Computer Modeling in Engineering & Sciences CMES, vol.123, no.2, pp.845-871, 2020

Extracting Sub-Networks from Brain Functional Network Using Graph Regularized Nonnegative Matrix Factorization

Zhuqing Jiao[1, *], Yixin Ji[1], Tingxuan Jiao[1] and Shuihua Wang[2, *]

Abstract: Currently, functional connectomes constructed from neuroimaging data have emerged as a powerful tool in identifying brain disorders. If one brain disease just manifests as some cognitive dysfunction, it means that the disease may affect some local connectivity in the brain functional network. That is, there are functional abnormalities in the sub-network. Therefore, it is crucial to accurately identify them in pathological diagnosis. To solve these problems, we proposed a sub-network extraction method based on graph regularization nonnegative matrix factorization (GNMF). The dynamic functional networks of normal subjects and early mild cognitive impairment (eMCI) subjects were vectorized and the functional connection vectors (FCV) were assembled to aggregation matrices. Then GNMF was applied to factorize the aggregation matrix to get the base matrix, in which the column vectors were restored to a common sub-network and a distinctive sub-network, and visualization and statistical analysis were conducted on the two sub-networks, respectively. Experimental results demonstrated that, compared with other matrix factorization methods, the proposed method can more obviously reflect the similarity between the common sub-network of eMCI subjects and normal subjects, as well as the difference between the distinctive sub-network of eMCI subjects and normal subjects, Therefore, the high-dimensional features in brain functional networks can be best represented locally in the low-dimensional space, which provides a new idea for studying brain functional connectomes.

Keywords: Brain functional network, sub-network, functional connectivity, graph regularized nonnegative matrix factorization (GNMF), aggregation matrix.

1 Introduction

Alzheimer's disease (AD), a progressive, irreversible neurodegenerative disease, is one of the most common types of dementia in life which accounts for 50% to 80% of dementia cases. So far, there is no effective clinical treatment for this disorder, which has seriously

[1] School of Information Science & Engineering, Changzhou University, Changzhou, 213164, China.
[2] School of Architecture Building & Civil engineering, Loughborough University, Loughborough, LE11 3TU, UK.
* Corresponding Authors: Zhuqing Jiao. Email: jzq@cczu.edu.cn; Shuihua Wang. Email: shuihuawang@ieee.org.
Received: 01 November 2019; Accepted: 20 November 2019.

affected the patients' daily life [Alzheimer's Association Calcium Hypothesis Workgroup (2017); Tobia, Hayashi, Ballard et al. (2017)]. Mild cognitive impairment (MCI) is an intermediate stage between the expected cognitive decline in normal aging and the more pronounced decline in dementia. On average, the probability of MCI conversion to AD is 10% to 15% annually, and the conversion rate exceeds 50% within 5 years [Gauthier, Reisberg, Zaudig et al. (2006); Wang, Wu, Hou et al. (2016)]. Due to high conversion rate of MCI and increasing life expectancy, MCI intervention through drug and non-drug approaches to reduce AD conversion rate has become a research focus [Wang, Du, Atangana et al. (2018); Lu, Lu and Zhang (2019)]. In the past decades, the accurate identification of early mild cognitive impairment (eMCI) has aroused great concern among researchers [Zhang, Dong, Liu et al. (2015)]. More and more studies show that there are very subtle changes in the brain of patients with eMCI, which mainly manifest in some abnormal functional connectivity [Jiao, Zou, Cao et al. (2014); Zhang, Wang and Sui (2018)]. Therefore, it is necessary to study the differences between eMCI and normal people through brain functional networks [Wee, Yap, Zhang et al. (2012); Chen, Ward, Xie et al. (2011); Zhou, Wang, Li et al. (2011)].

Nowadays, scientists utilize functional magnetic resonance imaging (fMRI) to divide brains into different voxels, and analyze different activity patterns through the response strength of voxels [Jie, Shen and Zhang (2014)]. In the studies of resting-state fMRI, many researchers pay attention to the brain functional networks based on graph theories [Stam (2010); Alzheimer's Association Calcium Hypothesis Workgroup (2017)], and their work mainly focus on the overall modes of the functional networks. The functions of human brain need to be realized by mutual cooperation between different brain regions, and the cooperative relations of brain regions can be demonstrated through functional connectivity. In recent years, feature learning for functional connectomes has been proved to be a valuable tool in characterizing and differentiating brain disorders from normal subjects. If a certain brain disease is just manifested as some cognitive dysfunction while other cognitive functions are normal, it indicates that the functional abnormalities of sub-networks which are composed of some brain regions [Liang, Zhou, Jiang et al. (2006); Li and Li (2016)]. Therefore, accurate identification of these sub-networks is crucial for disease diagnosis.

Previous studies often adopted some matrix factorization methods during numerical analysis, such as principal component analysis [Baumgartner, Ryner, Richter et al. (2000)], independent component analysis [Beckmann, Deluca, Devlin et al. (2005)], dictionary learning algorithm [Li, Zhu, Jiang et al. (2014)], clustering algorithm [Chen, Li, Zhu et al. (2013)], sparse coding [Lee, Tak and Ye (2011); Zhang, Li, Lv et al. (2013)], etc. Traditional matrix factorization ways allow the existences of negative factorized matrices, but it is difficult to explain negative values by its physical meaning in practical applications. Nonnegative matrix factorization (NMF) aims to find two low-dimensional non-negative matrices to approximately express the high-dimensional original matrix, and to characterize the internal structural information and characteristics

in the matrices [Shahnaz, Berry, Pauca et al. (2006)]. Since Lee et al. [Lee and Seung (1999)] proposed the NMF method based on a single target, many popular machine learning methods including NMF algorithm have been applied to various research fields. Padilla et al. [Padilla, Górriz, Ramírez et al. (2010)] used Support Vector Machine (SVM) to classify computed tomography (CT) images and applied NMF to AD diagnosis for the first time. Khambhati et al. [Khambhati, Sizemore, Betzel et al. (2018)] applied NMF to brain functional networks, and proposed a dynamic graphical architecture for modeling the activity patterns of functional connectivity. Ou et al. [Ou, Xie, Li et al. (2015)] extracted sub-networks from brain functional networks using divergence-based projective NMF and studied their connectivity patterns.

However, most machine learning algorithms are carried out in Euclidean space, and it is difficult to find the internal geometric structure of data space, resulting in the failure to effectively extract high-dimensional features from some data that can reflect the spatial structure in the network. Especially for a brain with complex spatial structure, extracting effective high-dimensional features is crucial for data clustering and classification [Stam (2010)]. Cai et al. [Cai, He, Han et al. (2011)] introduced the manifold concept into NMF, and proposed graph regularized nonnegative matrix factorization (GNMF) for unsupervised learning. In relevant studies, some features were extracted for factorization, such as matrix dispersion etc., which could not only cluster these features into more compact clusters, better reflect the geometric connections among the elements in a matrix, but also ensure the speed and efficiency of matrix factorization. In fact, brain functional networks change dynamically over time, which contain a lot of valuable information [Wang, Ren and Zhang (2017); Hutchison, Womelsdorf, Gati et al. (2013)]. Therefore, it is essential to accurately describe the actual state of functional connectivity, especially the parts of the functional network with abnormal connectivity patterns. We vectorized the dynamic functional networks of normal subjects and early mild cognitive impairment (eMCI) subjects respectively, and assembled the functional connectivity vectors (FCV) into aggregation matrices. Then we applied GNMF to splice the aggregation matrix to several sub-networks, and conducted visualization and statistical analysis on common sub-networks and distinctive sub-networks, respectively. Finally, the results of the proposed method were compared with those of other matrix factorization methods to verify the validity of GNMF on extracting sub-networks from brain functional networks.

2 Materials and methods

Our work used the resting-state fMRI data from ADNIGo and ADNI2 dataset of Alzheimer's Disease Neuroimaging Initiative (ADNI) (http://adni.loni.ucla.edu). A total of 62 subjects were involved in the experiment, including 32 eMCI subjects (16 males and 16 females, age 63±7 years) and 30 normal subjects (17 males and 13 females, age 65±10 years). Since two female eMCI subjects had head movements and other behaviors during the test, these two subjects were ignored. All subjects were obtained by the same scanning method using a 3.0 T Philips Achieva scanner, and subjects were required to

remain relax and awake during the scanning process. The scanning time for each subject was 7 minutes, with a large frame shift of more than 2.5 minutes (FD>0.5). Specific scanning parameters are as follows: Repeat time TR=3000 ms, Echo time TE=30 ms, Flip angle 80°, Time points=140, Imaging matrix size=64×64, Layer thickness=3.3 mm.

The raw fMRI data was preprocessed in DPARSF toolbox (http://rfmri.org/dparsf) of Matlab R2012a. The preprocessing operations include slice timing, realignment, spatial normalization, smoothing, detrend, filtering, etc. Each subject's brain was divided into 90 brain regions according to Anatomical Automatic Labeling (AAL) template, and the time series of each brain region were extracted [Tzourio-Mazoyer, Landeau, Papathanassiou et al. (2002)]. The filtering range is 0.01-0.08 Hz, the standardized bounding box is [-90, -126, -72; 90, 90, 108], and the voxel size is [3, 3, 3]. It takes a certain amount of time for both the machine and the subjects to enter a stable state. The first 3 time points are removed during preprocessing, and the remaining 137 time points are retained for subsequent analysis. The data of subjects with large head movements (translation >2 mm, rotation >2°) are removed after realigning.

The whole time series are divided into several overlapping sub-segments with the same window length by sliding time windows [Chen, Zhang, Gao et al. (2016); Chen, Zhang, Zhang et al. (2017)]. Then the dynamic functional networks are constructed by calculating the *Pearson* correlation coefficients between the time series of two brain regions in the same window. There are two main methods of vectorizing brain functional networks: (i) In view of the symmetry of correlation coefficient matrices, the columns in their upper triangular elements are assembled into one column to generate a FCV with the dimension of (89+1)×89/2=4005. (ii) The correlation coefficients between each brain region and all other brain regions are accumulated into one column to generate a 90-dimensional vector that represents the functional connectivity strength (FCS) of a brain region. The first method completely retains all the information in the networks and the process is reversible, but the computing speed is very slow. Although the second method loses some information in the networks, it has lower dimensions and is easy to calculate. Therefore, we used the first method to vectorize dynamic functional networks and the second method to analyze FCS.

NMF is a linear and non-negative approximate data description of a non-negative matrix [Pascual-Montano, Carmona-Saez, Chagoyen et al. (2006)]. For a given matrix $\boldsymbol{V}$ with a size of $m \times n$, m represents the sample characteristics and n represents the number of samples, that is, each column vector of the matrix represents a sample. The matrix $\boldsymbol{V}$ is factorized into a base matrix $\boldsymbol{W}$ with a size of $m \times r$ and a weight matrix $\boldsymbol{H}$ with a size of $r \times n$, where $\boldsymbol{H}$ is the projection matrix of the original matrix $\boldsymbol{V}$ on the base matrix $\boldsymbol{W}$, and the value of r is generally smaller than that of m and n. The dynamic brain functional network of subject i was vectorized to some FCVs from a mathematical point of view, and the spliced sub-aggregation matrix $\boldsymbol{V}_i$ can be approximately factorized as follows:

$$\boldsymbol{V_i} \approx h_{1,t}\boldsymbol{W}_1 + \ldots + h_{j,t}\boldsymbol{W}_j + \ldots + h_{r,t}\boldsymbol{W}_r = \left[\boldsymbol{W}_1 \ldots \boldsymbol{W}_j \ldots \boldsymbol{W}_r\right] \begin{bmatrix} h_{1,t} \\ \vdots \\ h_{j,t} \\ \vdots \\ h_{r,t} \end{bmatrix} \tag{1}$$

where $\boldsymbol{W}_j$ is the column vector corresponding to the j-th sub-network, $h_{j,t}$ is the weight $\boldsymbol{W}_j$, t is the total number of column vectors in $\boldsymbol{V}_i$, and r is the number of sub-networks.

Accordingly, the aggregation matrix spliced by the FCVs of a type of subjects can be approximately factorized as follows:

$$\boldsymbol{V} = \left[\boldsymbol{V}_1 \ldots \boldsymbol{V}_i \ldots \boldsymbol{V}_N\right] \approx \left[\boldsymbol{W}_1 \ldots \boldsymbol{W}_j \ldots \boldsymbol{W}_r\right] \begin{bmatrix} h_{1,1} & \cdots & h_{1,t} & \cdots & h_{1,n} \\ \vdots & \ddots & & \iddots & \vdots \\ h_{j,1} & & h_{j,t} & & h_{j,n} \\ \vdots & \iddots & & \ddots & \vdots \\ h_{r,1} & \cdots & h_{r,t} & \cdots & h_{r,n} \end{bmatrix} = \boldsymbol{WH} \tag{2}$$

where $\boldsymbol{W}$ is the base matrix composed of the column vectors corresponding to all sub-networks, $\boldsymbol{H}$ is the weight matrix composed of the weights of all subjects' column vectors, N is the number of subjects, and n is the total number of all subjects' column vectors. Thus, the aggregation matrix can also be considered as a weighted combination of all subjects' sub-networks.

The square of Euclidean distance between $\boldsymbol{V}$ and $\boldsymbol{WH}$ is taken as the objective of NMF:

$$\min_{\boldsymbol{W},\boldsymbol{H}} \|\boldsymbol{V} - \boldsymbol{WH}\|_F^2 \text{ s.t.} \boldsymbol{W} \geq 0, \boldsymbol{H} \geq 0 \tag{3}$$

where $\|\cdot\|$ is Frobenius norm. The updating criteria of the objective function of NMF are as follows:

$$\begin{cases} \boldsymbol{W}_{i,k} \leftarrow \boldsymbol{W}_{i,k} \dfrac{\left(\boldsymbol{V}\boldsymbol{H}^T\right)_{i,k}}{\left(\boldsymbol{WH}\boldsymbol{H}^T\right)_{i,k}} \\ \boldsymbol{H}_{k,j} \leftarrow \boldsymbol{H}_{k,j} \dfrac{\left(\boldsymbol{W}^T\boldsymbol{V}\right)_{k,j}}{\left(\boldsymbol{W}^T\boldsymbol{WH}\right)_{k,j}} \end{cases} \tag{4}$$

Both $\boldsymbol{W}$ and $\boldsymbol{V}$ obtained by NMF have non-negative characteristics and locality, so that the result after dimensionality reduction can partially represent the data of vector space [Brunet, Tamayo, Golub et al. (2004); Stam and Reijneveld (2007)]. In particular, $\boldsymbol{W}$ cannot retain as many neighbor structures in $\boldsymbol{V}$ as possible, that is, the geometric structure of data in $\boldsymbol{V}$ is not considered. This method ignores the intrinsic geometric structure of spatial data, which is crucial for clustering and classification. GNMF added graph regularization constraints on

the basis of NMF, which not only retained the local sparsity of NMF, but also enabled to maintain the similarity between the samples in $\boldsymbol{V}$ after data dimensionality reduction. The sample $\boldsymbol{h}_i$ in the i-th column of $\boldsymbol{H}$ can be viewed as a new expression of the sample $\boldsymbol{v}_i$ in the i-th column of $\boldsymbol{V}$ in $\boldsymbol{W}$. Assume that the samples $\boldsymbol{v}_i$ and $\boldsymbol{v}_j$ in $\boldsymbol{V}$ are adjacent samples, the samples $\boldsymbol{h}_i$ and $\boldsymbol{h}_j$ in $\boldsymbol{H}$ are also adjacent samples after mapping the two samples in $\boldsymbol{V}$ to the low-dimensional manifold. In order to make the factorized samples maintain the manifold characteristics of $\boldsymbol{V}$, we let $\boldsymbol{G}$ be the graph composed of the samples in $\boldsymbol{V}$. The objective function of GNMF is defined as:

$$\min_{\boldsymbol{W},\boldsymbol{H}} \|\boldsymbol{V}-\boldsymbol{W}\boldsymbol{H}\|_F^2+\lambda tr\left(\mathrm{HLH}^T\right)\mathrm{s.t.}\boldsymbol{W}\geq 0,\boldsymbol{H}\geq 0 \tag{5}$$

where tr represents trace, $\boldsymbol{L}$=$\boldsymbol{D}$-$\boldsymbol{K}$ is the Laplace matrix of $\boldsymbol{G}$, $\boldsymbol{K}$ is the weight matrix to measure the similarity between samples, and $\boldsymbol{D}$ is a diagonal matrix where each element on the diagonal is the sum of each row (or column) in $\boldsymbol{K}$, that is, $d_{ij}=\sum_i k_{ij}$; λ is the regularized parameter, and when it is equal to 0, GNMF degenerates to NMF.

The objective function measures the degree of smoothness of the space represented by data points in low dimensions. In view of this, the $\boldsymbol{K}$ introduced in GNMF can further highlight the difference between different samples and serve as a regularization term to optimize the objective function. The binary elements of $\boldsymbol{K}$ are as follows:

$$k_{ij}=\begin{cases}1, v_i\in \boldsymbol{N}_k\left(\boldsymbol{v}_j\right) or\, \boldsymbol{v}_j\in \boldsymbol{N}_k\left(\boldsymbol{v}_i\right)\\ 0, else\end{cases} \tag{6}$$

where $\boldsymbol{N}_k(v_i)$ and $\boldsymbol{N}_k\left(v_j\right)$ represent the sets of the k most similar samples of the i-th sample and the j-th sample in the feature space, respectively, which are called k-Nearest Neighbors.

The updating criteria of the objective function of GNMF are as follows:

$$\begin{cases} w_{ik}\leftarrow w_{ik}\dfrac{\left(\boldsymbol{V}\boldsymbol{H}^T\right)_{ik}}{\left(\boldsymbol{W}\boldsymbol{H}\boldsymbol{H}^T\right)_{ik}}\\ h_{kj}\leftarrow h_{kj}\dfrac{\left(\boldsymbol{W}^T\boldsymbol{V}+\lambda\boldsymbol{H}\boldsymbol{K}^T\right)_{kj}}{\left(\boldsymbol{W}^T\boldsymbol{W}\boldsymbol{H}+\lambda\mathrm{H}\boldsymbol{D}^T\right)_{kj}}\end{cases} \tag{7}$$

Assume that the original matrix $\boldsymbol{V}$ = [$\boldsymbol{v}_1$,…, $\boldsymbol{v}_i$,…, $\boldsymbol{v}_n$] is divided into k clustering clusters, and each cluster forms a base matrix $\boldsymbol{W}$. When the sample $\boldsymbol{v}_i$ belongs to the k-th cluster, the sample $\boldsymbol{h}_i$ in the weight matrix $\boldsymbol{H}$ is all 1, otherwise it is 0. Then the objective function of k-means clustering is:

$$\min_{\boldsymbol{V}\geq 0,\boldsymbol{H}\geq 0,\boldsymbol{H}^T\boldsymbol{H}=\boldsymbol{I}} \|\boldsymbol{V}-\boldsymbol{W}\boldsymbol{H}\|_F^2 \tag{8}$$

When $\boldsymbol{H}^T\boldsymbol{H}$=$\boldsymbol{I}$, NMF and k-means clustering are equivalent, so there a certain correlation among NMF, GNMF and k-means clustering [Li, Bu, Li et al. (2018)]. In dictionary

learning, the base matrix $\boldsymbol{W} \in \boldsymbol{R}^{m\times r}$ is used to sparse the unfactorized matrix, among which the weight matrix is $\boldsymbol{H} \in \boldsymbol{R}^{r\times n}$. The objective function of dictionary learning is:

$$\min_{\boldsymbol{W},\boldsymbol{H}} \|\boldsymbol{V} - \boldsymbol{W}\boldsymbol{H}\|_F^2 \, s.t. \forall i, \|\boldsymbol{h}_i\|_0 \leq T_0 \tag{9}$$

where $\boldsymbol{h_i}$ is the vector in the i-th row of the weight matrix, and T_0 is the sparsity constraint parameter.

Optimizing objective function is a least-square problem, which can be solved by singular value decomposition (SVD). The objective function of dictionary learning and the objective function of NMF algorithm are equivalent under certain conditions [Zhang and Li (2010)], so there is also a certain correlation between GNMF and dictionary learning.

The number of sub-networks is a very important parameter in the factorization of aggregation matrices. There are many methods to determine the number of clusters in traditional clustering, but there is no method to determine the optimal number of clusters in NMF. Brunet et al. [Brunet, Tamayo, Golub et al. (2004)] proposed a method about cophenetic correlation coefficient (CCC) to select the optimal value of r. CCC is defined as the correlation coefficient between the distance matrix $(\boldsymbol{I} - \overline{\boldsymbol{C}})$ corresponding to the consensus matrix $\overline{\boldsymbol{C}}$ and the linkage matrix of $\overline{\boldsymbol{C}}$ in the process of reordering it. Notably $\overline{\boldsymbol{C}}$ generating from multiple clustering is the average value of the adjacent matrix $\boldsymbol{C}$, a binary matrix. If the samples $\boldsymbol{v}_i$ and $\boldsymbol{v}_j$ belong to the same cluster in a certain clustering, the corresponding element in $\boldsymbol{C}$ is 1, otherwise 0. If the result of each clustering is the same, the elements in the consensus matrix $\overline{\boldsymbol{C}}$ can only be 0 or 1. The range of elements in $\overline{\boldsymbol{C}}$ is 0 to 1, it reflects the probability that the samples $\boldsymbol{v}_i$ and $\boldsymbol{v}_j$ are clustered to the same cluster. If a clustering result is stable, there will be no changes in $\overline{\boldsymbol{C}}$, and the elements of $\overline{\boldsymbol{C}}$ are close to 0 or 1. The value of CCC is 1 when all elements in $\overline{\boldsymbol{C}}$ are 0 or 1, and the value of CCC is less than 1 when the elements are between 0 and 1. Therefore, the optimal value of r is generally taken when CCC decreases and the blocking effect of $\overline{\boldsymbol{C}}$ is the most significant (the classification results are most stable).

The aggregation matrices of normal subjects and eMCI subjects were factorized using GNMF, respectively. Then, each column in the obtained base matrix is restored to a sub-network by converting the 90×90 correlation coefficient matrix into a 4005×1 column vector. Then the sub-network was transformed by Min-Max standard normalization [Jiao, Xia, Ming et al. (2019)], and the elements in the matrix were normalized to the interval of [0, 1]. The transformation function is as follows:

$$Corr^* = \frac{Corr - \text{Min}}{\text{Max} - \text{Min}} \tag{10}$$

where *Corr** is the correlation coefficient after transformation, *Corr* is the correlation coefficient before transformation, Max and Min are the maximum value and the minimum value among the correlation coefficients, respectively.

3 Results

Zhang et al. [Zhang, Zhang, Chen et al. (2017)] studied the influence of the width of sliding windows on the accuracy of MCI classification, and found that width of 70 performed best. In this study, the width of sliding windows was also set as 70, and the step length was set as 1. The size of all aggregation matrices was 4005×630. The number of sub-networks is selected in the range from r=2 to r=10, and the consensus matrices of the corresponding aggregation matrices are visualized respectively as shown in Fig. 1. When r=2, the blocking effect of the consensus matrix is the most obvious, and the elements are also close to 0 or 1, and the classification result is the most stable. From r=3, the blocking effect becomes unstable and has some elements between 0 and 1.

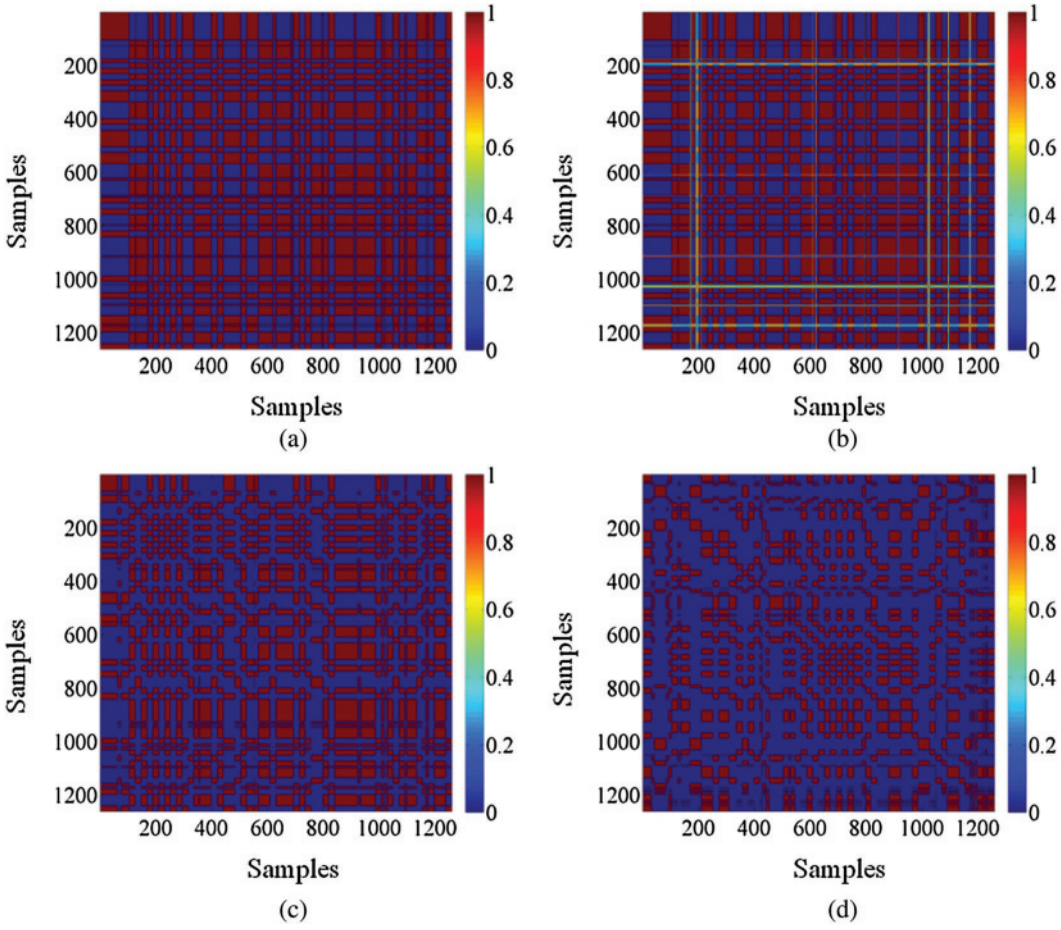

Figure 1: Consensus matrices of aggregation matrices for all subjects, where the horizontal and vertical coordinates all contain 1260 FCVs. (a) r=2, (b) r=3, (c) r=4, (d) r=5

GNMF was used to factorize the aggregation matrices corresponding to each type of subjects, extract the normalized sub-networks and visualize them. Fig. 2 illustrates the factorization process of an aggregation matrix which is expressed in formula (2).

The extracted sub-networks are visualized by BrainNet Viewer toolkit (http://www.nitrc.org/projects/bnv) in Matlab R2012a [Jiao, Ma, Wang et al. (2018)]. There are more functional connections in sub-networks, and the effect is not obvious after all of them are visualized. Therefore, a specific number of functional connections with the highest correlation coefficient are selected for visualization. Our preliminary experiments showed that visualization was the most obvious when the top 200 functional connectivities with the highest correlation coefficients were selected as significant functional connectivities. Then r=2 was selected as the number of sub-networks, and the sub-networks were extracted by GNMF. Fig. 3 shows the visualization of sub-networks extracted for normal subjects and eMCI subjects.

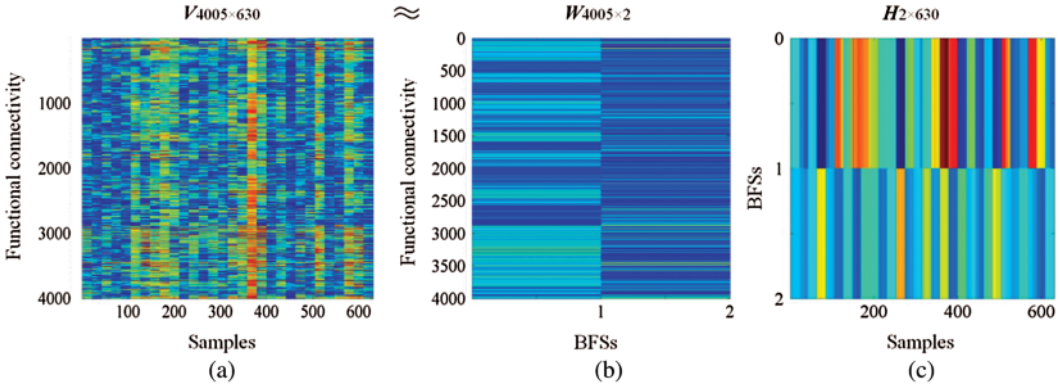

Figure 2: Factorization process of an aggregation matrix (a) Aggregation matrix, (b) Base matrix (c) Weight matrix

In Fig. 3, Normal-BFS#1 and Normal-BFS#2 represent the two sub-networks extracted from the functional network of normal subjects, and eMCI-BFS#1 and eMCI-BFS#2 represent the two sub-networks extracted from the functional network of eMCI subjects. Each column of the sub-networks was taken as a group of samples, and the similarity and difference test were performed on BFS#1 and BFS#2 by two-sample t-test. The binary parameter h was set to test whether the samples from each group came from the same distribution. h=0 indicates that the mean values of the two groups of samples from the sub-networks of normal subjects and eMCI subjects are equal, that is, they are from the same distribution. On the contra, h=1 indicates that the mean values of the two groups of samples from the sub-networks of normal subjects and eMCI subjects are different and not from the same distribution. The parameter p was set to represent the criteria for significant differences. p<0.05 indicated that there was a difference between

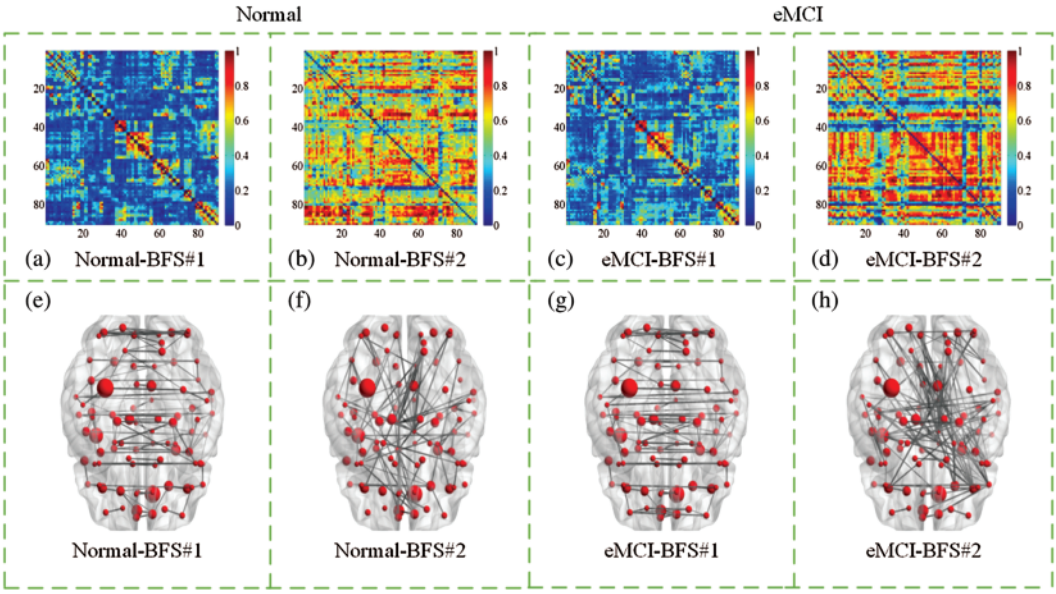

Figure 3: Visualization of the sub-networks extracted by GNMF

the samples from the sub-networks of normal subjects and eMCI subjects, and $p<0.01$ indicated that the difference between the samples from the sub-networks of normal subjects and eMCI subjects was extremely significant. Tab. 1 shows the two-sample t-test results of sub-networks extracted by GNMF.

There were 62 groups of samples in Normal-BFS#1 and eMCI-BFS#1 tested but only 47 groups of samples in Normal-BFS#2 and eMCI-BFS#2 tested when h was equal to 0. It indicates that most samples in the two BFS#1 were from the same distribution while most samples in the two BFS#2 were from different distributions. There were only 28 groups of samples in Normal-BFS#1 and eMCI BFS#1 tested and 43 groups of samples in Normal-BFS#1 and eMCI BFS#1 tested when h was equal to 1. It indicated that most

Table 1: Two-sample t-test results of the sub-networks extracted by GNMF

Index	Number of samples	
	Normal- BFS#1 and eMCI-BFS#1	**Normal-BFS#2 and eMCI-BFS#2**
h=0	62	47
h=1	28	43
$p<0.05$	28	43
$p<0.01$	19	32

samples in the two BFS#1 were from the same distribution while most samples in the two BFS#2 were from different distributions. There only 28 groups of samples in Normal-BFS#1 and eMCI BFS#1 were tested but 43 groups of samples in Normal-BFS#2 and eMCI-BFS#2 tested when p was less than 0.05. It indicated that there was little difference between the two BFS#1 while there was significant difference between the two BFS#2. There were only 19 groups of samples in BFS#1 and eMCI BFS#1 were tested but 32 groups of samples in Normal-BFS#2 and eMCI-BFS#2 were tested when p was less than 0.01. It indicates that there was extremely significant difference between the two BFS#2 while there was no significant difference between the two BFS#1.

Therefore, the four sub-networks were categorized into a similarity pair and a difference pair. The similar pair consisted of Normal BFS#1 and eMCI BFS#1, which were called common sub-networks and reflected the similar connectivity patterns in non-brain disease-related sub-networks. The difference pair consisted of Normal BFS#2 and eMCI BFS#2, which were called distinctive sub-networks and reflected the different connectivity patterns in brain disease-related sub-networks. Fig. 4 shows the visualization of the common sub-networks, where there are no significant differences between the common sub-networks of normal subjects and eMCI subjects. Moreover, the significant functional connectivities are mainly concentrated in their corresponding default networks, where the connectivity patterns are basically consistent [Jiao, Wang and Ma (2016); Jiao, Wang, Ma et al. (2017)]. However, in the common sub-network of eMCI subjects, some brain regions such as Occipital_Inf_L (IOG.L), Frontal_Mid_Orb_L (ORBmid.L) and Lingual_R (LING.R), etc., lacked connectivities with other brain regions.

Fig. 5 shows the visualization of the distinctive sub-networks. We can find that there is big difference in the density of functional connectivities between the distinctive sub-networks of normal subjects and eMCI subjects. It indicates that the connectivity patterns of distinctive

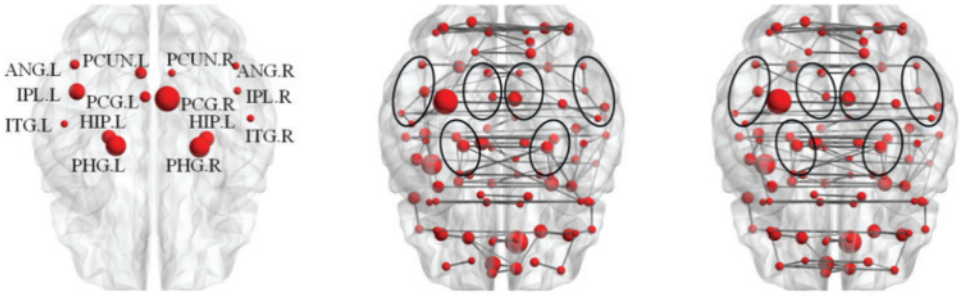

Figure 4: Visualization of the common sub-networks, (a) Brain regions in the default network, which are circled in (b) and (c), (b) Connectivity patterns of default network in the common sub-network of normal subjects, (c) Connectivity patterns of the default network in the common sub-network of eMCI subjects

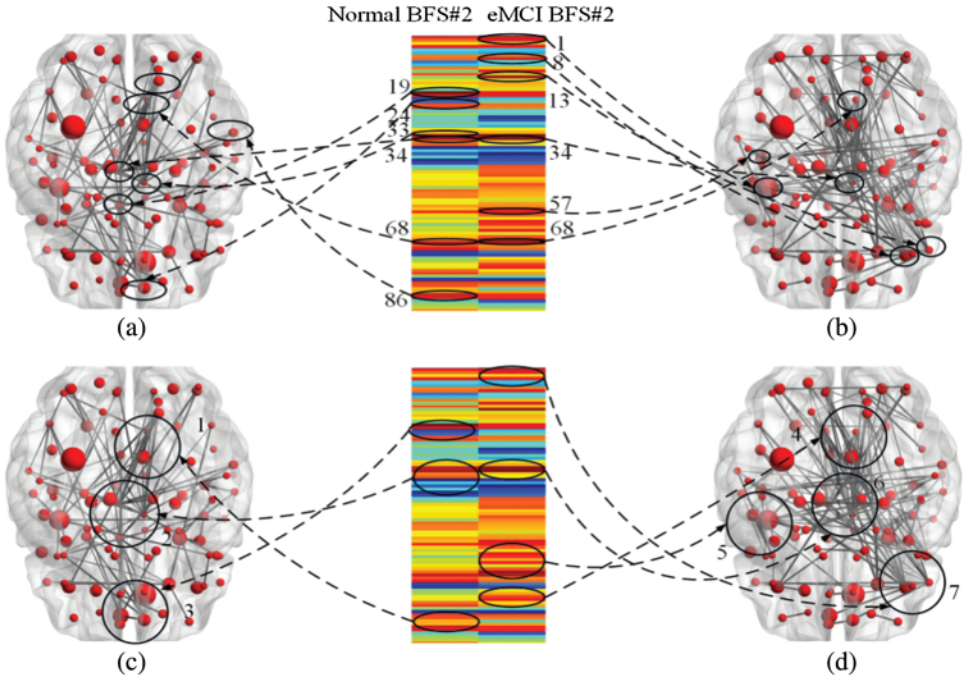

Figure 5: Visualization of the distinctive sub-networks. (a) Circled part is the brain regions of normal subjects with significantly enhanced FCS; (b) Circled part is the brain regions of eMCI subjects with significantly enhanced FCS; (c) and (d) Circled parts are the connectivity patterns of the brain regions with significantly different FCS between the distinctive sub-networks of normal subjects and eMCI subjects

sub-networks of eMCI subjects had been significantly changed compared with those of normal subjects. From the perspective of graph theory, FCS can be considered as the degree of a node which reflects the activity degree of a brain region to some extent. The color of the vectors representing FCS can more intuitively distinguish the difference between the connectivity patterns in the distinctive sub-networks of normal subjects and eMCI subjects.

In Figs. 5(a) and 5(b), the FCS in Precentral_L (PreCG.L), Frontal_Mid_R (MFG.R), Frontal_Inf_Tri_L (IFGtriang.L) and other brain regions of eMCI subjects are stronger than those in the same brain regions of normal subjects. While the FCS in Supp_Motor_Area_L (SMA.L), Frontal_Sup_Medial_R (SFGmed.R), Cingulum_Mid_L (DCG.L) and other brain regions of eMCI subjects are weaker than those in the same brain regions of normal subjects. Seven connectivity patterns of the brain regions with significant different FCS were found out from the distinctive sub-networks of normal subjects and eMCI subjects respectively, as shown in Figs. 5(c) and 5(d). Therefore, the connectivity patterns of brain region in default network plays a very important role

whether it is a common sub-network or a distinctive sub-network [Raichle, MacLeod, Snyder et al. (2001)]. Tabs. 2 and 3 show the brain regions with significantly enhanced FCS in the distinctive sub-networks of normal subjects and eMCI subjects, respectively.

Table 2: Brain regions with significantly enhanced FCS in distinctive sub-networks of normal subjects

Number	Brain region	Abbreviations (L: left R: right)	MNI coordinates		
			X (mm)	Y (mm)	Z (mm)
19	Supp_Motor_Area_L	SMA.L	-5.32	4.85	61.38
24	Frontal_Sup_Medial_R	SFGmed.R	9.10	50.84	30.22
33	Cingulum_Mid_L	DCG.L	-5.48	-14.92	41.57
34	Cingulum_Mid_R	DCG.R	8.02	-8.83	39.79
68	Precuneus_R	PCUN.R	9.98	-56.05	43.77
86	Temporal_Mid_R	TPOmid.R	57.47	-37.23	-1.47

Table 3: Brain regions with significantly enhanced FCS in distinctive sub-networks of eMCI subjects

Number	Brain region	Abbreviations (L: left R: right)	MNI coordinates		
			X (mm)	Y (mm)	Z (mm)
1	Precentral_L	PreCG.L	-38.65	-5.68	50.94
8	Frontal_Mid_R	MFG.R	37.59	33.06	34.04
13	Frontal_Inf_Tri_L	IFGtriang.L	-45.58	29.91	13.99
34	Cingulum_Mid_R	DCG.R	8.02	-8.83	39.79
57	Postcentral_L	PoCG.L	-42.46	-22.63	48.92
68	Precuneus_R	PCUN.R	9.98	-56.05	43.77

In the distinctive sub-network of normal subjects, FCS was significantly enhanced in some major brain regions on which cognitive functions were dependent, such as Frontal_Sup_Medial_R (SFGmed.R), Supp_Motor_Area_L (SMA.L), and Para_Hippocampal_L (PHG.L) in prefrontal lobe. FCS was also significantly enhanced in some major brain regions on which vision and hearing were dependent, such as Precuneus_R (PCUN.R) in occipital lobe and Temporal_Mid_R (TPOmid.R) in temporal lobe, which belong to parietal lobe. Moreover, these brain regions are the main parts of front parietal network (FPN). FCS was also significantly enhanced in some important brain regions on which language functions were dependent, such as Temporal_Pole_Mid_R (TPOmid.R) in

temporal lobe. However, FCS has begun to weaken in the brain regions related to cognitive functions of eMCI subjects. For example, the connectivity pattern of some brain regions in frontal lobe, such as Frontal_Mid_R (MFG.R), is quite different from that of normal subjects. Precentral_L (PreCG.L) and Frontal_Inf_Tri_L (IFGtriang.L) are the major brain regions to control behavioral movements. The FCSs of the two brain regions of eMCI subjects were abnormal compared with those of normal subjects, indicating that the behavior and movement of eMCI subjects are different from the normal subjects.

The top 200 functional connectivities with the highest correlation coefficients were selected as significant functional connectivities, and the number of sub-networks was set to r=2. NMF and method optimal direction (MOD) in dictionary learning were applied to factorize the aggregation matrices of normal subjects and eMCI subjects. Figs. 6 and 8 show the visualization of the sub-networks extracted by NMF and MOD, respectively. In addition, k-means clustering was also applied to cluster the aggregation matrices and calculate the average values of the column vectors clustered to the same clusters. Fig. 7 shows the visualization of the sub-networks extracted by k-means clustering.

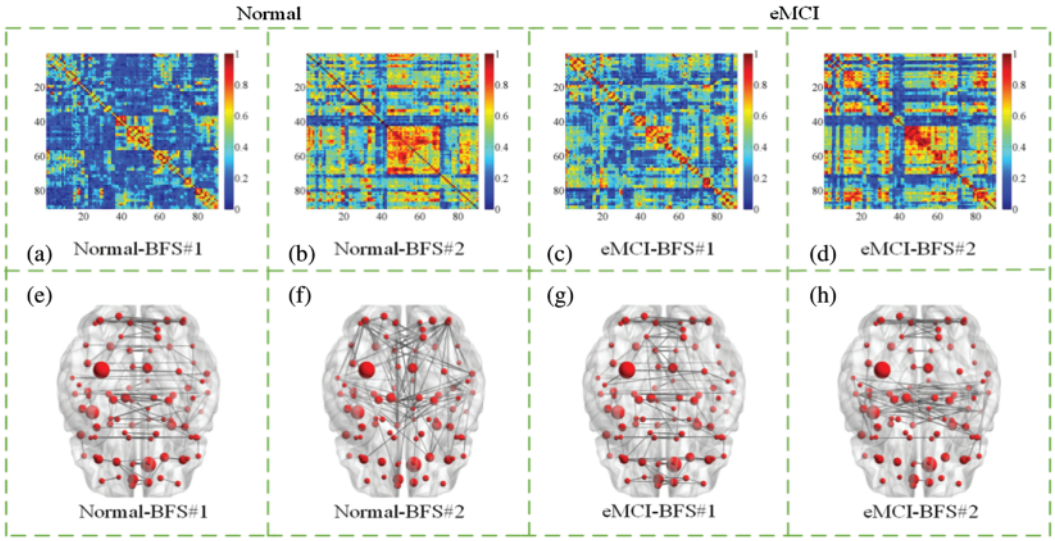

Figure 6: Visualization of the sub-networks extracted by NMF

From the above results, we find that the FCS of some brain regions, such as Precuneus_R (PCUN.R) and Frontal_Sup_Medial_R (SFGmed.R), in the two common sub-networks extracted by GNMF are significantly higher than those in the two common sub-networks extracted by NMF. However, there are no significant differences between the two common sub-networks and between the two distinctive sub-networks extracted by k-means clustering and MOD. Tabs. 4-6 show the results of two-sample t-test for the

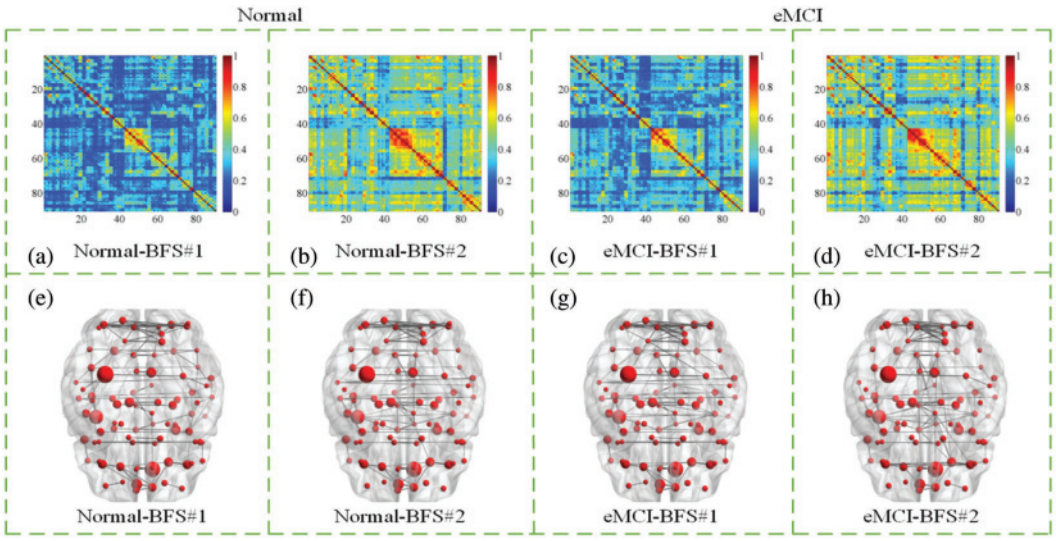

Figure 7: Visualization of the sub-networks extracted by *k*-means clustering

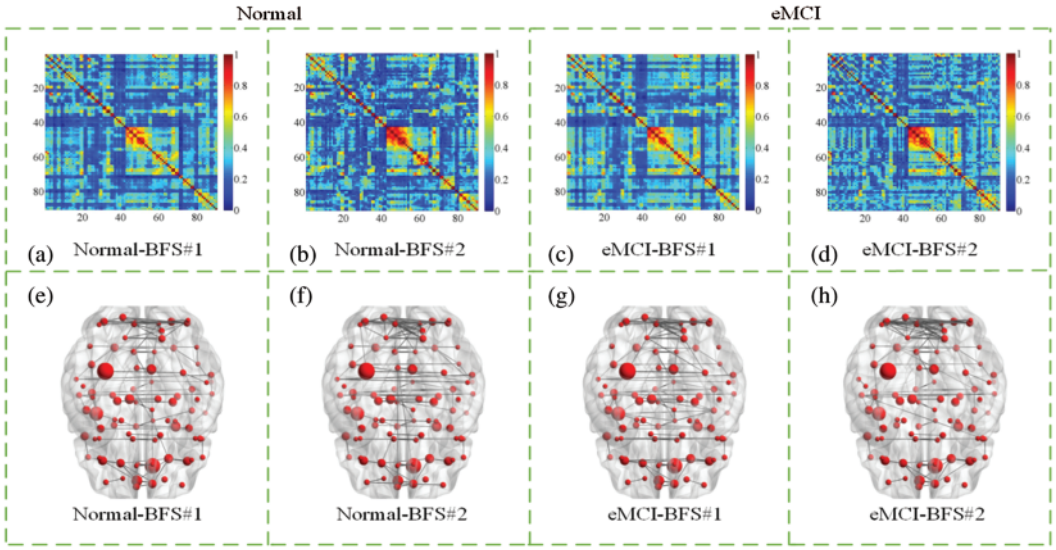

Figure 8: Visualization of the sub-networks extracted by MOD

common sub-networks and the distinctive sub-networks extracted by NMF, *k*-means clustering and MOD, respectively.

In Tabs. 4-6, there were more samples tested in the two common sub-networks when h was equal to 0. It indicated that the results of the above three methods could reflect the similarity

Table 4: Results of two-sample t-test for the sub-networks extracted by NMF

Index	Number of samples	
	Normal-BFS#1 and eMCI-BFS#1	**Normal-BFS#2 and eMCI-BFS#2**
h=0	31	52
h=1	59	38
p<0.05	59	38
p<0.01	51	28

Table 5: Results of two-sample t-test for the sub-networks extracted by k-means clustering

Index	Number of samples	
	Normal-BFS#1 and eMCI BFS#1	**Normal BFS#2 and eMCI-BFS#2**
h=0	42	62
h=1	48	28
p<0.05	48	28
p<0.01	27	18

Table 6: Results of two-sample t-test for the sub-networks extracted by MOD

Index	Number of samples	
	Normal-BFS#1 and eMCI-BFS#1	**Normal-BFS#2 and eMCI-BFS#2**
h=0	80	63
h=1	10	27
p<0.05	10	27
p<0.01	5	13

between the common sub-networks of normal subjects and eMCI subjects. However, there were only a small number of samples tested in the two common sub-networks and the two distinctive sub-networks when h was equal to 1. It indicated that most samples of the two common sub-networks were from the same distribution, while it was difficult to identify whether most samples of the two distinctive sub-networks come from different distribution. There were fewer samples tested in the two common sub-networks than those in the two distinctive sub-networks when p was less than 0.05. Especially in Tabs. 5 and 6, there were very few samples tested in the two common sub-networks and the two distinctive sub-networks when p was less than 0.05. It indicated that the results

obtained by the above three methods could hardly reflect the difference between the distinctive sub-networks of normal subjects and eMCI subjects. The number of samples tested in the two common sub-networks was even more than the number of samples tested in the two distinctive sub-networks when p was less than 0.01. It indicates that there was no significant difference between the distinctive sub-networks of normal subjects and eMCI subjects.

4 Discussions

The idea of key node is introduced to analyze the node attributes of sub-networks as a means of verifying the validity of extracting sub-networks by the above method [Aharon, Elad, Bruckstein et al. (2006); Jiao, Xia, Cai et al. (2018)]. A certain number of highly defined brain regions are key nodes, which can measure the importance of different nodes in sub-networks [Chang and Glover (2010)]. The top 10 brain regions with the highest degree were extracted as the key nodes in each sub-network. Figs. 9-12 shows the degrees of key nodes obtained by different methods, respectively, where the distance between the two polylines of the same color reflects the difference of the degrees of key between the sub-network of normal subjects and the sub-network of eMCI subjects.

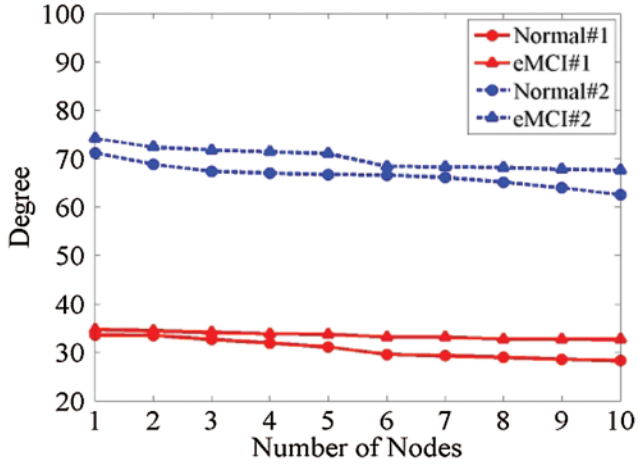

Figure 9: Degrees of key nodes by GNMF

In Figs. 9-12, the two red polylines are very close to each other and even have intersecting and overlapping parts, indicating that the degree of the key nodes in the common sub-network of the normal subjects is not much different from that of the key nodes in the common sub-network of the eMCI subjects. In Fig. 9, the blue-dot polyline is always under the blue-triangle polyline, that is, the degree of the key nodes of the normal subjects in the distinctive sub-network is always lower than that of the eMCI subjects, indicating that the connection pattern of some brain regions in the eMCI subjects is abnormal, reflecting the difference between the normal subjects and the distinctive sub-

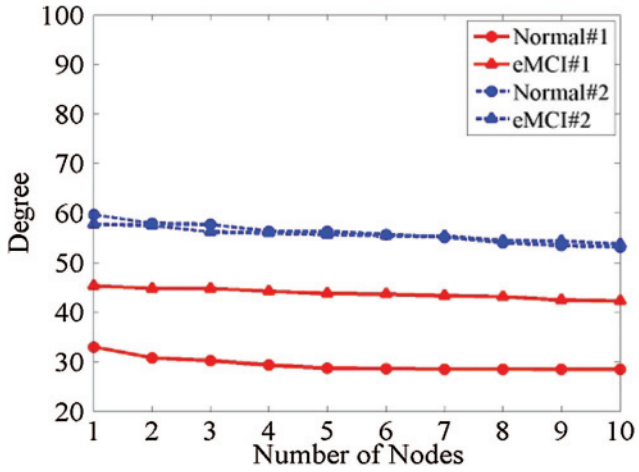

Figure 10: Degrees of key nodes by NMF

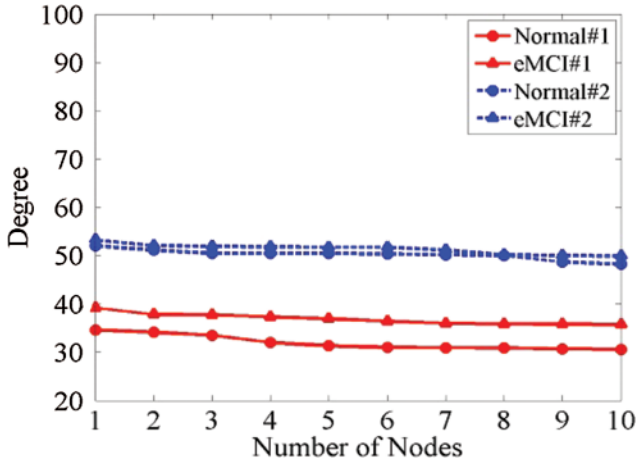

Figure 11: Degrees of key nodes by *k*-means clustering

network of the eMCI subjects. In Fig. 9, the distance between the two blue polylines is larger than that between the two blue polylines in other figures. However, the two blue polylines in Figs. 10-12 are very close. Especially in Figs. 11 and 12, the blue-dot polylines are below the blue-triangle polylines, that is, the results of *k*-means clustering and MOD are contrary to the results of the other two methods. The above results show that GNMF is more effective than other methods in extracting key nodes, and most of the extracted key nodes are consistent with the nodes with abnormal functional connectivity found by visualization.

With different weight matrix solutions, the objective function constructed will be different, resulting in the difference of the extracted sub-networks. We applied binarization, heat

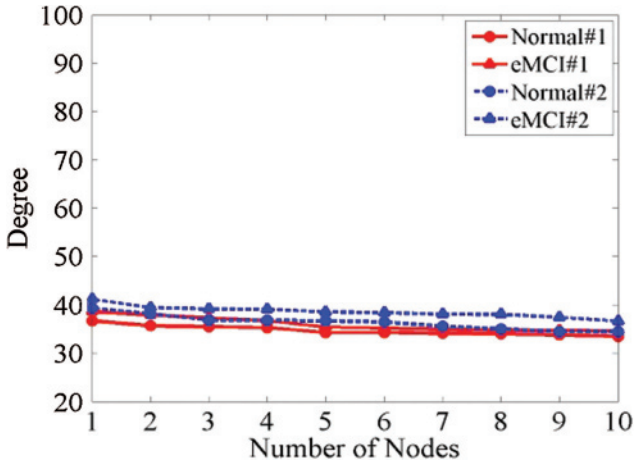

Figure 12: Degrees of key nodes by MOD

kernel function, and cosine to construct the weight matrix $\boldsymbol{K}$ of GNMF [Cai, He, Han et al. (2011)], and the degrees of the key nodes obtained are shown in Figs. 13-15, respectively. By comparing Figs. 13-15, it can be found that the distance between the two red polylines and the two blue broken lines is very small when using heat kernel function to construct the weight matrix. The distance between two blue broken lines is the largest when applying binarization to construct weight matrix. The distance between two red broken lines is the largest, and the distance between two blue broken lines is the smallest when using cosine to construct weight matrix. It shows that these weight matrix methods lead to different

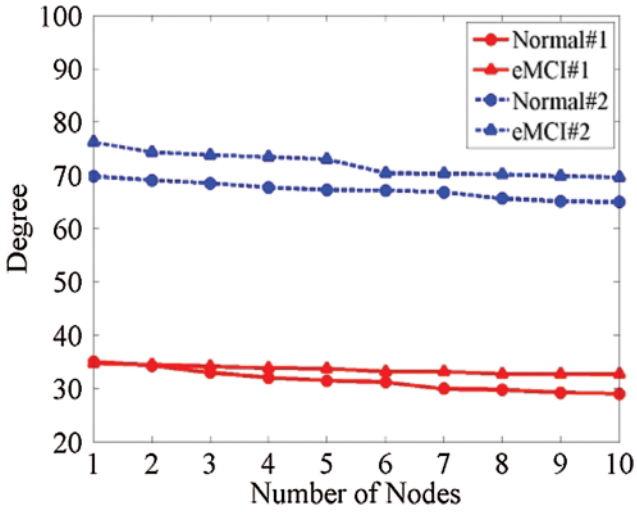

Figure 13: Degrees of key nodes by binarization

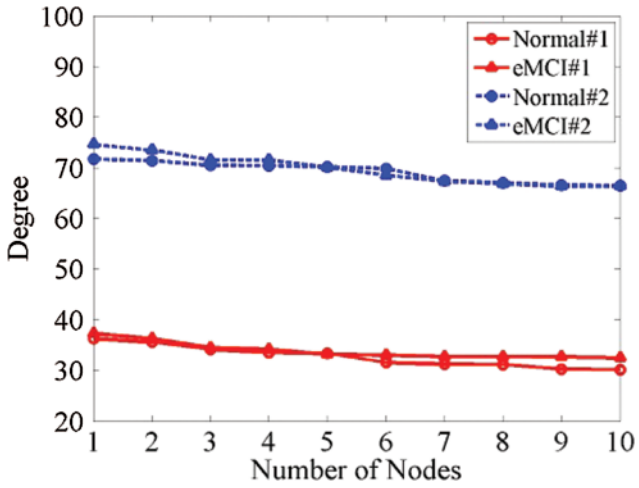

Figure 14: Degrees of key nodes by heat kernel function

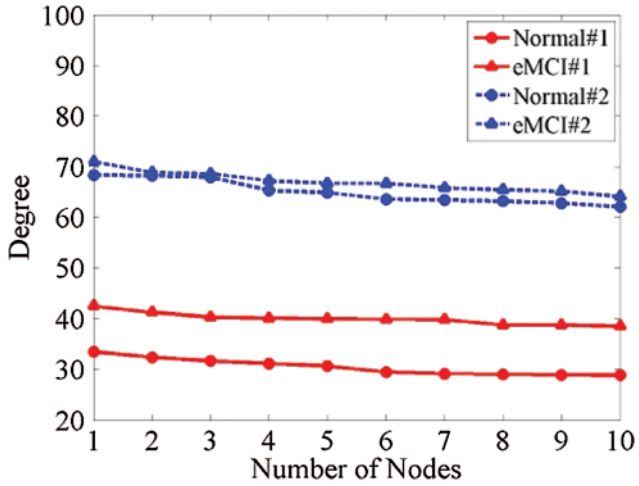

Figure 15: Degrees of key nodes by cosine

similarity between the common sub-networks and difference between the distinctive sub-networks. If it is just to illustrate the similarity between the common sub-networks, it is better to use heat kernel function. However, solving the weight matrix by binarization can simultaneously reflect the similarity between the common sub-networks and the difference between the distinctive sub-networks.

In GNMF, each sample finds out its k nearest neighbor samples and assigns values to the elements in K to form a weight matrix. If the value of k is too large, it may lead to the underfitting of k-Nearest Neighbor model, while if the value of k is too small, it may lead

to the over-fitting of *k*-Nearest Neighbor models. Therefore, different values of *k* have certain influence on the similarity between common sub-networks and the difference between specific sub-networks. Besides 5 is selected as the default value, 4, 5 and 7 values are selected for the experiment. The degrees of key nodes of different *k* values are shown in Figs. 16-19.

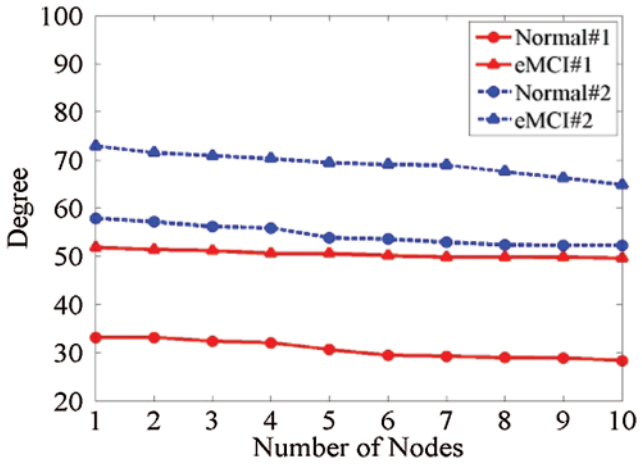

Figure 16: Degree of key nodes when *k*=2

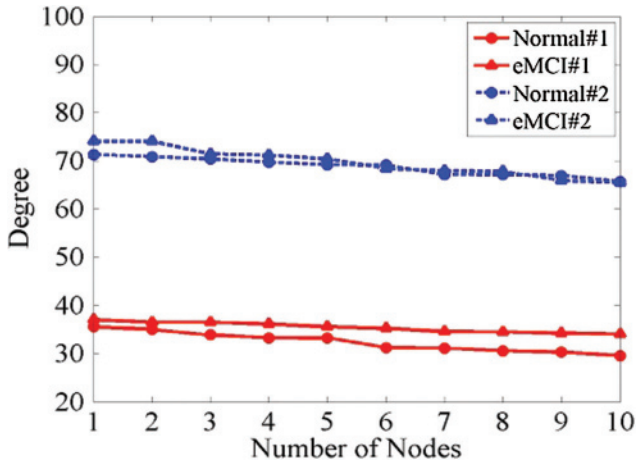

Figure 17: Degree of key nodes when *k*=4

The distance between the two red polylines is the largest when *k* is equal to 2, but the distance between the two red polylines is the smallest when *k* is equal to 5. When *k* is less than 5, as k gets larger and larger, the distances between the two red polylines and between the two blue polylines get smaller and smaller. This shows that when the

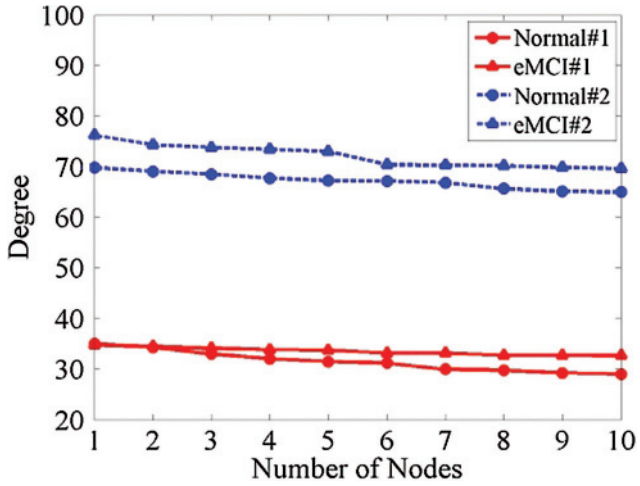

Figure 18: Degree of key nodes when k=5

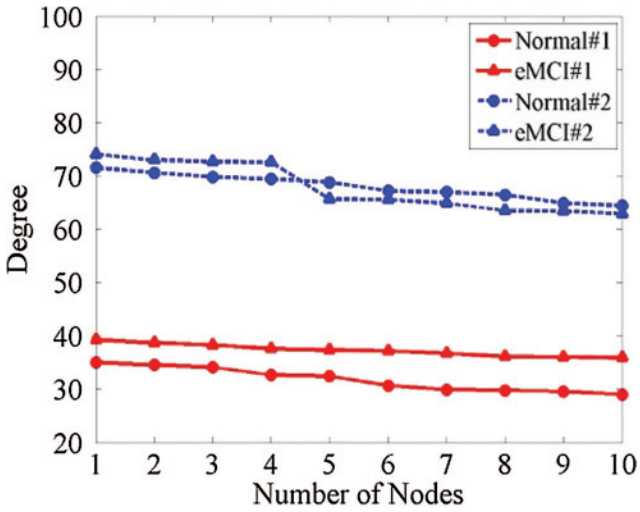

Figure 19: Degree of key nodes when k=7

number of adjacency points increases, the similarity between the two common sub-networks becomes larger and larger, while the difference between the two distinctive sub-networks becomes smaller and smaller. When the default value of k is 5, the two sub-networks obtained have the best effect and can more obviously reflect the similarity between the common sub-networks and the difference between the distinctive sub-networks.

5 Conclusions

Brain functional network is one of the most important technical ways to reveal the pathological mechanism of brain diseases. We propose a method on extracting sub-networks

from brain functional networks. The dynamic functional networks are vectorized and assembled into some aggregation matrices, and GNMF is used to factorize the aggregation matrix into several FCVs, which are restored to sub-networks. We carried out visualization, two-sample t-test and analysis of node attribute. Experimental results show that compared with other matrix factorization algorithms, the proposed method can more effectively analyze the similarity between common sub-networks and the difference between distinctive sub-networks of normal subjects and eMCI subjects. Although other algorithms can reveal the similarity between the common sub-networks, they cannot well reflect the differences between the distinctive sub-networks. Therefore, the dynamic sub-network extraction based on GNMF can not only provide research ideas for determining the core nodes in brain functional network, but also provide important theoretical basis for the analysis of the pathophysiological mechanism of eMCI. However, the process of GNMF is uncertain and can lead to multiple scenarios of factorization results, and this approach focuses on the extraction of data and metrics. As a result, the following research will turn to the use of machine learning algorithm [Sun and Zhang (2019); Yu, Zeng, Liu et al. (2019); Jiang and Zhang (2019)] to classify sub-networks and study the evolution of brain functional connectomes.

Funding Statement: Our work was supported by the National Natural Science Foundation of China (No. 51877013), (ZJ), (http://www.nsfc.gov.cn/), and the Natural Science Foundation of Jiangsu Province (No. BK20181463), (ZJ), (http://kxjst.jiangsu.gov.cn/). This work is also sponsored by Qing Lan Project of Jiangsu Province (no specific grant number), (ZJ), (http://jyt.jiangsu.gov.cn/).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

Aharon, M.; Elad, M.; Bruckstein, A.; Katz, Y. (2006): K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311-4322. DOI 10.1109/TSP.2006.881199.

Alzheimer's Association Calcium Hypothesis Workgroup. (2017): Calcium hypothesis of Alzheimer's disease and brain aging: a framework for integrating new evidence into a comprehensive theory of pathogenesis. *Alzheimer's & Dementia*, vol. 13, no. 2, pp. 178-182.e17. DOI 10.1016/j.jalz.2016.12.006.

Baumgartner, R.; Ryner, L.; Richter, W.; Summers, R.; Jarmasz, M. et al. (2000): Comparison of two exploratory data analysis methods for fMRI: fuzzy clustering *vs.* principal component analysis. *Magnetic Resonance Imaging*, vol. 18, no. 1, pp. 89-94. DOI 10.1016/S0730-725X(99)00102-2.

Beckmann, C. F.; Deluca, M.; Devlin, J. T.; Smith, D. S. M. (2005): Investigations into resting-state connectivity using independent component analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 360, no. 1457, pp. 1001-1013. DOI 10.1098/rstb.2005.1634.

Brunet, J. P.; Tamayo, P.; Golub, T. R.; Mesirov, J. P. (2004): Metagenes and molecular pattern discovery using matrix factorization. *Proceedings of the National Academy of Sciences*, vol. 101, no. 12, pp. 4164-4169. DOI 10.1073/pnas.0308531101.

Cai, D.; He, X.; Han, J.; Huang, T. S. (2011): Graph regularized nonnegative matrix factorization for data representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1548-1560. DOI 10.1109/TPAMI.2010.231.

Chang, C.; Glover, G. H. (2010): Time-frequency dynamics of resting-state brain connectivity measured with fMRI. *NeuroImage*, vol. 50, no. 1, pp. 81-98. DOI 10.1016/j.neuroimage.2009.12.011.

Chen, H. B.; Li, K. M.; Zhu, D. J.; Jiang, X.; Yuan, Y. X. et al. (2013): Inferring group-wise consistent multimodal brain networks via multi-view spectral clustering. *IEEE Transactions on Medical Imaging*, vol. 32, no. 9, pp. 1576-1586. DOI 10.1109/TMI.2013.2259248.

Chen, G.; Ward, B. D.; Xie, C.; Li, W. J.; Wu, Z. L. et al. (2011): Classification of Alzheimer disease, mild cognitive impairment, and normal cognitive status with large-scale network analysis based on resting-state functional MR imaging. *Radiology*, vol. 259, no. 1, pp. 213-221. DOI 10.1148/radiol.10100734.

Chen, X. B.; Zhang, H.; Gao, Y.; Wee, C. Y.; Li, G. et al. (2016): High-order resting-state functional connectivity network for MCI classification. *Human Brain Mapping*, vol. 37, no. 9, pp. 3282-3296. DOI 10.1002/hbm.23240.

Chen, X. B.; Zhang, H.; Zhang, L. C.; Shen, C.; Lee, S. W. et al. (2017): Extraction of dynamic functional connectivity from brain grey matter and white matter for MCI classification. *Human Brain Mapping*, vol. 38, no. 10, pp. 5019-5034. DOI 10.1002/hbm.23711.

Gauthier, S.; Reisberg, B.; Zaudig, M.; Petersen, R.; Ritchie, K. et al. (2006): Mild cognitive impairment. *Lancet*, vol. 367, no. 9518, pp. 1262-1270. DOI 10.1016/S0140-6736(06)68542-5.

Hutchison, R. M.; Womelsdorf, T.; Gati, J. S.; Everling, S.; Menon, R. S. (2013): Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques. *Human Brain Mapping*, vol. 34, no. 9, pp. 2154-2177. DOI 10.1002/hbm.22058.

Jiang, X. W.; Zhang, Y. D. (2019): Chinese sign language finger spelling recognition via six-layer convolutional neural network with leaky rectified linear units for therapy and rehabilitation. *Journal of Medical Imaging and Health Informatics*, vol. 9, no. 9, pp. 2031-2038. DOI 10.1166/jmihi.2019.2804.

Jiao, Z. Q.; Ma, K.; Wang, H.; Zou, L.; Zhang, Y. D. (2018): Research on node properties of resting-state brain functional networks by using node activity and ALFF. *Multimedia Tools and Applications*, vol. 77, no. 17, pp. 22689-22704. DOI 10.1007/s11042-017-5163-2.

Jiao, Z. Q.; Wang, H.; Ma, K. (2016): The connectivity measurement in complex directed networks by motif structure. *International Journal of Sensor Networks*, vol. 21, no. 3, pp. 197-204. DOI 10.1504/IJSNET.2016.078374.

Jiao, Z. Q.; Wang, H.; Ma, K.; Zou, L.; Xiang, J. B. et al. (2017): Effective connectivity in the default network using Granger causal analysis. *Journal of Medical Imaging and Health Informatics*, vol. 7, no. 2, pp. 407-415. DOI 10.1166/jmihi.2017.2029.

Jiao, Z. Q.; Xia, Z. W.; Cai, M.; Zou, L.; Xiang, J. B. et al. (2018): Hub recognition for brain functional networks by using multiple-feature combination. *Computers & Electrical Engineering*, vol. 69, pp. 740-745. DOI 10.1016/j.compeleceng.2018.01.010.

Jiao, Z. Q.; Xia, Z. W.; Ming, X. L.; Cheng, C.; Wang, S. H. (2019): Multi-scale feature combination of brain functional network for eMCI classification. *IEEE Access*, vol. 7, no. 1, pp. 74263-74273. DOI 10.1109/ACCESS.2019.2920978.

Jiao, Z. Q.; Zou, L.; Cao, Y.; Qian, N.; Ma, Z. H. (2014): Effective connectivity analysis of fMRI data based on network motifs. *Journal of Supercomputing*, vol. 67, no. 1, pp. 74263-74273. DOI 10.1007/s11227-013-1010-z.

Jie, B.; Shen, D. G.; Zhang, D. Q. (2014): Brain connectivity hyper-network for MCI classification. *Medical Image Computing and Computer-Assisted Intervention*, vol. 17, no. 2, pp. 724-732.

Khambhati, A. N.; Sizemore, A. E.; Betzel, R. F.; Bassett, D. S. (2018): Modeling and interpreting mesoscale network dynamics. *NeuroImage*, vol. 180, no. 8, pp. 337-349. DOI 10.1016/j.neuroimage.2017.06.029.

Lee, D. D.; Seung, H. S. (1999): Learning the parts of objects by non-negative matrix factorization. *Nature*, vol. 401, no. 6755, pp. 788-791. DOI 10.1038/44565.

Lee, K.; Tak, S.; Ye, J. C. (2011): A data-driven sparse GLM for fMRI analysis using sparse dictionary learning with MDL criterion. *IEEE Transactions on Medical Imaging*, vol. 30, no. 5, pp. 1076-1089.

Li, H. J.; Bu, Z.; Li, Y. L.; Zhang, Z. Y.; Chu, Y. C. et al. (2018): Evolving the attribute flow for dynamical clustering in signed networks. *Chaos, Solitons & Fractals*, vol. 110, pp. 20-27. DOI 10.1016/j.chaos.2018.02.009.

Li, H. J.; Li, H. Y. (2016): Scalably revealing the dynamics of soft community structure in complex networks. *Journal of Systems Science and Complexity*, vol. 29, no. 4, pp. 1071-1088. DOI 10.1007/s11424-015-4145-6.

Li, X.; Zhu, D. J.; Jiang, X.; Jin, C. F.; Zhang, X. et al. (2014): Dynamic functional connectomics signatures for characterization and differentiation of PTSD patients. *Human Brain Mapping*, vol. 35, no. 4, pp. 1761-1778. DOI 10.1002/hbm.22290.

Liang, M.; Zhou, Y.; Jiang, T. Z.; Liu, Z. N.; Tian, L. X. et al. (2006): Widespread functional disconnectivity in schizophrenia with resting-state functional magnetic resonance imaging. *NeuroReport*, vol. 17, no. 2, pp. 209-213. DOI 10.1097/01.wnr.0000198434.06518.b8.

Lu, S. Y.; Lu, Z. H.; Zhang, Y. D. (2019): Pathological brain detection based on AlexNet and transfer learning. *Journal of Computational Science*, vol. 30, pp. 41-47. DOI 10.1016/j.jocs.2018.11.008.

Ou, J. L.; Xie, L.; Li, X.; Zhu, D. J.; Terry, D. P. et al. (2015): Atomic connectomics signatures for characterization and differentiation of mild cognitive impairment. *Brain Imaging and Behavior*, vol. 9, no. 4, pp. 663-677. DOI 10.1007/s11682-014-9320-1.

Padilla, P.; Górriz, J. M.; Ramírez, J.; Lang, E. M.; Chaves, R. et al. (2010): NMF-based analysis of SPECT brain images for the diagnosis of Alzheimer's disease. *Hybrid Artificial Intelligence Systems: 5th International Conference, HAIS 2010, San Sebastián, Spain, June 23-25, 2010. Proceedings, Springer-Verlag*, Part I, pp. 468-475. DOI 10.1007/978-3-642-13769-3_57.

Pascual-Montano, A.; Carmona-Saez, P.; Chagoyen, M.; Tirado, F.; Carazo, J. M. et al. (2006): bioNMF: a versatile tool for non-negative matrix factorization in biology. *BMC Bioinformatics*, vol. 7, no. 1, pp. 366. DOI 10.1186/1471-2105-7-366.

Raichle, M. E.; MacLeod, A. M.; Snyder, A. Z.; Powers, W. J.; Gusnard, D. A. et al. (2001): A default mode of brain function. *Proceedings of the National Academy of Sciences*, vol. 98, no. 2, pp. 676-682. DOI 10.1073/pnas.98.2.676.

Shahnaz, F.; Berry, M. W.; Pauca, V. P.; Plemmons, R. J. (2006): Document clustering using nonnegative matrix factorization. *Information Processing & Management*, vol. 42, no. 2, pp. 373-386. DOI 10.1016/j.ipm.2004.11.005.

Stam, C. J. (2010): Characterization of anatomical and functional connectivity in the brain: a complex networks perspective. *International Journal of Psychophysiology*, vol. 77, no. 3, pp. 186-194. DOI 10.1016/j.ijpsycho.2010.06.024.

Stam, C. J.; Reijneveld, J. C. (2007): Graph theoretical analysis of complex networks in the brain. *Nonlinear Biomedical Physics*, vol. 1, no. 1, pp. 3. DOI 10.1186/1753-4631-1-3.

Sun, J. D.; Zhang, Y. D. (2019): High performance multiple sclerosis classification by data augmentation and AlexNet transfer learning model. *Journal of Medical Imaging and Health Informatics*, vol. 9, no. 9, pp. 2012-2021. DOI 10.1166/jmihi.2019.2692.

Tobia, M. J.; Hayashi, K.; Ballard, G.; Gotlib, I. H.; Waugh, C. E. (2017): Dynamic functional connectivity and individual differences in emotions during social stress. *Human Brain Mapping*, vol. 38, no. 12, pp. 6185-6205. DOI 10.1002/hbm.23821.

Tzourio-Mazoyer, N.; Landeau, B.; Papathanassiou, D.; Crivello, F.; Etard, O. et al. (2002): Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, vol. 15, no. 1, pp. 273-289. DOI 10.1006/nimg.2001.0978.

Wang, S. H.; Du, S. D.; Atangana, A.; Liu, A. J.; Lu, Z. Y. (2018): Application of stationary wavelet entropy in pathological brain detection. *Multimedia Tools and Applications*, vol. 77, no. 3, pp. 3701-3714. DOI 10.1007/s11042-016-3401-7.

Wang, X.; Ren, Y. S.; Zhang, W. S. (2017): Multi-task fused Lasso method for constructing dynamic functional brain network of resting-state fMRI. *Journal of Image and Graphics*, vol. 22, pp. 978-987.

Wang, K. C.; Wu, G. B.; Hou, X.; Wei, D. T.; Liu, H. S. et al. (2016): Segmentation and application of functional network from group to individual. *Science Bulletin*, vol. 61, pp. 3022-3035.

Wee, C. Y.; Yap, P. T.; Zhang, D. Q.; Denny, K.; Browndyke, J. N. et al. (2012): Identification of MCI individuals using structural and functional connectivity networks. *NeuroImage*, vol. 59, no. 3, pp. 2045-2056. DOI 10.1016/j.neuroimage.2011.10.015.

Yu, X.; Zeng, N. Y.; Liu, S.; Zhang, Y. D. (2019): Utilization of DenseNet201 for diagnosis of breast abnormality. *Machine Vision and Applications*, vol. 30, no. 7-8, pp. 1135-1144. DOI 10.1007/s00138-019-01042-8.

Zhang, Y.; Dong, Z.; Liu, A.; Wang, S.; Ji, G. et al. (2015): Magnetic resonance brain image classification via stationary Wavelet transform and generalized eigenvalue proximal support vector machine. *Journal of Medical Imaging and Health Informatics*, vol. 5, no. 7, pp. 1395-1403. DOI 10.1166/jmihi.2015.1542.

Zhang, Q.; Li, B. X. (2010). Discriminative K-SVD for dictionary learning in face recognition. *The Twenty-Third IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2010, San Francisco, CA, USA, June 13-18, 2010. Proceedings, IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 2691-2698. DOI 10.1109/CVPR.2010.5539989.

Zhang, S.; Li, X.; Lv, J. L.; Jiang, X.; Zhu, D. J. et al. (2013): Sparse representation of higher-order functional interaction patterns in task-based fMRI data. *Medical Image Computing and Computer-Assisted Intervention*, vol. 16, no. 3, pp. 626-634.

Zhang, Y. D.; Wang, S. H.; Sui, Y. X. (2018): Multivariate approach for Alzheimer's disease detection using stationary wavelet entropy and predator-prey particle swarm optimization. *Journal of Alzheimer's Disease*, vol. 65, no. 3, pp. 855-869. DOI 10.3233/JAD-170069.

Zhang, Y.; Zhang, H.; Chen, X. B.; Lee, S. W.; Shen, D. G. (2017): Hybrid high-order functional connectivity networks using resting-state functional MRI for mild cognitive impairment diagnosis. *Scientific Reports*, vol. 7, no. 1, pp. 6530. DOI 10.1038/s41598-017-06509-0.

Zhou, L. P.; Wang, Y. P.; Li, Y.; Yap, P. T.; Shen, D. G. (2011): Hierarchical anatomical brain networks for MCI prediction: revisiting volumetric measures. *PLoS One*, vol. 6, no. 7, pp. e21935. DOI 10.1371/journal.pone.0021935.