

Original Article

Optimization of Pipeline through Preprocessing Steps Sequence Alteration using Graph Theory for Resting State fMRI

Deepa Nath¹, Anil Hiwale², Nilesh Kurwale³

^{1,2}School of ECE, Dr. Vishwanath Karad, MIT World Peace University, Pune, MH, India

³Deenanath Mangeshkar Hospital, Pune, MH, India.

¹Corresponding Author : deepa.nath@mitwpu.edu.in

Received: 19 December 2022

Revised: 07 February 2023

Accepted: 12 March 2023

Published: 25 March 2023

Abstract - The resting-state functional magnetic resonance imaging (rs-fMRI) is an advanced imaging method with several benefits compared to functional magnetic resonance imaging (fMRI) techniques. Major benefits of rs-fMRI are easy receiving of the signal, needs minimum effort from the patient, and good at differentiating the functional areas in patients. Resting-state fMRI (rs-fMRI) identifies the amount of Blood Oxygen Level-called as BOLD. These are variations at frequencies smaller than 0.1 Hz to identify the functional variations of the brain. Usage of this BOLD level allows viewing resting-state networks (RSN). RSN are distinct areas of the brain which are spatially distributed and which confirm synchronous BOLD fluctuations at rest. In the pipeline for fMRI data, various preprocessing steps are carried out as a normal procedure for achieving a better quality of data. In this paper, these steps for preprocessing are altered/changed in the pipeline of CONN neuroimaging software, and its effects on graph theory outcomes are explored and discussed.

Keywords - Resting-state functional magnetic resonance (rs-fMRI), Functional Connectivity (FC), Blood Oxygen Level Dependent (BOLD).

1. Introduction

The areas with correlated activity can form functional networks at rest by using rs-fMRI [1] identification of the temporal correlation of naturally occurring BOLD signals [2]. These signals are found among the brain regions. Functional connectivity (FC) is a measure to identify the connection among the regions of interest with the help of neural connection [3-5]. Few neuroimaging softwares are developed in recent years which can perform analysis. These have been explored to check the brain patterns by looking out for functionally connected networks and thus exploring a more detailed examination of functional connectivity. Such studies are performed while the brain is at rest. They have shown the functional time series and coactivation levels among the brain areas that are functionally connected. They indicate the possible functional links between various brain regions at rest.

The brain of a human being is a highly complex grid of functionally interacting brain regions. It can depict deeper findings of different communications in the neural areas. It establishes a way to evaluate the connectivity patterns that are functionally linked and information that reveals human behavior and how these connections might have been altered in neurological disorders [4][5]. fMRI data needs to be preprocessed, which involves several steps for achieving a

better quality of signal and Signal to Noise ratio. Available Neuroimaging softwares has many steps for achieving preprocessing of the acquired signal. In this paper, one tries to explore the different steps sequence. Generally, a pipeline is suggested by the software. However, which step sequence following would lead to a better outcome is not explored much. Section 1 is the introductory part. Section 2 describes the fundamentals of rs-fMRI. Section 3 explains all the methods available for determining the functional correlations. Section 4 explains the workflow for fMRI data. Section 5 describes the steps of alteration carried out in the pipeline for preprocessing the data, and the results are discussed.

2. Fundamentals of rs-fMRI

Signal frequency variations are generated at the level below 0.1Hz in the rs-fMRI due to neural activity [1]. During the process of acquisition, several situations can occur. One such example is when a subject undergoes rs-fMRI, he/she can have possible movement of his/her head. This movement can trigger misleading outcomes during the processing of data [5][6]. These false results and other things can be corrected through the preprocessing steps involved in any neurological software for analysis. Thus, preprocessing and its steps are essential pre-requisite to the processing and analysis of the collected fMRI data.



For the processing and analysis of the collected fMRI data [6] data, several alternative methods are available. Each method has its pros and cons.

A summary of the steps involved in preliminary processing is elaborated in detail. Additionally, the methods applied for measuring this functional connectivity are discussed. Several programs like Analysis of Functional Neuro Images (AFNI), Free Surfer, Statistical Parametric Mapping (SPM), Free Surfer, DPABI Surf, CONN (Functional Connectivity Toolbox), and Configurable Pipeline for the Analysis of Connectomes (CPAC) are available for analysis. These listed softwares can process and improve acquisition quality by motion correction and noise removal. [7,8,9]

3. Methods for Determining the Functional Correlation

A signal coming just from a specific voxel or cluster of voxels is known as the seed or ROI, which is used as a source voxel to determine associations with other brain voxels.

Establishing the connections between various brain regions at rest requires a comprehensive understanding of the regions in the brain (for example, which seeds are functionally related). For computing functional correlation [6], several model-based and model-free approaches are involved in detecting the alterations in the human brain observed in rs-fMRI BOLD signals [3].

The model-based methods listed here are Seed Based on Connectivity and Region of Interest to Region of Interest. (ROI-ROI). The model-free techniques covered include clustering approaches, Graph Theory techniques, Independent Component Analysis, and Principal Component Analysis.

3.1. Model-Dependent Methods

3.1.1. Seed-Based Analysis

The first and simplest way is seed-based analysis, in which a seed is chosen as a reference and mapped with another seed (voxel) throughout the entire brain to produce a seed-based connectivity map. This is the simplest method; its implementation makes it the most popular approach for rs-fMRI.

The initial technique used by Biswal[1] et al. to identify the resting state networks was seed-based analysis. A seed-based analysis is simple, easy, and can give broad results. The seed-based approach focuses on selecting a seed region and attempting to connect it to the other seeds across the remainder of the brain. The Seed selection needs to be selected by an expert resource as the results depend on the seed choice. As shown in Figure 1, seed-based connectivity is observed in the brain map.



Fig. 1 Seed-Based Connectivity–Software used CONN [17] platform

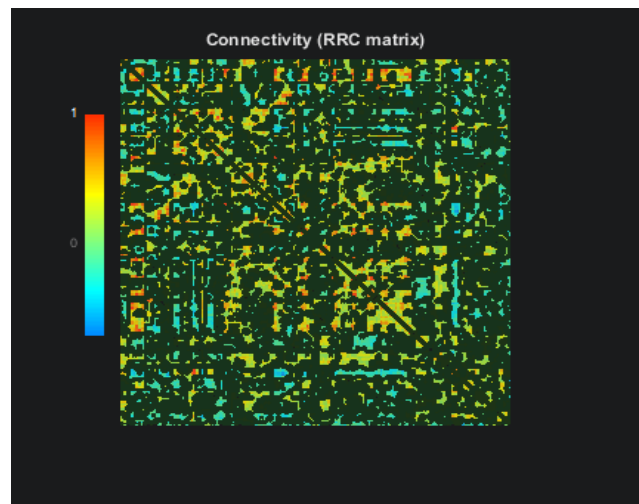


Fig. 2 Connectivity matrix with ROI to ROI implemented in CONN [17] Software

3.1.2. Region of Interest to Region of Interest (ROI –ROI Analysis)

ROI-to-ROI connectivity measures the connectivity among all groups of ROIs between a set of regions in the brain. The RRC matrix shown below in Figure 2 represents the level of functional connectivity amongst each pair of ROIS. When the simultaneous study of the entire network of connections is needed, such a method will be quite promising. [10]

3.2. Model-free Methods

3.2.1. Principal Component Analysis (PCA)

Calculating the principal components and using them to make changes based on the data is known as principal component analysis (PCA). Sometimes, only the top few principal components are used, and the remaining ones are ignored. The principal components are sorted out using mutual orthogonality between signal and noise to separate the fMRI data into signal and noise subspaces. It works on how to maximize the difference. Generally, a few directions where most of the variation occurs are interesting for analysis. [3]

3.2.2. Independent Component Analysis (ICA)

The ICA method is another most popular and widely used method, which uses a mathematical technique to get the statistical choice within its components [11]. As a result, various reliable resting-state networks are discovered in numerous distinct individual investigations employing ICA on rs-fMRI data. It focuses on maximizing the statistical impartiality among its components to detect functional connectivity. It is proven to give consistent results with good accuracy.

3.2.3. Graph Theory Processing

The graph method uses the nodes and edge theory, where the network behaves like node (ROI) sets with edges that reflect association among ROIs.[12] It works as an alternative approach for seed-based connectivity [13]. For Graph theory [11], the node in the brain corresponds to a region of interest, and the edges represent the association between them. The graph theory displays parameters like global efficiency, local efficiency, degree, cost, and many others. This network's overall information transfer efficiency is gauged by its global efficiency. Local efficiency is inversely proportional to the average path length between all neighbors. The length of the nodes' paths is what is referred to as the average path length. The clustering coefficient measures how closely nodes in a graph typically cluster together. (neighboring node connections) [16] These graph theory attributes help provide some statistical data for calculating the association among the nodes [15]. These metrics of graph theory, as listed above, like Global Efficiency and Local Efficiency, help to calculate the capability of a network to communicate information at the global and local levels, respectively.[16]

Global Efficiency

In a graph, say $G(N, K)$, the global efficiency is defined as where N are vertices and K are edges connecting the pair vertices.

$$G_{E_i} = \frac{\sum_{j \neq i} \frac{1}{F_{i,j}}}{N-1} \quad (1)$$

where G_{E_i} is global efficiency, F is the shortest path distance matrix, N is the number of nodes.

Local Efficiency

Local Efficiency is a metric that helps to calculate the average efficiency of information transfer within local subgraphs or neighborhoods. It is measured with respect to the shortest average path length within adjacent neighbors.

Clustering Coefficient

The clustering coefficient in graph theory measures how closely connected nodes in a graph tend to be. It measures the fraction of a node's neighbors. Here, considering then the equation for clustering coefficient C_{c_i} is

$$C_{c_i} = \sum_{j,k \in T_i} \frac{A_{j,k}^{(i)}}{d_i(d_i-1)} \quad (2)$$

where d is the degree of a node, and A is the adjacency matrix.

Cost

On a graph, the total cost of the edges that make up a path represents the path's cost. The cheapest path between two nodes in the path between them has the lowest cost. Network cost is an imperative metric that succinctly combines with network efficiency. This measures how expensive it is to build a network.

Average Path Length

The average distances between node i and every other node are what is known as the average path length.

$$L = \frac{\sum_{i \in N} L_i}{n} \quad (3)$$

Degree

In theory, for graph study, the degree of a vertex is determined by how many edges connect it.

3.2.4. Clustering Methods

This method performs division in the brain within a set of groups or regions based on functional connectivity linkage. Such methods are acceptable for determining the linkage in the distinct regions in the brain. [16]

These methods can describe the active regions and find networks in the brain and fMRI data proficiently without the need for previous data about activation arrays or trials.

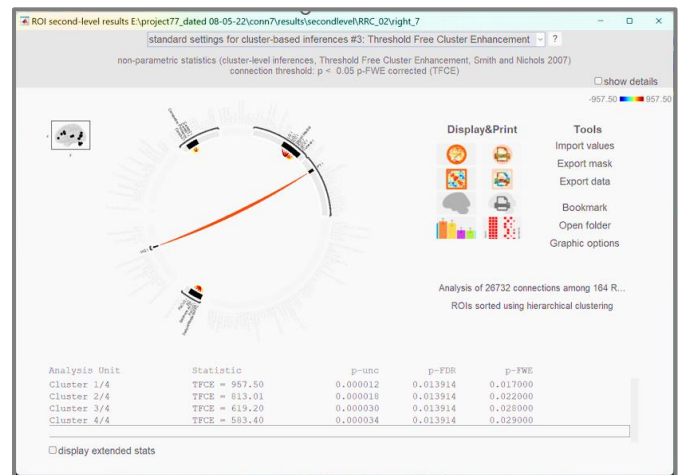


Fig. 3 Cluster-based results

4. fMRI Workflow

The general steps involved in this fMRI analysis pipeline means the steps involved for preprocessing the data, are summarized in the following chart. The steps involved in preprocessing fMRI are illustrated in Figure 4.

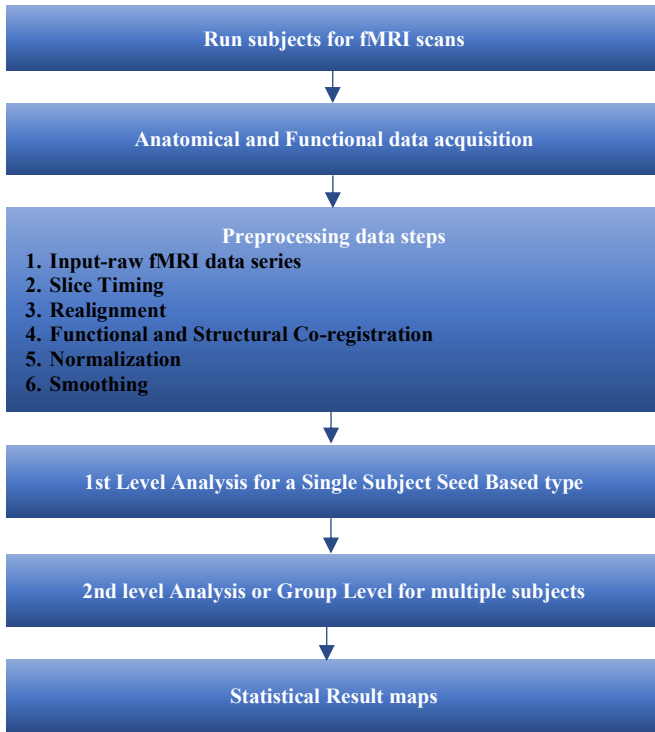


Fig. 4 fMRI Pipeline Workflow

Steps involved in preprocessing fMRI are:

4.1. Functional Realignment

Functional data is realigned, and it is unwrapped in this step. By using b-spline interpolation, all the given images are coregistered, and a reference template is used for resampling. The step focuses on the vulnerability of disturbances in images triggered by interactive movements. This is done by approximating the outcomes of the distortion zone concerning the motion of the head. Further, the resampling is performed on functional data to equalize the distorted area of the original template image. [17]

4.2. Slice-Timing Correction

Time-based dislocation among diverse portions of the functional data fetched via the fMRI capturing methods serially is extracted by using the slice-timing correction (STC) process. This is done as we keep moving the functional data on the timeline and for repeat sampling.

4.3. Outlier Identification

Here the representative data is identified, probably out of specification from the gathered universal BOLD signal and the number of scanned subject motions. The BOLD signal received above the time greater than five seconds and displacement per frame greater than 1.0 mm are termed probable out-of-specification data.

4.4. Segmentation in a Direct Manner and Normalization

The brain images obtained during rs-fMRI have images in functional and anatomical forms. Both are standardized into

the regular template like Montreal Neurological Institute (MNI space). This is a standard template. Here, the brain is divided further into areas like white, grey regions, and CSF tissue using SPM software. It performs the classification of tissues, subsequent tissue approximation, and probability maps (TPMs) of the standard functional/anatomical image from the intensity values and registration iteratively. Thus, it helps to estimate the nonlinear three-dimensional conversion best resembling the next image.

4.5. Functional Smoothing

The BOLD signal-to-noise ratio is improved by smoothing, using a three-dimensional convolution method that applies a Gaussian kernel with a dimension of 8mm full-width half maximum (FWHM). This, in turn, helps to decrease the effect of remaining roughness in functional and gyral composition among subjects.

Several neuroimaging Softwares like AFNI, FSL, Free surfer, fMRIPrep, and CONN [17] are available for performing these predefined steps like functional rearrangement and unwarp, improvement in slice-timing, identification of probable out-of-specification images, direct segmentation and normalization and functional smoothing of the images [18].

5. Experimentation Results

The acquired fMRI data must be processed to achieve a better Signal Noise Ratio (SNR). These processing steps are performed one after another to achieve a better output of the fMRI images. While using CONN [17] software, a default pipeline is specified in processing the output. In this paper, various results achieved on CONN software are displayed.

Figure 5 below shows the selection of the default pipeline-the sequence of processing the images -the standard steps that are followed for processing the fMRI images.

After selecting this option, the default pipeline is implemented, and processing is carried out as stated in the fMRI Workflow in Figure 4.

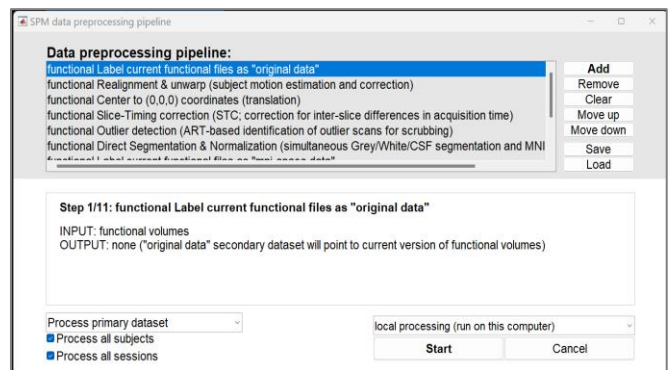


Fig. 5 Example of CONN's preprocessing GUI [17]

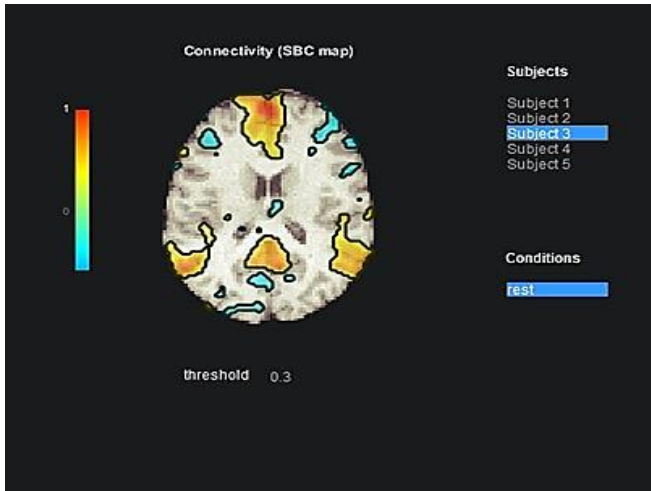


Fig. 6 CONN's [17] default pipeline SBC connectivity maps by using Default pipeline direct normalization method by using Strategy A.

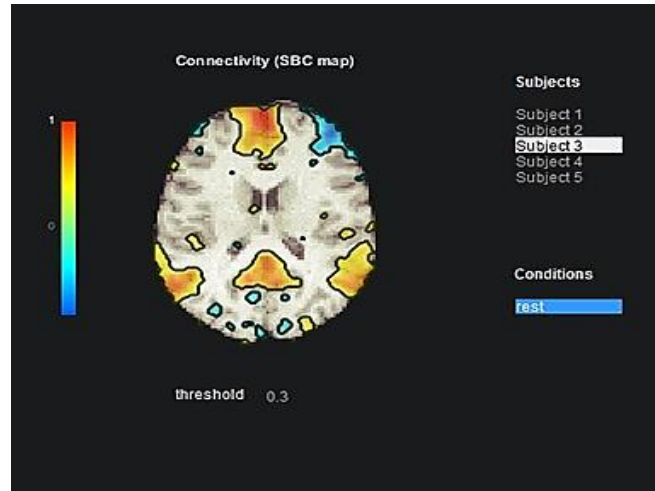


Fig. 7 Sequence of steps altered in CONN [17] default pipeline- SBC connectivity maps by using Strategy B

Figure 6 demonstrates the results achieved by implementing the default pipeline; results are displayed with a certain p-value from CONN GUI [17] with a threshold value of 0.3, 1st level analysis results obtained, as shown in Figure 6. Results show functional connectivity achieved using seed-based connectivity.

5.1. Different Strategies using Sequence Alterations

Different strategies [19] (preprocessing steps sequence as shown in Figure 4, fMRI workflow diagram, block 3) were applied to the same subject to see the effect of alteration/change of sequence of preprocessing steps on the measurement metrics to graph theory measures.

5.1.1. Strategy A

In this pipeline method, the realignment step is performed first, followed by slice timing correction. Direct segmentation and normalization are applied in this step.

5.1.2. Strategy B

Here, in this pipeline, the slice timing step is performed first, followed by the realignment step. The sequence of execution is altered.

5.1.3. Strategy C

In this step, first, realignment is done, followed by slice timing like Strategy A. However, indirect segmentation and normalization are applied. Indirect segmentation and normalizations mean that the functional and anatomical datasets use the same non-linear transformation, estimated using only the structural data.

Experimentation is done in the above-mentioned strategies and tried altering the step sequences for evaluating the performance.

Strategy A output, when done on CONN [17], as shown in Figure 6, is with the standard default pipeline steps followed.

Normally in default pipeline steps, firstly, functional realignment is performed. However, here in Strategy B- firstly, the slice-timing correction is done, followed by the functional realignment step.

Further steps like outlier identification, direct segmentation and normalization, and functional smoothing are kept as the standard ones. While doing so, no significantly different results could be seen.

The results are displayed in Figure 7. If we try comparing Figure 6 and Figure 7, the results look the same.

For Strategy C, one major change deployed is the indirect normalization step. The functional and anatomical data are normalized independently, each using its own non-linear transformation for data from its origin to a common MNI-space in a direct normalization procedure.

On the other hand, in an indirect normalization procedure, both the functional and anatomical datasets use the same non-linear transformation, which is estimated using only the structural data.

Table 1. Average values of different parameters for different preprocessing steps

Method	Global efficiency	Local efficiency	Cost	Average Path Length	Degree	Clustering Co-efficient
Strategy -A	0.496933	0.764214	0.247444	1.759636	40	0.530746
Strategy- B	0.477607	0.772052	0.222086	1.784774	36	0.549282
Strategy- C	0.512849	0.758258	0.143149	2.205747	23	0.526241

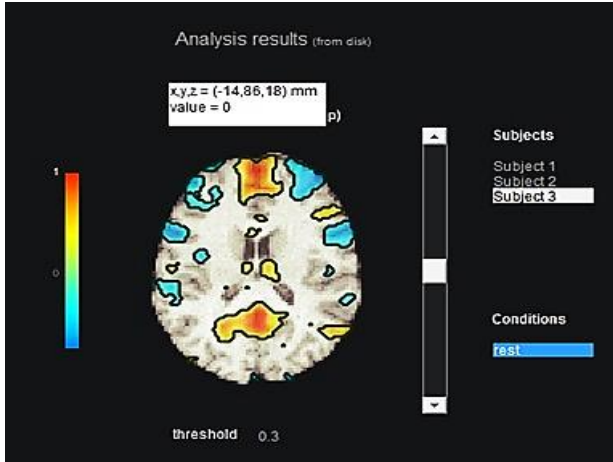


Fig. 8 Output Results by implementing indirect normalization in CONN [17] by using Strategy C

This is significantly better in terms of resolution and improved contrast of tissues. Rest all the further pipeline steps are the same as the default one. The only important precondition for implementing this method is that high-quality Field maps (fmaps) are necessary. Field maps (fmaps) information is related to the static magnetic field distribution for getting precise geometric distortion in MRI. These fmaps are needed for susceptibility distortion correction during the functional realignment & unwarp step.

The important point to note in Strategy C is that after maintaining the same threshold (p-value) at 0.3 for the same subject and the same seed- Connectivity varies significantly with Indirect Normalization [20]. Results implemented by Strategy C are shown in Figure 8.

In Figure 6,7, we cannot see many variations in the output; however, in Figure 8, we can see a significant difference is attained by implementing Strategy C. Apart from the connectivity differences observed, graph theory statistical measures also differ for the three strategies applied.

5.2. Graph Theory Results

Here dataset is taken from the open-neuro forum for resting-state fMRI analysis. The average of 46 subjects' data is compared in the given table on the parameters of graph

theory like global efficiency, clustering coefficient, average path length, local efficiency, cost and degree.

Based on data in Table 1 derived from the graph theory results using CONN software, it is observed that the parameters like Global Efficiency, Local Efficiency, Cost, Average Path length, Degree, and Clustering Coefficient in different strategies applied by changing the preprocessing step sequence, shows statistically significant different outcomes. Values vary significantly. In the case of Strategy C –Global efficiency seems to be improved, while in the case of Strategy, A Average path length is lesser, and the degree is higher in the same method. These results need to be further validated with clinical evidence for further inferences. The future scope for this research is validating the results obtained using CONN [17] from an expert neurosurgeon. Further, these results with the help of machine learning [21] for fMRI can be made easier for the classification of neurological diseases.

6. Conclusion

While performing preprocessing steps on the fMRI images, generally, the method followed is applying normalization in a direct manner. Indirect normalization seems to be a better alternative approach which is used to get significant outcomes as compared to the direct method [19] when done through visualization. They all were implemented on CONN [17] software. This approach can be used for checking the resting state networks for different neurological diseases like Epilepsy, Alzheimer's, Parkinson and other brain disorders [22]. To summarize both Direct and Indirect Normalization methods with different strategies, preprocessing steps sequences were altered. To further continue unique method was tried to implement, and substantial differences in measuring parameters of the brain through graph theory results were observed. From Table 2, global efficiency changed by 3% in Strategy Cover the direct method Strategy A. Similarly, the Degree metric seems to be altered by 40%, and the Cost metric also looks altered by 2%. The same needs to be further clinically validated.

Acknowledgments

fMRI analysis -the idea was initiated by Nilesh Kurwale. The work done was supervised and administered by Anil Hiwale and Nilesh Kurwale.

References

- [1] B. Biswal et al., "Functional Connectivity in the Motor Cortex of Resting Human Brain using Echo-planar MRI," *Magnetic Resonance in Medicine*, vol. 34, no. 4, pp. 537–541, 1995. [CrossRef] [Google Scholar] [Publisher link]
- [2] S.R. Bhagyashree et al., "Review on Technological Aspects of Magnetic Resonance Imaging and Functional Magnetic Resonance Imaging," *International Journal of Engineering Trends and Technology*, vol. 59, no. 2, pp. 96-104, 2018. [CrossRef] [Google Scholar] [Publisher link]
- [3] Mikail Rubinov, and Olaf Sporns, "Complex Network Measures of Brain Connectivity: Uses and Interpretations," *NeuroImage*, vol. 52, no. 3, pp.1059–1069, 2010. [CrossRef] [Google Scholar] [Publisher link]
- [4] K.A. Smith et al., "Resting State fMRI: A Review on Methods in Resting State Connectivity Analysis and Resting State Networks,"

Neuroradiology Journal, vol. 30, no. 4, pp. 305–317, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher link](#)]

- [5] Jose M. Soares et al., “A Hitchhiker’s Guide to Functional Magnetic Resonance Imaging,” *Frontiers in Neuroscience*, vol. 10, 2016. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher link](#)]
- [6] Martijn P. van den Heuvel, and Hilleke E. Hulshoff Pol, “Exploring the Brain Network: A Review on Resting State fMRI Functional Connectivity,” *European Neuropsychopharmacology*, vol. 20, no. 8, pp. 519-534, 2010. . [[CrossRef](#)] [[Publisher link](#)]
- [7] Carl D. Hacker et al., “Resting-state Network Mapping in Neurosurgical Practice: A Review,” *Neurosurgical Focus*, vol. 47, no. 6, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher link](#)]
- [8] Chao-Gan Yan, Xin-Di Wang, and Bin Lu, “DPABISurf: Data Processing & Analysis for Brain Imaging on Surface,” *Science Bulletin*, vol. 66, no. 24, pp. 2453-2455, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher link](#)]
- [9] Victoria L. Morgan et al., “Comparison of fMRI Statistical Software Packages and Strategies for Analysis of Images Containing Random and Stimulus-Related Motion,” *Computerized Medical Imaging and Graphics*, vol. 31, no. 6, pp. 436-446, 2007. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher link](#)]
- [10] Ali-Mohammad Golestani, and Bradley G. Goodyear, “Regions of Interest for Resting State fMRI Analysis Determined by Inter-voxel Cross Correlation,” *Neuroimage*, vol. 56, no. 1, pp. 246-251, 2011. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher link](#)]
- [11] Fabrizio Esposito et al., “Independent Component Model of the Default-mode Brain Function: Combining Individual-level and Population-level Analyses in Resting-state fMRI,” *Magnetic Resonance Imaging*, vol. 26, no. 7, pp. 905–913, 2008. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher link](#)]
- [12] Ed Bullmore, and Olaf Sporns, “Complex Brain Networks: Graph Theoretical Analysis of Structural and Functional Systems,” *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186-198, 2009. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher link](#)]
- [13] John D. Medaglia, “Graph Theoretic Analysis of Resting State Functional MR Imaging,” *Neuroimaging Clinics of North America*, vol. 27, no. 4, pp. 593-607, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher link](#)]
- [14] Jinhui Wang, Xinian Zuo, and Yong He, “Graph-based Network Analysis of Resting-state Functional MRI,” *Frontiers in Systems Neuroscience*, vol. 4, 2010. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher link](#)]
- [15] H. Lv et al., “Resting-State Functional MRI: Everything That Nonexperts Have Always Wanted to Know,” *AJNR American Journal of Neuroradiology*, vol. 39, no. 8, pp. 1390-1399, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher link](#)]
- [16] Bertrand Thirion et al., “Which fMRI Clustering Gives Good Brain Parcellations?,” *Frontiers in NeuroScience*, vol. 8, 2014. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher link](#)]
- [17] Susan Whitfield-Gabrieli, and Alfonsa Nieto-Castanon, “Conn: A Functional Connectivity Toolbox for Correlated and Ant Correlated Brain Networks,” *Brain Connectivity*, vol. 2, no. 3, pp. 125-141, 2012. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher link](#)]
- [18] Jonathan D. Power et al., “Spurious but Systematic Correlations in Functional Connectivity MRI Networks Arise from Subject Motion,” *Neuroimage*, vol. 59, no. 3, pp. 2142-2154, 2012. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher link](#)]
- [19] Fatma Gargouri et al., “The Influence of Preprocessing Steps on Graph Theory Measures Derived from Resting State Computational fMRI,” *Frontiers in Computational Neuroscience*, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher link](#)]
- [20] Vince D. Calhoun et al., “The Impact of T1 versus EPI Spatial Normalization Templates for fMRI Data Analyses,” *Human brain mapping*, vol. 38, no. 11 pp. 5331–5342, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher link](#)]
- [21] Ashwini S. Savanth, and P.A. Vijaya, “Artificial Neural Networks for fMRI Data Analysis: A Survey,” *International Journal of Engineering Trends and Technology*, vol. 49, no. 8, pp. 487-494, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher link](#)]
- [22] D. Nath, A. Hiwale and N. Kurwale, "Resting-State Functional Connectivity Analysis for diagnosis of Epilepsy patients," *2021 2nd International Conference for Emerging Technology (INCET)*, Belagavi, India, pp. 1-5, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher link](#)]