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

A. Noroozi, M. Rezghi

Abstract—Recently machine learning methods has gained lots of publicity among researchers in order to analyze brain images such as Resting-State Functional Magnetic Resonance Imaging(rs-fMRI) to obtain a better understanding of the brain and its related disease such as Alzheimers disease. Finding the common patterns caused by a brain disorder through analyzing the functional connectivity(FC) network along with discriminating brain diseases from normal controls have traditionally been two main goals in studying rs-fMRI data. The majority of techniques for finding an FC, calculate the FC matrix for each subject and then use simple techniques in order to combine them to obtain general functional connectivity. Also, the state-of-the-art classification techniques for finding subjects with brain disorders, also rely on calculating an FC for each subject, vectorize them and then feed them to the classifier. Considering these problems and based on multi-dimensional nature data, we have come up with a novel tensor framework in which the FC calculation for each class is done without the need to construct the FC for each sample, also this framework allows us to reduce the dimensionality, and create a novel discriminant function that avoids vectorization in any step and uses the test data in the training process without forcing any prior knowledge about its label to 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.

Index Terms—Alzheimers disease (AD) classification, Functional Connectivity, Tensor, High Order Singular Value Decomposition, Dimension Reduction.

I. INTRODUCTION

ALZHEIMERS Alzheimers disease (AD) is a progressive neurodegenerative disorder with a long pre-morbid asymptomatic period which affects millions of elderly individuals worldwide[1]. 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 [2]. The predominant clinical symptoms of AD include a decline in some important brain cognitive and intellectual abilities, such as memory, thinking, and reasoning. Precise diagnosis of AD, especially at its early warning stage: early Mild Cognitive Impairment (eMCI), enables treatments to delay or even avoid such disorders [3].

Functional Magnetic Resonance Imaging (fMRI)[4], which is a non-invasive brain imaging technique, is vastly used by researchers in order to monitor brain activities [5]. An fMRI sample is naturally a 4D tensor consisting of 3D time-varying

A. Noroozi, Department of Computer Science, Tarbiat Modares University, Tehran-Iran. e-mail: Alinoroozics@gmail.com.
M. Rezghi, Department of Computer Science, Tarbiat Modares University, Tehran-Iran. e-mail: Rezghi@modares.ac.ir.

voxels, and each voxel contains an intensity value that is proportional to the strength of the Blood Oxygenation Level Dependent(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, focuses on the low-frequency $(< 0.1Hz)$ oscillations of BOLD signal, which presents the underlying neuronal activation patterns of brain regions. rs-fMRI is usually used in order to analyze brain diseases like AD or Autism[6], [7].

Since each fMRI volume 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)[8]. By putting together the average time-series for all the ROIs, the ith volume would become $X_i \in \mathbb{R}^{T \times R}, i = \{1, 2, \cdots, S\}$ in which R, T and S are the number of ROIs, time points and samples respectively.

There are two major studies associated with rs-fMRI data: finding common brain disorders caused by diseases like Alzheimer's, Autism, schizophrenia and etc. And more recently detecting patients with brain disorders using classification techniques [9], [10]. Due to the high dimensionality of data and the nature of diseases like eMCI which does not show any reliable clinical symptoms, researchers moved towards advanced machine learning techniques in order to achieve more reliable analysis [11].

A powerful tool that is commonly used in order to achieve aforementioned goals is the Functional Connectivity(FC) network. FC is a $region \times region$ matrix $\bar{X}$ in which $\bar{x}_{ij}$ represents the functional connectivity 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 [12]. Recent studies have shown that some brain disorders like AD could alter the way that some brain regions interact with each other. For example, compared with healthy subjects, AD patients have been found decreased functional connectivity between the hippocampus and other brain regions, and MCI patients have been observed increased functional connectivity between the frontal lobe and other brain regions[5]. So, Finding an

FC that highlights the patterns caused by a disease, i.e. a **General** functional connectivity, has been a common goal in the rs-fMRI study for a long time. Several approaches exist to find common patterns among different brain scans. Data-driven methods such as PCA have been proposed for this task [13]. But ultimately most of them rely on calculating a network for each volume which may overlook the role of noises or outliers within the data[14].

In recent years FCs are also used as features in classification. So, instead of using X_i as the ith sample, corresponding FC i.e. $\bar{X}_i$ is used as a feature. Although FCs show promising results, they bring their own challenges. The computational cost of FC is usually high and also its quality massively affects the performance of the learning process. Also, Since the conventional classifiers like **S**upport **V**ector **M**achine(SVM) and 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 which produce poor performance due to the phenomena known as the curse of dimensionality. Alongside the curse of dimensionality, vectorization also destroys potential information that is 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[15].

In this paper, based on high order tensor decomposition, we have created a framework in which the aforementioned goals i.e. finding a general FC and detecting a disorder via 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. The majority of connectivity patterns detected by this method have been observed and studied in several separated types of research which shows 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 also outperforms state of the art eMCI classification methods.

Viewing each class as a tensor allows us to work with *time* and *region* features separately but simultaneously. This multilinear view ables us to design a proper dimension reduction relative to the nature of each feature along with a discriminant function based on linear regression on latent space of samples that uses the test data to enhance the quality of the training set without forcing any a prior knowledge to the classifier, a task which is not possible through well known classifiers like SVM, logistic regression or k-NN. It is also notable that the proposed discriminant function directly works with the X_is as features. Having the FC calculation step omitted in classification not only heavily affects the computational performance of the method, but it also saves us from the trouble of FCs which will be discussed in the next section.

Multilinear approaches have been used before in order to analyze fMRI data. For example [16], considers dynamic whole-brain functional connectivity 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. [17] uses Multilinear PCA in order to classify fMRI data by Subject and Motor Task. [18] and [19] also uses tensors for analyzing brain connectivity networks. Recently [20] and [21] proposed a multilinear method for voxel-wise analysis of the resting state fMRI volumes.

To verify our approach, we conduct an extensive experimental study on rs-fMRI data from the benchmark 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 functional connectivity patterns.

II. PROPOSED FMRI ANALYSIS FRAMEWORK BASED ON HOSVD

In the following section we will describe the proposed method based on HOSVD for Classification and General Functional Connectivity construction.

A. A Classification of Region-Time data based on HOSVD

For reducing the computational complexity and reducing the occurrence of overfitting, especially for data with large features relative to samples (Like fMRI data), using dimensionality reduction techniques is inevitable. Although each sample $X \in \mathbb{R}^{T \times R}$ has two different time and region features, the classical methods like PCA, SVD and etc. only works on vectorized version ($x = vec(X)$) of such data. Although this approach is easy to deploy, it has several drawbacks like the occurrence of Curse of dimensionality and mixing up different features. Recently multilinear dimension reduction methods like MPCA, GLRAM [15], [22] have been proposed that could work with multi-dimensional data, without folding them into vectors. In these methods, there is a freedom to select specific reduction for each kind of feature. In this section, we will use a well-known tensor decomposition named HOSVD for both dimension reduction and classification of fMRI data.

Let tensors $\mathcal{X}^{(i)} \in \mathbb{R}^{T \times R \times S_i}$, consists 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

$$\mathcal{X}^{(i)} = \left(U^{(i)}, V^{(i)}, W^{(i)} \right) \cdot \mathcal{S}^{(i)}, \qquad (1)$$

is known as **H**igher **O**rder **S**ingular **V**alue **D**ecomposition(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 and $\mathcal{S}^{(i)}$ is the corresponding core tensor [23]. 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 times. 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. On the other hand, the last columns of these singular matrices,

have more fluctuations and are usually associated with the noisy parts of their corresponding fibers[23]. 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:

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

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[15].

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 (2) becomes

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

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 $\overline{\mathcal{S}}^{(i)} = \mathcal{S}^{(i)}(1:k_1^i, 1:k_2^i, :)$. So if a test data like $X \in \mathbb{R}^{T \times R}$ belongs to the ith class its natural to expect that its projected version into principle region and times spaces, spanned by $U_{k_1^i}, V_{k_2^i}$, i.e,

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

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

$$Z^{(i)} \approx \sum_{k=1}^{S_i} \lambda_k^i \overline{\mathcal{S}}^{(i)}(:,:,k). \qquad (3)$$

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 (3). 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 $\overline{\mathcal{S}}^{(i)}(:,:,k)$ could be used in (3). In this viewpoint each test data X would be assigned to the l^{th}class, if

$$r_l = \min_{i=1,2} r_i,$$

where

$$r_i = \min_{\lambda^i} \| Z^{(i)} - \sum_{k=1}^{k_3^i} \lambda_k^i \overline{\mathcal{S}}^{(i)}(:,:,k) \|, \quad \lambda^i = \begin{pmatrix} \lambda_1^i \\ \vdots \\ \lambda_{s_i} \end{pmatrix} \qquad (4)$$

shows the reconstruction error of the projected version of X in the ith class. The minimization (4) is a simple least square problem that could be solved easily.

The proposed method has an interesting property which allows us to enhance the classification performance by using the test data in the training process without forcing any prior knowledge to the classifier. We have shown that the principal properties of the ith class is reflected in $\overline{\mathcal{S}}^{(i)}$, we also know that the first slices of this tensor, i.e. slices with lower indices, have the most role in reconstructing the main parts of this class i.e. the signal parts. The same reasoning also leads to the conclusion that slices with higher indices are responsible for the possible noise in this class. Now consider that the test data X is added to data set $\mathcal{X}^{(i)}$ of the ith class. So the new data set will be $\widetilde{\mathcal{X}} \in \mathbb{R}^{T \times R \times (S_i+1)}$,

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

If X belongs to the ith class, then in the decomposition of $\widetilde{\mathcal{X}}^{(i)}$, X would reinforce the first slices of the core tensor. On the other hand, if X does not belong to the ith class, HOSVD would naturally consider it as noise, 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, i.e. slices with higher indices.

Remember that the last slices of the core tensor are discarded in the dimension reduction process. As a result, if X does not belong to the ith class, it would not be involved in the classification process, on the other hand, if X belongs to this class it would affect the first slices of the core tensor and thus would lead into smaller reconstruction error. So if we add X to all classes before the decomposition process, the reconstruction error for the right class would be less than other classes. Note that since X is added to all classes, without sneaking any information about its label to the training process this technique is legit. After computing the HOSDVD of this new data sets $\widetilde{\mathcal{X}}^{(i)}$, we apply the method in (4) for classification.

The proposed method have three main advantages: 1) There is no necessity for constructing FC for each sample, 2) each feature could be treated separately and simultaneously and 3) the test sample could be used in order to enhance the classifier.

B. General Functional Connectivity Calculations based on HOSVD

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. Viewing each region as a slice would allow us to consider its behavior in all times and across all samples, which itself allows us to shed more light on common properties and ignore individual differences that are highly possible due to the presence of noise and outliers. Thus, 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:

$$\mathcal{Y}^{(i)}(:,l,:) = \left(U_{k_1^i}^{(i)\mathsf{T}}, W_{k_3}^{(i)\mathsf{T}}\right)_{1,3} \cdot \mathcal{X}^{(i)}(:,l,:). \quad (5)$$

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

$$\mathbb{R}^{k_1^i \times R \times k_3^i} \ni \mathcal{Y}^{(i)} = \left(U_{k_1^i}^{(i)\mathsf{T}}, W_{k_3^i}^{(i)\mathsf{T}}\right)_{1,3} \cdot \mathcal{X}^{(i)} \quad (6)$$

denote all reduced regions of the ith class. By this structure and substituting the HOSVD decomposition of $\mathcal{X}^{(i)}$ in (6), we obtain

$$\mathcal{Y}^{(i)} = \left([I_{k_1^i} \quad 0], V, [I_{k_3^i} \quad 0]\right) \cdot \mathcal{S}^{(i)}$$
$$= (V)_2 \cdot \mathcal{S}^{(i)}(1:k_1^i,:,1:k_3^i)$$

thus

$$\mathcal{Y}^{(i)}(:,k,:) = \sum_{k'}^{R} V^{(i)}(k,k')\bar{\mathcal{C}}^{(i)}(:,k',:)$$
$$= \left(V^{(i)}(k,:)\right)_2 \cdot \mathcal{C}^{(i)} \quad (7)$$

in which

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

The equation (7), 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 (7) as new features for the lth region. Mathematically this means each region in the ith class could be represented by a new feature vector $V(l,1;k_3^i) \in \mathbb{R}^{k_3^i}$.

The main benefit of this approach is that each the region could be represented only be a vector with size k_3^i, instead of a large time-sample matrix. Having each region be represented via 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.

III. EXPERIMENTAL STUDY

A. Data Acquisition and Experimental Settings

Rs-fMRI data of 196 subjects were downloaded from the ADNI website[1]. 44 eMCI and 38 NC subjects was obtained as it was in [24], except that the ROIs within the Cerebellum and Vermis are not excluded in our study. The IDs of the 82 (38 NC and 44 early MCI) subjects are provided in the supplementary material.

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

B. Classification

Almost every subject in ADNI dataset has several scans. Usually, random scan data is selected and enters the processing step[24]. 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 prediction accuracy and other classification evaluation methods. 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 (i.e. Same patient with the different scan), we have selected 18 different permutations of data and test two state of the art **eMCI** classification methods on them: **HON**[25] and **k-SICE**[24]. To make full use of the limited subjects, a leave-one-out procedure is used for training and test. That is, each sample is reserved for the test in turn, while the remaining samples are used for training. We have use five evaluation measures: accuracy (ACC), sensitivity (SEN), Youdens index(YI), F-score, and balanced accuracy (BAC). In this article, we treat the eMCI samples as positive class and the NC samples as negative class.

1) Classification performance: The classification accuracy measure(ACC), After fine-tuning the input parameter set for each method, show 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. i.e. in 88.8% of datasets, 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 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 are calculated. The std. of accuracy for the proposed method is 0.64 times less than HON and 1.73 times less than k-SICE method. Similar results also hold for other classification measures.

Figure (1) show the performance of these three methods in all five measurements. Some statistical information about these plots is also included in the embedded table. As it can be seen in this figure, similar to the accuracy, the proposed method in overall works much better than HON and k-SICE. For a better Demonstration, table (I) 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 i.e HON and SICE show similar results in average.

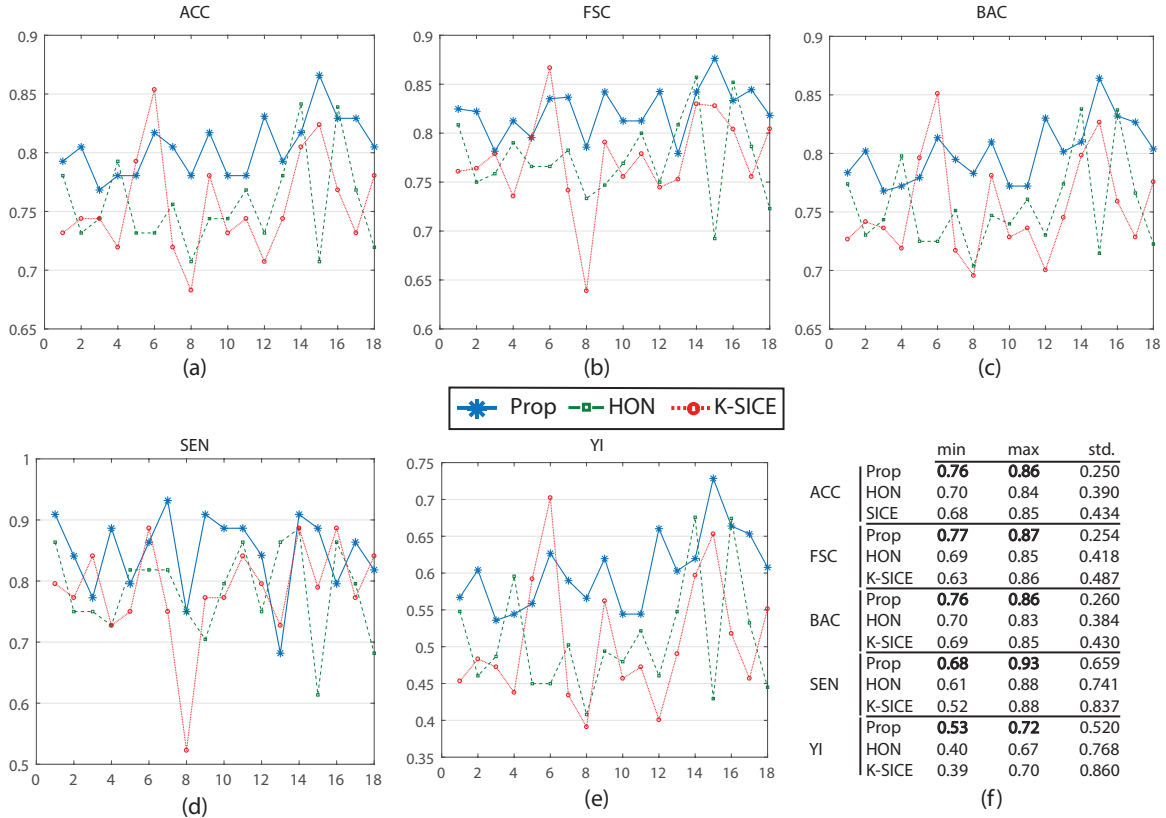One other key aspect of the proposed classifier is that it

Fig. 1. Comparison of proposed method(Prop) with K-SICE and HON applied on 18 different dataset permutations in five different classification evaluation measures. Figures *a* through *e* show *accuracy*, *F-Score*, *balanced accuracy*, *sensitivity* and *Youden Index* respectively along with the max, min and standard deviation of each one presented at the embedded table (f).

TABLE I
THE AVERAGE OF DIFFERENT CLASSIFICATION MEASUREMENTS IN ALL DATASET PERMUTATIONS IN %

Method	ACC	F-Score	SEN	SPE	YI	BAC
k-SICE	75.57	77.36	78.50	72.19	50.69	75.34
HON	75.66	77.44	78.40	72.48	50.89	75.44
Proposed	**80.43**	**82.20**	**84.60**	**75.59**	**60.20**	**80.09**

works significantly faster that the other two, specially in the training process. Our method is more than 600 times faster than HON and 20 times faster that SICE. Having a huge execution time especially affects the parameter selection for HON, since it uses cross-validation procedure in order to find the optimal parameters which itself require several runs of the algorithm.

C. Functional connectivity Network

The vector features for both Normal and eMCI classes was obtained via the proposed method as it is described in (II-B). 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(2). The nodes of D shows the ROIs according to the AAL atlas. The size of each node is proportional to its graph clustering coefficient, i.e. 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 decreasing in activity between the corresponding nodes in eMCI subjects and vice versa, 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 i.e. ROIs with higher activities does not necessarily establish strong connections with the other nodes. As an obvious example, higher activities in Lingual gyrus(ROI index: 47,48)[26], Calcarine sulcus(ROI index: 43, 44)[27], [28], Supplementary motor area(ROI index: 19,20)[28], [29] and Temporal_mid_L(ROI index: 85)[30] are easily detectable. The majority of ROIs located in frontal lobe also shows rather high activities comparing to normal subjects[31], [5].

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 functional connectivity between the two corresponding ROIs. The difference 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)
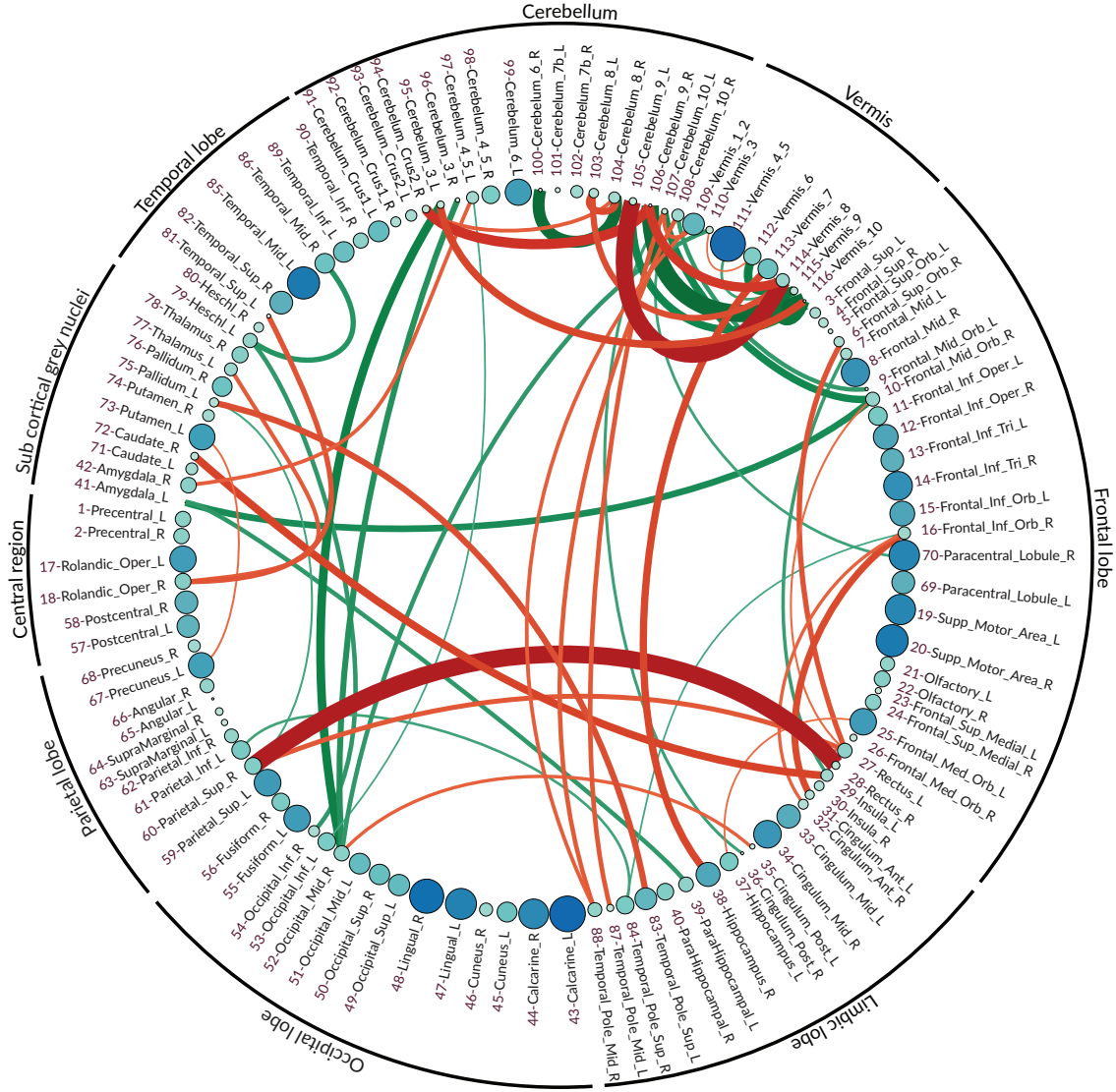
Fig. 2. The difference graph. This graph is obtained via subtracting the functional connectivity of eMCI subjects from normal subjects. Each circle represents a ROI in AAL atlas and the color and size of each circle is proportional to the graph clustering coefficient of the difference graph. red = more activity in EMCI, green: less.

[32], [33], Frontal_Inf_Orb_R(ROI index: 16 in Frontal lobe) and Cingulum_Ant(ROI index: 31,32 in Limbic lobe)[34], 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 Grey Nuclei)[35]. It can also be seen that within activities in frontal lobe also increased in patients with eMCI[36]. There is a decrease in connectivity between Amygdala_L(ROI index: 41 in Sub Cortical Grey Nuclei) with Frontal_Mid_Orb_R(ROI index: 10 in Sub Frontal lobe) and ParaHippocampal_L(ROI index: 39 in Sub Limbic lobe)[37]. 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[38].

Regarding the Cerebellum and Vermis: In fMRI data analysis and especially in Alzheimer's disease studies, ROIs within the Cerebellum and Vermis are usually excluded since their role was regarded as insignificant[39], [40]. Recent studies have shown that the traditional assumption that Cerebral area is essential only for 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[41].

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 functional connectivity 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.

IV. CONCLUSION

The majority of classification techniques uses the vectorized version of data as the input of the discriminant function. As the number of the dimensions grows, i.e. when data is naturally a high order tensor, this vectorization become problematic and it would highly affect the performance. Taking advantage of the techniques designed for the tensors, we have developed a framework for fMRI data analysis in which the following objectives: *Classification* and *General Functional Connectivity network* are achieved via a singe HOSVD of the input tensors.

Extensive studies on the rs-fMRI provided by ADNI shows the superiority of the proposed framework in both classification and functional connectivity. The obtained FC network not only acknowledge the previous discovered connections but also reveals new connectivity patterns previously unknown. The framework proposed in this paper can be easily extended to other studies involved with high order data.

REFERENCES

[1] R. J. Caselli, E. Reiman, D. Osborne, J. Hentz, L. Baxter, J. Hernandez *et al.*, "Longitudinal changes in cognition and behavior in asymptomatic carriers of the apoe e4 allele," *Neurology*, vol. 62, no. 11, pp. 1990–1995, 2004.

[2] R. Brookmeyer, E. Johnson, K. Ziegler-Graham, and H. M. Arrighi, "Forecasting the global burden of alzheimers disease," *Alzheimer's & dementia*, vol. 3, no. 3, pp. 186–191, 2007.

[3] T. Musha, H. Matsuzaki, Y. Kobayashi, Y. Okamoto, M. Tanaka, and T. Asada, "Eeg markers for characterizing anomalous activities of cerebral neurons in nat (neuronal activity topography) method," *IEEE Transactions on biomedical Engineering*, vol. 60, no. 8, pp. 2332–2338, 2013.

[4] A. Golby, G. Silverberg, E. Race, S. Gabrieli, J. O'Shea, K. Knierim *et al.*, "Memory encoding in alzheimer's disease: an fmri study of explicit and implicit memory," *Brain*, vol. 128, no. 4, pp. 773–787, 2005.

[5] E. L. Dennis and P. M. Thompson, "Functional brain connectivity using fmri in aging and alzheimers disease," *Neuropsychology review*, vol. 24, no. 1, pp. 49–62, 2014.

[6] N. Leonardi, J. Richiardi, M. Gschwind, S. Simioni, J.-M. Annoni, M. Schluep *et al.*, "Principal components of functional connectivity: a new approach to study dynamic brain connectivity during rest," *NeuroImage*, vol. 83, pp. 937–950, 2013.

[7] V. L. Cherkassky, R. K. Kana, T. A. Keller, and M. A. Just, "Functional connectivity in a baseline resting-state network in autism," *Neuroreport*, vol. 17, no. 16, pp. 1687–1690, 2006.

[8] M. L. Stanley, M. N. Moussa, B. Paolini, R. G. Lyday, J. H. Burdette, and P. J. Laurienti, "Defining nodes in complex brain networks," *Frontiers in computational neuroscience*, vol. 7, p. 169, 2013.

[9] Y. Du, Z. Fu, and V. D. Calhoun, "Classification and prediction of brain disorders using functional connectivity: promising but challenging," *Frontiers in neuroscience*, vol. 12, 2018.

[10] F. de Vos, M. Koini, T. M. Schouten, S. Seiler, J. van der Grond, A. Lechner *et al.*, "A comprehensive analysis of resting state fmri measures to classify individual patients with alzheimer's disease," *Neuroimage*, vol. 167, pp. 62–72, 2018.

[11] R. Cuingnet, E. Gerardin, J. Tessieras, G. Auzias, S. Lehéricy, M.-O. Habert *et al.*, "Automatic classification of patients with alzheimer's disease from structural mri: a comparison of ten methods using the adni database," *neuroimage*, vol. 56, no. 2, pp. 766–781, 2011.

[12] K. J. Friston, "Functional and effective connectivity: a review," *Brain connectivity*, vol. 1, no. 1, pp. 13–36, 2011.

[13] N. Leonardi, J. Richiardi, M. Gschwind, S. Simioni, J.-M. Annoni, M. Schluep *et al.*, "Principal components of functional connectivity: a new approach to study dynamic brain connectivity during rest," *NeuroImage*, vol. 83, pp. 937–950, 2013.

[14] J. A. Contreras, A. Avena-Koenigsberger, S. L. Risacher, J. D. West, E. Tallman, B. C. McDonald *et al.*, "Resting state network modularity along the prodromal late onset alzheimer's disease continuum," *NeuroImage: Clinical*, vol. 22, p. 101687, 2019.

[15] S. Ahmadi and M. Rezghi, "A novel extension of generalized low-rank approximation of matrices based on multiple-pairs of transformations," *arXiv preprint arXiv:1808.10632*, 2018.

[16] N. Leonardi and D. Van De Ville, "Identifying network correlates of brain states using tensor decompositions of whole-brain dynamic functional connectivity," in *2013 International Workshop on Pattern Recognition in Neuroimaging*. IEEE, 2013, pp. 74–77.

[17] S. W. Park, "Multifactor analysis for fmri brain image classification by subject and motor task," *Electrical and computer engineering technical report*, 2011.

[18] E. Al-sharoa, M. Al-khassaweneh, and S. Aviyente, "Tensor based temporal and multilayer community detection for studying brain dynamics during resting state fmri," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 3, pp. 695–709, 2018.

[19] A. Ozdemir, M. Villafane-Delgado, D. C. Zhu, M. A. Iwen, and S. Aviyente, "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)*. IEEE, 2017, pp. 6299–6303.

[20] L. He, C.-T. Lu, H. Ding, S. Wang, L. Shen, P. S. Yu *et al.*, "Multi-way multi-level kernel modeling for neuroimaging classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 356–364.

[21] G. Ma, L. He, C.-T. Lu, P. S. Yu, L. Shen, and A. B. Ragin, "Spatio-temporal tensor analysis for whole-brain fmri classification," in *Proceedings of the 2016 SIAM International Conference on Data Mining*. SIAM, 2016, pp. 819–827.

[22] J. Ye, "Generalized low rank approximations of matrices," *Machine Learning*, vol. 61, no. 1-3, pp. 167–191,

2005.

[23] M. Rezghi, "A novel fast tensor-based preconditioner for image restoration," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4499–4508, 2017.

[24] J. Zhang, L. Zhou, L. Wang, and W. Li, "Functional brain network classification with compact representation of sice matrices," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 6, pp. 1623–1634, 2015.

[25] X. Chen, H. Zhang, Y. Gao, C.-Y. Wee, G. Li, D. Shen *et al.*, "High-order resting-state functional connectivity network for mci classification," *Human brain mapping*, vol. 37, no. 9, pp. 3282–3296, 2016.

[26] Y. He, L. Wang, Y. Zang, L. Tian, X. Zhang, K. Li *et al.*, "Regional coherence changes in the early stages of alzheimers disease: a combined structural and resting-state functional mri study," *Neuroimage*, vol. 35, no. 2, pp. 488–500, 2007.

[27] A. Bakkour, J. C. Morris, D. A. Wolk, and B. C. Dickerson, "The effects of aging and alzheimer's disease on cerebral cortical anatomy: specificity and differential relationships with cognition," *Neuroimage*, vol. 76, pp. 332–344, 2013.

[28] A. A. Brewer and B. Barton, "Visual cortex in aging and alzheimer's disease: changes in visual field maps and population receptive fields," *Frontiers in psychology*, vol. 5, p. 74, 2014.

[29] J.-H. Jacobsen, J. Stelzer, T. H. Fritz, G. Chételat, R. La Joie, and R. Turner, "Why musical memory can be preserved in advanced alzheimers disease," *Brain*, vol. 138, no. 8, pp. 2438–2450, 2015.

[30] M. Kosicek and S. Hecimovic, "Phospholipids and alzheimers disease: alterations, mechanisms and potential biomarkers," *International journal of molecular sciences*, vol. 14, no. 1, pp. 1310–1322, 2013.

[31] C. Salvatore, A. Cerasa, P. Battista, M. C. Gilardi, A. Quattrone, and I. Castiglioni, "Magnetic resonance imaging biomarkers for the early diagnosis of alzheimer's disease: a machine learning approach," *Frontiers in neuroscience*, vol. 9, p. 307, 2015.

[32] A. M. Brickman, L. B. Zahodne, V. A. Guzman, A. Narkhede, I. B. Meier, E. Y. Griffith *et al.*, "Reconsidering harbingers of dementia: progression of parietal lobe white matter hyperintensities predicts alzheimer's disease incidence," *Neurobiology of aging*, vol. 36, no. 1, pp. 27–32, 2015.

[33] J. De Reuck, F. Auger, N. Durieux, V. Deramecourt, C. Cordonnier, F. Pasquier *et al.*, "Topography of cortical microbleeds in alzheimers disease with and without cerebral amyloid angiopathy: a post-mortem 7.0-tesla mri study," *Aging and disease*, vol. 6, no. 6, p. 437, 2015.

[34] D. Perani, M. Farsad, T. Ballarini, F. Lubian, M. Malpetti, A. Fracchetti *et al.*, "The impact of bilingualism on brain reserve and metabolic connectivity in alzheimer's dementia," *Proceedings of the National Academy of Sciences*, vol. 114, no. 7, pp. 1690–1695, 2017.

[35] R. Watson, S. J. Colloby, A. M. Blamire, and J. T. OBrien, "Subcortical volume changes in dementia with lewy bodies and alzheimer's disease. a comparison with

healthy aging," *International psychogeriatrics*, vol. 28, no. 4, pp. 529–536, 2016.

[36] S. Cai, L. Huang, J. Zou, L. Jing, B. Zhai, G. Ji *et al.*, "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*, vol. 10, no. 2, p. e0115573, 2015.

[37] M. Ortner, L. Pasquini, M. Barat, P. Alexopoulos, T. Grimmer, S. Förster *et al.*, "Progressively disrupted intrinsic functional connectivity of basolateral amygdala in very early alzheimers disease," *Frontiers in neurology*, vol. 7, p. 132, 2016.

[38] R. M. Steketee, E. E. Bron, R. Meijboom, G. C. Houston, S. Klein, H. J. Mutsaerts *et al.*, "Early-stage differentiation between presenile alzheimers disease and frontotemporal dementia using arterial spin labeling mri," *European radiology*, vol. 26, no. 1, pp. 244–253, 2016.

[39] E. J. Sanz-Arigita, M. M. Schoonheim, J. S. Damoiseaux, S. A. Rombouts, E. Maris, F. Barkhof *et al.*, "Loss of small-worldnetworks in alzheimer's disease: graph analysis of fmri resting-state functional connectivity," *PloS one*, vol. 5, no. 11, p. e13788, 2010.

[40] D. Zhang, Y. Wang, L. Zhou, H. Yuan, D. Shen, A. D. N. Initiative *et al.*, "Multimodal classification of alzheimer's disease and mild cognitive impairment," *Neuroimage*, vol. 55, no. 3, pp. 856–867, 2011.

[41] H. I. Jacobs, D. A. Hopkins, H. C. Mayrhofer, E. Bruner, F. W. van Leeuwen, W. Raaijmakers *et al.*, "The cerebellum in alzheimers disease: evaluating its role in cognitive decline," *Brain*, vol. 141, no. 1, pp. 37–47, 2017.