



Ensemble Classification of Alzheimer's Disease and Mild Cognitive Impairment Based on Complex Graph Measures from Diffusion Tensor Images

Ashkan Ebadi^{1†}, Josué L. Dalboni da Rocha^{2†}, Dushyanth B. Nagaraju³, Fernanda Tovar-Moll^{4,5}, Ivanei Bramati⁴, Gabriel Coutinho^{4,5}, Ranganatha Sitaram^{6,7*} and Parisa Rashidi^{1*}

OPEN ACCESS

Edited by:

Pedro Antonio Valdes-Sosa,
Joint China Cuba Lab for Frontiers
Research in Translational
Neurotechnology, Cuba

Reviewed by:

Alexis Roche,
Siemens Healthcare/CHUV,
Switzerland
Claudio Babiloni,
Sapienza University of Rome, Italy

*Correspondence:

Ranganatha Sitaram
rsitaram@uc.cl
Parisa Rashidi
parisa.rashidi@ufl.edu

[†]These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 19 April 2016

Accepted: 26 January 2017

Published: 28 February 2017

Citation:

Ebadi A, Dalboni da Rocha JL,
Nagaraju DB, Tovar-Moll F, Bramati I,
Coutinho G, Sitaram R and Rashidi P
(2017) Ensemble Classification of
Alzheimer's Disease and Mild
Cognitive Impairment Based on
Complex Graph Measures from
Diffusion Tensor Images.
Front. Neurosci. 11:56.
doi: 10.3389/fnins.2017.00056

¹ Department of Biomedical Engineering, University of Florida, Gainesville, FL, USA, ² Brain and Language Lab, Department of Clinical Neuroscience, University of Geneva, Geneva, Switzerland, ³ Department of Computer and Information Science and Engineering, University of Florida, Gainesville, FL, USA, ⁴ D'Or Institute for Research and Education (IDOR), Rio de Janeiro, Brazil, ⁵ Institute for Biomedical Sciences, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil, ⁶ Institute for Biological and Medical Engineering, Schools of Engineering, Biology and Medicine, and Department of Psychiatry and Section of Neuroscience, Pontificia Universidad Católica de Chile, Santiago, Chile, ⁷ Laboratory for Brain-Machine Interfaces and Neuromodulation, Pontificia Universidad Católica de Chile, Santiago, Chile

The human brain is a complex network of interacting regions. The gray matter regions of brain are interconnected by white matter tracts, together forming one integrative complex network. In this article, we report our investigation about the potential of applying brain connectivity patterns as an aid in diagnosing Alzheimer's disease and Mild Cognitive Impairment (MCI). We performed pattern analysis of graph theoretical measures derived from Diffusion Tensor Imaging (DTI) data representing structural brain networks of 45 subjects, consisting of 15 patients of Alzheimer's disease (AD), 15 patients of MCI, and 15 healthy subjects (CT). We considered pair-wise class combinations of subjects, defining three separate classification tasks, i.e., AD-CT, AD-MCI, and CT-MCI, and used an ensemble classification module to perform the classification tasks. Our ensemble framework with feature selection shows a promising performance with classification accuracy of 83.3% for AD vs. MCI, 80% for AD vs. CT, and 70% for MCI vs. CT. Moreover, our findings suggest that AD can be related to graph measures abnormalities at Brodmann areas in the sensorimotor cortex and piriform cortex. In this way, node redundancy coefficient and load centrality in the primary motor cortex were recognized as good indicators of AD in contrast to MCI. In general, load centrality, betweenness centrality, and closeness centrality were found to be the most relevant network measures, as they were the top identified features at different nodes. The present study can be regarded as a "proof of concept" about a procedure for the classification of MRI markers between AD dementia, MCI, and normal old individuals, due to the small and not well-defined groups of AD and MCI patients. Future studies with larger samples of subjects and more sophisticated patient exclusion criteria are necessary toward the development of a more precise technique for clinical diagnosis.

Keywords: Alzheimer's disease, ensemble classification, diffusion tensor images, machine learning, graph measures

INTRODUCTION

Alzheimer's disease (AD), a neurodegenerative disorder characterized by progressive dementia, is the seventh leading cause of death in the United States (Heron and National Center for Health Statistics, 2009). AD is the most common form of dementia and currently affects over five million Americans. This number will grow to as many as 14 million by year 2050 (Brookmeyer et al., 2007). AD may affect people in different ways, but the most common first symptom is the inability to remember new information (Burns and Iliffe, 2009). The progression of AD toward other brain regions causes disruption of the daily life, changes in personality, and withdrawal from work and social activities. In advanced stages, the individual is unable to communicate with others and to recognize close relatives. Finally, difficulty executing motor tasks precedes death (Alzheimer's Association, 2011). There is currently no cure for AD, while most drugs only alleviate the symptoms (Jack et al., 2008).

The transitional stage in which the patient is not considered normal, but does not meet the criteria for dementia is called Mild Cognitive Impairment (MCI). MCI consists of heterogeneous symptomatology, and includes prodromal AD stages as well as mild stages of other types of dementing disorders (Dubois and Albert, 2004). Prodromal AD brains have partial similarities to more severe AD brains, while mild stages of other types of dementing disorders have different features by which it would be possible to distinguish non-AD MCI patients from the AD patients.

According to the International Working Group-2 (IWG-2) criteria by Dubois et al. (2014), the diagnostic biomarkers for AD are the clinical phenotype (typical or atypical presentations of AD related neuropsychological deficits), as well as the pathophysiologic presence of extracellular amyloid plaques and intra-neuronal neurofibrillary tangles, as defined by Alzheimer (1907), which are associated with synaptic loss, functional neurotransmission deficits, and neuronal death. The pathophysiologic markers of AD can be accessed by positron emission tomography (PET; Drzezga et al., 2003) and cerebrospinal fluid tests (Tapiola et al., 2009). Topographic biomarkers such as F-2-fluoro-2-deoxy-D-glucose (PET-FDG; Bohnen et al., 2014) and Magnetic Resonance Imaging (MRI) hippocampal volumetry (Sarazin et al., 2010) are proposed to track disease progression.

The diagnosis of MCI due to AD is possible nowadays (Alzheimer's Association, 2011; Dubois et al., 2014), based on cognitive impairments with mild impact on the daily activities, and considering some excluding clinical criteria. The diagnosis is also possible based on positive physio-pathologic markers of AD such as an abnormal level of amyloid beta and/or tau in the cerebrospinal fluid (CSF), or an abnormal load of amyloid beta and/or tau in the brain as revealed by PET. Parallel early evidences of AD are reduction of brain metabolism in the parietal, temporal and hippocampal regions measured by FDG-PET, and hippocampal atrophy revealed by structural MRI.

In recent years, structural MRI experiments (e.g., Wolz et al., 2011; Casanova et al., 2012; Lillemark et al., 2014) have revealed

prospective biomarkers with top achievements reaching values superior to 80%. However, it is not yet clear if structural MRI is able to detect earlier biomarkers of AD compared to other MRI modalities, such as functional and diffusion MRI. Several studies have investigated the use of resting-state functional Magnetic Resonance Imaging (rs-fMRI), a non-invasive method for automatic diagnosis of brain diseases (e.g., Chen et al., 2011; Brier et al., 2012; Tang et al., 2013; Hoekzema et al., 2014; Zeng et al., 2014). Chen et al. (2011) achieved accuracy slightly higher than 80%. In that approach, Pearson's correlation coefficient (r) of pairwise regions of interests (ROIs) was used for distinguishing AD patients from healthy subjects in a group analysis. However, given that the correlation coefficient represents only a linear relationship, it follows that if the underlying ground truth is not linear, the result of such an analysis will be inaccurate. Furthermore, the above approach cannot be applied to diagnosing individual patients in a clinical setting as the correlation analysis is calculated on group data. Hence, to study the complexities of brain networks and to identify brain disorders (e.g., AD) only by studying the rs-fMRI correlation between different brain regions may not be sufficient. A more advanced approach is necessary to study the complexities of brain networks, and the diagnosis of single subject data.

Complex networks can be analyzed efficiently using graph theory. Graph theory models each brain region as a node and the relationship between two regions as an edge. Recent studies have focused on the use of graph theoretical measures on the structural brain networks (Bassett et al., 2008; Bassett and Bullmore, 2009; Bullmore and Sporns, 2009; Lynall et al., 2010; Wang et al., 2010; Várkuti et al., 2011; Zhang et al., 2011; Khazaei et al., 2015). Depending on the type of data, various graphs can be constructed. For example, representations of neural networks can be constructed using microscopic data (e.g., Chatterjee and Sinha, 2007), where nodes represent neurons and edges axons connecting neurons to each other. Recent studies have shown that graph theoretical measures are vital in identifying network measures of psychiatric and neurological diseases. Graph theory not only can be used to study various network properties such as small-world property or efficiency of the information transfer, but it can also be employed in medical applications and disease diagnosis. For example, Bassett et al. (2008) used graph theoretic methods to show that patients suffering from Alzheimer's disease and schizophrenia have abnormal network configurations.

Recently, machine learning has been used in the detection of diseases by recognizing physiological patterns (biomarkers) of healthy and pathological conditions (Teipel et al., 2008; Mapstone et al., 2014). There has been a growing interest in applying machine learning techniques on DTI data of Alzheimer's patients (Hahn et al., 2013). Recent studies have derived connection matrices as well as graph metrics from DTI data (Lo et al., 2010) of Alzheimer's patients. Dyrba et al. (2013) report a machine learning approach for discriminating between Alzheimer's disease and healthy controls using fractional anisotropy (FA) values as input features, achieving 80% of accuracy. However, these studies have used information such as Pearson's product moment correlation coefficient between brain regions (Wang et al., 2006; Chen et al., 2011),

regional homogeneity (ReHo), and amplitude of low-frequency fluctuations (ALFF; Dai et al., 2012; Zhang et al., 2012), and a limited number of network measures as discriminant features (Li et al., 2013). Additionally, these studies considered only global network measures but ignored local network measures. Local network measures that are indicators of individual network elements (such as nodes or links) typically quantify connectivity profiles associated with these elements and reflect the way in which these elements are embedded in the network. Hence, one major focus of the current research is to identify more effective local and global structural properties of the brain network to enhance the performance of pattern classification.

The purpose of current study is to implement a system that efficiently classifies AD patients from MCI patients and healthy subjects (CT). We analyzed DTI data of 15 AD patients, 15 MCI patients, and 15 healthy volunteers, and performed pattern classification of the three categories of subjects based on graph measures. We considered pair-wise class combinations of subjects and defined three separate classification tasks, i.e., AD-CT, AD-MCI, and CT-MCI. An ensemble classification module was used to perform the classification tasks. A limitation of this study is the small sample of amnesic MCI, as those could be a mix of people with AD and other dementing disorders such as dementia with Lewy body dementia. Due to the limited sample size, this study can be regarded as a “proof of concept.”

MATERIALS AND METHODS

The study was performed in three stages: data collection and processing, feature extraction, and classification. Data acquisition was performed in an MR scanner, and data processing included a preprocessing phase containing realignment, coregistration, normalization, and segmentation of the data. Fractional anisotropy based tractography, and feature extraction was performed based on graph theory by calculating local network centrality measures (Freeman, 1977). Pattern classification was performed by the method of ensemble classification (Rokach, 2010).

Data Collection and Processing

Subjects and Data Acquisition

Forty-five adults consisting of 15 AD patients, 15 MCI patients, and 15 healthy volunteers were recruited for DTI data acquisition. DTI scans were acquired in the Institute D'or (Rio de Janeiro, Brazil) on a Philips Achieva 3.0 Tesla MR scanner, using a spin echo (SE) sequence with the following parameters: repetition time (TR) = 5620 ms, echo time (TE) = 65 ms, flip angle = 90°, acquisition matrix = 96 × 96 with a spatial resolution of 2.5 × 2.5 mm, 60 transversal slices with thickness = 2.5 mm, 32 gradient directions, and *b*-value 1000 s/mm².

The adults in study were referred for neuropsychological evaluation by their physicians, because of memory complaints to discriminate among normal aging, MCI, or dementia. Participants with tumor, stroke, traumatic brain injury, or hydrocephalus were excluded from the experiment. AD and MCI diagnoses were made by consensus among a trained, a neuropsychologist and a psychiatrist, based on DSM-IV

criteria, MRI overview, clinical data and neuropsychological tests. The AD patients had mild dementing disorders and the supportive tests for first AD diagnosis were performed in parallel to the MR/DTI acquisition. AD diagnoses were performed considering NINCDS-ADRDA criteria (Knopman et al., 2001). All individuals underwent a comprehensive evaluation for diagnostic propose, including the following tests, whose quantitative results are available in the Appendix (**Table A2**) in Supplementary Material:

1. Mini Mental State Evaluation (MMSE),
2. Span (digit and spatial) forward and backward,
3. Clock Drawing Test (CDT),
4. Verbal Fluency (semantic and letter),
5. Family Pictures,
6. Geriatric Depression Scale (Yesavage et al., 1983),
7. Memory Assessment Complaints Questionnaire (Crook et al., 1992).

The MCI patients were all from the amnesic category. We did not apply the Hachinski Ischemic Scale. We used Petersen criteria (Petersen et al., 1997, 2001) for MCI diagnosis, including:

1. Memory problems,
2. Objective memory disorder,
3. Absence of other cognitive disorders or repercussions on daily life,
4. Normal general cognitive function,
5. Absence of dementia.

Healthy volunteers were selected by matching age and education level to the MCI patients and AD patients (**Table 1**). A two-tailed independent *t*-test was used to determine whether the differences in age and educational level between the groups were statistically significant. The results show *P*-values between 0.9 and 1.0 (**Table 2**), which confirms no significant difference (Panagiotakos, 2008) between the groups. The connectivity regions of participants' DTI data were defined in terms of Brodmann areas. A Brodmann area is a region of the cerebral cortex in the human brain, defined by cytoarchitectonic and histological analysis of the structure and organization of cells (Brodmann, 1909).

For each subject, T1-weighted structural images and DTI images were acquired by a gradient recalled echo (GRE) scanning sequence, with the following parameters: TR = 7.16 ms, TE = 3.41 ms, flip angle = 8°, acquisition matrix = 480 × 480 with spatial resolution 0.5 × 0.5 mm, and 340 sagittal slices with thickness of 0.5 mm.

Data Processing

The DTI data was preprocessed with realignment, coregistration, and normalization using the software, SPM8 (Friston, 1996), and segmentation by using the software, DSI Studio (Yeh and Tseng, 2011). FA was calculated for each voxel of each subject's brain volume, generating a FA map for each subject. FA is commonly considered to be an indicator of structural brain connectivity, since it is sensitive to the axonal structure (Basser and Pierpaoli, 1996). Important factors influencing FA is the integrity of axons and their myelin sheaths (Damoiseaux et al., 2009). Subsequently,

TABLE 1 | The 45 participants of the study and their corresponding class, sex, age and educational level.

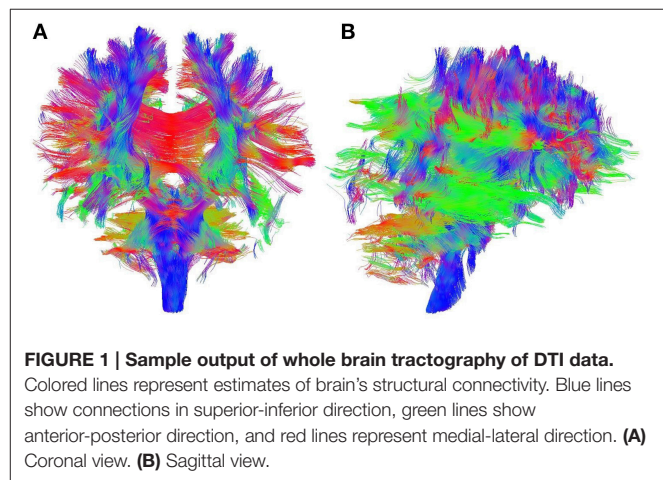
The 15 control volunteers				The 15 Mild Cognitive impairment patients				The 15 Alzheimer's disease patients			
Subject ID	Age (years)	Education (years)	Sex	Subject ID	Age (years)	Education (years)	Sex	Subject ID	Age (years)	Education (years)	Sex
01	85	11	Female	16	80	08	Female	31	74	08	Female
02	81	08	Female	17	81	04	Female	32	73	15	Female
03	76	17	Female	18	68	15	Female	33	77	04	Female
04	74	11	Female	19	73	15	Male	34	68	15	Female
05	74	04	Female	20	88	08	Male	35	68	15	Male
06	78	11	Female	21	84	11	Female	36	71	11	Female
07	74	15	Male	22	70	20	Male	37	74	20	Male
08	83	15	Male	23	67	04	Male	38	84	11	Male
09	62	11	Female	24	71	11	Female	39	75	11	Male
10	67	11	Female	25	71	20	Female	40	86	15	Female
11	77	08	Female	26	80	14	Female	41	73	11	Female
12	78	11	Female	27	77	08	Female	42	80	15	Male
13	72	11	Female	28	70	15	Male	43	61	04	Male
14	61	21	Male	29	66	15	Female	44	72	15	Female
15	77	15	Male	30	69	11	Female	45	81	11	Female
Mean	74.6	12.0	73%F 27% M	Mean	74.3	11.9	67% F 33%M	Mean	74.5	12.1	60% F 40%M
Standard deviation	6.9	4.1	–	Standard deviation	6.8	5.0	–	Standard deviation	6.5	4.3	–

TABLE 2 | P-values by age and educational level on binary discrimination.

Binary discrimination	Age(p-value)	Educational level(p-value)
Control vs. MCI	0.916	1.000
Control vs. Alzheimer	0.957	0.931
MCI vs. Alzheimer	0.957	0.938

tractography (Le Bihan and Breton, 1985), a 3D modeling technique that visually represents neural tracts using DTI data, was performed to obtain the Connectome (Sporns et al., 2005). Tractography (Figure 1) was performed deterministically by using the entire brain as the seed with the toolbox, DSI Studio (Kreher et al., 2008; Yeh et al., 2013). This process was performed according to the following parameters: fractional anisotropy (FA) threshold = 0.1, number of seed points = 1,000,000, maximum angle = 60°, step size = 1.25 mm, length constraint = 25–100 mm, and no spatial smoothing¹. Connectivity matrices were obtained and stored accordingly. We considered 41 Brodmann areas according to the coordinates established in DSI studio (Yeh et al., 2013). In a connectivity matrix, rows and columns of the matrix represent different Brodmann areas. Each cell of the matrix represents a distinct connection between two Brodmann areas corresponding to specific row and column. Figure 2 shows

¹We used 1,000,000 seeds as starting fiber points for each individual brain. This produces a potential maximum of 1,000,000 deterministic fibers per brain (in practice we have much less due to minimum length constraint threshold). Using 1,000,000 fibers can balance the deterministic tractography for all the participants.

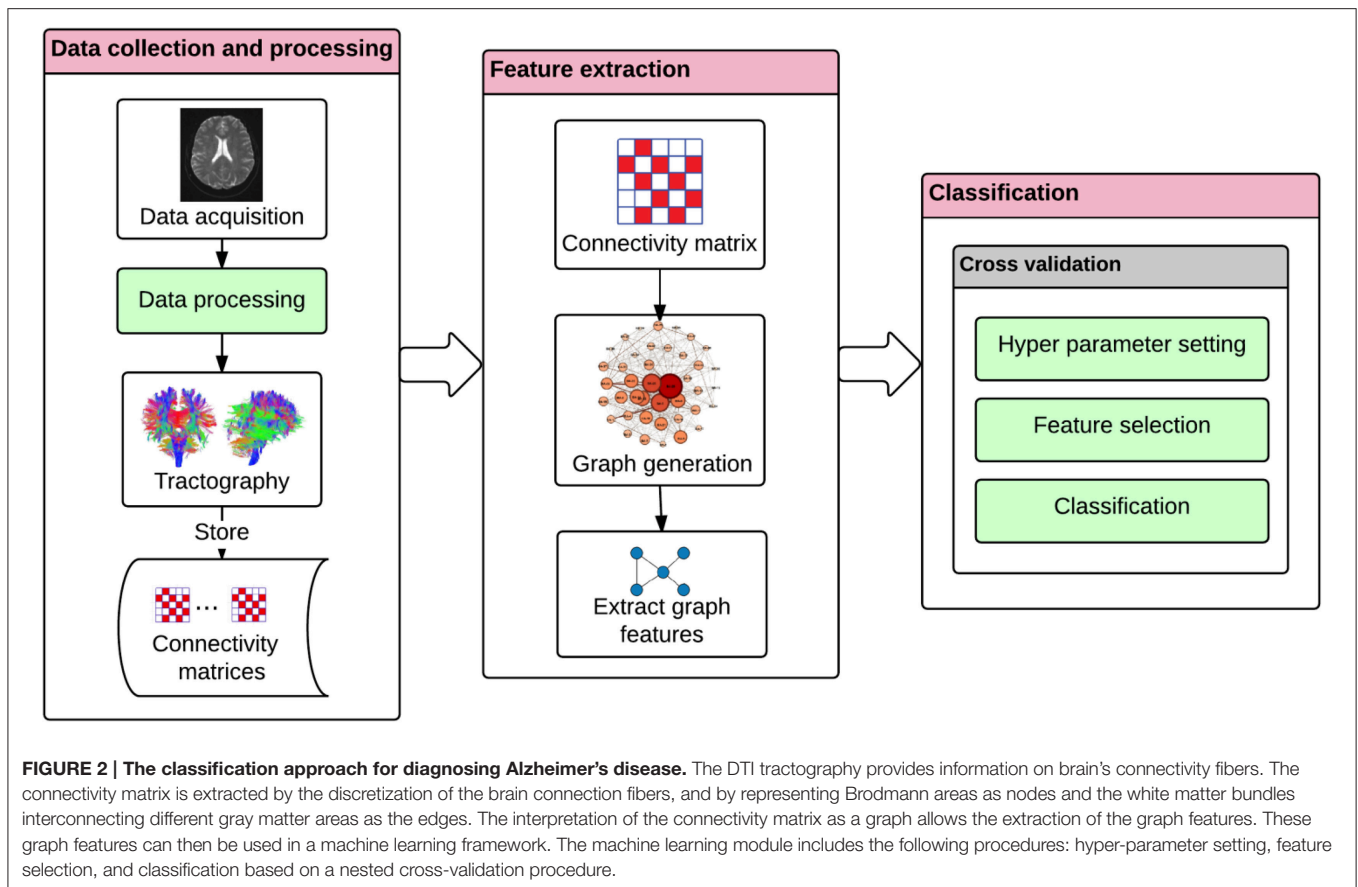


the different stages of our approach. Each block in the figure represents a stage in the classification procedure.

Feature Extraction

Graph Generation and Construction of Structural Brain Networks

In graph theory, a graph is defined as a set of nodes and edges (Bullmore and Sporns, 2009). Using connectivity matrices obtained above, we generated the corresponding graphs where the nodes are brain regions and edges are the bundles of nerves connecting those regions (Gong et al., 2009). In this manner, a graph was generated for each subject using the



connectivity matrices generated in the previous data processing. The graphs we constructed were all undirected, i.e., the nodes are connected together by bidirectional edges. Brodmann areas were represented as network nodes. Based on the 41 Brodmann areas considered in this analysis, each network was composed by a total of 41 nodes. An edge was marked between two nodes, representing two Brodmann areas, if there was at least one fiber connecting them. Otherwise these two nodes were not considered directly connected, although they could still be connected via another node or nodes. We used the weight (W_{ij}) value of the edge to describe the strength of the connectivity of Brodmann areas i and j . We constructed the connectivity matrix as a weighted matrix for each subject. As there were 41 Brodmann areas, the maximum number of all possible edges was 820 [$N(N-1)/2$; $N = 41$]. The number of nodes (41) was the same for each subject since each brain was parcellated using the same scheme. However, subjects differed in the structure of their graph in terms of both the number and weighting on the edges. The numbers of edges were likely to differ for different graphs as the result of varying strengths and presence/absence of the connectivity of Brodmann areas in the subjects. **Figure 3** shows the graph of an MCI patient as an example.

Extracting Graph Features

The network properties that we computed included the following: closeness centrality, betweenness centrality,

eigenvector centrality, Katz centrality, hyperlink-induced topic search (HITS) centrality, degree centrality, clustering coefficient centrality, load centrality, and efficiency and node redundancy coefficient (**Figure 4**). The definition of the different network measures are presented below.

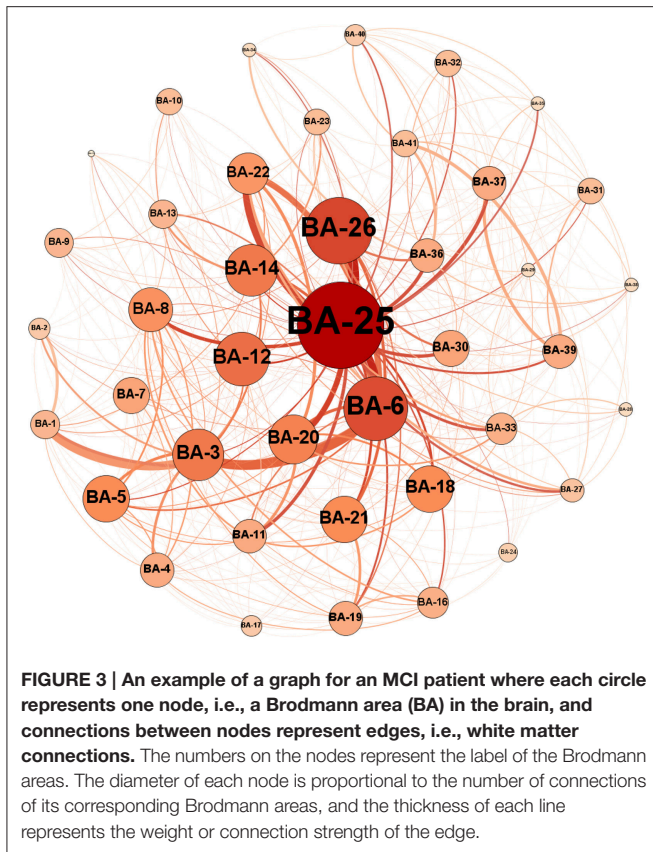
Closeness Centrality

Closeness centrality indicates how close a node is to all the other nodes in a network (Wasserman and Faust, 1994). The centrality of a given node is the sum of the geodesic distances (shortest paths from one node to other nodes). Closeness centrality describes the extent of influence of a node on the network. Closeness can be regarded as a measure of how long it will take to spread information from node i to all other nodes sequentially. Consider a connected weighted undirected graph $G = (V, E)$ with n vertices and m edges ($|V| = n$, $|E| = m$). Let $d(v, u)$ denote the length of the shortest path between v and u . The closeness centrality C_v (Sabidussi, 1966) of node v is defined in Equation (1):

$$C_v = \frac{n-1}{\sum_{u \in V} d(v, u)} \quad (1)$$

Betweenness centrality

Betweenness centrality is a measure of the number of times a node is seen to lie along the shortest path between two other



nodes (Brandes, 2001). It is equal to the number of shortest paths from all vertices to all others that pass through that node. A node with high betweenness centrality has a large influence on the transfer of items through the network, under the assumption that item transfer follows the shortest paths. Studies have illustrated that the brain uses the shortest path during information processing to save time and energy (Klyachko and Stevens, 2003). Hence, we selected closeness and betweenness centrality as measures reflecting the shortest paths in our model. For a graph $G = (V, E)$, betweenness centrality $C_B(v)$ (Brandes, 2001) is given in Equation (2).

$$C_B(v) = \sum_{s,t:s \neq t \neq v} \frac{\sigma_v(s,t)}{\sigma(s,t)} \quad (2)$$

Here $\sigma(s, t)$ is the number of shortest paths from s to t , and $\sigma_v(s, t)$ is the number of shortest paths from s to t that pass through v .

Eigenvector centrality

Eigenvector centrality assigns relative scores to all nodes in the network based on the concept that connections to high-scoring nodes contribute more to the score of the node in question than connections to low-scoring nodes. Eigenvector centrality is calculated by assessing how well-connected an individual is to the parts of the network with the greatest connectivity (Gould, 1967). Recent studies have shown that eigenvector centrality identifies the most prominent regions in a network (Binnewijzend et al.,

2014). For a weighted, undirected graph $G = (V, E)$ with n vertices and m edges, connectivity matrix A , Eigen value λ and corresponding eigenvector x satisfying the condition $Ax = \lambda x$, the eigenvector centrality $C_e(v)$ (Newman, 2008) of node v is given in Equation (3):

$$C_e(v) = \frac{1}{\lambda(\max)} \sum_{i=1}^n A(v, i) x_i \quad (3)$$

The above equation represents the v^{th} entry in the eigenvector x corresponding to the largest eigenvalue $\lambda(\max)$.

Katz centrality

Katz centrality measures the degree of influence of a node in a network. Unlike typical centrality measures which consider only the shortest path (the geodesic) between a pair of actors, Katz centrality measures influence by taking into account the total number of walks between a pair of nodes. Katz centrality computes the relative influence of a node within a network by measuring the number of the immediate neighbors (first degree nodes) and also all other nodes in the network that connect to the node under consideration through these immediate neighbors (Katz, 1953). It is closely related to eigenvector centrality. For a connected, weighted and undirected graph $G = (V, E)$ with n vertices, m edges and connectivity matrix A , Katz centrality $C_k(v)$ (Junker and Schreiber, 2011) of node v is given in Equation (4).

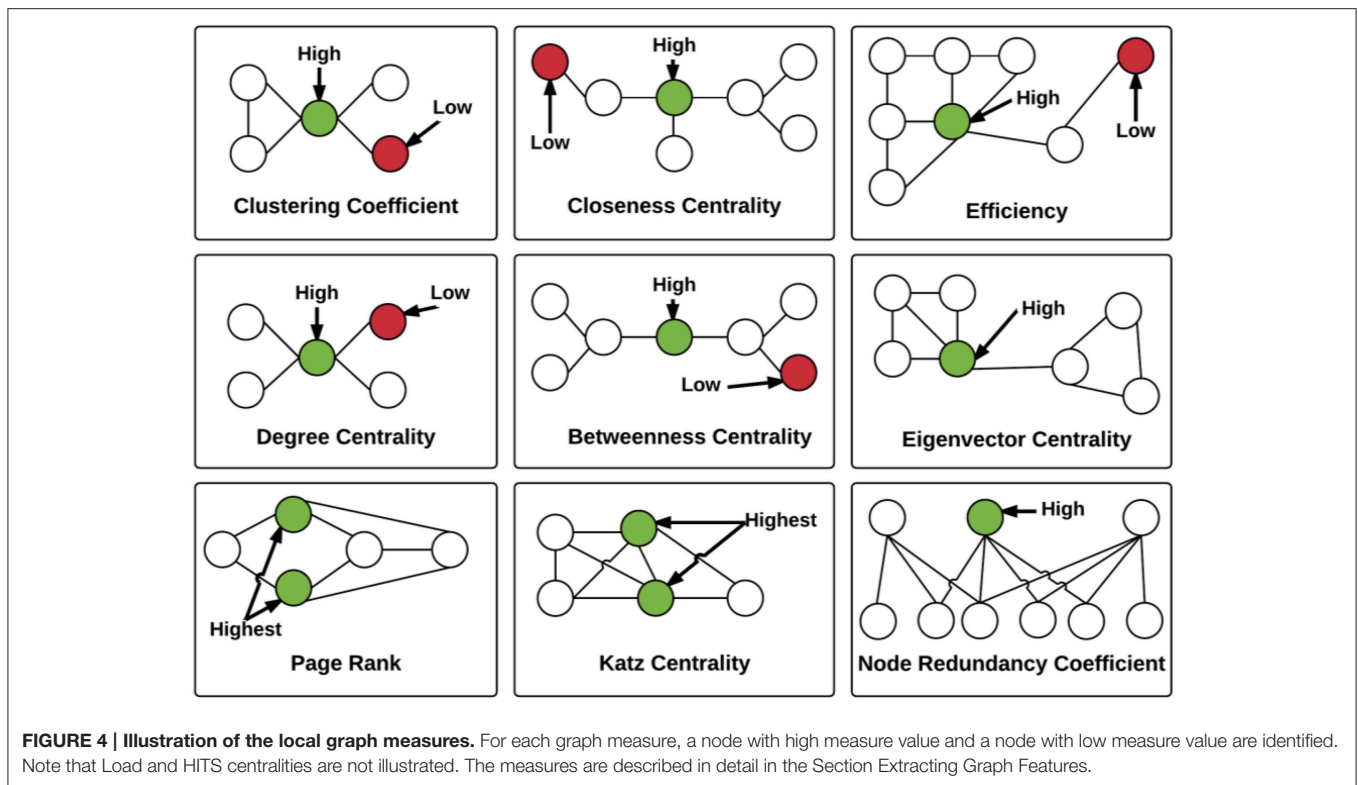
$$C_k(v) = \sum_{k=1}^{\infty} \sum_{u=1}^n \alpha^k (A^k)_{uv} \quad (4)$$

In Equation (4), α , the attenuation factor, is a value which is chosen to be smaller than the reciprocal of the absolute value of the largest eigenvalue of adjacency matrix A . The powers of A indicate the presence or absence of edges between two nodes through intermediaries. For instance in matrix A^4 if the element $(a_1, a_6) = 1$, indicates that through some first and second degree neighbors of node 1, nodes 1, and 6 are connected.

HITS centrality

HITS centrality measures the hub and authority centrality scores of a valued network. Hubs and authorities are a natural generalization of eigenvector centrality. There are two scores for each node in a network, a hub and an authority score. A high hub actor points to many good authorities and a high authority actor receives from many good hubs. The authority score of a vertex is therefore proportional to the sum of the hub scores of the vertices on the in-coming ties and the hub score is proportional to the authority scores of the vertices on the out-going ties. Theoretically, consider a connected, weighted and undirected graph $G = (V, E)$ with n vertices and m edges ($|V| = n$, $|E| = m$). The HITS algorithm (Von Ahn, 2008) consists of a series of iterations with the following steps:

1. Authority update rule: For a given node v , the authority update score is the summation of hub scores of each node which point



to node and is denoted in Equation (5):

$$\text{auth}(v) = \sum_{i=1}^n \text{hub}(i) \tag{5}$$

- Hub update rule: For a given node v , update of its hub score is the summation of authority scores of all nodes pointing to v which is denoted in Equation (6):

$$\text{hub}(v) = \sum_{i=1}^n \text{auth}(i) \tag{6}$$

Degree centrality

Degree (also called valency) of a node in undirected graphs is the total amount of edges directly connecting that node to any other node. It is given by the sum of numbers of edges connecting the nodes. A node with a higher degree of connectivity may be more relevant to the network in terms of number of connections, but it does not measure the importance of each of those connections. Degree has its neurobiological interpretation: a region with high degree interacts structurally and functionally with many other regions in a brain network (Sporns and Kötter, 2004). For a connected weighted undirected graph $G = (V, E)$ with n vertices and m edges ($|V| = n, |E| = m$), the degree centrality $C_D(v)$ (Freeman, 1979) of node v is defined in Equation (7).

$$C_D(v) = \frac{\text{deg}(v)}{n-1} \tag{7}$$

In Equation (7), $\text{deg}(v)$ is the number of edges incident upon a node.

Clustering coefficient

The Clustering coefficient of a node in an undirected graph is a measure that quantifies the fraction of direct connections between the nearest neighbor nodes that exist out of all possible direct connections amongst those nearest neighbor nodes (Watts and Strogatz, 1998). It quantifies the presence of clusters or groups within a network as a measure of functional segregation in the brain, and denotes ability for specialized processing to occur within densely interconnected groups of brain regions. Therefore, it was also added to our model. For a connected, weighted and undirected graph $G = (V, E)$ with n vertices and m edges ($|V| = n, |E| = m$), the clustering coefficient C_v (Watts and Strogatz, 1998) of node v is given in Equation (8).

$$C_v = \frac{2e}{k(k-1)} \tag{8}$$

In Equation (8), k is the number of neighbors of v , and e is the number of connected pairs between all neighbors of v . Clustering coefficient is a ratio N/M , where N is the number of edges between the neighbors of v , and M is the maximum number of edges that could possibly exist between the neighbors of v .

Efficiency

The Efficiency of a node in a network is a measure of how efficiently, in terms of path length, information can be exchanged with other nodes. Thus, the efficiency is inversely related to the shortest path length between the nodes and is used to evaluate

how easily a node can be reached from other nodes. Therefore, the efficiency of a node is the inverse of the harmonic mean of the distances to other nodes (Latora and Marchiori, 2001).

Page rank

Page rank is a well-known algorithm, originally used by Google to rank websites for their search engine (Sullivan, 2007). Page rank is another proxy for evaluating the importance of a node in a network, by assigning a probability of visiting that node after many steps and extent of time. Theoretically, first the standard adjacency matrix is normalized such that the columns of the matrix sum to 1. Next, page rank values are obtained as the values in the eigenvector with the highest corresponding eigenvalue of the normalized adjacency matrix (Page et al., 1999). In our case, page rank can highlight the Brodmann areas with higher number of external links, and the ones that are more frequently addressed. That is, if a Brodmann area is receiving more links from other areas, it might have more important role within the brain.

Load centrality

The load centrality of node i in a network is calculated based on the fraction of all the shortest paths that pass through the node i (Goh et al., 2001). Load centrality is different from betweenness centrality, although they are both related to the shortest path (Brandes, 2008). Theoretically, load centrality is defined based on a hypothetical flow process. It is assumed that each node in a network sends a unit amount of a given commodity to every other node in the network, regardless of any edge or node capacity limit. Using a priority system, the commodity is passed to the neighbors of the node that are closest to the target destination. If there is a tie, i.e., more than one node candidate, the commodity is divided equally among them. The load of a node is calculated as the total amount of commodity passing through it, within the whole process (Goh et al., 2001). Load centrality is a potential alternative to betweenness centrality which provides a complementary view over the flow structures in the network.

Node redundancy coefficient

The redundancy coefficient of node i is calculated based on the fraction of pairs of neighbors of node i that are also connected to other nodes in the network. This measure captures the nodes that are of lower importance and can be induced by other nodes in the network. For any node i , the node redundancy coefficient of node i , RC_i , is defined as in Equation (9).

$$RC_i = \frac{\left| \left\{ \{j, k\} \subseteq N(i), \ni i' \neq i, (i', j) \in E \text{ and } (i', k) \in E \right\} \right|}{\frac{N(i)(|N(i)| - 1)}{2}} \quad (9)$$

In Equation (9), $N(i)$ is the set of neighbor nodes for node i , and E is the set of edges of the network (Latapy et al., 2008). The redundancy coefficient ranges from 0 to 1, where larger numbers for a node mean higher redundancy. Despite the differences in definitions, node redundancy coefficient can be considered as a generalization of clustering coefficient to squares, i.e., C_4 , introduced by Lind et al. (2005), which calculates the probability of a situation in which a node in the network has two neighbors

where these two neighbors have another neighbor in common. If a network is fully connected (where all nodes are connected to all other nodes), redundancy coefficient will be extremely high, since a damage in removal of a node does not affect the network, due to the existence of parallel pathways. In case of brain networks, it could be useful to evaluate the damage caused by illness known to affect gray matter connections such as Schizophrenia.

The above mentioned 11 network properties were computed for each node (i.e., 41 Brodmann areas) in the graph to form a vector of 451 features. Feature selection was then performed on the feature vector.

Classification Model

We used nested cross validation for tuning the hyper-parameters of the model, selecting the features, and evaluating the model. In machine learning, hyper-parameter tuning refers to the problem of finding the best set of parameters for a model (Bergstra and Bengio, 2012). As seen in **Figure 5**, the module automatically performs an exhaustive search over various parameter values for an estimator and finds the best performing set of parameters for the given estimator, using the grid search approach (Chang and Lin, 2011). In this approach, a (multi-dimensional) grid of parameters is defined for the estimator and the cross validation technique is used for searching over the grid, and identifying the best performing set of parameters. After finding the best set of parameters for the estimator, we selected the best set of features automatically; using another five-fold nested cross validation module, and performed the classification tasks, i.e., AD-CT, AD-MCI, MCI-CT. In five-fold cross validation, the classification is performed in five iterations, separately for each pair-wise class combination of patients, i.e., AD-CT, AD-MCI, MCI-CT. Each iteration was divided into three steps: feature selection/hyper-parameter tuning, classifier training and classification. Feature selection/hyper-parameter tuning and classifier training was performed using $(5 - 1 = 4)$ data folds per class, leaving out one-fold per class for subsequent classification. This process was done five times in order to apply classification to all the subjects. We implemented a five-fold nested cross validation (NCV) as the training phase included setting the hyper-parameters of the model and selecting the features. Through NCV (**Figure 5**), an outer cross validation was used to evaluate the accuracy and performance of an inner cross validation in which the hyper-parameters were tuned or the features were selected for the learning process. Thus, NCV helps to obtain an unbiased

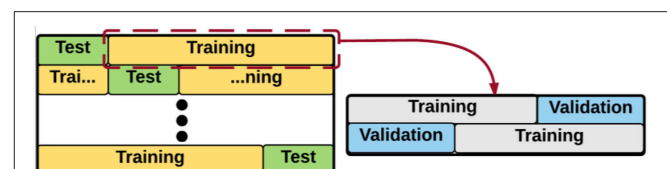


FIGURE 5 | Representation of the Nested Cross Validation (NCV). An outer cross validation module is used to evaluate the accuracy/error of an inner cross validation. In the inner cross validation module, the hyper-parameters are tuned or the features are selected.

estimation of the true error and performance of the estimator (Varma and Simon, 2006).

Hyper-Parameter Setting

Hyper-parameter tuning is a crucial step in the model selection procedure. The task is to determine the best set of hyper-parameters for a given learning algorithm, through optimizing a performance measure, e.g., accuracy or error (Bergstra and Bengio, 2012). We used grid search method for hyper-parameter tuning. In the grid search method, a search is performed on a grid of parameters created based on a pre-defined subset of hyper-parameter space of a learner, in an aim to find the set of hyper-parameters that maximizes the performance of the learning algorithm. The grid search was implemented as part of a five-fold nested cross validation module, explained before, through which the accuracy of the learner was evaluated.

Feature Selection

Feature (attribute) selection is an approach to identify relevant features and reduce the noise in order to increase signal-to-noise ratio and reduce overfitting, by constructing a generalized model through the selections of a subset of features from the original feature set (Bermingham et al., 2015). The central assumption when using a feature selection technique is that the data contains many redundant or irrelevant features. Redundant features are those which are repeated and provide no extra information compared to the currently selected features, while irrelevant features are those which do not provide any useful information. Feature selection techniques are often used in domains where there are many features and comparatively few samples, as is our problem. We implemented a nested cross validation module to test various feature selection methods for all the three classification tasks, i.e., AD-CT, AD-MCI, and CT-MCI. Amongst all the tested feature selection approaches, the K -best features outperformed others. In K -best feature selection, the features are ranked based on their power in performing the classification, and then the top K features are selected for the given estimator. The features were ranked based on the ANOVA F -value between features. The number of top features, i.e., K , should be manually defined. We tested various K s and evaluated the classification accuracy in each of the three classification tasks, **Figure 6**, shows the classification accuracy vs. the best number of features to include in the classifier (K), for different classification tasks. As seen, $K = 430$ maximizes the classification accuracy in both AD-CT and AD-MCI classification tasks. And, $K = 110$ was found to be the best number of top features to include in the MCI-CT classifier. The features were selected based on their score and p -value. The selected top- K features for each of the classification tasks were considered in the respective classifier, which will be explained in detail in the next part.

Classification

As mentioned, we considered pair-wise class combinations of subjects and defined three separate classification tasks, i.e., AD-CT, AD-MCI, and CT-MCI. After tuning the hyper-parameters and selecting the most informative features, the classification

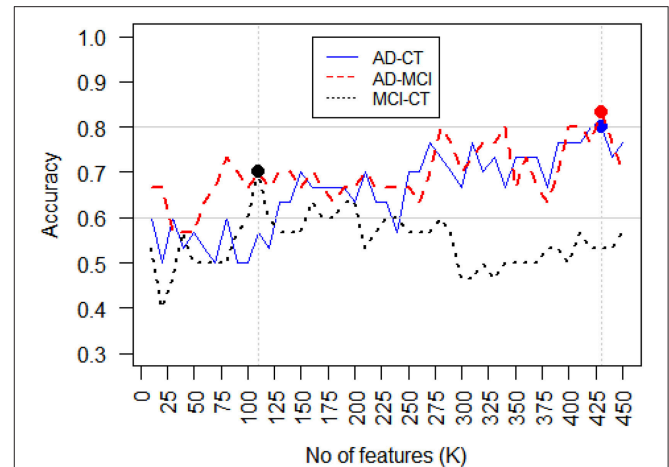


FIGURE 6 | Classification accuracy vs. feature selection threshold, i.e., the number of features selected (K). For AD-CT and AD-MCI classification task, the best $K = 430$, features maximize the accuracy of the classifier. The best K -value found for MCI-CT classification was $K = 110$.

was done using the ensemble classification method. In ensemble learning, multiple classifiers are used such that the ensemble classification system outperforms all of the constituent classifiers (Opitz and Maclin, 1999; Dietterich, 2000). Ensemble classifiers become even more advantageous if classifiers with different decision boundaries are used, allowing for more flexibility through promoting diversity among the models (Brown et al., 2005). The ensemble learning is very similar to human behavior in making important decisions. Especially in case of medical diagnosis, humans prefer to increase the reliability of their decision through asking the opinion of various doctors (Sesmero et al., 2015). We simulated the same process in our system by generating a diverse set of base classifiers whose decision borders complement each other, and combining their outputs such that the accuracy of the classification was improved (**Figure 7**). Thus, the ensemble learner was built in two phases: (1) generating a set of diverse base classifiers, and (2) combining the decisions made by the base classifiers to obtain one decision.

The diversity of the base classifiers is of high importance. The diversity of two given classifiers is high if they result in errors at different samples (Sesmero et al., 2015). If a pair of classifiers is not diverse, then the decision made by each of them might be similar, therefore, the final ensemble decision will not improve. But if the decision borders of a pair of classifiers complement each other, we can expect an improved performance for the ensemble classifier. Although the diversity of base learners is a mandatory condition for a good ensemble, there is no consensus on quantifying the diversity (Kuncheva and Whitaker, 2003). We considered pair-wise combinations of the base classifiers candidates and calculated the Q metric as a proxy for diversity. Equation (10) shows the definition of the Q metric (Sesmero et al., 2015).

$$Q = \frac{N^{11}N^{00} - N^{01}N^{10}}{N^{11}N^{00} + N^{01}N^{10}}. \quad (10)$$

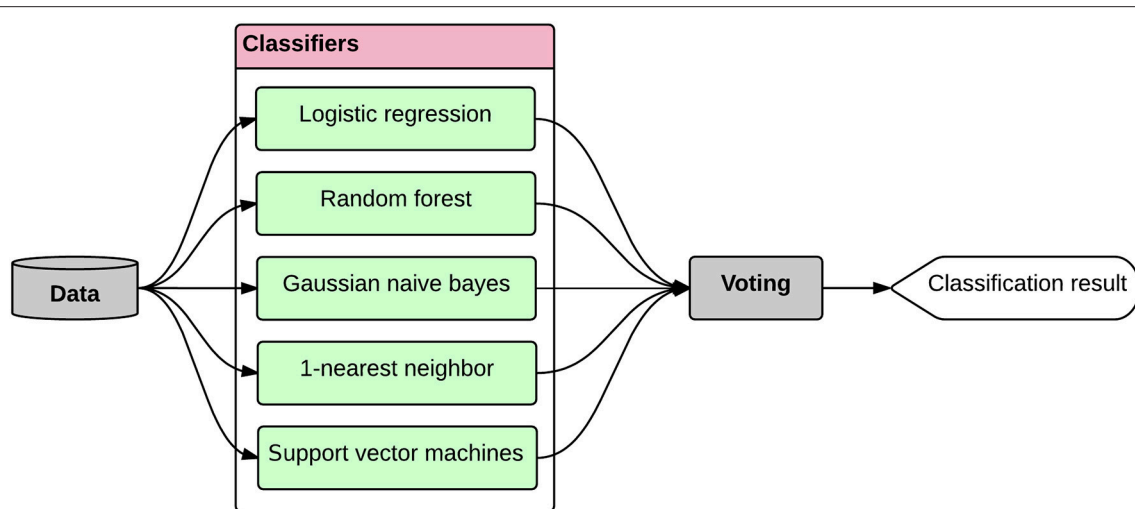


FIGURE 7 | The ensemble classifier design. The data is provided to a set of diverse base classifiers and their decisions are then combined through a voting procedure to obtain the final classification result. The whole process was performed within a nested five-fold cross validation module.

In Equation (10), N^{ij} is the number of samples that were classified correctly ($i = 1$) or incorrectly ($i = 0$) by the first classifier in the pair, and correctly ($j = 1$) or incorrectly ($j = 0$) by the second classifier in the pair. The lower the Q is for a pair of base classifiers, the more diversity the pair has. We tested various combinations of base learners and found that the combination of logistic regression, random forest, gaussian naïve bayes, 1-nearest neighbor, and support vector machines performs the best for the defined classification tasks. We then used voting for combining the decisions made by the base classifiers and obtaining the final classification result (Figure 7). In voting approach, the final classification decision on a new sample, i.e., $C(x)$, is made through voting on all the base classifiers (BC_i), each having a weight of w_i , as stated in Equation (11) (Rokach, 2010).

$$C(x) = \text{sign} \left(\sum_{i=1}^n w_i \cdot BC_i \right). \quad (11)$$

In Equation (11), n is the number of base classifiers (in our case, $n = 5$). The preliminary weights of the base classifiers were first generated within the hyper-parameter setting procedure. We then refined the weights by testing various combinations of weights and determined the best weighting set for each of the classification tasks. Intuitively, we assigned larger weights to more accurate classifiers, or set of classifiers with complementary decision borders. The entire classification process was performed within a nested five-fold cross validation module, as explained before in the cross validation section. The classification module generates a positive (+1) or negative (−1) label for each subject corresponding to the input labels, e.g., AD = +1 and CT = −1 in AD-CT task. After that, this output can be compared with the input label (neuropsychological diagnosis) in order to measure the success level of the approach. Success in classification is achieved when the output label from the

classifier is equivalent to the input label. Accuracy of classification is the percentage of correctly predicted (or detected) output labels for all subjects from both positive and negative input classes.

RESULTS

We evaluated our model on 45 subjects, comprising 15 AD patients, 15 MCI patients, and 15 healthy volunteers. A total number of 451 features was calculated for each subject from which 430, 430, and 110 features were selected as the most informative features for AD-CT, AD-MCI, and MCI-CT classification tasks, respectively. Table 3 shows the performance metrics calculated for the base classifiers as well as the ensemble technique with and without feature selection module. The feature selection module was considered for all the base classifiers during the evaluation stage, as it improved their performance. As seen in Table 3, the ensemble framework with feature selection outperforms all the other listed classifiers, and shows a promising performance. The ensemble framework with feature selection resulted in a classification accuracy of 83.3% for AD vs. MCI, 80% for AD vs. CT, and 70% for MCI vs. CT. Accuracy of the classification was defined as the number of correct predictions divided by the total number of predictions. We also checked for the recall metric defined as the number of true-positives divided by the number of true-positives and false-negatives, i.e., $\text{Recall} = TP / (TP + FN)$. According to Table 3, the recall is 80, 80, and 50% for AD-CT, AD-MCI, and MCI-CT classification tasks, respectively. Recall can be considered as a measure of completeness of a classifier such that a low recall can indicate the high number of false-negatives. Recall can be also regarded as a sensitivity measure since it evaluates the effectiveness of a classifier in identifying the positive labels (Sokolova and Lapalme, 2009). Finally, we checked the F-1 score of the classifiers. The F-1 score is useful

TABLE 3 | Performance results of the base classifiers as well as the ensemble in AD-CT, AD-MCI, and MCI-CT classification tasks.

Base Classifiers	Classification Task								
	AD-CT			AD-MCI			MCI-CT		
	Acc.	Recall	F-1	Acc.	Recall	F-1	Acc.	Recall	F-1
Logistic Regression	0.733	0.867	0.767	0.7	0.8	0.729	0.667	0.667	0.667
Gaussian Naïve Bayes	0.567	0.6	0.566	0.6	0.6	0.584	0.633	0.6	0.58
Support Vector Machines	0.767	0.8	0.779	0.7	0.733	0.72	0.533	0.533	0.518
Random Forests	0.43	0.2	0.26	0.633	0.467	0.491	0.567	0.533	0.542
1-Nearest Neighbor	0.633	0.533	0.58	0.667	0.467	0.531	0.6	0.6	0.578
Ensemble	0.733	0.733	0.731	0.767	0.733	0.758	0.5	0.467	0.48
Ensemble with Feature Selection	0.8	0.8	0.798	0.833	0.8	0.825	0.7	0.5	0.667

All the listed classifiers were tested within a five-fold cross validation module.

Feature selection was used for all the base classifiers as it improved their performance.

The best results for each column were bolded.

TABLE 4 | Accuracy metric 95% confidence intervals for the ensemble classifier with feature selection in AD-CT, AD-MCI, and MCI-CT classification tasks.

		Classification task		
		AD-CT	AD-MCI	MCI-CT
Median [95% Confidence Interval]	Accuracy	0.75 [0.743, 0.757]	0.767 [0.759, 0.774]	0.667 [0.662, 0.671]
	Recall	0.8 [0.792, 0.808]	0.73 [0.718, 0.742]	0.6 [0.589, 0.611]
	Specificity	0.67 [0.617, 0.723]	0.67 [0.626, 0.714]	0.43 [0.384, 0.476]
	F-1	0.77 [0.763, 0.777]	0.74 [0.729, 0.751]	0.56 [0.553, 0.567]
	ROC AUC	0.76 [0.754, 0.767]	0.78 [0.775, 0.785]	0.56 [0.557, 0.563]

if we want to select a classification model such that it holds a balance between precision and recall. Theoretically, the F-1 score is defined as the harmonic mean of precision and recall, i.e., $2(Precision * Recall)/(Precision + Recall)$, in which the precision is the number of true-positives divided by true-positives and false-positives. As it is seen in **Table 3**, the F-1 score of the ensemble framework with feature selection is 79.8, 82.5, and 66.7% for AD-CT, AD-MCI, and MCI-CT classification tasks.

We further investigated the performance of the ensemble classifier with feature selection by calculating the confidence intervals of various performance metrics. We ran the cross validated ensemble classifier for 100 times for each performance metric, i.e., accuracy, recall (sensitivity), specificity, F-1 score, and receiver operating characteristic area under the curve (ROC AUC), and each of the classification tasks, i.e., AD-CT, AD-MCI, and MCI-CT, and stored the result. Next, we checked whether the performance metric is normally distributed. As expected, we observed that none of the metrics for the classification tasks are normally distributed. Therefore, we used the statistical median for calculating the confidence intervals². **Table 4** presents the results for 95% confidence intervals. It shows that we can be 95% confident that the true median of the population is in the

range of [0.743, 0.757], [0.759, 0.774], and [0.662, 0.671] for AD-CT, AD-MCI, and MCI-CT classification tasks, respectively. Although the highest accuracy is still observed for AD-MCI classification task, it is very close to the confidence interval of AD-CT classification. Additionally, it was observed that the model is the best in detecting the positive events for AD-CT classification with recall of 0.8.

We checked for the top-5 features that were detected within the ensemble classification process for each of the classification tasks, separately. For AD-CT classification task, betweenness centrality at Brodmann Area 2 (primary somatosensory cortex), eigenvector centrality at Brodmann Area 1 (primary somatosensory cortex), load centrality at Brodmann Areas 2 (primary somatosensory cortex) and 27 (piriform cortex), and closeness centrality at Brodmann Area 1 (primary somatosensory cortex) were detected as the top-5 most important features. Katz centrality at Brodmann Area 3 (primary somatosensory cortex), degree and closeness centrality at Brodmann Area 5 (somatosensory association cortex), node redundancy coefficient and load centrality at Brodmann Area 4 (primary motor cortex) were found as the most important features for AD-MCI classification. For MCI-CT classification, we observed the hit centrality, page rank, betweenness centrality, and load centrality at Brodmann Area 6 (premotor cortex) along with hub centrality at Brodmann Area 1 (primary somatosensory cortex) to be the most important features.

²To calculate the confidence interval for a mean, we either need a large sample size or a normally distributed population (Dean and Dixon, 1951; Moore, 2007).

DISCUSSION

Our results indicate that complex graph measures can be used as makers to diagnose Alzheimer's disease. Moreover, these results indicate that Alzheimer's disease affects Brodmann areas in the sensorimotor cortex and piriform cortex, where we found abnormalities in our measures. The piriform cortex, whose load centrality has been shown as a good discriminative feature for Alzheimer's disease diagnosis by our results, has an important role in olfactory perception (Howard et al., 2009). The piriform cortex is located at the junction of the temporal and frontal lobes and is the neighbor of the entorhinal cortex (Mai et al., 2008). The entorhinal cortex is known in the literature as one of the first affected areas in Alzheimer disease (Khan et al., 2013). Additionally, the primary motor cortex is significantly involved in late and terminal stages of Alzheimer's disease (Suva et al., 1999).

In our study, node redundancy coefficient and load centrality in the primary motor cortex were recognized as good indicators of Alzheimer's disease in contrast to MCI. In relation to these findings in the primary motor cortex (and in other brain regions), it is important to point out that abnormalities in node redundancy coefficient and load centrality do not mean variations in the connectivity at that exact location. Degree is the only network measure that is directly correlated to connectivity changes in the specific node (in this case primary motor cortex). As Degree is not one of the abnormal measures, it indicates that abnormal connectivity patterns at primary motor cortex are not a relevant feature for AD diagnosis. The relevant network measures revealed here (node redundancy coefficient and load centrality) indicate that connectivity abnormalities on AD are occurring in parallel (in circuit terms) in brain pathways that cross primary motor cortex, which together generate these network patterns found at the primary motor cortex. In other words, the primary motor cortex does not represent the primary effect, rather than the secondary effect of the pathology.

Our findings identified measures in the primary sensory cortex (betweenness centrality, eigenvector centrality, load centrality, closeness centrality, and Katz centrality), somatosensory association cortex (degree and closeness centralities) and primary motor cortex (node redundancy coefficient and load centrality) as good discriminative features for Alzheimer's disease diagnosis, while measures at the premotor cortex (HITS centrality, page rank, betweenness centrality, and load centrality) were identified as good discriminative features for MCI diagnosis³.

Therefore, load centrality, betweenness centrality and closeness centrality were found to be the most relevant network measures in our study, as they were the top identified features at different nodes. Other measures, as eigenvector centrality, Katz centrality, degree centrality, node redundancy coefficient, HITS centrality, page rank and hub centrality were also recognized as relevant features for classifications

among Alzheimer's disease patients, MCI patients, and healthy controls.

The accuracy value achieved in our study can be compared to what other researches performed using other modalities of MRI, such as structural (Wolz et al., 2011; Casanova et al., 2012; Lillemark et al., 2014), functional (Chen et al., 2011) and voxel-based diffusion (Dyrba et al., 2013). All those studies, as well as ours, achieved accuracies in the range of 80–90%. While structural MRI can reveal localized deformations, diffusion MRI based network analysis has the potential to help the development of new pathologic brain segmentation atlases grouping brain matter according to their network patterns. In this way, our paper shows that the use of graph measures as feature of diffusion MRI data can run in parallel to the application of other modalities toward finding biomarkers of AD. The combination of features from these different modalities may considerably increase the potential of the AD diagnosis. Therefore, the development of methods which efficiently combines these multimodal features is a field to be explored by next studies. Nevertheless, our results indicate that complex graph measures may effectively be used to diagnose Alzheimer's disease.

A limitation of this study is the small sample size (15 subjects for each of the three classes). Therefore, this study is a “proof of concept” about a procedure for the classification of MRI markers between AD dementia, MCI, and normal old individuals. The reliability of these results can be tested later in a complementary study based on a larger sample size (at least 100 subjects per class). It is also important to remark that the small sample of amnesic MCI could be a mix of people with AD and other dementing disorders, such as dementia with Lewy body dementia. Due to these facts, the moderate classification between control subjects and MCI in this study does not mean an ability of the procedure to disentangle between normality and the prodromal stages of AD. However, although the sample size is small, it is quite encouraging that the proposed ensemble classification framework which included a well-tuned feature selection component and was validated within a five-fold cross validation module, resulted in quite good classification performance. We expect larger sample would help the system to better distinguish between the AD and non-AD groups. A longitudinal study can assign at what time-point in the evolution of the disease each region is more likely to be affected by Alzheimer's disease. Moreover, longitudinal studies may have the potential to reveal first evidences of prodromal Alzheimer's disease. Finally, supervised classification approach requires a set of labeled observations (Bishop, 2007), thus making it highly dependent on the ground truth provided by the diagnosis of the clinicians. Assuming that there is a possibility of misdiagnosis, unsupervised classification techniques, such as K-Means clustering (Lloyd, 1982; Jain et al., 1999), can be used to improve the reliability of classifications.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of ‘International Guidelines for Biomedical

³Please see **Table A1** for statistical comparison of the key DTI biomarkers at the group level.

Research involving human beings, Declaration of Helsinki' and 'Resolution 196/96, Brazilian National Health Council'. Being a retrospective survey of clinical files, patients did not sign an informed consent, in accordance to Brazilian regulations. The protocol was approved by the 'Ethics Committee of D'Or Institute for Research and Education'.

REFERENCES

- Alzheimer, A. (1907). Über eine eigenartige Erkrankung der Hirnrinde. *Allg Zeitschr Psychiatr.* 64, 146–148.
- Alzheimer's Association (2011). 2011 Alzheimer's disease facts and figures. *Alzheimers Dement.* 7, 208–244. doi: 10.1016/j.jalz.2011.02.004
- Basser, P. J., and Pierpaoli, C. (1996). Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *J. Magn. Reson. B* 111, 209–219. doi: 10.1006/jmrb.1996.0086
- Bassett, D. S., and Bullmore, E. T. (2009). Human brain networks in health and disease. *Curr. Opin. Neurol.* 22, 340–347. doi: 10.1097/WCO.0b013e32832d93dd
- Bassett, D. S., Bullmore, E., Verchinski, B. A., Mattay, V. S., Weinberger, D. R., and Meyer-Lindenberg, A. (2008). Hierarchical organization of human cortical networks in health and schizophrenia. *J. Neurosci.* 28, 9239–9248. doi: 10.1523/JNEUROSCI.1929-08.2008
- Bergstra, J., and Bengio, Y. (2012). Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* 13, 281–305.
- Bermingham, M. L., Pong-Wong, R., Spiliopoulou, A., Hayward, C., Rudan, I., Campbell, H., et al. (2015). Application of high-dimensional feature selection: evaluation for genomic prediction in man. *Sci. Rep.* 5:10312. doi: 10.1038/srep10312
- Binnewijzend, M. A., Adriaanse, S. M., Flier, W. M., Teunissen, C. E., Munck, J. C., Stam, C. J., et al. (2014). Brain network alterations in alzheimer's disease measured by eigenvector centrality in fMRI are related to cognition and CSF biomarkers. *Hum. Brain Mapp.* 35, 2383–2393. doi: 10.1002/hbm.22335
- Bishop, C. (2007). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. New York, NY: Springer.
- Bohnen, N. I., Djang, D. S., Herholz, K., Anzai, Y., and Minoshima, S. (2014). Effectiveness and safety of 18F-FDG PET in the evaluation of dementia: a review of the recent literature. *J. Nucl. Med.* 53, 59–71. doi: 10.2967/jnumed.111.096578
- Brandes, U. (2001). A faster algorithm for betweenness centrality. *J. Math. Sociol.* 25, 163–177. doi: 10.1080/0022250X.2001.9990249
- Brandes, U. (2008). On variants of shortest-path betweenness centrality and their generic computation. *Soc. Netw.* 30, 136–145. doi: 10.1016/j.socnet.2007.11.001
- Brier, M. R., Thomas, J. B., Snyder, A. Z., Benzinger, T. L., Zhang, D., Raichle, M. E., et al. (2012). Loss of intranetwork and internetwork resting state functional connections with Alzheimer's disease progression. *J. Neurosci.* 32, 8890–8899. doi: 10.1523/JNEUROSCI.5698-11.2012
- Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde*. Leipzig: Johann Ambrosius Barth.
- Brookmeyer, R., Johnson, E., Ziegler-Graham, K., and Arrighi, H. M. (2007). Forecasting the global burden of Alzheimer's disease. *Alzheimer Dement.* 3, 186–191. doi: 10.1016/j.jalz.2007.04.381
- Brown, G., Wyatt, J., Harris, R., and Yao, X. (2005). Diversity creation methods: a survey and categorisation. *Inf. Fusion* 6, 5–20.
- Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198. doi: 10.1038/nrn2575
- Burns, A., and Iliffe, S. (2009). Alzheimer's disease. *BMJ* 338:b158. doi: 10.1136/bmj.b158
- Casanova, R., Hsu, F.-C., Espeland, M. A., and Alzheimer's Disease Neuroimaging Initiative (2012). Classification of structural MRI images in Alzheimer's disease from the perspective of ill-posed problems. *PLoS ONE* 7:e44877. doi: 10.1371/journal.pone.0044877

AUTHOR CONTRIBUTIONS

Conceiving and designing the experiments: AE, JD, RS, PR. Performing algorithmic experiments: AE, JD, DN. Analyzing the data: AE. Data/materials: JD, FT, IB, GC. Writing of the manuscript: AE, JD, DN, RS, PR.

- Chang, C. C., and Lin, C. J. (2011). LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* 2:27. doi: 10.1145/1961189.1961199
- Chatterjee, N., and Sinha, S. (2007). Understanding the mind of a worm: hierarchical network structure underlying nervous system function in *C. elegans*. *Prog. Brain Res.* 168, 145–153. doi: 10.1016/S0079-6123(07)68012-1
- Chen, G., Ward, B. D., Xie, C., Li, W., Wu, Z., Jones, J. L., et al. (2011). Classification of Alzheimer disease, mild cognitive impairment, and normal cognitive status with large-scale network analysis based on resting-state functional MR imaging. *Radiology* 259, 213–221. doi: 10.1148/radiol.10100734
- Crook, T. H., Feher, E. P., and Larrabee, G. J. (1992). Assessment of memory complaint in age-associated memory impairment: the MAC-Q. *Int. Psychogeriatr.* 4, 165–176. doi: 10.1017/S1041610292000991
- Dai, Z., Yan, C., Wang, Z., Wang, J., Xia, M., Li, K., et al. (2012). Discriminative analysis of early Alzheimer's disease using multi-modal imaging and multi-level characterization with multi-classifier (M3). *Neuroimage* 59, 2187–2195. doi: 10.1016/j.neuroimage.2011.10.003
- Damoiseaux, J. S., Smith, S. M., Witter, M. P., Sanz-Arigita, E. J., Barkhof, F., Scheltens, P., et al. (2009). White matter tract integrity in aging and Alzheimer's disease. *Hum. Brain Mapp.* 30, 1051–1059. doi: 10.1002/hbm.20563
- Dean, R. B., and Dixon, W. J. (1951). Simplified statistics for small numbers of observations. *Anal. Chem.* 23, 636–638. doi: 10.1021/ac60052a025
- Dietterich, T. G. (2000). "Ensemble methods in machine learning," in *International Workshop on Multiple Classifier Systems*, Vol. 1857, eds J. Kittler and F. Roli (Berlin; Heidelberg: Springer), 1–5.
- Drzezga, A., Lautenschlager, N., Siebner, H., Riemenschneider, M., Willoch, F., Minoshima, S., et al. (2003). Cerebral metabolic changes accompanying conversion of mild cognitive impairment into Alzheimer's disease: a PET follow-up study. *Eur. J. Nucl. Med. Mol. Imaging* 30, 1104–1113. doi: 10.1007/s00259-003-1194-1
- Dubois, B., and Albert, M. L. (2004). Amnesic MCI or prodromal Alzheimer's disease? *Lancet Neurol.* 3, 246–248. doi: 10.1016/S1474-4422(04)00710-0
- Dubois, B., Feldman, H. H., Jacova, C., Hampel, H., Molinuevo, J. L., Blennow, K., et al. (2014). Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *Lancet Neurol.* 13, 614–629. doi: 10.1016/S1474-4422(14)70090-0
- Dyrba, M., Ewers, M., Wegrzyn, M., Kilimann, I., Plant, C., Oswald, A., et al. (2013). Robust automated detection of microstructural white matter degeneration in Alzheimer's disease using machine learning classification of multicenter DTI data. *PLoS ONE* 8:e64925. doi: 10.1371/journal.pone.0064925
- Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry* 40, 35–41. doi: 10.2307/3033543
- Freeman, L. C. (1979). Centrality in social networks conceptual clarification. *Soc. Netw.* 1, 215–239. doi: 10.1016/0378-8733(78)90021-7
- Friston, K. J. (1996). "Statistical parametric mapping and other analysis of functional imaging data," in *Brain Mapping: The Methods*, eds A. W. Toga and J. C. Mazziotta (San Diego, CA: Academic Press), 363–385.
- Goh, K. I., Kahng, B., and Kim, D. (2001). Universal behavior of load distribution in scale-free networks. *Phys. Rev. Lett.* 87:278701. doi: 10.1103/PhysRevLett.87.278701
- Gong, G., He, Y., Concha, L., Lebel, C., Gross, D. W., Evans, A. C., et al. (2009). Mapping anatomical connectivity patterns of human cerebral cortex using *in vivo* diffusion tensor imaging tractography. *Cereb. Cortex* 19, 524–536. doi: 10.1093/cercor/bhn102
- Gould, P. R. (1967). On the geographical interpretation of eigenvalues. *Trans. Inst. Br. Geogr.* 42, 53–86. doi: 10.2307/621372

- Hahn, K., Myers, N., Prigarin, S., Rodenacker, K., Kurz, A., Förstl, H., et al. (2013). Selectively and progressively disrupted structural connectivity of functional brain networks in alzheimer's disease—Revealed by a novel framework to analyze edge distributions of networks detecting disruptions with strong statistical evidence. *Neuroimage* 81, 96–109. doi: 10.1016/j.neuroimage.2013.05.011
- Heron, M., and National Center for Health Statistics (US). (2009). *Deaths: Final Data for 2006*, Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics.
- Hoekzema, E., Carmona, S., Ramos-Quiroga, J. A., Richarte Fernández, V., Bosch, R., Soliva, J. C., et al. (2014). An independent components and functional connectivity analysis of resting state fMRI data points to neural network dysregulation in adult ADHD. *Hum. Brain Mapp.* 35, 1261–1272. doi: 10.1002/hbm.22250
- Howard, J. D., Plailly, J., Grueschow, M., Haynes, J. D., and Gottfried, J. A. (2009). Odor quality coding and categorization in human posterior piriform cortex. *Nat. Neurosci.* 12, 932–938. doi: 10.1038/nn.2324
- Jack, C. R. Jr., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., et al. (2008). The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J. Magn. Reson. Imaging* 27, 685–691. doi: 10.1002/jmri.21049
- Jain, A. K., Murty, M. N., and Flynn, P. J. (1999). Data clustering: a review. *ACM Comput. Surv.* 31, 264–323. doi: 10.1145/331499.331504
- Junker, B. H., and Schreiber, F. (2011). *Analysis of Biological Networks*. Hoboken, NJ: John Wiley & Sons.
- Katz, L. (1953). A new status index derived from sociometric analysis. *Psychometrika* 18, 39–43. doi: 10.1007/BF02289026
- Khan, U. A., Liu, L., Provenzano, F. A., Berman, D. E., Profaci, C. P., Sloan, R., et al. (2013). Molecular drivers and cortical spread of lateral entorhinal cortex dysfunction in preclinical Alzheimer's disease. *Nat. Neurosci.* 17, 304–311. doi: 10.1038/nn.3606
- Khazaei, A., Ebrahimzadeh, A., and Babajani-Feremi, A. (2015). Identifying patients with Alzheimer's disease using resting-state fMRI and graph theory. *Clin. Neurophysiol.* 126, 2132–2141. doi: 10.1016/j.clinph.2015.02.060
- Klyachko, V. A., and Stevens, C. F. (2003). Connectivity optimization and the positioning of cortical areas. *Proc. Natl. Acad. Sci. U.S.A.* 100, 7937–7941. doi: 10.1073/pnas.0932745100
- Knopman, D. S., DeKosky, S. T., Cummings, J. L., Chuit, H., Corey-Bloom, J., Relkin, N., et al. (2001). Practice parameter: diagnosis of dementia (an evidence-based review). *Neurology* 56, 1143–1153. doi: 10.1212/WNL.56.9.1143
- Kreher, B., Schnell, S., Mader, I., Il'yasov, K., Hennig, J., Kiselev, V. G., et al. (2008). Connecting and merging fibers: pathway extraction by combining probability maps. *NeuroImage* 43, 81–89. doi: 10.1016/j.neuroimage.2008.06.023
- Kuncheva, L. I., and Whitaker, C. J. (2003). Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy. *Mach. Learn.* 51, 181–207. doi: 10.1023/A:1022859003006
- Latapy, M., Magnien, C., and Del Vecchio, N. (2008). Basic notions for the analysis of large two-mode networks. *Soc. Netw.* 30, 31–48. doi: 10.1016/j.socnet.2007.04.006
- Latora, V., and Marchiori, M. (2001). Efficient behavior of small-world networks. *Phys. Rev. Lett.* 87:198701. doi: 10.1103/PhysRevLett.87.198701
- Le Bihan, D., and Breton, E. (1985). Imagerie de diffusion *in-vivo* par résonance magnétique nucléaire. *Comptes-Rendus de l'Acad. des Sci.* 93, 27–34.
- Li, Y., Qin, Y., Chen, X., and Li, W. (2013). Exploring the functional brain network of Alzheimer's disease: based on the computational experiment. *PLoS ONE* 8:e73186. doi: 10.1371/journal.pone.0073186
- Lillemark, L., Sørensen, L., Pai, A., Dam, E. B., Nielsen, M., and for Alzheimer's Disease Neuroimaging Initiative (2014). Brain region's relative proximity as marker for Alzheimer's disease based on structural MRI. *BMC Med. Imaging* 14:21. doi: 10.1186/1471-2342-14-21
- Lind, P. G., González, M. C., and Herrmann, H. J. (2005). Cycles and clustering in bipartite networks. *Phys. Rev. E* 72:056127. doi: 10.1103/physreve.72.056127
- Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Trans. Inform. Theory* 28, 129–137. doi: 10.1109/TIT.1982.1056489
- Lo, C. Y., Wang, P. N., Chou, K. H., Wang, J., He, Y., and Lin, C. P. (2010). Diffusion tensor tractography reveals abnormal topological organization in structural cortical networks in Alzheimer's disease. *J. Neurosci.* 30, 16876–16885. doi: 10.1523/JNEUROSCI.4136-10.2010
- Lynall, M. E., Bassett, D. S., Kerwin, R., McKenna, P. J., Kitzbichler, M., Muller, U., et al. (2010). Functional connectivity and brain networks in schizophrenia. *J. Neurosci.* 30, 9477–9487. doi: 10.1523/JNEUROSCI.0333-10.2010
- Mai, J. K., Paxinos, G., and Voss, T. (2008). *Atlas of the Human Brain*. Amsterdam: Elsevier.
- Mapstone, M., Cheema, A. K., Fiandaca, M. S., Zhong, X., Mhyre, T. R., MacArthur, L. H., et al. (2014). Plasma phospholipids identify antecedent memory impairment in older adults. *Nat. Med.* 20, 415–418. doi: 10.1038/nm.3466
- Moore, D. S. (2007). *The Basic Practice of Statistics*, Vol. 2. New York, NY: WH Freeman.
- Newman, M. E. (2008). The mathematics of networks. *New Palgrave Encyclopedia Econ.* 2, 1–12. doi: 10.1057/978-1-349-95121-5_2565-1
- Opitz, D., and Maclin, R. (1999). Popular ensemble methods: an empirical study. *J. Artif. Intell. Res.* 11, 169–198. doi: 10.1613/jair.614
- Page, L., Sergey, B., Rajeev, M., and Terry, W. (1999). *The PageRank Citation Ranking: Bringing Order to the Web*. Stanford InfoLab.
- Panagiotakos, D. B. (2008). Value of p-value in biomedical research. *Open Cardiovasc. Med. J.* 2, 97–99. doi: 10.2174/1874192400802010097
- Petersen, R. C., Doody, R., Kurz, A., Mohs, R. C., Morris, J. C., Rabins, P. V., et al. (2001). Current concepts in mild cognitive impairment. *Arch. Neurol.* 58, 1985–1992. doi: 10.1001/archneur.58.12.1985
- Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik R. J., Kokmen E., and Tangelos E. G. (1997). Aging, memory, and mild cognitive impairment. *Int. Psychogeriatr.* 9(Suppl. 1), 65–69. doi: 10.1017/s1041610297004717
- Rokach, L. (2010). Ensemble-based classifiers. *Artif. Intell. Rev.* 33, 1–39. doi: 10.1007/s10462-009-9124-7
- Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika* 31, 581–603. doi: 10.1007/BF02289527
- Sarazin, M., Chauviré, V., Gerardin, E., Colliot, O., Kinkingnéhun, S., de Souza, L. C., et al. (2010). The amnesic syndrome of hippocampal type in Alzheimer's disease: an MRI study. *J. Alzheimers. Dis.* 22, 285–294. doi: 10.3233/JAD-2010-091150
- Sesmero, M. P., Ledezma, A. I., and Sanchis, A. (2015). Generating ensembles of heterogeneous classifiers using stacked generalization. *Wiley Interdiscipl. Rev. Data Min. Knowl. Discov.* 5, 21–34. doi: 10.1002/widm.1143
- Snedecor, G. W., and Cochran, W. G. (1989). *Statistical Methods, 8th Edn.*, Amsterdam: Iowa State University Press Iowa.
- Sokolova, M., and Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Inform. Process. Manage.* 45, 427–437. doi: 10.1016/j.ipm.2009.03.002
- Sporns, O., and Kötter, R. (2004). Motifs in brain networks. *PLoS Biol.* 2:e369. doi: 10.1371/journal.pbio.0020369
- Sporns, O., Tononi, G., and Kötter, R. (2005). The human connectome: a structural description of the human brain. *PLoS Comput. Biol.* 1:e42. doi: 10.1371/journal.pcbi.0010042
- Sullivan, D. (2007). *What Is Google PageRank? A Guide For Searchers & Webmasters*. Search Engine Land. Available online at: <http://searchengineland.com/what-is-google-pagerank-a-guide-for-searchers-webmasters-11068>
- Suva, D., Favre, I., Kraftsik, R., Esteban, M., Lóbrin, A., and Miklossy, J. (1999). Primary motor cortex involvement in Alzheimer disease. *J. Neuropathol. Exp. Neurol.* 58, 1125–1134. doi: 10.1097/00005072-199911000-00002
- Tang, Y., Jiang, W., Liao, J., Wang, W., and Luo, A. (2013). Identifying individuals with antisocial personality disorder using resting-state FMRI. *PLoS ONE* 8:e60652. doi: 10.1371/journal.pone.0060652
- Tapiola, T., Alafuzoff, I., Herukka, S. K., Parkkinen, L., Hartikainen, P., Soininen, H., et al. (2009). Cerebrospinal fluid β -amyloid 42 and tau proteins as biomarkers of Alzheimer-type pathologic changes in the brain. *Arch. Neurol.* 66, 382–389. doi: 10.1001/archneur.2008.596
- Teipel, S. J., Meindl, T., Grinberg, L., Heinsen, H., and Hampel, H. (2008). Novel MRI techniques in the assessment of dementia. *Eur. J. Nucl. Med. Mol. Imaging* 35(Suppl. 1), S58–S69. doi: 10.1007/s00259-007-0703-z
- Várkuti, B., Cavusoglu, M., Kullik, A., Schiffler, B., Veit, R., Yilmaz, Ö., et al. (2011). Quantifying the link between anatomical connectivity, gray matter volume and regional cerebral blood flow: an integrative MRI study. *PLoS ONE* 6:e14801. doi: 10.1371/journal.pone.0014801

- Varma, S., and Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics* 7:91. doi: 10.1186/1471-2105-7-91
- Von Ahn, L. (2008). *Hubs and Authorities: Science of the Web Course Notes*. Pittsburgh, PA: Carnegie Mellon University.
- Wang, K., Jiang, T., Liang, M., Wang, L., Tian, L., Zhang, X., et al. (2006). "Discriminative analysis of early Alzheimer's disease based on two intrinsically anti-correlated networks with resting-state fMRI," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Berlin; Heidelberg: Springer), 340–347.
- Wang, L., Metzak, P. D., Honer, W. G., and Woodward, T. S. (2010). Impaired efficiency of functional networks underlying episodic memory-for-context in schizophrenia. *J. Neurosci.* 30, 13171–13179. doi: 10.1523/JNEUROSCI.3514-10.2010
- Wasserman, S., and Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge, UK: Cambridge University Press.
- Watts, D. J., and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature* 393, 440–442.
- Wolz, R., Julkunen, V., Koikkalainen, J., Niskanen, E., Zhang, D. P., Rueckert, D., et al. (2011). Multi-method analysis of MRI images in early diagnostics of Alzheimer's disease. *PLoS ONE* 6:e25446. doi: 10.1371/journal.pone.0025446
- Yeh, F. C., and Tseng, W. Y. (2011). NTU-90: a high angular resolution brain atlas constructed by q-space diffeomorphic reconstruction. *Neuroimage* 58, 91–99. doi: 10.1016/j.neuroimage.2011.06.021
- Yeh, F.-C., Verstynen, T. D., Wang, Y., Fernández-Miranda, J. C., and Tseng, W.-Y. I. (2013). Deterministic diffusion fiber tracking improved by quantitative anisotropy. *PLoS ONE* 8:e80713. doi: 10.1371/journal.pone.0080713
- Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., et al. (1983). Development and validation of a geriatric depression screening scale: a preliminary report. *J. Psychiat. Res.* 17, 37–49.
- Zeng, L., Shen, H., Liu, L., and Hu, D. (2014). Unsupervised classification of major depression using functional connectivity MRI. *Hum. Brain Mapp.* 35, 1630–1641. doi: 10.1002/hbm.22278
- Zhang, J., Wang, J., Wu, Q., Kuang, W., Huang, X., He, Y., et al. (2011). Disrupted brain connectivity networks in drug-naive, first-episode major depressive disorder. *Biol. Psychiatry* 70, 334–342. doi: 10.1016/j.biopsych.2011.05.018
- Zhang, Z., Liu, Y., Jiang, T., Zhou, B., An, N., Dai, H., et al. (2012). Altered spontaneous activity in Alzheimer's disease and mild cognitive impairment revealed by regional homogeneity. *Neuroimage* 59, 1429–1440. doi: 10.1016/j.neuroimage.2011.08.049

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Ebadi, Dalboni da Rocha, Nagaraju, Tovar-Moll, Bramati, Coutinho, Sitaram and Rashidi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX

We statistically compared the key DTI biomarkers at the group level, i.e., AD vs. MCI, AD-CT, and MCI-CT. In particular, we performed the two-sample *t*-test (Snedecor and Cochran, 1989), for each of the key DTI biomarkers (e.g., primary somatosensory cortex-betweenness centrality) in each group (e.g., AD-CT group), to determine if the mean of the examined biomarker is equal for both samples (e.g., AD vs. CT). The results are shown in **Table A1**.

AD diagnoses were performed considering NINCDS-ADRDA criteria (Knopman et al., 2001). All individuals underwent a comprehensive evaluation for diagnostic propose, which included Mini Mental State Evaluation (MMSE), Span (digit and spatial) forward and backward, Clock Drawing Test (CDT), Verbal Fluency (semantic and letter), Family Pictures, Geriatric Depression Scale (Yesavage et al., 1983), Memory Assessment Complaints Questionnaire (Crook et al., 1992). The quantitative results for these tests are in the **Table A2**.

TABLE A1 | Statistical comparison of the key DTI biomarkers at the group level.

	Classification task					
	AD-CT		AD-MCI		MCI-CT	
	AD-CT Key Biomarker	<i>t</i> -statistic	AD-MCI Key Biomarkers	<i>t</i> -statistic	MCI-CT Key Biomarkers	<i>t</i> -statistic
Null Hypothesis: There is no difference between the average of the key DTI biomarker in both samples.	Primary somatosensory cortex -betweenness centrality	-2.41*	Primary somatosensory cortex -katz centrality	-1.59	Premotor cortex-HITS centrality	-3.6**
	Primary somatosensory cortex-eigenvector centrality	-2.37*	Somatosensory association cortex-degree centrality	-1.54	premotor cortex -page rank	-3.44**
	Primary auditory cortex-load centrality	-2.36*	Somatosensory association cortex-closeness centrality	-1.51	Premotor cortex -betweenness centrality	-2.55*
	Primary somatosensory cortex -load centrality	2.26*	Primary motor cortex-node redundancy	-1.49	Premotor cortex -load centrality	-2.54*
	Primary somatosensory cortex -closeness centrality	-2.22*	Primary motor cortex-load centrality	1.42	Primary somatosensory cortex-HITS centrality	-2.32*

* $p < 0.05$, ** $p < 0.01$.

TABLE A2 | Quantitative results for diagnostic propose of AD and amnesic MCI.

Patient	CDT	Span forward	Span backward	Fluency	MMSE	Crook	Yesavage	Family pictures	Diagnosis
Sub 01	10	4	4	16	28	30	6	30	Normal
Sub 02	10	5	4	12	26	21	2	29	Normal
Sub 03	10	6	4	14	26	32	5	26	Normal
Sub 04	10	6	4	21	25	35	10	26	Normal
Sub 05	10	5	3	8	22	29	1	24	Normal
Sub 06	10	4	2	11	21	31	4	20	Normal
Sub 07	10	5	4	24	29	27	0	30	Normal
Sub 08	10	6	3	12	29	21	4	29	Normal
Sub 09	10	3	3	19	29	31	8	27	Normal
Sub 10	10	4	4	12	26	29	11	30	Normal
Sub 11	8	7	5	19	25	29	7	25	Normal
Sub 12	10	5	3	11	29	26	4	23	Normal
Sub 13	10	4	4	16	28	33	3	27	Normal
Sub 14	10	6	5	21	29	27	0	27	Normal
Sub 15	9	4	3	11	25	23	5	24	Normal
Sub 16	10	5	4	15	26	31	6	28	Amnesic MCI
Sub 17	10	5	4	13	27	35	12	30	Amnesic MCI
Sub 18	10	5	4	10	27	31	13	30	Amnesic MCI
Sub 19	10	4	3	13	30	26	4	22	Amnesic MCI
Sub 20	8	4	4	6	20	30	2	27	Amnesic MCI
Sub 21	4	6	4	9	25	27	3	18	Amnesic MCI
Sub 22	10	5	4	19	27	28	1	30	Amnesic MCI
Sub 23	10	5	3	13	24	34	6	22	Amnesic MCI
Sub 24	10	4	2	17	27	23	4	24	Amnesic MCI
Sub 25	10	6	3	18	29	31	3	26	Amnesic MCI
Sub 26	10	4	2	10	24	25	3	23	Amnesic MCI
Sub 27	7	4	4	16	25	24	8	17	Amnesic MCI
Sub 28	2	5	3	13	28	26	11	17	Amnesic MCI
Sub 29	10	6	4	10	26	21	11	18	Amnesic MCI
Sub 30	10	4	3	13	23	27	12	15	Amnesic MCI
Sub 31	9	4	3	9	17	24	4	16	AD dementia
Sub 32	7	4	4	10	21	35	8	13	AD dementia
Sub 33	3	3	0	6	14	28	0	13	AD dementia
Sub 34	5	4	3	9	25	26	2	23	AD dementia
Sub 35	10	7	4	15	25	30	4	15	AD dementia
Sub 36	10	5	4	10	26	30	3	19	AD dementia
Sub 37	10	5	3	9	24	22	0	21	AD dementia
Sub 38	3	5	3	5	23	25	1	15	AD dementia
Sub 39	2	5	2	5	15	9	2	16	AD dementia
Sub 40	10	4	2	11	22	32	7	15	AD dementia
Sub 41	4	4	3	9	21	21	0	19	AD dementia
Sub 42	10	3	3	7	23	23	3	13	AD dementia
Sub 43	10	5	3	14	28	35	7	21	AD dementia
Sub 44	3	4	2	0	12	23	2	16	AD dementia
Sub 45	9	4	3	12	25	34	4	20	AD dementia