frontal lobe also increased in patients with eMCI (Cai et al., 2015). There is a decrease in connectivity between Amygdala_L (ROI index: 41 in sub-cortical gray nuclei) with Frontal_Mid_Orb_R (ROI index: 10 in sub-frontal lobe) and ParaHippocampal_L (ROI index: 39 in sub-limbic lobe) (Ortner et al., 2016). The connectivity between Heschl_L (ROI index: 79 in temporal lobe) and two ROIs Temporal_Mid_R (ROI index: 86 also in temporal lobe) and Occipital_Inf_R (ROI index: 54 in occipital lobe) also decreased in eMCI (Steketee et al., 2016).

5.3.1. Regarding the Cerebellum and Vermis

In fMRI data analysis and especially in AD studies, ROIs within the cerebellum and vermis are usually excluded since their role was regarded as insignificant (Sanz-Arigita et al., 2010; Zhang et al., 2011). Recent studies have shown that the traditional assumption that cerebral area is essential only to the coordination of voluntary motor activity and motor learning is not valid and indicates the significant role of the cerebellum in nervous system function, cognition, and emotion (Jacobs et al., 2017).

As it can be seen in the difference graph that we obtained, ROIs within cerebellum and vermis are highly active and both their Intra and interconnections are noticeable. There is increased FC between the limbic lobe, especially Hippocampus_R, Temporal_Pole_Mid (ROI index: 38, 87, 88)

and cerebral areas in eMCI patients. Also, the connectivity between occipital lobe, especially occipital_mid_R (ROI index: 52), the frontal lobe, especially in frontal_mid_orb (ROI index: 9,10), and cerebral areas seems to decrease in patients with eMCI.

6. CONCLUSION

In this article, we proposed a tensor framework for eMCI diagnosis and FC construction. There are two main issues associated with rs-fMRI analysis and in particular eMCI diagnosis. The first is that the majority of state-of-the-art fMRI classification techniques use the FC matrix as the feature for their discriminant function; hence, they have to deal with many challenges that are associated with FC calculations. The second comes from the fact that FC networks are among the best tools for studying brain activities, but the stationary and dynamic FC conflict and the fact that the majority of methods belonging to these paradigms work only with one sample would lead to vastly different brain networks. Therefore, we developed a tensor framework, which is able to directly use the samples in classification without the need for any FC calculations and is also able to calculate a general FC network that considers the time-varying nature of rs-fMRI signals since it works in the data-dependent HOSVD-domain and is able to consider all subjects within a class in order to obtain these connectivities. The proposed method is not only fast, but it also outperforms state-of-the-art techniques.

One possible drawback of this framework is the need for HOSVD calculation for both classes in each test phase. Although this problem is negligible in eMCI classification (since the number of samples is not high), it could be time consuming for larger datasets. In order to resolve this issue, incremental HOSVD calculations may be deployed that will accelerate the calculations.

DATA AVAILABILITY STATEMENT

All datasets presented in this study can be downloaded through ADNI website (http://adni.loni.usc.edu). Additional information about the subjects used in this study can be found in the **Supplementary Material**.

AUTHOR CONTRIBUTIONS

AN and MR contributed to the design and implementation of the research, to the analysis of the results, and to the writing of the manuscript. Both authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fninf. 2020.581897/full#supplementary-material

REFERENCES

Ahmadi, S., and Rezghi, M. (2020). Generalized low-rank approximation of matrices based on multiple transformation pairs. *Pattern Recogn.* 108:107545. doi: 10.1016/j.patcog.2020.107545

Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., and Calhoun, V. D. (2014). Tracking whole-brain connectivity dynamics in the resting state. *Cereb. Cortex* 24, 663–676. doi: 10.1093/cercor/bhs352

Al-sharoa, E., Al-khassaweneh, M., and Aviyente, S. (2018). Tensor based temporal and multilayer community detection for studying brain dynamics during resting state fMRI. *IEEE Trans. Biomed. Eng.* 66, 695–709. doi: 10.1109/TBME.2018.2854676

Arsigny, V., Fillard, P., Pennec, X., and Ayache, N. (2006). Log-euclidean metrics for fast and simple calculus on diffusion tensors. *Magn. Reson. Med.* 56, 411–421. doi: 10.1002/mrm.20965

Bakkour, A., Morris, J. C., Wolk, D. A., and Dickerson, B. C. (2013). The effects of aging and Alzheimer's disease on cerebral cortical anatomy: specificity and differential relationships with cognition. *Neuroimage* 76, 332–344. doi: 10.1016/j.neuroimage.2013.02.059

Bi, X., and Wang, H. (2019). Early Alzheimer's disease diagnosis based on EEG spectral images using deep learning. *Neural Netw.* 114, 119–135. doi: 10.1016/j.neunet.2019.02.005

Brewer, A. A., and Barton, B. (2014). Visual cortex in aging and Alzheimer's disease: changes in visual field maps and population receptive fields. *Front. Psychol.* 5:74. doi: 10.3389/fpsyg.2014.00074

Brickman, A. M., Zahodne, L. B., Guzman, V. A., Narkhede, A., Meier, I. B., Griffith, E. Y., et al. (2015). Reconsidering harbingers of dementia: progression of parietal lobe white matter hyperintensities predicts Alzheimer's disease incidence. *Neurobiol. Aging* 36, 27–32. doi: 10.1016/j.neurobiolaging.2014.07.019

Brookmeyer, R., Johnson, E., Ziegler-Graham, K., and Arrighi, H. M. (2007). Forecasting the global burden of Alzheimer's disease. *Alzheimers Dement.* 3, 186–191. doi: 10.1016/j.jalz.2007.04.381

Cai, S., Huang, L., Zou, J., Jing, L., Zhai, B., Ji, G., et al. (2015). Changes in thalamic connectivity in the early and late stages of amnestic mild cognitive impairment: a resting-state functional magnetic resonance study from ADNI. *PLoS ONE* 10:e0115573. doi: 10.1371/journal.pone.0115573

Caselli, R. J., Reiman, E., Osborne, D., Hentz, J., Baxter, L., Hernandez, J., et al. (2004). Longitudinal changes in cognition and behavior in asymptomatic carriers of the apoe e4 allele. *Neurology* 62, 1990–1995. doi: 10.1212/01.WNL.0000129533.26544.BF

Chandra, A., Valkimadi, P.-E., Pagano, G., Cousins, O., Dervenoulas, G., Politis, M., et al. (2019). Applications of amyloid, tau, and neuroinflammation pet imaging to Alzheimer's disease and mild cognitive impairment. *Hum. Brain Mapp.* 40, 5424–5442. doi: 10.1002/hbm.24782

Chang, C., and Glover, G. H. (2010). Time-frequency dynamics of resting-state brain connectivity measured with fMRI. *Neuroimage* 50, 81–98. doi: 10.1016/j.neuroimage.2009.12.011

Chao-Gan, Y., and Yu-Feng, Z. (2010). DPARSF: a MATLAB toolbox for "pipeline" data analysis of resting-state fMRI. *Front. Syst. Neurosci.* 4:13. doi: 10.3389/fnsys.2010.00013

Chen, X., Zhang, H., Gao, Y., Wee, C.-Y., Li, G., Shen, D., et al. (2016). High-order resting-state functional connectivity network for MCI classification. *Hum. Brain Mapp.* 37, 3282–3296. doi: 10.1002/hbm.23240

Colclough, G. L., Woolrich, M. W., Harrison, S. J., López, P. A. R., Valdes-Sosa, P. A., and Smith, S. M. (2018). Multi-subject hierarchical inverse covariance modelling improves estimation of functional brain networks. *Neuroimage* 178, 370–384. doi: 10.1016/j.neuroimage.2018.04.077

Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehéricy, S., Habert, M.-O., et al. (2011). Automatic classification of patients with Alzheimer's disease

from structural MRI: a comparison of ten methods using the ADNI database. *Neuroimage* 56, 766–781. doi: 10.1016/j.neuroimage.2010.06.013

De Lathauwer, L., De Moor, B., and Vandewalle, J. (2000). A multilinear singular value decomposition. *SIAM J. Matrix Anal. Appl.* 21, 1253–1278. doi: 10.1137/S0895479896305696

De Reuck, J., Auger, F., Durieux, N., Deramecourt, V., Cordonnier, C., Pasquier, F., et al. (2015). Topography of cortical microbleeds in Alzheimer's disease with and without cerebral amyloid angiopathy: a post-mortem 7.0-tesla MRI study. *Aging Dis.* 6:437. doi: 10.14336/AD.2015.0429

De Silva, V., and Lim, L.-H. (2008). Tensor rank and the ill-posedness of the best low-rank approximation problem. *SIAM J. Matrix Anal. Appl.* 30, 1084–1127. doi: 10.1137/06066518X

de Vos, F., Koini, M., Schouten, T. M., Seiler, S., van der Grond, J., Lechner, A., et al. (2018). A comprehensive analysis of resting state fMRI measures to classify individual patients with Alzheimer's disease. *Neuroimage* 167, 62–72. doi: 10.1016/j.neuroimage.2017.11.025

Dennis, E. L., and Thompson, P. M. (2014). Functional brain connectivity using fMRI in aging and Alzheimer's disease. *Neuropsychol. Rev.* 24, 49–62. doi: 10.1007/s11065-014-9249-6

Du, Y., Fu, Z., and Calhoun, V. D. (2018). Classification and prediction of brain disorders using functional connectivity: promising but challenging. *Front. Neurosci.* 12:525. doi: 10.3389/fnins.2018.00525

Folch, J., Petrov, D., Ettcheto, M., Abad, S., Sánchez-López, E., Luisa García, M., et al. (2016). Current research therapeutic strategies for alzheimer's disease treatment. *Neural Plast.* 2016:8501693. doi: 10.1155/2016/8501693

Foti, N. J., and Fox, E. B. (2019). Statistical model-based approaches for functional connectivity analysis of neuroimaging data. *Curr. Opin. Neurobiol.* 55, 48–54. doi: 10.1016/j.conb.2019.01.009

Friston, K. J. (2011). Functional and effective connectivity: a review. *Brain Connect.* 1, 13–36. doi: 10.1089/brain.2011.0008

Gao, Z., Chung, J., Abdelrazek, M., Leung, S., Hau, W. K., Xian, Z., et al. (2019). Privileged modality distillation for vessel border detection in intracoronary imaging. *IEEE Trans. Med. Imaging* 39, 1524–1534. doi: 10.1109/TMI.2019.2952939

Golby, A., Silverberg, G., Race, E., Gabrieli, S., O'Shea, J., Knierim, K., et al. (2005). Memory encoding in Alzheimer's disease: an fMRI study of explicit and implicit memory. *Brain* 128, 773–787. doi: 10.1093/brain/awh400

Han, J., Ji, X., Hu, X., Zhu, D., Li, K., Jiang, X., et al. (2013). Representing and retrieving video shots in human-centric brain imaging space. *IEEE Trans. Image Process.* 22, 2723–2736. doi: 10.1109/TIP.2013.2256919

Handwerker, D. A., Roopchansingh, V., Gonzalez-Castillo, J., and Bandettini, P. A. (2012). Periodic changes in fMRI connectivity. *Neuroimage* 63, 1712–1719. doi: 10.1016/j.neuroimage.2012.06.078

He, L., Lu, C.-T., Ding, H., Wang, S., Shen, L., Yu, P. S., et al. (2017). "Multi-way multi-level kernel modeling for neuroimaging classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Honolulu, HI), 356–364. doi: 10.1109/CVPR.2017.724

He, Y., Wang, L., Zang, Y., Tian, L., Zhang, X., Li, K., et al. (2007). Regional coherence changes in the early stages of Alzheimer's disease: a combined structural and resting-state functional MRI study. *Neuroimage* 35, 488–500. doi: 10.1016/j.neuroimage.2006.11.042

Hindriks, R., Adhikari, M. H., Murayama, Y., Ganzetti, M., Mantini, D., Logothetis, N. K., et al. (2016). Can sliding-window correlations reveal dynamic functional connectivity in resting-state fMRI? *Neuroimage* 127, 242–256. doi: 10.1016/j.neuroimage.2015.11.055

Huang, S., Li, J., Sun, L., Ye, J., Fleisher, A., Wu, T., et al. (2010). Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation. *Neuroimage* 50, 935–949. doi: 10.1016/j.neuroimage.2009.12.120

Jacobs, H. I., Hopkins, D. A., Mayrhofer, H. C., Bruner, E., van Leeuwen, F. W., Raaijmakers, W., et al. (2017). The cerebellum in Alzheimer's disease: evaluating its role in cognitive decline. *Brain* 141, 37–47. doi: 10.1093/brain/awx194

Jacobsen, J.-H., Stelzer, J., Fritz, T. H., Chételat, G., La Joie, R., and Turner, R. (2015). Why musical memory can be preserved in advanced Alzheimer's disease. *Brain* 138, 2438–2450. doi: 10.1093/brain/awv135

Jie, B., Zhang, D., Gao, W., Wang, Q., Wee, C.-Y., and Shen, D. (2013). Integration of network topological and connectivity properties for neuroimaging classification. *IEEE Trans. Biomed. Eng.* 61, 576–589. doi: 10.1109/TBME.2013.2284195

Kam, T.-E., Zhang, H., Jiao, Z., and Shen, D. (2019). Deep learning of static and dynamic brain functional networks for early MCI detection. *IEEE Trans. Med. imaging* 39, 478–487. doi: 10.1109/TMI.2019.2928790

Kazeminejad, A., and Sotero, R. C. (2019). Topological properties of resting-state fMRI functional networks improve machine learning-based autism classification. *Front. Neurosci.* 12:1018. doi: 10.3389/fnins.2018.01018

Kosicek, M., and Hecimovic, S. (2013). Phospholipids and Alzheimer's disease: alterations, mechanisms and potential biomarkers. *Int. J. Mol. Sci.* 14, 1310–1322. doi: 10.3390/ijms14011310

Leonardi, N., Richiardi, J., Gschwind, M., Simioni, S., Annoni, J.-M., Schluep, M., et al. (2013). Principal components of functional connectivity: a new approach to study dynamic brain connectivity during rest. *Neuroimage* 83, 937–950. doi: 10.1016/j.neuroimage.2013.07.019

Leonardi, N., and Van De Ville, D. (2013). "Identifying network correlates of brain states using tensor decompositions of whole-brain dynamic functional connectivity," in *2013 International Workshop on Pattern Recognition in Neuroimaging* (Philadelphia, PA: IEEE), 74–77. doi: 10.1109/PRNI.2013.28

Leonardi, N., and Van De Ville, D. (2015). On spurious and real fluctuations of dynamic functional connectivity during rest. *Neuroimage* 104, 430–436. doi: 10.1016/j.neuroimage.2014.09.007

Li, R., Rui, G., Chen, W., Li, S., Schulz, P. E., and Zhang, Y. (2018). Early detection of Alzheimer's disease using non-invasive near-infrared spectroscopy. *Front. Aging Neurosci.* 10:366. doi: 10.3389/fnagi.2018.00366

Li, Y., Liu, J., Huang, J., Li, Z., and Liang, P. (2018). Learning brain connectivity sub-networks by group-constrained sparse inverse covariance estimation for Alzheimer's disease classification. *Front. Neuroinform.* 12:58. doi: 10.3389/fninf.2018.00058

Liu, F., Choi, D., Xie, L., and Roeder, K. (2018). Global spectral clustering in dynamic networks. *Proc. Natl. Acad. Sci. U.S.A.* 115, 927–932. doi: 10.1073/pnas.1718449115

Lv, H., and Wang, R. (2019). Denoising 3D magnetic resonance images based on low-rank tensor approximation with adaptive multirank estimation. *IEEE Access* 7, 85995–86003. doi: 10.1109/ACCESS.2019.2924907

Ma, G., He, L., Lu, C.-T., Yu, P. S., Shen, L., and Ragin, A. B. (2016). "Spatio-temporal tensor analysis for whole-brain fMRI classification," in *Proceedings of the 2016 SIAM International Conference on Data Mining* (SIAM) (Miami, FL), 819–827. doi: 10.1137/1.9781611974348.92

Musha, T., Matsuzaki, H., Kobayashi, Y., Okamoto, Y., Tanaka, M., and Asada, T. (2013). EEG markers for characterizing anomalous activities of cerebral neurons in NAT (neuronal activity topography) method. *IEEE Trans. Biomed. Eng.* 60, 2332–2338. doi: 10.1109/TBME.2013.2255101

Ng, B., Varoquaux, G., Poline, J. B., and Thirion, B. (2013). "A novel sparse group Gaussian graphical model for functional connectivity estimation," in *International Conference on Information Processing in Medical Imaging* (Asilomar, CA: Springer), 256–267. doi: 10.1007/978-3-642-38868-2_22

Nguyen, D. T., Ryu, S., Qureshi, M. N. I., Choi, M., Lee, K. H., and Lee, B. (2019). Hybrid multivariate pattern analysis combined with extreme learning machine for Alzheimer's dementia diagnosis using multi-measure rs-fMRI spatial patterns. *PLoS ONE* 14:e0212582. doi: 10.1371/journal.pone.0212582

Ortner, M., Pasquini, L., Barat, M., Alexopoulos, P., Grimmer, T., Förster, S., et al. (2016). Progressively disrupted intrinsic functional connectivity of basolateral amygdala in very early Alzheimer's disease. *Front. Neurol.* 7:132. doi: 10.3389/fneur.2016.00132

Ozdemir, A., Villafane-Delgado, M., Zhu, D. C., Iwen, M. A., and Aviyente, S. (2017). "Multi-scale higher order singular value decomposition (MS-HOSVD) for resting-state fMRI compression and analysis," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (New Orleans, LA: IEEE), 6299–6303. doi: 10.1109/ICASSP.2017.7953368

Ozdemir, O., Russell, R. L., and Berlin, A. A. (2019). A 3D probabilistic deep learning system for detection and diagnosis of lung cancer using low-dose CT scans. *IEEE Trans. Med. Imaging* 39, 1419–1429. doi: 10.1109/TMI.2019.2947595

Park, S. W. (2011). *Multifactor Analysis for fMRI Brain Image Classification by Subject and Motor Task*. Electrical and computer engineering technical report. Pittsburgh, PA: Carnegie Mellon University.

Perani, D., Farsad, M., Ballarini, T., Lubian, F., Malpetti, M., Fracchetti, A., et al. (2017). The impact of bilingualism on brain reserve and metabolic connectivity in Alzheimer's dementia. *Proc. Natl. Acad. Sci. U.S.A.* 114, 1690–1695. doi: 10.1073/pnas.1610909114

Pervaiz, U., Vidaurre, D., Woolrich, M. W., and Smith, S. M. (2020). Optimising network modelling methods for fMRI. *Neuroimage* 211:116604. doi: 10.1016/j.neuroimage.2020.116604

Reid, A. T., Headley, D. B., Mill, R. D., Sanchez-Romero, R., Uddin, L. Q., Marinazzo, D., et al. (2019). Advancing functional connectivity research from association to causation. *Nat. Neurosci.* 22, 1751–1760. doi: 10.1038/s41593-019-0510-4

Rezghi, M. (2017). A novel fast tensor-based preconditioner for image restoration. *IEEE Trans. Image Process.* 26, 4499–4508. doi: 10.1109/TIP.2017.2716840

Rövid, A., Szeidl, L., and Várlaki, P. (2013). "Data representation in HOSVD-DCT based domain," in *2013 IEEE 17th International Conference on Intelligent Engineering Systems (INES)* (San Jose, CA: IEEE), 103–106. doi: 10.1109/INES.2013.6632791

Salvatore, C., Cerasa, A., Battista, P., Gilardi, M. C., Quattrone, A., and Castiglioni, I. (2015). Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: a machine learning approach. *Front. Neurosci.* 9:307. doi: 10.3389/fnins.2015.00307

Sanz-Arigita, E. J., Schoonheim, M. M., Damoiseaux, J. S., Rombouts, S. A., Maris, E., Barkhof, F., et al. (2010). Loss of 'small-world' networks in Alzheimer's disease: graph analysis of fMRI resting-state functional connectivity. *PLoS ONE* 5:e13788. doi: 10.1371/journal.pone.0013788

Smith, S. M. (2012). The future of fMRI connectivity. *Neuroimage* 62, 1257–1266. doi: 10.1016/j.neuroimage.2012.01.022

Sra, S. (2012). "A new metric on the manifold of kernel matrices with application to matrix geometric means," in *Advances in Neural Information Processing Systems*, Vol. 25, eds F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Red Hook, NY: Curran Associates, Inc.), 144–152. Available online at: https://proceedings.neurips.cc/paper/2012/file/98dce83da57b0395e163467c9dae521b-Paper.pdf

Stanley, M. L., Moussa, M. N., Paolini, B., Lyday, R. G., Burdette, J. H., and Laurienti, P. J. (2013). Defining nodes in complex brain networks. *Front. Comput. Neurosci.* 7:169. doi: 10.3389/fncom.2013.00169

Steketee, R. M., Bron, E. E., Meijboom, R., Houston, G. C., Klein, S., Mutsaerts, H. J., et al. (2016). Early-stage differentiation between presenile Alzheimer's disease and frontotemporal dementia using arterial spin labeling MRI. *Eur. Radiol.* 26, 244–253. doi: 10.1007/s00330-015-3789-x

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978

van de Leemput, S. C., Prokop, M., van Ginneken, B., and Manniesing, R. (2019). Stacked bidirectional convolutional lstms for deriving 3D non-contrast ct from spatiotemporal 4D CT. *IEEE Trans. Med. Imaging* 39, 985–996. doi: 10.1109/TMI.2019.2939044

Waller, L., Brovkin, A., Dorfschmidt, L., Bzdok, D., Walter, H., and Kruschwitz, J. D. (2018). Graphvar 2.0: a user-friendly toolbox for machine learning on functional connectivity measures. *J. Neurosci. Methods* 308, 21–33. doi: 10.1016/j.jneumeth.2018.07.001

Watson, R., Colloby, S. J., Blamire, A. M., and O'Brien, J. T. (2016). Subcortical volume changes in dementia with Lewy bodies and Alzheimer's disease. A comparison with healthy aging. *Int. Psychogeriatr.* 28, 529–536. doi: 10.1017/S1041610215001805

Wee, C.-Y., Yang, S., Yap, P.-T., and Shen, D. (2016). Sparse temporally dynamic resting-state functional connectivity networks for early MCI identification. *Brain Imaging Behav.* 10, 342–356. doi: 10.1007/s11682-015-9408-2

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer* 3, 32–35. doi: 10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3

Zalesky, A., Fornito, A., Cocchi, L., Gollo, L. L., and Breakspear, M. (2014). Time-resolved resting-state brain networks. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10341–10346. doi: 10.1073/pnas.1400181111

Zhang, D., Wang, Y., Zhou, L., Yuan, H., and Shen, D. (2011). Multimodal classification of Alzheimer's disease and mild cognitive impairment. *Neuroimage* 55, 856–867. doi: 10.1016/j.neuroimage.2011.01.008

Zhang, J., Zhou, L., Wang, L., and Li, W. (2015). Functional brain network classification with compact representation of sice matrices. *IEEE Trans. Biomed. Eng.* 62, 1623–1634. doi: 10.1109/TBME.2015.2399495

Zhang, S., Tong, H., Xu, J., and Maciejewski, R. (2019). Graph convolutional networks: a comprehensive review. *Comput. Social Netw.* 6:11. doi: 10.1186/s40649-019-0069-y

Zhang, T., Guo, L., Li, K., Jing, C., Yin, Y., Zhu, D., et al. (2012). Predicting functional cortical ROIs via DTI-derived fiber shape models. *Cereb. Cortex* 22, 854–864. doi: 10.1093/cercor/bhr152

Zhou, Z., Chen, X., Zhang, Y., Hu, D., Qiao, L., Yu, R., et al. (2020). A toolbox for brain network construction and classification (brainnetclass). *Hum. Brain Mapp.* 41, 2808–2826. doi: 10.1002/hbm.24979