Computer Methods and Programs in Biomedicine 201 (2021) 105954

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

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

Deep sparse graph functional connectivity analysis in AD patients using fMRI data

Hessam Ahmadi[a], Emad Fatemizadeh[b,*], Ali Motie-Nasrabadi[c]

[a] Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran
[b] School of Electrical Engineering, Sharif University of Technology, Tehran, Iran
[c] Biomedical Engineering Department, Shahed University, Tehran, Iran

ARTICLE INFO

Article history:
Received 13 May 2020
Accepted 22 January 2021

Keywords:
Functional connectivity
Graph sparsification
Thresholding
Autoencoders
Spectral sparsification

ABSTRACT

Functional magnetic resonance imaging (fMRI) is a non-invasive method that helps to analyze brain function based on BOLD signal fluctuations. Functional Connectivity (FC) catches the transient relationship between various brain regions usually measured by correlation analysis. The elements of the correlation matrix are between -1 to 1. Some of them are very small values usually related to weak and spurious correlations due to noises and artifacts. They can not be concluded as real strong correlations between brain regions and their existence could make a misconception and leads to fake results. It is crucial to make a conclusion based on reliable and informative correlations. In order to eliminate weak correlations, thresholding is a common method. In this routine, by adjusting a threshold the values below the threshold turn to zero and the rest remains. In this paper, in addition to thresholding, two other methods including spectral sparsification based on Effective Resistance (ER) and autoencoders are investigated for sparsing the correlation matrices. Autoencoders are based on deep learning neural networks and ER considers the network as a resistive circuit. The fMRI data of the study correspond to Alzheimer's patients and control subjects. Graph global measures are calculated and a non-parametric permutation test is reported. Results show that the autoencoder and spectral sparsification achieved more distinctive brain graphs between healthy and AD subjects. Also, more graph global features were significantly different from these two methods due to better elimination of weak correlations and preserve more informative ones. Regardless of the sparsification method features including average strength, clustering, local efficiency, modularity, and transitivity are significantly different (P-value=0.05). On the other hand, the measures radius, diameter, and eccentricity showed no significant differences in none of the methods. In addition, according to three different methods, the brain regions show fragile and solid FCs are determined.

1. Introduction

Alzheimer's disease (AD) is a dynamic illness that causes brain cells to degenerate and perish. An individual with AD will create extreme memory debilitation and lose the capacity to complete regular assignments [1]. Although AD drugs may briefly improve manifestations or moderate the disease progression there is no specific treatment that cures AD or decreases the sickness procedure in the brain [2]. The prevalence of AD is very challenging. since 2000, dying from AD have risen by 89 percent while those from coronary illness have diminished [3]. According to the mortality rate (sixth-leading cause of death in the US [4]) and severeness of the AD, early and precise diagnosis of AD is very crucial.

Diagnosis of AD from structural imaging such as Magnetic Resonance Imaging (MRI) is common in literature. In a recent study [5] by use of Shearlet Transform (ST) feature extraction technique and K-Nearest Neighbor (KNN) classifier, the accuracy of 98.48% is provided. Another routine to investigate AD is Diffusion Tensor Imaging (DTI). In [6] by use of DTI data, the different stages of AD were classified by employing a deep learning algorithm called Convolutional Neural Networks (CNN), and the accuracy of 92.6% is achieved for four classes. Using deep learning improves the results effectively. Duc et al. [7] proposed functional 3-dimensional (3-D) independent component spatial maps for use as features and deep CNN as a classifier to distinguish AD subjects from healthy ones. Deep learning methods also have been used to analyze other biomedical signals such as Electrocardiograms (ECG) [8,9]. The ac-

* Corresponding author.
E-mail addresses: hessam.ahmadi@srbiau.ac.ir (H. Ahmadi), fatemizadeh@sharif.edu (E. Fatemizadeh), nasrabadi@shahed.ac.ir (A. Motie-Nasrabadi).

https://doi.org/10.1016/j.cmpb.2021.105954
0169-2607/

curacy in both mentioned studies reaches more than 98%. It is worthwhile mentioning that the use of deep learning is not limited to medical applications and is very prevalent in other topics such as [10].

A very reliable and prevalent method to investigate AD is the functional MRI (fMRI) technique. In the fMRI session, brain activity measures based on changes related to blood flow [11]. At brain regions where neurons are activated, the bloodstream to that area likewise increases. In ordinary fMRI, brain activity is estimated based on low-frequency signals called Blood-Oxygen-Level-Dependant (BOLD) [12]. A common application is to investigate the dynamics of fMRI signals which is done in a recent study and the accuracy of 88.52% is provided [13]. Another main application of fMRI in brain research is to calculate and analysis of brain connectivities (anatomical, functional, effective). Based on connectivities, the hypothesis is that healthy brains show a pattern for connectivities, and for other people who suffer from mental and brain diseases, their brain often shows disruption in the connectivities. This routine can understand the abnormality and it is useful for early detection of illnesses such as AD which started several years before the clinical signs reveal [14]. Functional Connectivity (FC) relates the dependencies of different regions in the brain which are not connected structurally but are functionally connected. It tends to be characterized as the temporal relationships between's spatially remote neurophysiological occasions [15]. Commonly, FC is calculated among the fMRI time series from different ROIs or Voxels [16].

A brain graph theory system is a numerical portrayal of the genuine brain design that comprises of many nodes and connections (edges) intervened between them. Nodes normally speak to brain areas, while connections speak to anatomical, functional, or effective extracts [17]. According to correlation analysis, the edges of the graphs show the value of the correlation which is between −1 to 1. The sign of the correlation shows the direction and the value shows the strength of the connectivity between two different brain regions. It should be noted that small values could happen also because of recording device or physiological and experimental noise and artifacts [18]. It is common to use a threshold in order to sparse the graphs and remains the higher correlations which implying strong FC's, in this manner making it computationally increasingly tractable [19,20], In many studies [21-26] the analysis are reported with several thresholds. Although this is a common approach, it makes it difficult to compare the results of different studies. Another drawback is that there is no specific physiological meaning for thresholding. Despite all the issues, for brain graph sparsification, thresholding is a popular procedure [27]. Also, another approach for thresholding is the fixed density that jams the objective density of links [28]. Bordier et al. [20] proposed a method established in statistical material science, to distinguish a sparsification threshold that augments data on the graph modular structure in order to overcome the problem of calculating the optimal threshold. Luo et al. [29] were employed fMRI data of Chinese people to investigate brain functional connectivity modification to explore AD. For sparsification step thresholding was used and the value is characterized as the proportion of real links number to the most extreme potential links. Also, ar recent study [30], investigates the effect of weighted or binary graphs in fMRI of AD patients. Another recent study [31] proposed a computationally productive technique to find the optimal threshold based on the projection of the data to its highest covariance and variance state.

Sparsification is a fundamental step in neuroimaging data connectivity analysis. Although the most frequent method for sparsification is thresholding, has its own drawbacks (as mentioned above) and there is no robust and gold standard algorithm to select the optimal threshold. Accordingly, in this paper, to overcome the limitation of thresholding two other methods are presented. Most of the studies are employed deep learning algorithms for classification but in this research, a deep learning method is employed to sparse the connectivity matrices called sparse autoencoders. Also, to perform a comprehensive investigation, another approach for graph sparsification is proposed called Effective Resistance (ER).

The paper is organized as follows: In the Material and Methods, the database and preprocessing are explained. Secondly, the graph theory and features are presented and then the ER method and autoencoders are elaborated in this section. Next, the outcomes are reported in the Result section. Then, the results are interpreted in the Discussion part and finally, the conclusion section is presented.

Table 1
Demographic information of the data.

Group	No. (Male/Female)	Age	CDR score	MMSE score
AD	(15/13)	75.9 ± 4.8	1.03±0.54	21.58±3.16
Healthy	(14/14)	74.8 ± 5.7	0.16±0.05	28.94±2.51

2. Material and methods

Analyzing the effect of different methods for graph sparsification is the main idea of this research, accordingly in this section, the fMRI data and the preprocessing method are presented. Also, the sparsification methods and brain graph generating are explained.

2.1. Data and preprocessing

One of the publicly available datasets of AD is the Alzheimer's Disease Neuroimaging Initiative (ADNI) which started collecting data since 2004 [32]. The data of this study is from ADNI2 which started in 2011 and last for 5 years [33]. The data for this research consists of 2 groups (AD and healthy) and each of them has 28 subjects. All the scans are taken by a 3 tesla Philips device and based on the ADNI protocol [34]. The parameters are: TR/TE 3000/30 msec, flip angle = 80, 3.3125 mm slice thickness, 48 slices, and the functional volumes were 140. In Table 1, the information of the data is shown. The CDR and MMSE scores in Table 1 are related to Clinical Dementia Rating and Mini-Mental State Exam respectively. They are used for AD and dementia clinical testing.

The Data Processing Assistant for rs-fMRI (DPARSF) toolbox is used for preprocessing [35]. The functions of this toolbox were based on SPM and REST toolboxes. The slice timing correction was performed utilizing the last cut as the reference. According to longitude fMRI recording, correcting head motion is necessary. Thus, rigid body registration is performed. To be able to compare among different subjects, scans are standardized to a reference brain because the anatomy of the brain is somewhat unique in the individuals. In this way, the subjects were standardized to Montreal Neurological Institute (MNI) atlas. By utilizing a Gaussian kernel with FWHM= 4 mm, the fMRI data was smoothed. As to hold the low-recurrence motions, the data were filtered by a band-pass filter (0.01–0.08 Hz) and detrended. The Automated Anatomical Labeling (AAL) atlas was utilized to segment the data by standard ROIs [36]. The AAL parcellates the brain into 116 specific locales.

2.2. Graph theory and features

Complex networks such as brain networks can be modeled with the aid of graph theory. Voxels or ROIs of the brain correspond to nodes and the connectivity between regions refers to links between two ROIs or voxels. In this research, an ROI based (116 ROI according to the AAL atlas) approach is used. The signal of each ROI was gained by averaging of all voxels inside that ROI. Edges

Table 2
Graph measures and formula.

Graph Measure	Formula
Degree	–
Eccentricity	$Ecc = \max\{d_G(x,y)\}$
Strength	$Str(\upsilon) = \sum_{\upsilon \in V} E$
Radius	$R = \min\{ECC\}$
Diameter	$D = \max\{ECC\}$
Characteristic Path Length	$L = \frac{\sum_{x,y \in V,(G)} d_G(x,y)}{n(n-1)}$
Global Efficiency	$E_{glob}(G) = \frac{E(G)}{E(G^{Ideal})}$
Local Efficiency	$E_{loc}(G) = \frac{1}{n}\sum_{x \in G} E(G_x)$
Clustering	$C = \frac{\text{Number of closed triplets}}{\text{number of all triplets}}$
Modularity	$M = \frac{1}{l}\sum_{x,y}[A_{x,y} - \frac{k_x k_y}{l}]\delta_{x,y}$
Transitivity	$T = \frac{3*\text{number of triangles}}{\text{number of connected triples of nodes}}$
Small-Worldness	$\sigma = \frac{C/C_r}{L/L_r}$

of the brain graph were computed using the Pearson correlation analysis and finally, the weighted undirected brain graphs are constructed. Graph measures can be computed over the FC matrices that show the integration and segregation of the brain and based on the features can conclude the effect of different diseases on the FCs in the brain. The features and their formulas are listed in Table 2.

In Table 2, the graph is shown by $G = (V, E)$where Gcorresponds to a graph in which Erefers to links and Vrelated to the nodes. The distance between the x and y nodes was shown by $d_G(x,y)$. It also $n$refers to the number of vertices. $A_{x,y}$ refers to the connectivity matrix, and l related to the number of links. If the two hubs are from one network, the $\delta_{x,y}$is 0 and finally the C_r and L_r relate to an identical random graph. The average efficiency of a graph is $E(G) = \frac{1}{n(n-1)}\sum_{x \neq y \in G}\frac{1}{p(x,y)}$ where $p(x, y)$is the shortest path length between $x$and y. Furthermore, the G^{Ideal} relates to a diagram of vertices wherein every single imaginable edge is available.

2.3. Autoencoders

Autoencoders are commonly based neural networks which often used to code the input data efficiently. In the first part which is called encoder, the network, codes the input data (outputs of the first hidden layer are the code), and in the decoder part, the network extracts the input again from the code [30]. Consider the following equations:

$$h = f(x)encoder$$

$$r = g(h)decoder$$

$$g(f(x)) = \widehat{x}$$

The equation $h = f(x)$ can rewrite as $h = \sigma(Wx + b)$where $W$is the weight matrix, σ an activation function, and b represents the bias vector.

During the training phase the reconstruction error which evaluates the differences between input data and the output is minimized:

$$L(x, g(f(x)))$$

Where L is a cost function that trying to minimize the difference between $g(f(x))$ and x through functions such as mean squared error. A schematic of an autoencoder is depicted in Fig. 1. X is the input layer. Z is the result of the encoder section that codes the input (coding layer). X' is the output of the decoder part and tries to represent the input.

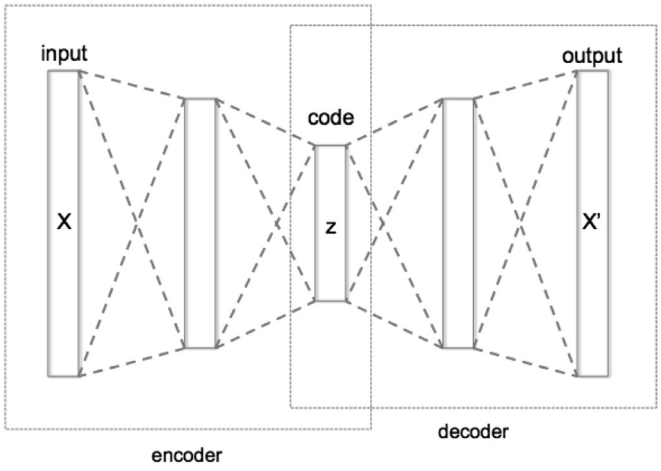

Fig. 1. A schematic of an autoencoder. The encoder and decoder are illustrated separately [37].

One of the main applications of autoencoders is to sparse the input data [38]. In this case, a term such $\Omega(h)$ is added to the cost function.:

$$L(x, g(f(x))) + \Omega(h)$$

The sparsity term forces the output of the encoder to be more sparse in comparison to input data. For optimization, Kullback-Leibler (KL) divergence is used [39]. According to it:

$$\widehat{\rho}_j = \frac{1}{m}\sum_{i=1}^{m}\left[h_j(x_i)\right]$$

The above equation computes the activation over the hidden layer j for the input x_i and averages it for m inputs. In fact, by obliging the neuron's activation over a gathering of data, the neurons urging to fire only for several inputs and it means that although the network has a large structure most of the neurons are inactive. In this criterion whatever the $\widehat{\rho}$ is close to zero, it is better, so according to the method the constraint has been compelled to $\widehat{\rho} = \rho$where ρ is the variable for sparsity. As the KL-divergence calculates probability distributions variation, ρ can be explained as Bernoulli random variable distribution and it can be written as follows:

$$\beta\sum_{j=1}^{s}\left[\rho\log\frac{\rho}{\widehat{\rho}_j} + (1-\rho)\log\frac{1-\rho}{1-\widehat{\rho}_j}\right]$$

Where s corresponds to the hidden layer neurons, j implying to hidden units and β is a parameter that controls the sparsity. In the above optimization $\widehat{\rho}$ is going to be penalized if it differs from ρ remarkably. This function is zero if $\widehat{\rho} = \rho$ and is going to increase based on the differences between $\widehat{\rho}$ and ρ.

If there is more than one hidden layer (deep network), the network is going to be stacked autoencoder. In the stacked networks the training phase is a bit different from ordinary Multi-Layer Perceptron (MLP). Firstly the network has been trained with one hidden layer after that the second one is going to be trained and the first hidden layer takes the input layer role and finally the hidden layers merged and make the stacked autoencoder. In this paper, a stacked autoencoder has been used with two hidden layers [40].

After trying different specifications, the characteristics of the stacked autoencoder in this research were set to four layers including two layers for the encoder and two other layers for the decoder part. The parameters of the autoencoder are exhibited in Table 3. 70% of the data were used for training, 10% for validation, and the rest for testing randomly.

Table 3
The autoencoder specifications and parameters.

Number of layers	4
Number of neurons	150 and 116
Maximum epoch	200
Initial weights	Randomly selected
Transfer function	Logistic sigmoid function
Regularizer terms	0.002 and 0.004
Sparsity term	4

2.4. Spectral sparsifier and effective resistance

The graph G can be presented by the following relation:

$$G = (V, E, w)$$

Where, *V, E, w*are the vertices, links, and weights, respectively. The graph was assumed as a resistor network (an electrical circuit) so that, every link $e \in E$between $(i, j)$is related to a resistor,$r_{i,j} = 1/w$ ohm, and the energy of graph G is obtained as follows:

$$\varepsilon_G(x) = \sum_{i,j \in E} (x_i - x_j)^2.$$

The ER of a link is characterized as the potential distinction that must be applied to i and j to drive one unit of current through the graph (potentials $x : V \to \mathbb{R}$). Then. ER $R_G(e)$ is measured as follows:

$$R_G(e) = {x_{i,j}}^T \ell_G x_{i,j}$$

Where ℓ_G is the Moore-Penrose pseudo-inverse of the Laplacian matrix, that is equal to -1 at the node j, and is equal to 1 at the node i, and also is equal to 0 for the rest. Also, $x_i$is the magnitude of $x$for vertex$i$.

Spielman and Teng [41] have introduced the spectral similarity that has an important role in solving Laplacian systems. As we know the Laplacian form of G is:

$$x^T L_G x = \sum_{i,j \in E} w_{i,j} (x_i - x_j)^2$$

A $(1 \pm \varepsilon)$-spectral sparsifier H approximates G if for every subset $S \subset V$, the sum of weights for links leaving S is protected. In other words, H and G are spectrally similar if:

$$(1 - \varepsilon) x^T L_H x \le x^T L_G x \le (1 + \varepsilon) x^T L_H x$$

According to Loewner ordering on matrices:

$(1 - \varepsilon)L_H \le L_G \le (1 + \varepsilon)L_H$, $x^T A x \ge 0$ if A is positive semi-definite for all $x \in \mathbb{R}^n$Networks that are spectrally similar offer numerous arithmetical properties. The ER between all sets of vertices is comparable in spectrally similar networks. Spielman and Srivastava [42] have proved that every graph has a spectral sparsifier that can be obtained by the ER method through a random process. They have prepared the samples from edges of graph G proportional to their ER [43]. In the present paper, in addition to the autoencoder method, the ER procedure was used based on [42] the algorithm to provide a sparse representation for correlation matrices.

Accordingly, the flowchart of the research is shown in Fig. 2.

3. Results

The fMRI data were preprocessed and the time-series were extracted based on the AAL atlas. In the next step, the Pearson correlation coefficients of the data were calculated to obtain the FC among the brain ROIs and to generate the weighted undirected connectivity graphs. Then, the matrices were sparsed by thresholding, autoencoders, and ER methods that were fully explained in the Material and Method section. It is worthwhile mentioning that all

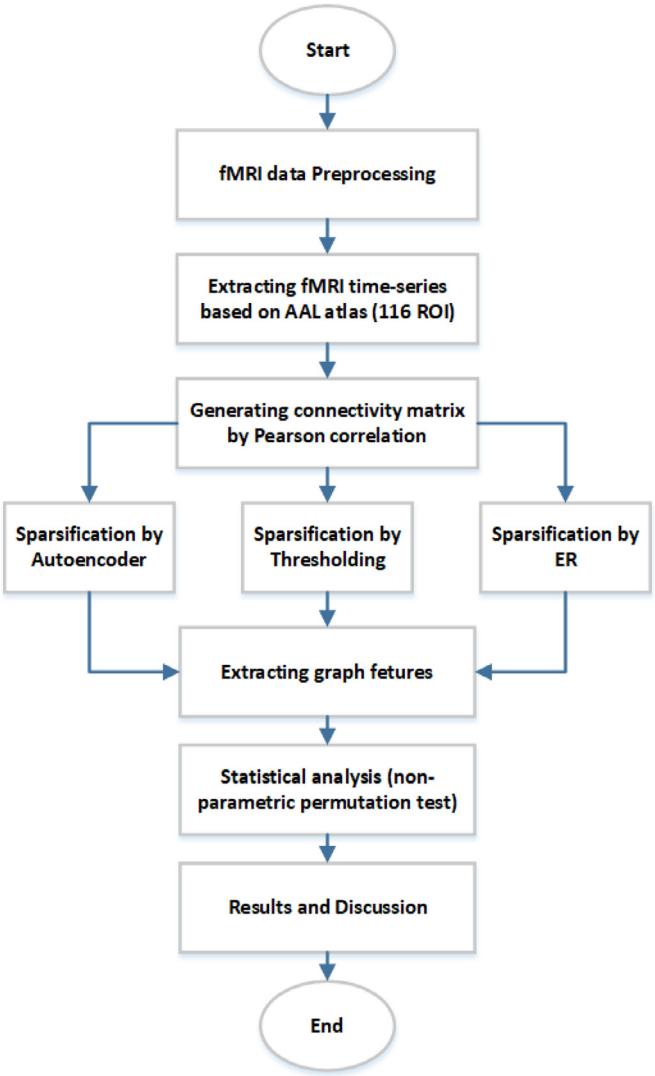

Fig. 2. The overview of the study.

the processing was done in Matlab2018a and the hardware specifications are as follows: CPU: Intel Core i7 - 8550 U, RAM: 16 GB DDR 4. Fig. 3 shows the connectivity matrix of a subject.

As can be seen in Fig. 3, all the elements of the matrix have a value that some of them are weak and may be spurious because of experimental noise. The weak correlations should be removed to obtain a matrix with strong FCs. For this purpose, thresholding can be used. As shown in Fig. 4, the above matrix has been thresholded by 0.5. It is worthwhile to mention that, as the goal was analyzing the functional connectivity and the direction of the correlations was not important in this analysis, so the absolute value of the elements was considered before thresholding.

The matrix shown in Fig. 4 has more sparsity than that in Fig. 3. This sparsity makes further computations very fast and also the researcher is sure about the results because they are based on the strong FCs and the weaker links have no role in the results but finding an optimal threshold is a problem for this method. Also, performing the same threshold is not very appropriate in different subjects because as the people got older, the FCs become weaker and the brain loses its segregation [44].

In addition to thresholding, a stacked spars-autoencoder was used to sparse the input matrix (correlation matrix). In the training phase, the autoencoder learns to make a more sparse representation of the input data. Seventy percent of the data was selected

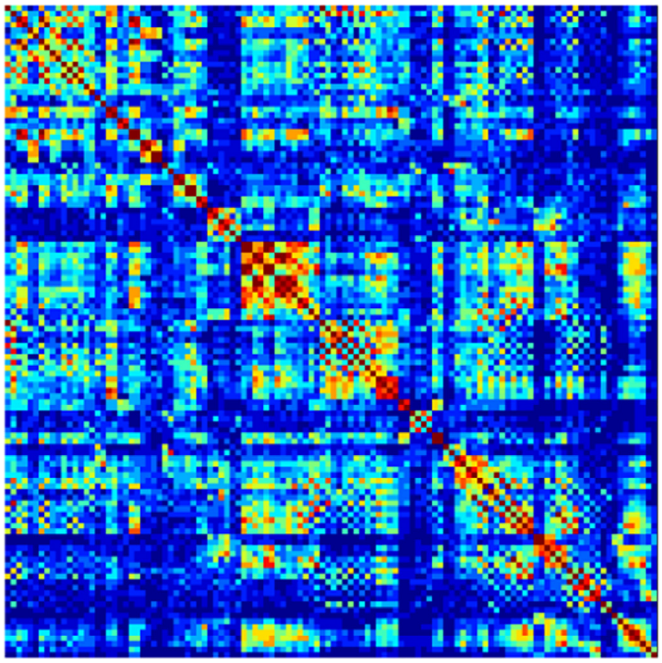

Fig. 3. A 116×116 matrix. The elements are from −1 (Dark Blue) to 1 (Dark red) showing the correlation (FC) between the brain ROIs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

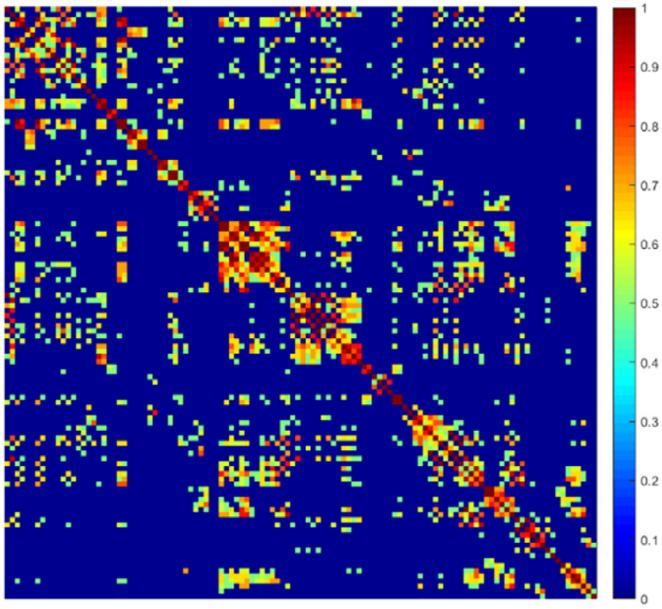

Fig. 4. The thresholded matrix by 0.5. The strong FCs have remained.

randomly and was used for training. As illustrated in Fig. 5, a matrix has been sparsed by a threshold of 0.5 and the autoencoder. The difference between the results is also shown in Fig.5.

The primary matrix had 13,456 elements (116×116). After sparsification, its elements were reduced to 1306 by the thresholding method and 1636 by an autoencoder. Also, the differences between the two methods were 378 elements. Moreover, the same comparison was done between the thresholding, spectral sparsification, and ER algorithm as shown in Fig. 6.

As shown in Figs. 5 and 6, the results of the three sparsification methods are different. The graph measures were computed from the results of sparsification methods for AD and healthy groups to compare them. Besides, the non-parametric permutation test was done to determine which measures are significantly different and are capable to discriminate between AD and healthy groups. Table 4 shows the results of the statistical analysis.

Regardless of the sparsification method, there was a significant difference in the features average strength, clustering, local efficiency, modularity, and transitivity (P-value=0.05). It can be concluded that these features are highly influenced by the AD; therefore they are more appropriate features for FC analysis in the AD. There were no significant differences in the features such as radius, diameter, and eccentricity (that were calculated based on radius and diameter).

The two features including average degree and global efficiency showed no significant differences using the thresholding method but they were significantly different using the other methods showing that the autoencoder and spectral sparsification methods are more capable to distinguish between AD and healthy subjects in comparison with simple thresholding.

Comparing between the autoencoder and spectral sparsification method revealed that, the Characteristic Path Length (CPL) and small-worldness features showed significant differences in the spectral sparsification method, so this method is more reliable and appropriate for sparsification in FC analysis of AD patients.

Moreover, when the P-value was considered as 0.01, only two features including clustering and transitivity showed significant differences in the thresholding method while the average strength, clustering, transitivity, and modularity showed significant differences in the autoencoder method and the average strength, global and local efficiencies, clustering, and transitivity also showed significant differences in the spectral sparsification method confirming the above-mentioned results again.

For a better visual display, Fig. 7 shows the binary brain graphs generated by different sparsification methods.

As depicted in Fig. 6, the graphs generated by spectral sparsification and autoencoder methods are more capable to distinguish between a patient with AD and a healthy subject in comparison with the thresholding method.

After sparsification, several ROIs had the most changes (lost their connections) and also several ROIs were intact approximately and had kept their connections. The numbers of these ROIs are listed in Table 4 based on the AAL atlas and each sparsification method. It is worthwhile to mention that the sparsed ROIs listed in Table 5 have lost more than 90% of their connections and the intact ROIs have kept more than 90% of their connections after the sparsification process.

According to Table 5, the ROIs that were common in all three methods were highlighted. It can be concluded that these ROIs have kept or lost their connections regardless of the sparsification methods. In the AD group, the ROIs 28 and 41 lost more than 90% of their edges while, in the control group, these ROIs were not sparsed like the AD group. It can be interpreted that, AD affects the functional connectivity in ROIs 28 and 41. Fig. 8 shows the highlighted ROIs.

4. Discussion

The current study was designed to analyze the effect of the sparsification method on distinguishing between AD patients and healthy subjects. For this purpose, after preprocessing the fMRI data and extracting the time-series, the Pearsons Correlation Coefficients between different brain ROIs were calculated and the weighted undirected connectivity graphs were generated. Next, the connectivity matrices were sparsed using three different methods including thresholding (0.5), autoencoder, and spectral sparsification. It is worthwhile to mention that the sparsification step is needed due to the existence of weak and spurious edges related to

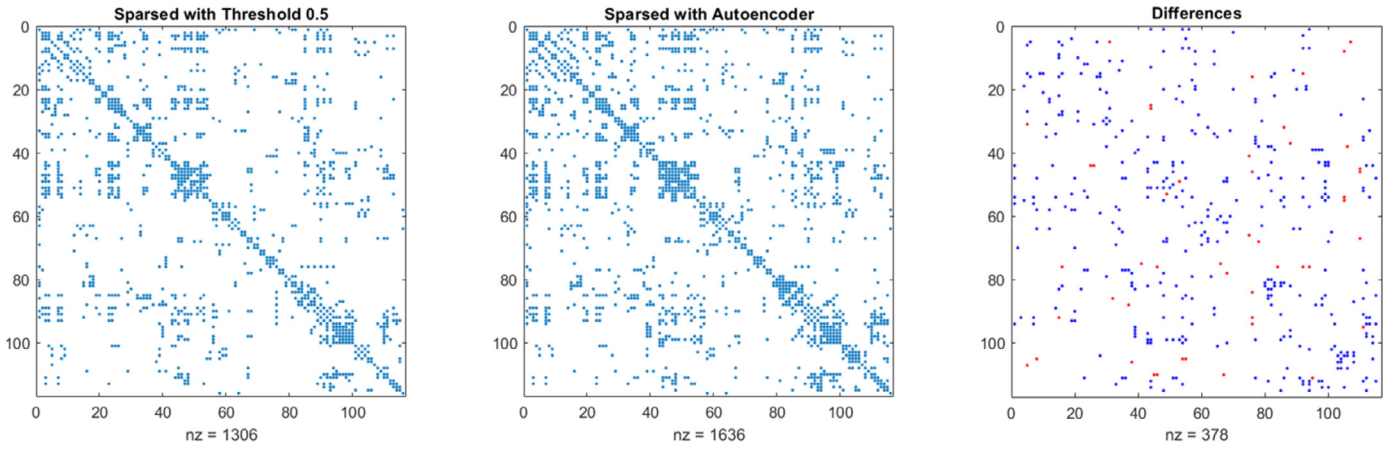

Fig. 5. A correlation matrix was sparsed by thresholding and autoencoder. Also, the results were binarized to facilitate the comparison. The 'nz' below each diagram corresponds to non-zero elements. The 'red' circles in the 'Differences' chart indicate the elements that are available in thresholding but not in the autoencoder method and conversely, the 'blue' circles show the elements that are available in the autoencoder method but not in thresholding.

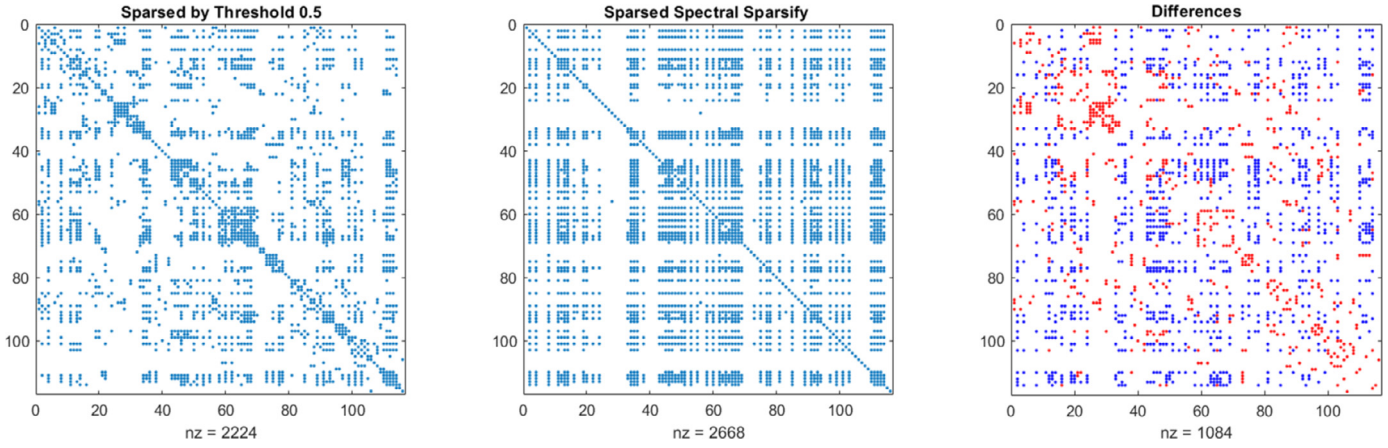

Fig. 6. A correlation matrix was sparsed by thresholding and spectral sparsification. The Results were binarized and 'nz' is related to non-zero elements. The red and blue circles show similar information to Fig. 5.

Table 4
The P-values of non-parametric permutation test between healthy and AD subjects in three different sparsification methods. The P-values of less than 0.05 are highlighted.

Measures	Thresholding (0.5)	Autoencoder	Spectral Sparsification
Av. Degree	0.0666	**0.0032**	**0.0167**
Av. Strength	**0.0170**	**0.0076**	**0.0040**
Radius	0.3875	0.0604	0.3423
Diameter	0.2014	0.2474	0.2009
Eccentricity	0.4013	0.0612	0.3024
Char. Path Length	0.1029	0.4109	**0.0430**
Global Efficiency	0.0900	**0.0136**	**0.0082**
Local Efficiency	**0.0369**	**0.0480**	**0.0023**
Clustering	**0.0028**	**0.0027**	**0.0007**
Transitivity	**0.0024**	**0.0022**	**0.0024**
Modularity	**0.0183**	**0.0038**	**0.0387**
Small-Worldness	0.1397	0.3291	**0.0367**

the noises, and as it makes other processing steps more tractable. According to graph global measures and non-parametric permutation test, the features including average strength, clustering, local efficiency, modularity, and transitivity were significantly different (P-value=0.05) in all three sparsification methods. It can be interpreted that these features are more appropriate for FC analysis in AD patients because they can discriminate the AD patients from the healthy controls regardless of the sparsification method. Moreover, the two features including global efficiency and average degree were significantly different in spectral sparsification and autoencoder methods showing that these methods are more powerful tools to discriminate the AD subjects from healthy ones in comparison with simple thresholding. Among these two methods, CPL and small-worldness measures also showed significant differences in the spectral sparsification. As shown in the brain graphs, the FCs in spectral sparsification is more different between AD and healthy subjects, which leads to making more features that are significantly different. As a consequence, the results of the spectral sparsifica-

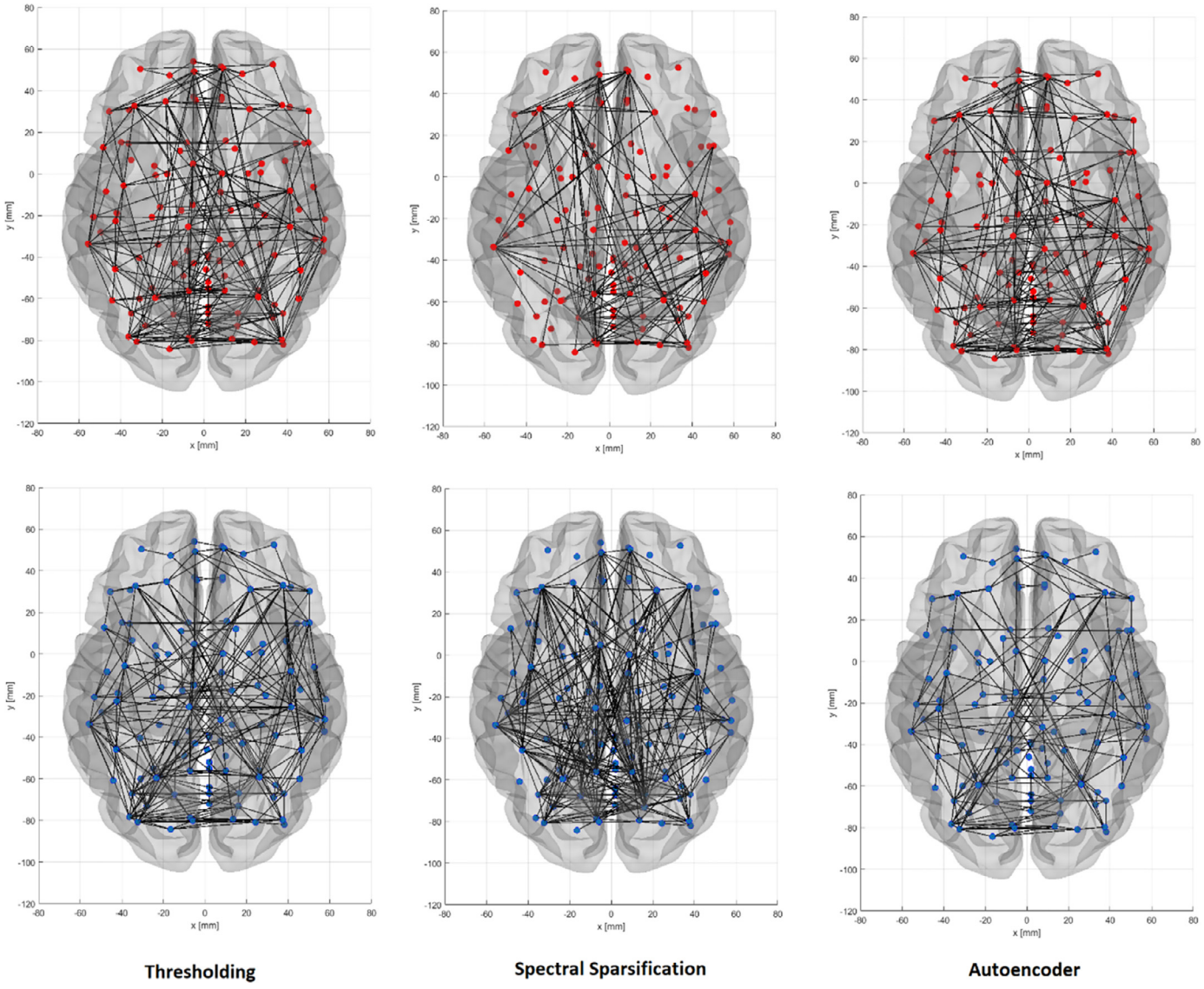

Fig. 7. Binary graphs generated from different sparsification methods. The above figures (with red nodes) are related to the AD subject and the below figures (with blue nodes) correspond to the healthy subject.

Table 5
The ROI numbers that had the highest lost and maintained connections based on the AAL atlas. The highlighted numbers are common among the three methods.

AD group						Control group					
Sparsed ROIs			Intact ROIs			Sparsed ROIs			Intact ROIs		
AE	SS	Thr	AE	SS	Thr	AE	SS	Thr	AE	SS	Thr
28	18	28	3	3	23	22	6	22	19	20	19
41	22	41	43	43	43	41	9	42	47	47	47
79	26	75	44	44	44	42	10	95	48	48	48
108	28	79	47	47	46	72	17	107	85	85	85
109	29	108	–	–	47	107	21	109	86	86	86
116	38	109	–	–	–	109	22	116	91	89	91
–	41	116	–	–	–	116	42	–	–	90	–

tion method were the best among the three methods followed by those obtained from the autoencoder and thresholding. Besides, as selecting the optimal threshold has always been a complex issue it should be mentioned that, in the autoencoder and spectral sparsification methods, the procedures did not need any assumption or operator adjustment like selecting an optimal threshold by visualizing the results of different values. On the other hand, the three features including radius, diameter, and eccentricity computed based on radius and diameter showed no significant differences in all the methods, which can be interpreted that these features are not proper for analyzing the FCs in AD patients. In comparison to previous studies, Bordier et al. [20] proposed a data-driven method. Although the algorithm overcomes the limitations it is based on graph modules and does not rely on total edge density but proposed methods in this study are based on the whole network. Another valuable study [31] suggested a spatial sparsification. The idea is, before calculating the connectivity matrices or any other processing, the voxels which demonstrate more BOLD activation take into account. Since the presented approach, sparsify the data before any further processing it is about 4 times faster. The autoencoder and ER approaches are reliable and more effective tools to replace ordinary thresholding and does not affect any other parts of the pipeline. Regardless of the sparsification method, several ROIs were intact after the sparsification process. These ROIs showed powerful FC. Also, several ROIs lost most of their connections in all the sparsification methods showing that they were not critical ROIs in this analysis.

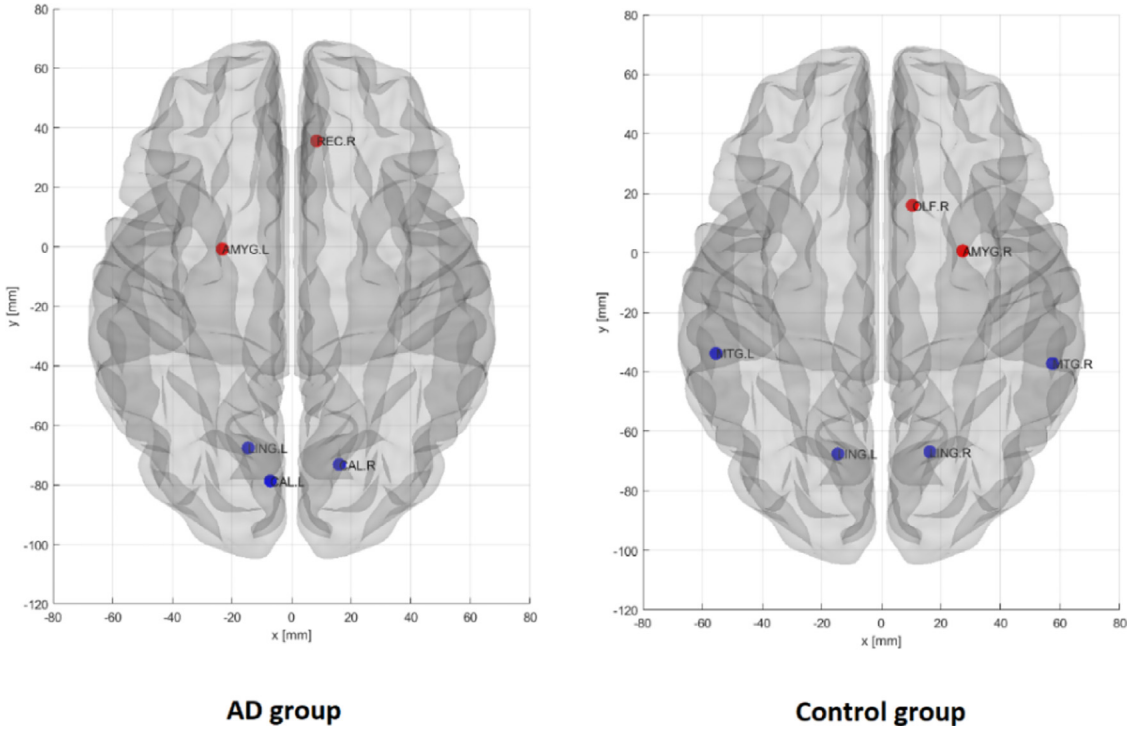

Fig. 8. The red circles are related to the most sparsed ROIs in the groups and the blue circles show the most intact ROIs in both groups (the brain view is dorsal axial).

5. Conclusion

In the present study, two methods are proposed for brain graph sparsification to overcome the limitations of ordinary thresholding. Based on the results, the sparsed graph that came from the proposed method reveals more distinction between groups. It can be concluded that the outcomes of the algorithms are more noiseless and contains more reliable information that can exhibit more discrimination. Although the presented approaches have more potential their computational costs are high and the implementations are more complex in comparison to simple thresholding. More research on different types of deep learning algorithms and also find the optimal structure of neural networks are recommended. Besides, considering the physiological properties in the sparsification process will lead to interesting and more authentic results.

Declaration of Competing Interest

This statement is to certify that all authors have seen and approved the manuscript being submitted and have no conflict of interest.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cmpb.2021.105954.

References

[1] B. Dubois, et al., Preclinical Alzheimer's disease: definition, natural history, and diagnostic criteria, Alzheimer's Dementia 12 (3) (2016) 292–323.
[2] A. s. Association, in: 2018 Alzheimer's Disease Facts and Figures, 14, Alzheimer's & Dementia, 2018, pp. 367–429.
[3] M. Yusufov, L.L. Weyandt, I. Piryatinsky, Alzheimer's disease and diet: a systematic review, Int. J. Neurosci. 127 (2) (2017) 161–175.
[4] J. Weuve, L.E. Hebert, P.A. Scherr, D.A. Evans, Prevalence of Alzheimer disease in US states, Epidemiology 26 (1) (2015) e4-e6.
[5] U.R. Acharya, et al., Automated detection of Alzheimer's disease using brain MRI images–a study with various feature extraction techniques, J. Med. Syst. 43 (9) (2019) 302.
[6] A. De, A.S. Chowdhury, DTI based Alzheimer disease classification with rank modulated fusion of CNNs and random forest, Expert Syst. Appl. (2020) 114338.
[7] N.T. Duc, S. Ryu, M.N.I. Qureshi, M. Choi, K.H. Lee, B. Lee, 3D-deep learning based automatic diagnosis of Alzheimer's disease with joint MMSE prediction using resting-state fMRI, Neuroinformatics 18 (1) (2020) 71–86.
[8] P. Pławiak, U.R. Acharya, Novel deep genetic ensemble of classifiers for arrhythmia detection using ECG signals, Neural Comput. Appl. 32 (15) (2020) 11137–11161.
[9] M. Hammad, P. Pławiak, K. Wang, U.R. Acharya, ResNet-Attention model for human authentication using ECG signals, Expert Syst. (2020) e12547.
[10] P. Pławiak, M. Abdar, U.R. Acharya, Application of new deep genetic cascade ensemble of SVM classifiers to predict the Australian credit scoring, Appl. Soft Comput. 84 (2019) 105740.
[11] M.D. Fox, M.E. Raichle, Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging, Nat. Rev. Neurosci. 8 (9) (2007) 700.
[12] C. Caballero-Gaudes, R.C. Reynolds, Methods for cleaning the BOLD fMRI signal, Neuroimage 154 (2017) 128–149.
[13] Z. Gao, Y. Feng, C. Ma, K. Ma, Q. Cai, A.s.D.N. Initiative, Disrupted Time-Dependent and Functional Connectivity Brain Network in Alzheimer's Disease: a Resting-State fMRI Study Based on Visibility Graph, Current Alzheimer Res. 17 (1) (2020) 69–79.
[14] A. Khazaee, A. Ebrahimzadeh, A. Babajani-Feremi, Identifying patients with Alzheimer's disease using resting-state fMRI and graph theory, Clin. Neurophysiol. 126 (11) (2015) 2132–2141.
[15] M.P. Van Den Heuvel, H.E.H. Pol, Exploring the brain network: a review on resting-state fMRI functional connectivity, Eur. Neuropsychopharmacol. 20 (8) (2010) 519–534.
[16] T.O. Laumann, et al., On the stability of BOLD fMRI correlations, Cereb. Cortex 27 (10) (2016) 4719–4732.
[17] E. Bullmore, O. Sporns, Complex brain networks: graph theoretical analysis of structural and functional systems, Nat. Rev. Neurosci. 10 (3) (2009) 186.
[18] J.C. Reijneveld, S.C. Ponten, H.W. Berendse, C.J. Stam, The application of graph theoretical analysis to complex networks in the brain, Clin. Neurophysiol. 118 (11) (2007) 2317–2331.
[19] B.R. Logan, D.B. Rowe, An evaluation of thresholding techniques in fMRI analysis, Neuroimage 22 (1) (2004) 95–108.
[20] C. Bordier, C. Nicolini, A. Bifone, Graph analysis and modularity of brain functional connectivity networks: searching for the optimal threshold, Front. Neurosci. 11 (2017) 441.
[21] S. Chiang, et al., Time-dependence of graph theory metrics in functional connectivity analysis, Neuroimage 125 (2016) 601–615.

[22] J. delEtoile, H. Adeli, Graph theory and brain connectivity in Alzheimer's disease, Neuroscientist 23 (6) (2017) 616–626.
[23] T.-Y. Chang, et al., Graph theoretical analysis of functional networks and its relationship to cognitive decline in patients with carotid stenosis, J. Cerebral Blood Flow Metabol. 36 (4) (2016) 808–818.
[24] S.H. Hojjati, A. Ebrahimzadeh, A. Khazaee, A. Babajani-Feremi, A.s.D.N. Initiative, Predicting conversion from MCI to AD by integrating rs-fMRI and structural MRI, Comput. Biol. Med. 102 (2018) 30–39.
[25] J. Xiang, et al., Graph-based network analysis of resting-state fMRI: test-retest reliability of binarized and weighted networks, Brain Imag. Behav. (2019) 1–12.
[26] J. Benito-León, et al., Graph theory analysis of resting-state functional magnetic resonance imaging in essential tremor, Hum. Brain Mapp. (2019).
[27] Y. He, A. Evans, Graph theoretical modeling of brain connectivity, Curr. Opin. Neurol. 23 (4) (2010) 341–350.
[28] M.-.E. Lynall, et al., Functional connectivity and brain networks in schizophrenia, J. Neurosci. 30 (28) (2010) 9477–9487.
[29] Y. Luo, et al., Alterations of brain networks in Alzheimer's disease and mild cognitive impairment: a resting state fMRI study based on a population-specific brain template, Neuroscience (2020).
[30] H. Ahmadi, E. Fatemizadeh, A.M. Nasrabadi, A comparative study of the effect of weighted or binary functional brain networks in fMRI data analysis, Front. Biomed. Technol. 7 (3) (2020) 159–168.
[31] D. Dash, V. Abrol, A.K. Sao, B. Biswal, Spatial sparsification and low rank projection for fast analysis of multi-subject resting state fMRI data, in: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), IEEE, 2018, pp. 1280–1283.
[32] S.G. Mueller, et al., The Alzheimer's disease neuroimaging initiative, Neuroimag. Clinics 15 (4) (2005) 869–877.
[33] X. Hua, et al., MRI-based brain atrophy rates in ADNI phase 2: acceleration and enrichment considerations for clinical trials, Neurobiol. Aging 37 (2016) 26–37.
[34] C.R. Jack Jr, et al., The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods, J. Magnetic Reson. Imag.: Off. J. Int. Soc. Magn. Reson. Med. 27 (4) (2008) 685–691.
[35] C. Yan, Y. Zang, DPARSF: a MATLAB toolbox for" pipeline" data analysis of resting-state fMRI, Front. Syst. Neurosci. 4 (2010) 13.
[36] N. Tzourio-Mazoyer, et al., Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain, Neuroimage 15 (1) (2002) 273–289.
[37] W. C. contributors. "File: autoencoder structure.png." Wikimedia Commons, the free media repository. https://commons.wikimedia.org/w/index.php?title=File:Autoencoder_structure.png&oldid=422961315 (accessed 7 December 2020 15:39 UTC.
[38] A. Ng, Sparse autoencoder, CS294A Lecture notes 72 (2011) 1–19 2011.
[39] L. Zhang, Y. Lu, B. Wang, F. Li, Z. Zhang, Sparse Auto-encoder with Smoothed $$ l_1 $$ Regularization, Neural Process. Lett. 47 (3) (2018) 829–839.
[40] P. Baldi, Autoencoders, unsupervised learning, and deep architectures, in: Proceedings of ICML workshop on unsupervised and transfer learning, 2012, pp. 37–49.
[41] D.A. Spielman, S.-.H. Teng, Spectral sparsification of graphs, SIAM J. Comput. 40 (4) (2011) 981–1025.
[42] D.A. Spielman, N. Srivastava, Graph sparsification by effective resistances, SIAM J. Comput. 40 (6) (2011) 1913–1926.
[43] R. Arora, J. Upadhyay, On Differentially Private Graph Sparsification and Applications, in: Advances in Neural Information Processing Systems, 2019, pp. 13378–13389.
[44] F. Liem, L. Geerligs, J.S. Damoiseaux, and D.S. Margulies, "Functional Connectivity in Aging," 2019.