

# FMRI Functional Connectivity Evaluation in Alzheimer's Stages: Linear and Non-Linear Approaches

Hessam Ahmadi (✉ [hessam.ahmadi@srbiau.ac.ir](mailto:hessam.ahmadi@srbiau.ac.ir))

Islamic Azad University Science and Research Branch <https://orcid.org/0000-0002-7910-6283>

Emad Fatemizadeh

Sharif University of Technology <https://orcid.org/0000-0001-5017-7136>

Ali Motie Nasrabadi

Shahed University

---

## Research Article

**Keywords:** fMRI, functional connectivity, linear and non-linear, graph theory, Alzheimer's disease

**Posted Date:** February 13th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-189491/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# fMRI functional connectivity evaluation in Alzheimer's stages: Linear and Non-linear approaches

Hessam Ahmadi<sup>1</sup>, Emad Fatemizadeh<sup>2\*</sup><sup>i</sup>, Ali Motie-Nasrabadi<sup>3</sup>

1- Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

2- School of Electrical Engineering, Sharif University of Technology, Tehran, Iran, Email: fatemizadeh@sharif.edu

3- Biomedical Engineering Department, Shahed University, Tehran, Iran

## Abstract

Neuroimaging data analysis reveals the underlying interactions in the brain. It is essential, yet controversial, to choose a proper tool to manifest brain functional connectivity. In this regard, researchers have not reached a definitive conclusion between the linear and non-linear approaches, as both have pros and cons. In this study, to evaluate this concern, the functional Magnetic Resonance Imaging (fMRI) data of different stages of Alzheimer's disease are investigated. In the linear approach, the Pearson Correlation Coefficient (PCC) is employed as a common technique to generate brain functional graphs. On the other hand, for non-linear approaches, two methods including Distance Correlation (DC) and the kernel trick are utilized. By the use of the three mentioned routines and graph theory, functional brain networks of all stages of Alzheimer's disease (AD) are constructed and then sparsed. Afterwards, graph global measures are calculated over the networks and a non-parametric permutation test is conducted. Results reveal that the non-linear approaches have more potential to discriminate groups in all stages of AD. Moreover, the kernel trick method is more powerful in comparison to the DC technique. Nevertheless, AD degenerates the brain functional graphs more at the beginning stages of the disease. At the first phase, both functional integration and segregation of the brain degrades, and as AD progressed brain functional segregation further declines. The most distinguishable feature in all stages is the clustering coefficient that reflects brain functional segregation.

**Keywords:** *fMRI, functional connectivity, linear and non-linear, graph theory, Alzheimer's disease*

## Introduction

Investigation of the effect of neurological disease on the brain has attracted researchers in the field. Through different types of imaging such as Computed Tomography (CT) or Magnetic Resonance Imaging (MRI), structural changes in the brain have been studied but the illness's effect on brain function was not as clear as the structural effect. During recent decades, by the invention of methods such as functional MRI (fMRI) or Positron Emission Tomography (PET), the number of brain function studies has been ever increasing. fMRI consists of several MRI scans performed every couple of seconds to follow alterations in brain oxygen consumption. Therefore, fMRI is a non-invasive method that captures low-frequency oscillations called Blood-Oxygen-Level-Dependent (BOLD) signal (Uludag, Ugurbil, & Berliner, 2015). The nature of fMRI signals like

<sup>i</sup> - Corresponding author ([fatemizadeh@sharif.edu](mailto:fatemizadeh@sharif.edu)) (+98912) 3504238).

Electroencephalograms (EEG) is derived from the brain function. Although brain behavior is an open-debating issue, some scientists prefer to model and analyze the brain as a linear system. On the other hand, approaching the brain as a linear system is simplistic and offers non-linear approaches that are more similar to brain nature.

A common tool to analyze fMRI signals is graph theory. Here, the brain regions or voxels are modeled as graph nodes. The links between the nodes are made using fMRI signals. Accordingly, the edges are representative of the relationship between brain regions. The most prevalent approach to model the brain functional connectivity is the Pearson Correlation Coefficient (PCC). PCC measures the correlation between two fMRI signals and the result is a number between -1 to 1. The sign shows the direction of the connectivity and the magnitude shows its strength (Li, Guo, Nie, Li, & Liu, 2009). The point is that PCC captures only the linear dependency of two time-series which is a simple assumption of the brain's relationships. However, ample research shows that the brain's signals including fMRI demonstrate non-linear behavior the PCC is a widespread and reliable method for brain functional connectivity analysis (Anzellotti, Fedorenko, Kell, Caramazza, & Saxe, 2017; Hlinka, Paluš, Vejmelka, Mantini, & Corbetta, 2011). Recently, a study has been conducted on a non-linear alternative for PCC based on fMRI time-series of Alzheimer's Disease (AD). In this study (Ahmadi, Fatemizadeh, & Motie-Nasrabadi, 2020a), they used the kernel trick which a polynomial kernel to increase the dimension of the input space and perform the PCC calculation in a new space. The PCC in the new space is equivalent to non-linear relations in the primary space. In another research, the Kernel Canonical Correlation (KCC) was employed to analyze fMRI and EEG data (Yang et al., 2018). Among all controversies about linear methods or non-linear ones, in 2005, Gabor proposed a method called Distance Correlation (DC) to overcome PCC restrictions based on the limitations of PCC in capturing non-linear dependencies. DC quantifies both linear and non-linear dependencies between two signals (Székely, Rizzo, & Bakirov, 2007). The results show that DC is more powerful than PCC for measuring the relationship between the two vectors.

Neurological diseases affect brain structural and functional connectivity. AD is a destructive and progressive neurological disease discovered in 1906 by Dr. Alois Alzheimer. Although more than a century has passed since the first AD case, there is still no definitive and effective treatment. Studies showed that AD has different stages called Early and Late Mild Cognitive Impairment (EMCI and LMCI) and it may take up to a decade for acute clinical symptoms to appear (Mysterud, 2019). Since there is no specific treatment to return the patient to normal mental health, early detection is vital. Several experiments revealed that the brain suffered from atrophy in AD (Pini et al., 2016). Patients are classified from normal subjects according to the structural changes widely in the literature (Khagi, Kwon, & Lama, 2019; Rathore, Habes, Iftikhar, Shacklett, & Davatzikos, 2017). Also, several biomarkers have been identified (Frisoni et al., 2017). Functional connectivity information was used to distinguish the AD subjects and predict conversion from MCI to AD by Khazaee et al. in recent years and they achieved more than 96% accuracy (Hojjati, Ebrahimzadeh, Khazaee, Babajani-Feremi, & Initiative, 2018). Another interesting approach is to combine

structural data with functional data in order to make a generalized insight into AD. They reached 56% accuracy for the three-class classification (Hojjati, Ebrahimzadeh, & Babajani-Feremi, 2019). As mentioned before, most studies used PCC to generate brain graphs or used non-linear approaches individually. Although there are studies such as (Hessam Ahmadi, 2021) to compare different correlation methods, no study has been conducted on PCC and robust non-linear methods such as DC and kernel-based. On the other hand, as mentioned, AD has a nature that makes it crucial to understand what happens in the brain's functional connectivity as AD progressed. Several studies based on linear or non-linear approaches have been conducted such as (Ahmadi, Fatemizadeh, & Motie-Nasrabadi, 2020b). In the present study, to overcome the limitations of previous analyses via PCC as the most accepted linear method and kernel-based and DC methods as non-linear tools, functional connectivity is employed to analyze fMRI data of AD. Furthermore, to consider the gradual nature of AD and also perform a precise experiment, three distinct conditions are explored including Healthy subjects vs. EMCI, EMCI vs. LMCI and, LMCI vs. AD. The goal of this study is to clarify what exactly happens to brain functional graphs as AD progressed. Moreover, by combining the outcomes of three different methods, one can arrive at reliable generalizations.

The rest of the article is arranged as follows: In the materials and methods section, the fMRI data are introduced and the steps of preprocessing are described. Then, there is a sub-section called Correlation Methods which consists of PCC, kernel-based, and DC definitions and relative equations. Afterwards, graph theory and statistical tests are described. In the Results section, the outcomes of analyses are elaborately reported through tables and figures. In the Discussion, the results are interpreted. Finally, concluding remarks are presented in the last section.

## Materials and Methods

In this section, the utilized data and tools are presented separately and each of them is explained completely. The steps of the research are summarized in Fig 1 as a block diagram.

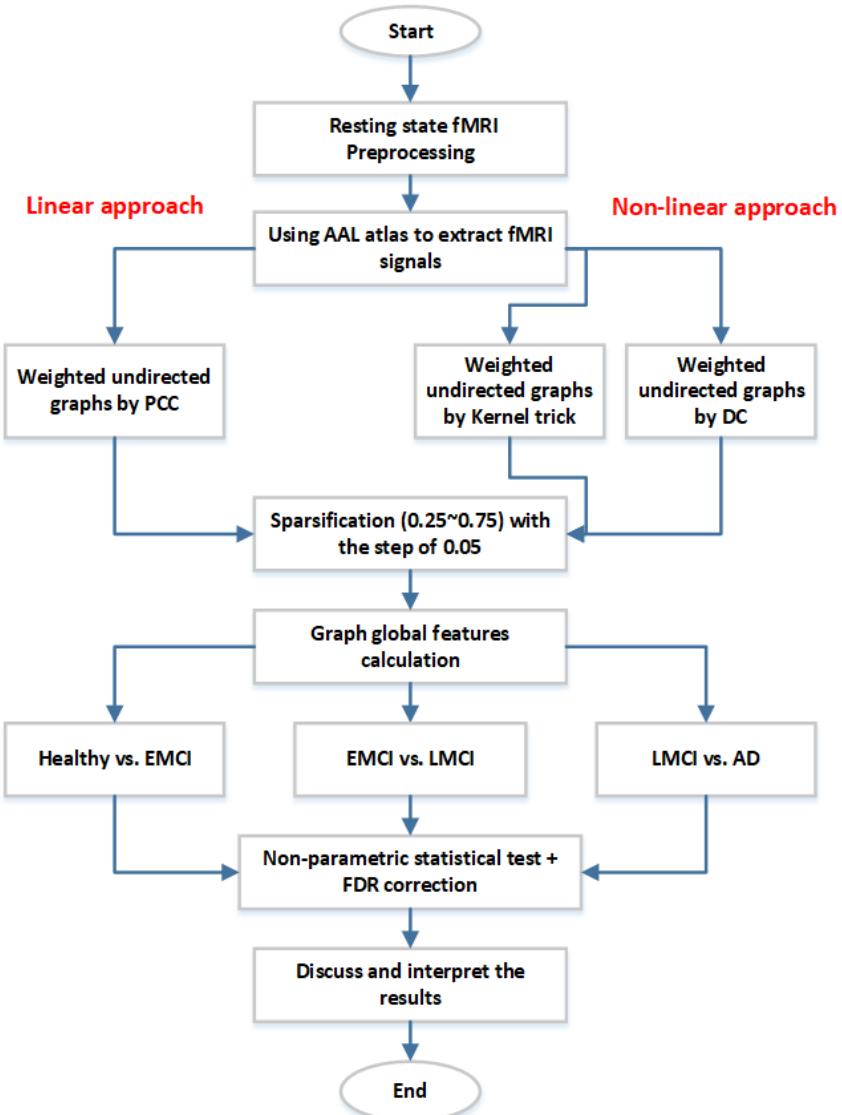


Fig 1- The research flowchart.

## Data and Preprocessing

The fMRI data were collected from the second phase of the Alzheimer's Disease Neuroimaging Initiative (ADNI) project which contains healthy subjects and all stages of AD (EMCI, LMCI, AD) (Petersen et al., 2010). The selected cases are aged-matched and the mental examination scores including Mini-Mental State Examination (MMSE) and Clinical Dementia Rating (CDR) were checked. Each fMRI data contains 140 volumes with a Repetition Time (TR) of 3000 msec. Also, the Echo Time (TE), flip angle, and slice thickness were 30 msec, 80 degrees, and 3.3125 mm respectively. The information is summarized in Table 1.

Table 1- fMRI data information

Group	No. (Male/Female)	Age	CDR score	MMSE score
<b>Control</b>	(10/8)	75.9±4.3	0	30±0
<b>EMCI</b>	(10/10)	75±2.4	0.5±0	27.3±1.1
<b>LMCI</b>	(8/8)	75.3±4.3	0.56±0.16	24.5±1.9
<b>AD</b>	(8/10)	76.2±5.6	1.27±0.53	19. 4±2.5

For preprocessing the data, according to ADNI, there was no need to remove first time-points. The rest of the preprocessing steps were implemented as depicted in Fig. 2. Data Processing Assistant for Resting-State fMRI (DPARSF) (Chao-Gan & Yu-Feng, 2010) was used to implement preprocessing, and Automated Anatomical Labeling (AAL) (Tzourio-Mazoyer et al., 2002) was employed to extract Regions of Interest (ROI) signals. All the other processing were conducted utilizing Matlab 2018a software.

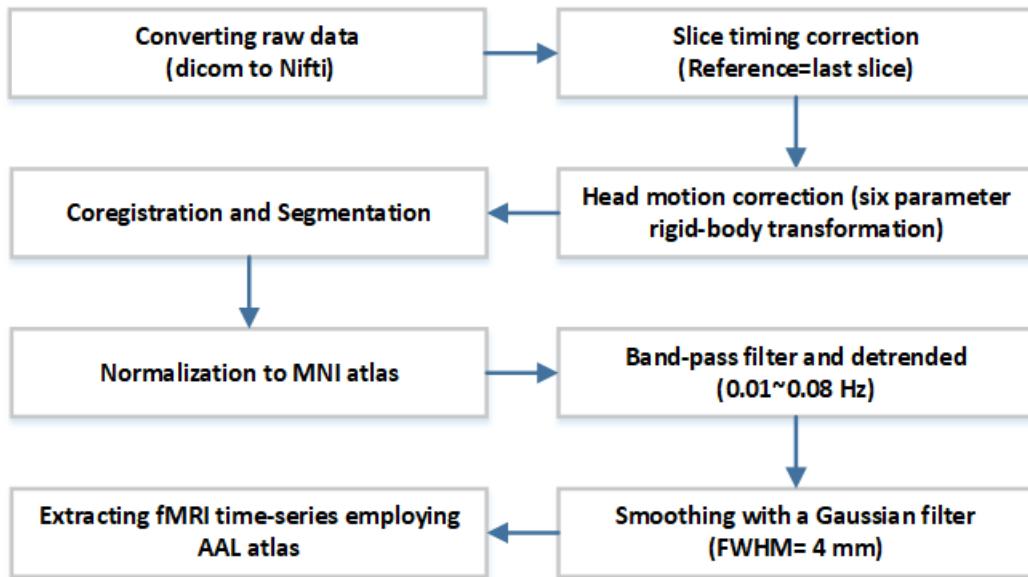


Fig 2- The preprocessing steps.

## Correlation Methods

### 1- Pearson Correlation Coefficient (PCC)

In statistics, the correlation coefficients are used to measure the dependency of two vectors. PCC is the most popular method for this assessment but only addresses the linear relationship. If two variables have a total positive correlation the PCC is +1 and -1 corresponds to the total negative correlation. A value 0 exhibits no correlation. The PCC formula is:

$$\rho(X_i, X_j) = \frac{Cov(X_i, X_j)}{\sigma_{X_i} \sigma_{X_j}}$$

Where  $\sigma_{X_i}$  and  $\sigma_{Y_j}$  are the representative of the standard deviation of vector  $X_i$  and  $Y_j$ . Also  $Cov$  is the covariance of the vectors (Benesty, Chen, Huang, & Cohen, 2009).

## 2- Kernel trick and PCC

In kernel trick, by using kernel functions, the input data is mapped to a new space. The linear calculations in the new space are equivalent to the non-linear computations in the primary space (Alam, Calhoun, & Wang, 2018; T. Hofmann, Schölkopf, & Smola, 2008). Assuming that the data is  $x$  and the  $\varphi$  is the corresponded transformation, the kernel trick is as follows:

$$K(x_i, x_j) = x_i^T x_j$$

$$\varphi : x \rightarrow \varphi(x)$$

$$K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j)$$

Where  $\varphi$  is the nonlinear transformation,  $x_i, x_j$  are the two variables, and  $K$  is the kernel function. According to Mercer theorem,  $K$  has to be a positive definite (M. Hofmann, 2006). By the use of the kernel trick, one can evaluate the inner product of two signals without knowing the  $\varphi(x)$  (Kung, 2014).

If  $X_i, X_j$  are two signals, the covariance between them is defined as:

$$Cov(X_i, X_j) = (X_i - \mu) \cdot (X_j - \nu)$$

where  $\mu$  and  $\nu$  are the means of two  $X_i, X_j$  signals respectively and ‘.’ is the dot product. The above covariance is the Pearson Covariance (Towsley, Pakianathan, & Douglass, 2011). Now, the PCC is defined as (Towsley et al., 2011):

$$\rho(X_i, X_j) = \frac{Cov(X_i, X_j)}{\sqrt{Cov(X_i, X_i)Cov(X_j, X_j)}}$$

Consequently, the PCC is rewritten regarding dot product; therefore, based on (Scholkopf & Smola, 2001) the kernel trick is applicable as follows:

$$K\{\rho(X_i, X_j)\} = \frac{K\{(X_i, X_j)\}}{\sqrt{K\{(X_i, X_i)\}K\{(X_j, X_j)\}}},$$

As it was shown, the PPC can not uncover the nonlinear relations, so rather than figuring the PCC in the primary space, kernel trick is utilized to process PCC in the new space. This is identical to nonlinear relationships in the primary space. In this study based on (Ahmadi et al., 2020a), the polynomial kernel function has opted.

### 3- Distance Correlation (DC)

Based on the limitations of PCC for the evaluation of non-linear dependencies, DC was introduced in 2005 to quantify non-linear relationships (Székely et al., 2007). Assuming that  $(X_m, Y_m), m = 1, 2, \dots, n$  are two vectors. In this method, the distance matrix is defined:

$$a_{e,f} = \|X_e - X_f\|, e, f = 1, 2, \dots, n$$

$$b_{e,f} = \|Y_e - Y_f\|, e, f = 1, 2, \dots, n$$

Where  $\|\cdot\|$  is the Euclidean distance. Also,  $A_{e,f}$  and  $B_{e,f}$  can be described as:

$$A_{e,f} = a_{e,f} - \bar{a}_e - \bar{a}_f + \bar{a}_{..}$$

$$B_{e,f} = b_{e,f} - \bar{b}_e - \bar{b}_f + \bar{b}_{..}$$

Where  $\bar{a}_e$  is the mean of the  $e^{th}$  row and  $\bar{a}_f$  is the mean of the  $f^{th}$  column. The mean correlation distance is  $\bar{a}_{..}$ .  $B$  and  $A$  are defined identically. Finally, the arithmetic average of the product of  $A_{e,f}$  and  $B_{e,f}$  is the distance covariance:

$$dCov_n^2(X, Y) = \frac{1}{n^2} \sum_{e=1}^n \sum_{f=1}^n A_{e,f} B_{e,f}$$

Accordingly, the DC is defined:

$$dCor(X, Y) = \frac{dCov(X, Y)}{\sqrt{dVar(X)dVar(Y)}}$$

Where  $dVar$  is distance variance and computed similar to  $dCov$  as mentioned above.

## Graph Theory

Regardless of the region-based or voxel-based analysis, fMRI processing incorporated large data. Since many voxels or regions are available, calculating the correlation among all pairwise of them provides a large amount of information. A practical approach to overcome this issue is graph theory. The nodes of a graph represent brain regions or voxels. Also, functional or effective connectivity can be considered as the graph's links.  $G = (V, E)$  is used to show a graph, and  $V$   $E$  denote the nodes (brain regions) and edges (connectivity) respectively (Sporns, 2018). To

eliminate the weak and spurious edges that are not representative of real and strong correlations in the brain, a sparsification step is considered. The sparsification step makes the binarized graphs out of weighted ones (Logan & Rowe, 2004). Since the identification of an optimal threshold is still a controversy, in this research the functional graphs are sparsed from the threshold of 0.25 up to 0.75 with the step of 0.05 for a comprehensive investigation.

By employing graph theory, many measures can be defined to reflect the characteristics of the brain graph. A healthy brain network shows functional integration and segregation. These properties make the information flow in the brain efficiently and flexibly. Neurological disease such as AD degrades brain networks and affects these properties. The utilized graph features are introduced in Table 2.

Table 2- Graph features and explanations (Harris, Hirst, & Mossinghoff, 2008; Humphries & Gurney, 2008; Kepner & Gilbert, 2011; Rubinov & Sporns, 2010).

Metric	Formula	Definition
Degree	-	number of edges connected to a node
Radius	$R = \min \{ECC\}$	-
Diameter	$D = \max \{ECC\}$	-
Eccentricity	$Ecc = \max \{d_G(x, y)\}$	The maximal distance between a particular node and some other center point
Characteristic Path Length (CPL)	$L = \frac{\sum_{x,y \in V(G)} d_G(x, y)}{n(n-1)}$	The average distance between a node to others
Global Efficiency	$E_{glob}(G) = \frac{E(G)}{E(G^{FC})}$	Average of the inverse shortest path length
Local Efficiency	$E_{loc}(G) = \frac{1}{n} \sum_{x \in G} E(G_x)$	Global efficiency of a node. calculated on the node's neighbors
Clustering	$C = \frac{\text{Number of closed triplets}}{\text{number of all triplets}}$	A fraction of available triangles around a node
Modularity	$M = \frac{1}{l} \sum_{x,y} \left[ A_{x,y} - \frac{k_x k_y}{l} \right] \delta_{x,y}$	The degree to which a graph can be partitioned into obviously isolated networks
Transitivity	$T = \frac{3 * \text{number of triangles}}{\text{number of connected triples of nodes}}$	The ratio of the total number of triangles to the number of triplets
Small-Worldness	$\sigma = \frac{C}{C_r} / \frac{L}{L_r}$	A small-world graph has a comparative trademark path length as an irregular graph with a similar degree conveyance yet is fundamentally more clustered

$n$  is the number of nodes,  $l$  is the number of links, and  $d_G(x, y)$  is the distance between the  $x$  and  $y$ .  $G^{FC}$  is a representative of a fully connected graph.  $A_{x,y}$  shows the connectivity matrix. The  $\delta_{x,y}$  is 0 if the two vertices are from one module, otherwise, it is 1.

$E(G) = \frac{1}{n(n-1)} \sum_{x \neq y \in G} \frac{1}{p(x,y)}$  demonstrates the average efficiency and  $p(x,y)$  corresponds to the shortest path length between  $x$  and  $y$ .  $C_r$  and  $L_r$  relates to an identical irregular graph.

## Statistical analysis

In neuroimaging data processing, a non-parametric permutation test has been widely used and recommended. It is based on bootstrapping and also by employing random subsets of the data, the results are validated. In this paper, the number of permutations adjusted to 5000 times, and the significance level is considered as %5 (P-Value < 0.05). It is worthwhile mentioning that due to the multiple comparisons and to control type I error, the False Discovery Rate (FDR) is applied (Nichols & Holmes, 2002).

## Results

The purpose of this study is to investigate brain functional graph changes during stages of AD by employing linear and non-linear methods. Since the generating methods are different, the brain graphs vary in structure. Fig 3. depicts the brain graphs in control subjects utilizing three routines. Notably, for more appropriate visualization the graphs are sparsed with an 0.75 threshold.

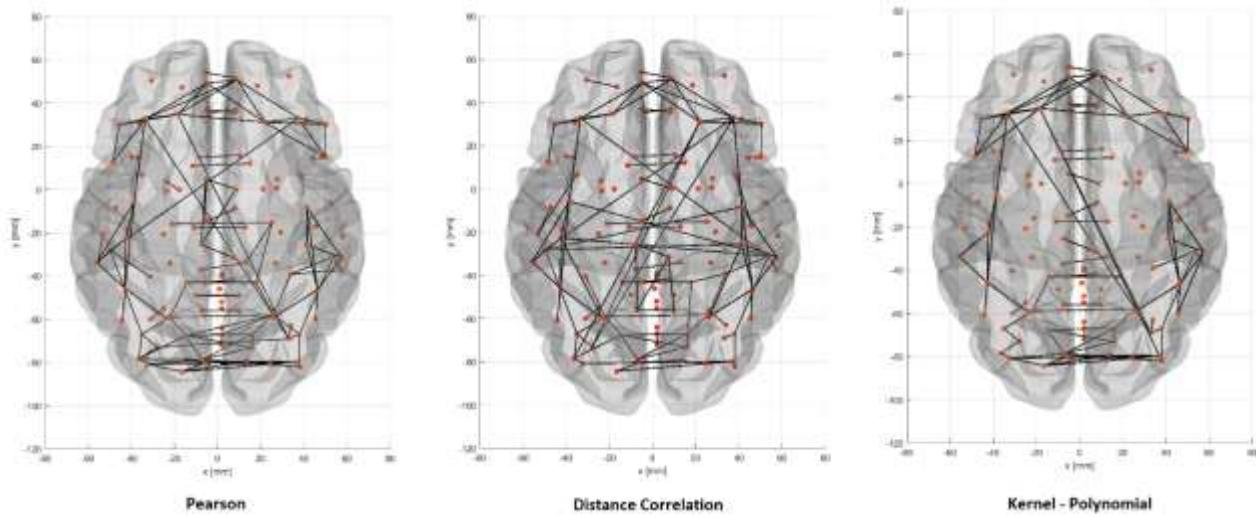


Fig 3- Demonstration of brain networks utilizing different methods in control subjects. All the graphs are sparsed with an 0.75 threshold.

As displayed in Fig. 3, each method computes and predicts functional connectivity differently. However, the variation of all the methods exhibits more inter modular functional connectivity in

Occipital and Frontal areas. Also, in the same threshold, DC shows more functional connectivity in comparison to other methods.

In the first analysis, brain functional graphs of healthy subjects and EMCI groups are compared in different thresholds by permutation test. In Table 3, the result for “Modularity” is exhibited:

Table 3- P-Values of modularity feature obtained from permutations statistical test between healthy and EMCI groups. The boldface numbers are significantly different (P-Value < 0.05).

Modularity			
Threshold	PCC	DC	Polynomial
<b>0.25</b>	0.2136	0.1272	0.0706
<b>0.3</b>	0.2068	0.0894	<b>0.0366</b>
<b>0.35</b>	0.2566	0.1002	<b>0.0422</b>
<b>0.4</b>	0.2954	0.1110	<b>0.0308</b>
<b>0.45</b>	0.3756	0.1388	<b>0.0426</b>
<b>0.5</b>	0.4766	0.1260	0.0554
<b>0.55</b>	0.4122	0.0808	<b>0.0486</b>
<b>0.6</b>	0.2952	<b>0.0378</b>	<b>0.0248</b>
<b>0.65</b>	0.2066	<b>0.0270</b>	<b>0.0386</b>
<b>0.7</b>	0.1332	<b>0.0234</b>	<b>0.0478</b>
<b>0.75</b>	0.0870	<b>0.0200</b>	0.0922

According to Table 3, regardless of the selected threshold, PCC shows no significant differences. In other words, from the point of view of PCC, the modularity feature of functional brain graphs in EMCI subjects is almost the same as healthy subjects. On the other hand, the DC method shows significant changes in the modularity feature in several thresholds. Although other thresholds show no significant difference the P-Values are much smaller in comparison to PCC. The kernel-based method also shows significant changes. In comparison to DC, the kernel-based method exhibits more significant difference, and has more power to discriminate between healthy vs. EMCI subjects. To have a comprehensive understanding of Table 3, Fig. 4 illustrates the P-Values distribution.

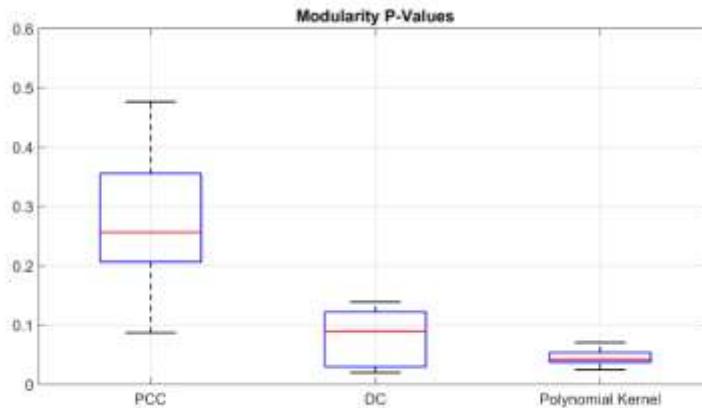


Fig 4- Distribution of P-Values of modularity feature in healthy vs. EMCI subjects.

To avoid many tables (one table for every feature), the results are summarized in Table 4 which shows the number of significant differences for all the features in each method.

Table 4- Number of significant differences (P-Value < 0.05) for every feature in each method between healthy and EMCI subjects.

<b>Measure \ Method</b>	<b>PCC</b>	<b>DC</b>	<b>Polynomial</b>
<b>Degree</b>	4	4	5
<b>Radius</b>	0	0	1
<b>Diameter</b>	1	1	2
<b>Eccentricity</b>	1	6	7
<b>CPL</b>	1	4	4
<b>Global Efficiency</b>	6	9	8
<b>Local Efficiency</b>	6	6	8
<b>Clustering</b>	6	6	7
<b>Transitivity</b>	5	7	7
<b>Modularity</b>	0	4	8
<b>Small-Worldness</b>	0	7	7
<b>Total</b>	<b>30</b>	<b>54</b>	<b>64</b>

Table 4 explains that DC is more powerful than PCC and kernel trick is the most discriminant method in total. Regardless of the method features such as global and local efficiencies and clustering are the most distinguishable measures and these graph properties of the brain are the most affected characteristics when healthy subjects turn into EMCI.

Sparsification is a major issue in graph analysis. Table 5 represents the effect of different thresholds. The arrays are the number of significant differences in each threshold for all the features.

Table 5- Number of significant differences (P-Value < 0.05) for each feature in all the thresholds between healthy and EMCI subjects.

Measure \ Threshold	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75
<b>Degree</b>	0	1	1	1	1	1	1	2	2	2	1
<b>Radius</b>	0	0	0	0	0	1	0	0	0	0	0
<b>Diameter</b>	0	0	1	1	1	1	0	0	0	0	0
<b>Eccentricity</b>	2	2	2	2	2	2	1	0	1	0	0
<b>CPL</b>	0	1	2	2	2	1	1	0	0	0	0
<b>Global Efficiency</b>	3	2	3	3	2	1	2	2	2	2	1
<b>Local Efficiency</b>	2	3	2	2	1	1	1	2	2	1	2
<b>Clustering</b>	2	2	3	2	2	2	2	1	1	1	1
<b>Transitivity</b>	3	3	2	2	2	3	2	1	0	0	1
<b>Modularity</b>	0	1	1	1	1	0	1	2	2	2	1
<b>Small-Worldness</b>	2	2	2	2	2	2	2	0	0	0	0
<b>Total</b>	<b>14</b>	<b>17</b>	<b>19</b>	<b>18</b>	<b>16</b>	<b>15</b>	<b>13</b>	<b>10</b>	<b>10</b>	<b>8</b>	<b>7</b>

As it is shown in Table 5, there is no exact pattern for thresholds and the feature behaviors are nonidentical. On the other hand, the thresholds from 0.3 to 0.4 are the optimal values for discrimination between classes in total. For instance, the modularity feature extracted from EMCI (absence of difference in none of the thresholds) is illustrated in Fig. 5 as the threshold changed.

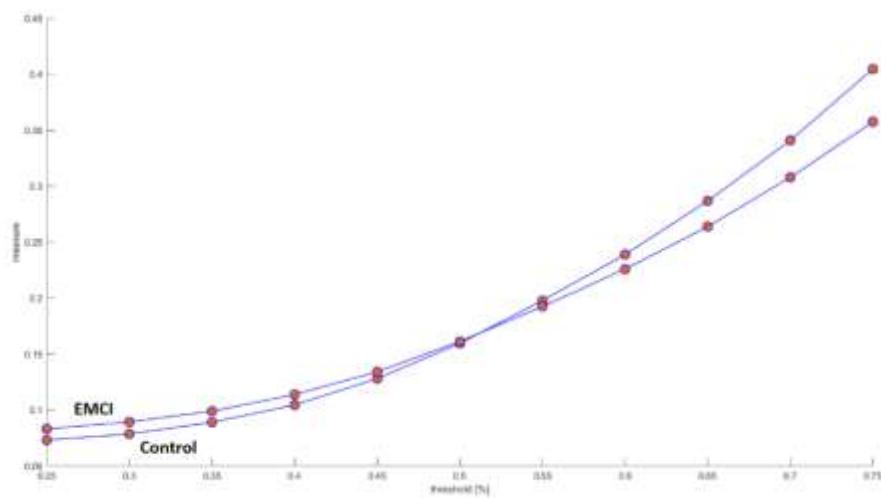


Fig 5- Variation of modularity measure in PCC analysis as thresholds increased.

Fig. 4 depicts that the average values of modularity from PCC analysis are nearly identical in healthy (control) and EMCI subjects. Although the increment of threshold makes alterations according to the statistical test, they are not significant.

As the disease progressed, the EMCI subjects convert to LMCI and then AD. The approach for investigating the EMCI vs. LMCI and LMCI vs. AD is the same as above. To summarize, the outputs are given in Table 6 as follows:

Table 6- Number of significant differences (P-Value < 0.05) for each feature in each method.

Measure	EMCI vs. LMCI			LMCI vs. AD		
	PCC	DC	Polynomial	PCC	DC	Polynomial
<b>Degree</b>	0	2	4	0	1	3
<b>Radius</b>	1	1	1	0	0	1
<b>Diameter</b>	0	1	1	1	2	2
<b>Eccentricity</b>	0	2	2	0	0	0
<b>CPL</b>	2	4	6	0	0	2
<b>Global Efficiency</b>	1	6	7	0	0	1
<b>Local Efficiency</b>	0	5	6	1	1	1
<b>Clustering</b>	5	9	11	3	4	7
<b>Transitivity</b>	0	6	6	0	2	6
<b>Modularity</b>	0	6	7	0	1	2
<b>Small-Worldness</b>	0	0	0	0	0	0
<b>Total</b>	<b>9</b>	<b>42</b>	<b>51</b>	<b>5</b>	<b>11</b>	<b>25</b>

According to Table 6, in EMCI vs. LMCI analysis, again the non-linear methods exhibit more power to distinguish the groups. Also, the Kernel-based method shows better performance than the DC method. There is the same pattern in the LMCI vs. AD examination. Between EMCI and LMCI, clustering is the most discriminative feature, then modularity, CPL, transitivity, and efficiencies illustrate significant differences. Between LMCI and AD, the clustering is the most distinguishable measure and other metrics show no significant changes. It is worthwhile mentioning that the effect of thresholding is as same as before (healthy vs. EMCI). The optimal threshold for better discrimination is approximately 0.3 to 0.4.

## Discussion

In this study, the whole-brain functional graphs are generated by the use of PCC, kernel trick, and DC. Afterwards, the graphs are going to be sparsed from the threshold of 0.25 to 0.75 (the step is 0.05), and the extracting feature step is implemented according to Table 2. Finally, through a non-

parametric permutation test, three different comparisons (Healthy subjects vs. EMCI, EMCI vs. LMCI and, LMCI vs. AD) are made to reveal- which method can clarify the differences properly and what exactly happens to brain functional graphs as AD progressed. Table 7 summarizes the results.

Table 7- Summary of the analyses.

<b>Parameter \ Groups</b>	<b>CN vs. EMCI</b>	<b>EMCI vs. LMCI</b>	<b>LMCI vs. AD</b>
<b>Most discriminative features</b>	Global and local efficiency, clustering, transitivity	Clustering	Clustering
<b>Least discriminative features</b>	Radius, diameter	Radius, eccentricity, diameter	Radius, eccentricity, small-worldness, CPL, global efficiency
<b>Number of significant difference based on PCC</b>	30	9	5
<b>Number of significant difference based on DC</b>	54	42	11
<b>Number of significant difference based on Kernel</b>	64	51	25
<b>Optimal threshold</b>	0.3 to 0.4	0.3 to 0.4	0.3 to 0.4

PCC analysis reveals the linear dependencies and has limitations for non-linear relationships. According to the non-linear behavior of the brain, PCC has the lowest ability to discriminate the groups based on graph features extracted from fMRI signals. Non-linear approaches have the better discriminative capability and show far more significant changes. Between them, the Kernel-based method is more powerful. As a consequence, the non-linear approaches are suggested for brain fMRI analysis. It is worthwhile mentioning that DC has no assumption but in the use of the Kernel-based method, selecting the optimal kernel function is important. In this study, the polynomial kernel was chosen according to X.

An interesting and important result shown in Table 7 is the rate of change in the various stages of the disease. In this regard, in both linear and non-linear strategies, most changes in the brain functional graphs are in the first stage of the disease. As AD progressed, the rate of variations is decreased till the last stage. This pattern is the same in all three correlation methods. Therefore, early detection of AD is crucial. The variation is completely gradual and the minimum changes belong to LMCI subjects converted to AD. Accordingly, in the first stage of the disease, there are most discriminative features including global and local efficiency, clustering, and transitivity. These features reflect both brain functional integration and segregation. As a result, in the first stage of AD, brain functions degrade significantly. As the disease progressed and EMCI subjects turn to LMCI, the clustering metric which represents the brain functional segregation has the most

significant alterations. Since features exhibiting the functional integration have less modification, the overall functional decay is also less evident in comparison to the first phase. In the last step, LMCI to AD, despite fewer alterations clustering demonstrates the most significant changes. Hence, functional segregation still declines in the last stage. In summary, with the beginning of Alzheimer's, the rate of variation in the brain functional graph is high and degenerates both brain functional integration and segregation. With further progression, the rate is declined, functional segregation is affected and the pattern remains the same until the last step. On the other hand, in every stage, some features express the least significant changes. Radius and diameter which are the minimum and maximum of eccentricity are had the least discriminative potential in all stages. These metrics display the distance of a node to a specific node. Although AD degenerates the brain functional graphs, with no significant differences there are routes to pass between two specific nodes. In other words, nodes or ROI's are not completely and significantly isolated. It may originate in the plasticity and flexibility behaviors of the brain (human body) confronting problems and pathologic circumstances. In confirmation of previous results as AD progressed the number of features exhibit no significant changes are increased.

To investigate the effect of thresholding in fMRI connectivity analysis, different thresholds were evaluated. The findings demonstrate no exact patterns as threshold modify. Nevertheless, by increasing the threshold the graphs become more sparse, and accordingly, the computational costs are lower. As a trade-off among computational cost meaningful characteristics and features, and eliminating weak and spurious links, the threshold of 0.3 to 0.4 are suggested. Therefore, the best discrimination efficiency between groups in all the three analyses belong to thresholds of 0.3 to 0.4.

## Conclusion

Although non-linear approaches are more complex to implement due to the non-linear nature of the brain, they are strongly suggested. Kernel analysis is a powerful tool wherever it can be applied but choosing the optimal kernel function is a challenge. Although there are algorithms for the optimal function, the most prevalent routine is still trial and error. Since the brain function declines more rapidly at the beginning phase of AD, and there are no specific treatments, early detection is of great importance. In this regard, nodal analysis of brain regions is highly recommended to reveal the most affected areas of the brain and understand how the AD degenerates functional brain graphs in detail.

## Declarations

- Ethical Approval:** Not applicable
- Consent to Participate:** Not applicable
- Consent to Publish:** Not applicable
- Authors Contributions:** Emad Fatemizadeh is the supervisor of the study and hypothesized the idea. Ali Motie Nasrabadi is the advisor of the project and Hessam Ahmadi implement it. Hessam Ahmadi wrote the draft and the other authors complete and correct it.
- Funding:** There is no funding for the study
- Competing Interests:** Not applicable
- Availability of data and materials:** The data and the codes are available upon request
- Conflict of interests:** All authors have seen and approved the manuscript being submitted and have no conflict of interest.

## References

- Ahmadi, H., Fatemizadeh, E., & Motie-Nasrabadi, A. (2020a). fMRI functional connectivity analysis via kernel graph in Alzheimer's disease. *Signal, Image and Video Processing*, 1-9.
- Ahmadi, H., Fatemizadeh, E., & Motie-Nasrabadi, A. (2020b). Identifying brain functional connectivity alterations during different stages of Alzheimer's disease. *International Journal of Neuroscience*, 1-13.
- Alam, M. A., Calhoun, V. D., & Wang, Y.-P. (2018). Identifying outliers using multiple kernel canonical correlation analysis with application to imaging genetics. *Computational Statistics & Data Analysis*, 125, 70-85.
- Anzellotti, S., Fedorenko, E., Kell, A. J., Caramazza, A., & Saxe, R. (2017). Measuring and modeling nonlinear interactions between brain regions with fMRI. *bioRxiv*, 074856.
- Benesty, J., Chen, J., Huang, Y., & Cohen, I. (2009). Pearson correlation coefficient. In *Noise reduction in speech processing* (pp. 1-4): Springer.
- Chao-Gan, Y., & Yu-Feng, Z. (2010). DPARSF: a MATLAB toolbox for "pipeline" data analysis of resting-state fMRI. *Frontiers in systems neuroscience*, 4.
- Frisoni, G. B., Boccardi, M., Barkhof, F., Blennow, K., Cappa, S., Chiotis, K., . . . Gietl, A. (2017). Strategic roadmap for an early diagnosis of Alzheimer's disease based on biomarkers. *The Lancet Neurology*, 16(8), 661-676.
- Harris, J. M., Hirst, J. L., & Mossinghoff, M. J. (2008). *Combinatorics and graph theory* (Vol. 2): Springer.
- Hessam Ahmadi, E. F., Ali Motie-Nasrabadi (2021). A Comparative Study of Correlation Methods in Functional Connectivity Analysis using fMRI Data of Alzheimer's Patients. *Journal of Biomedical Physics and Engineering*.
- Hlinka, J., Paluš, M., Vejmelka, M., Mantini, D., & Corbetta, M. (2011). Functional connectivity in resting-state fMRI: is linear correlation sufficient? *Neuroimage*, 54(3), 2218-2225.
- Hofmann, M. (2006). Support vector machines-kernels and the kernel trick. *Notes*, 26(3).
- Hofmann, T., Schölkopf, B., & Smola, A. J. (2008). Kernel methods in machine learning. *The annals of statistics*, 1171-1220.

- Hojjati, S. H., Ebrahimzadeh, A., & Babajani-Feremi, A. (2019). Identification of the early stage of Alzheimer's disease using structural MRI and resting-state fMRI. *Frontiers in neurology*, 10, 904.
- Hojjati, S. H., Ebrahimzadeh, A., Khazaee, A., Babajani-Feremi, A., & Initiative, A. s. D. N. (2018). Predicting conversion from MCI to AD by integrating rs-fMRI and structural MRI. *Computers in biology and medicine*, 102, 30-39.
- Humphries, M. D., & Gurney, K. (2008). Network 'small-world-ness': a quantitative method for determining canonical network equivalence. *PloS one*, 3(4), e0002051.
- Kepner, J., & Gilbert, J. (2011). *Graph algorithms in the language of linear algebra*: SIAM.
- Khagi, B., Kwon, G. R., & Lama, R. (2019). Comparative analysis of Alzheimer's disease classification by CDR level using CNN, feature selection, and machine-learning techniques. *International Journal of Imaging Systems and Technology*, 29(3), 297-310.
- Kung, S. Y. (2014). *Kernel methods and machine learning*: Cambridge University Press.
- Li, K., Guo, L., Nie, J., Li, G., & Liu, T. (2009). Review of methods for functional brain connectivity detection using fMRI. *Computerized Medical Imaging and Graphics*, 33(2), 131-139.
- Logan, B. R., & Rowe, D. B. (2004). An evaluation of thresholding techniques in fMRI analysis. *Neuroimage*, 22(1), 95-108.
- Mysterud, I. (2019). Superb book about Alzheimer's disease. In.
- Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human brain mapping*, 15(1), 1-25.
- Petersen, R. C., Aisen, P., Beckett, L. A., Donohue, M., Gamst, A., Harvey, D. J., . . . Toga, A. (2010). Alzheimer's disease neuroimaging initiative (ADNI): clinical characterization. *Neurology*, 74(3), 201-209.
- Pini, L., Pievani, M., Bocchetta, M., Altomare, D., Bosco, P., Cavedo, E., . . . Frisoni, G. B. (2016). Brain atrophy in Alzheimer's disease and aging. *Ageing research reviews*, 30, 25-48.
- Rathore, S., Habes, M., Iftikhar, M. A., Shacklett, A., & Davatzikos, C. (2017). A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. *Neuroimage*, 155, 530-548.
- Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*, 52(3), 1059-1069.
- Scholkopf, B., & Smola, A. J. (2001). *Learning with kernels: support vector machines, regularization, optimization, and beyond*: MIT press.
- Sporns, O. (2018). Graph theory methods: applications in brain networks. *Dialogues in clinical neuroscience*, 20(2), 111.
- Székely, G. J., Rizzo, M. L., & Bakirov, N. K. (2007). Measuring and testing dependence by correlation of distances. *The annals of statistics*, 35(6), 2769-2794.
- Towsley, A., Pakianathan, J., & Douglass, D. H. (2011). Correlation angles and inner products: Application to a problem from physics. *ISRN Applied Mathematics*, 2011.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., . . . Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, 15(1), 273-289.
- Uludag, K., Ugurbil, K., & Berliner, L. (2015). *fMRI: from nuclear spins to brain functions* (Vol. 30): Springer.
- Yang, B., Cao, J., Zhou, T., Dong, L., Zou, L., & Xiang, J. (2018). Exploration of neural activity under cognitive reappraisal using simultaneous eeg-fmri data and kernel canonical correlation analysis. *Computational and mathematical methods in medicine*, 2018.

# Figures

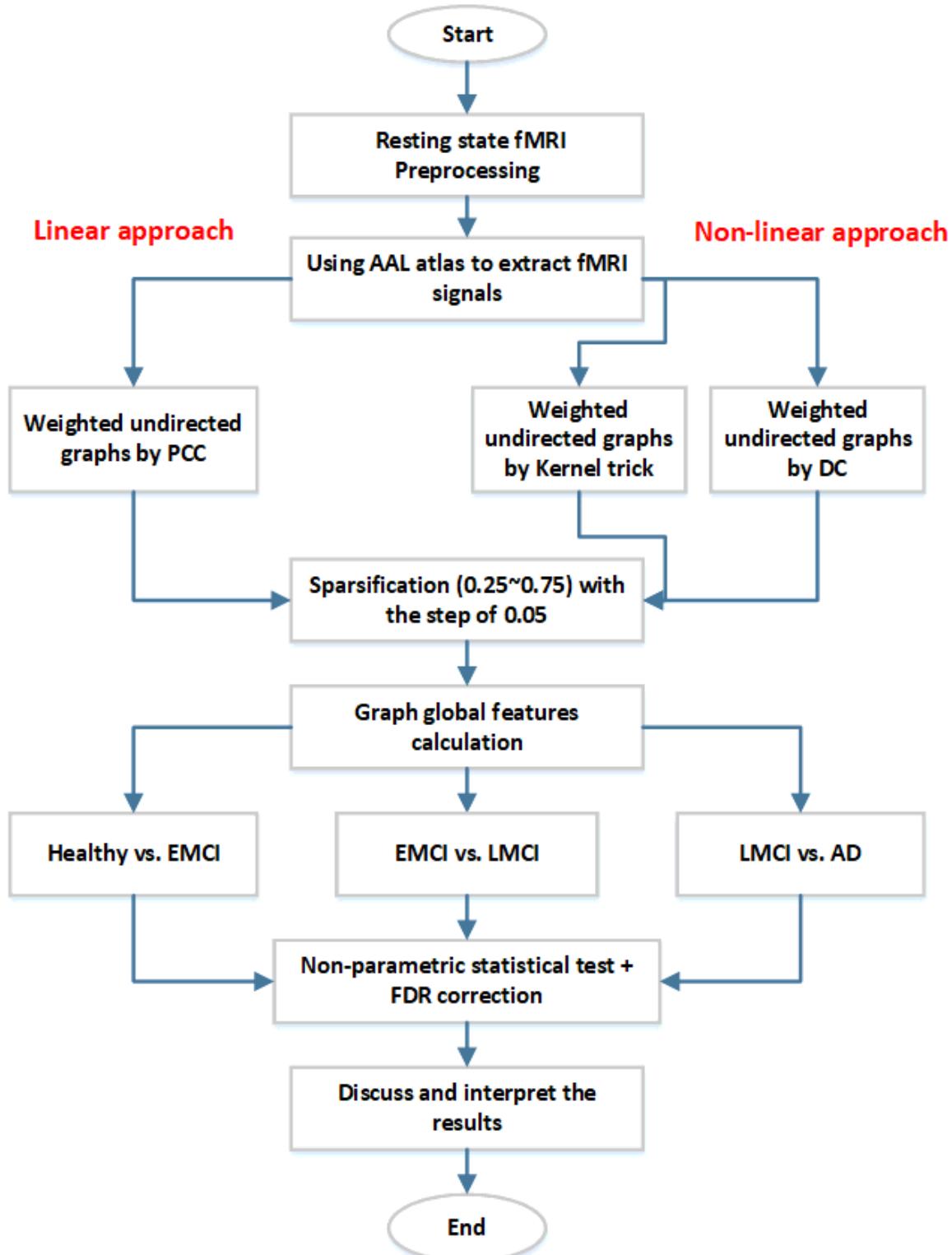
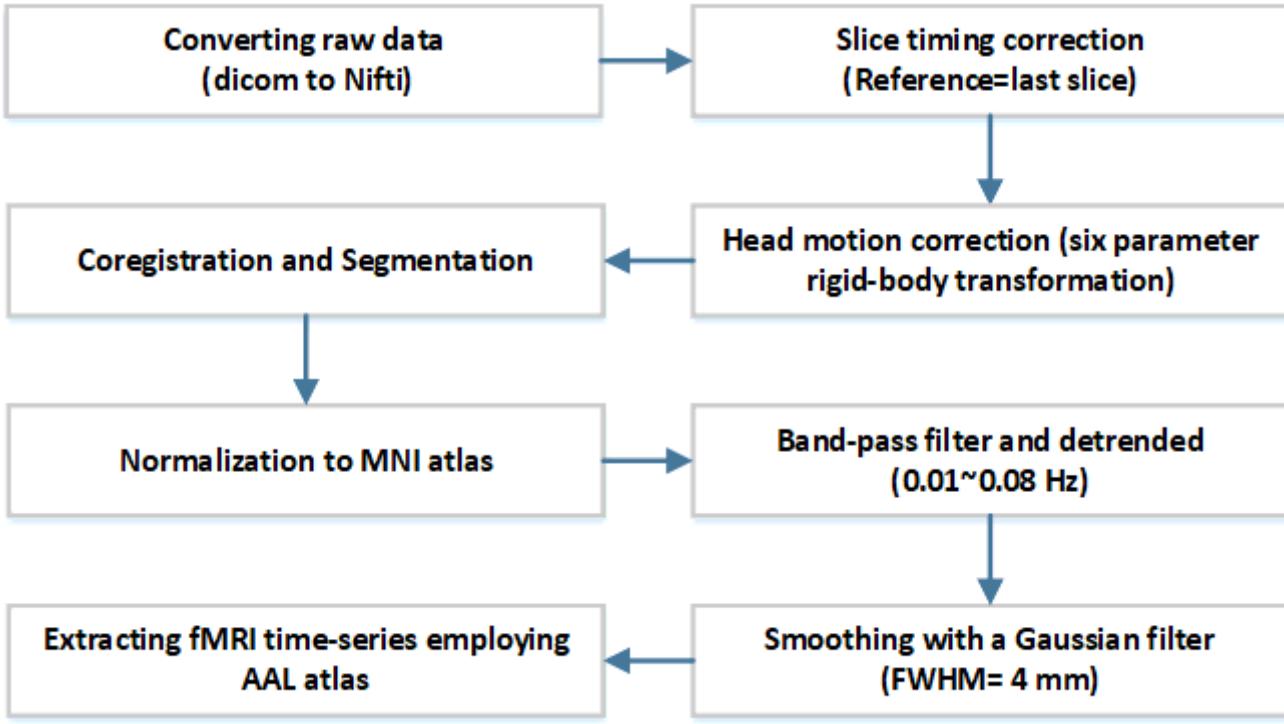


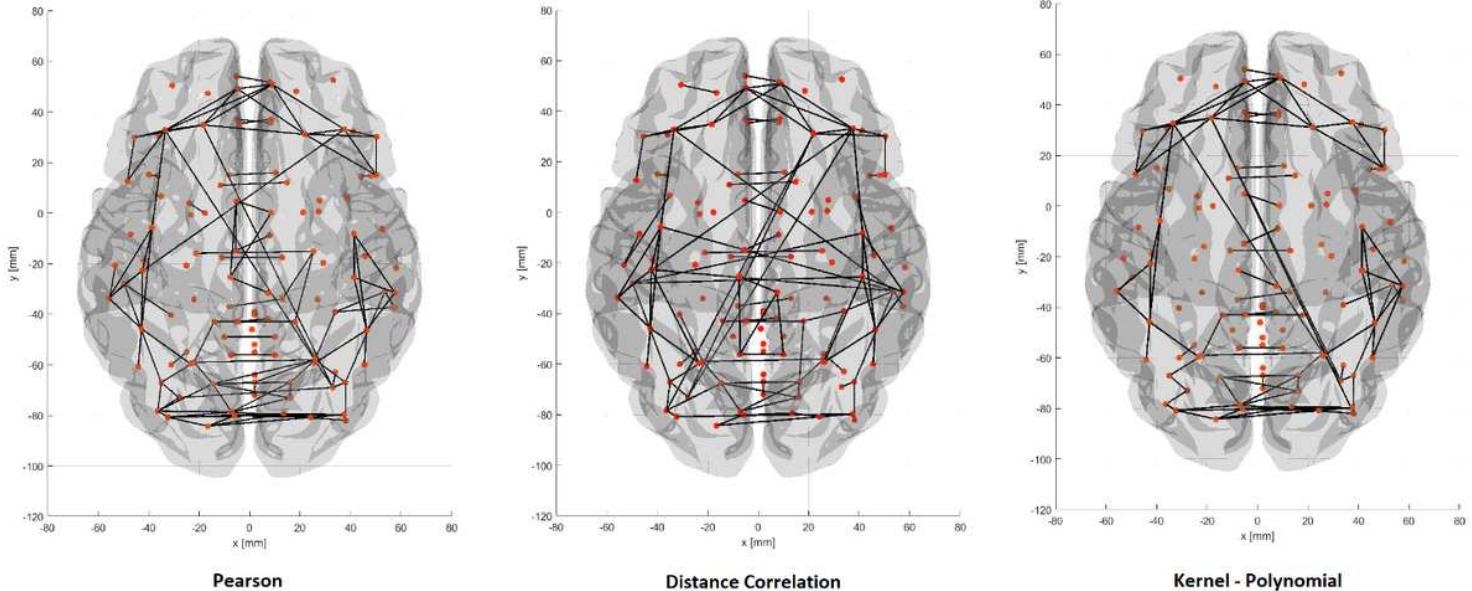
Figure 1

The research flowchart.



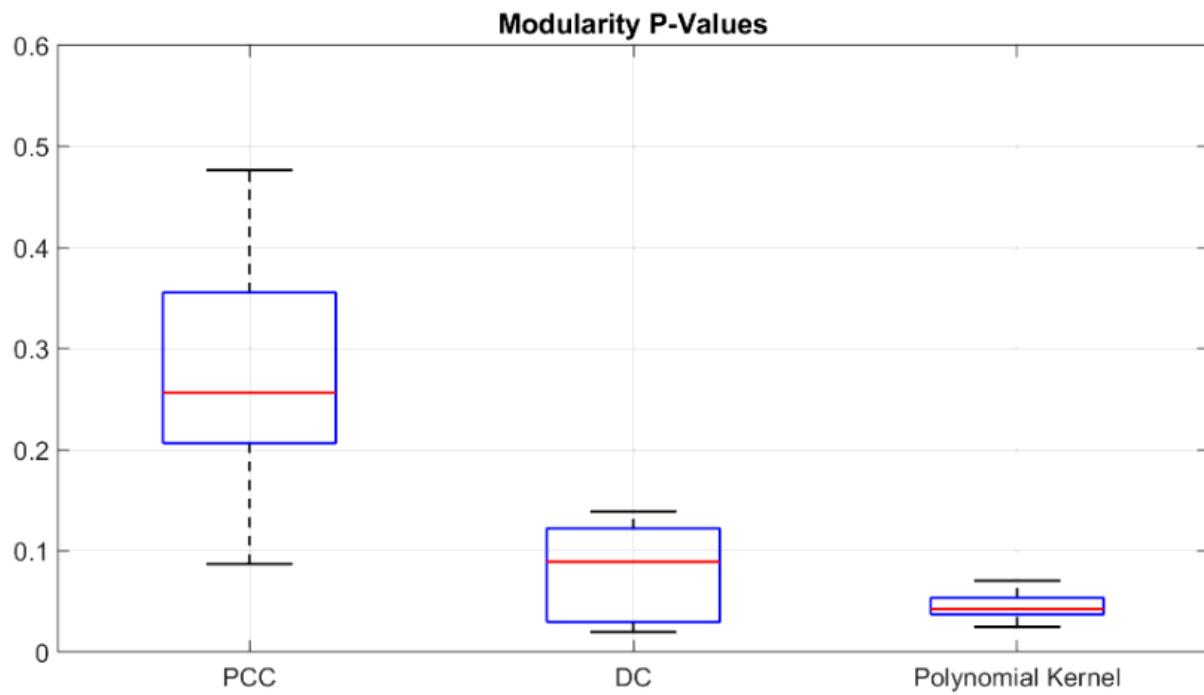
**Figure 2**

The preprocessing steps.



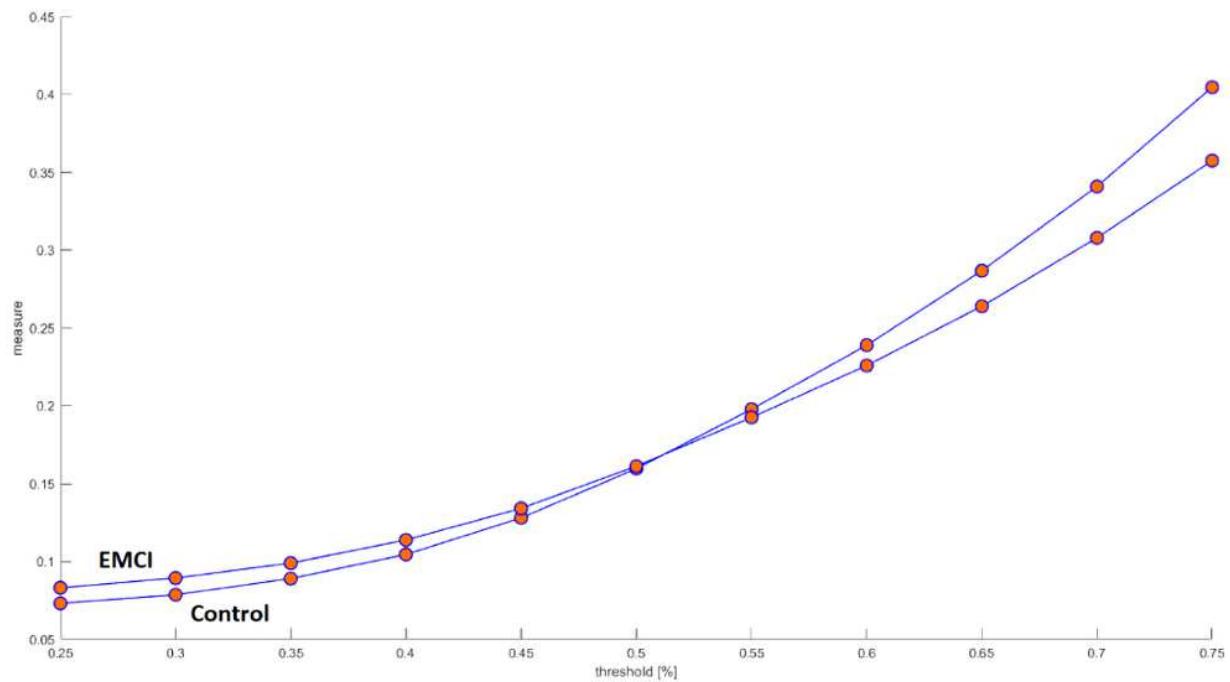
**Figure 3**

Demonstration of brain networks utilizing different methods in control subjects. All the graphs are sparsed with an 0.75 threshold.



**Figure 4**

Distribution of P-Values of modularity feature in healthy vs. EMCI subjects.



**Figure 5**

Variation of modularity measure in PCC analysis as thresholds increased.