Multiplex Networks for Early Diagnosis of Alzheimer's Disease

Nicola Amoroso [1,2], Marianna La Rocca [3]*, Stefania Bruno [4], Tommaso Maggipinto [1,2], Alfonso Monaco [2], Roberto Bellotti [1,2†], and Sabina Tangaro [2†] for the Alzheimer's Disease Neuroimaging Initiative [‡]

[1] Dipartimento Interateneo di Fisica "M. Merlin", Università degli studi di Bari "A. Moro", Bari, Italy, [2] Dipartimento Interateneo di Fisica "M. Merlin", Istituto Nazionale di Fisica Nucleare, Sezione di Bari, Bari, Italy, [3] Laboratory of Neuro Imaging, USC Stevens Neuroimaging and Informatics Institute, Keck School of Medicine of USC, University of Southern California, Los Angeles, CA, United States, [4] Blackheath Brain Injury Rehabilitation Centre, London, United Kingdom

Analysis and quantification of brain structural changes, using Magnetic Resonance Imaging (MRI), are increasingly used to define novel biomarkers of brain pathologies, such as Alzheimer's disease (AD). Several studies have suggested that brain topological organization can reveal early signs of AD. Here, we propose a novel brain model which captures both intra- and inter-subject information within a multiplex network approach. This model localizes brain atrophy effects and summarizes them with a diagnostic score. On an independent test set, our multiplex-based score segregates (i) normal controls (NC) from AD patients with a 0.86 ± 0.01 accuracy and (ii) NC from mild cognitive impairment (MCI) subjects that will convert to AD (cMCI) with an accuracy of 0.84 ± 0.01. The model shows that illness effects are maximally detected by parceling the brain in equal volumes of 3,000 mm^3 ("patches"), without any *a priori* segmentation based on anatomical features. The multiplex approach shows great sensitivity in detecting anomalous changes in the brain; the robustness of the obtained results is assessed using both voxel-based morphometry and FreeSurfer morphological features. Because of its generality this method can provide a reliable tool for clinical trials and a disease signature of many neurodegenerative pathologies.

Keywords: multiplex networks, machine learning, diagnosis support system, Alzheimer's disease, mild cognitive impairment, magnetic resonance imaging (MRI), brain Connectivity

1. INTRODUCTION

Alzheimer's disease (AD) is a progressive, neurodegenerative disease accounting for most cases of dementia after the age of 65. It is expected that over 115 million people will develop AD by 2050 (Alzheimer's Association, 2018). Illness related brain changes can be detected *in vivo* with Magnetic Resonance Imaging (MRI) and neuroimaging has been playing an increasingly important role for the diagnosis of neurodegenerative disorders (Bron et al., 2015; Wei et al., 2016; Lebedeva et al., 2017) to the extent that it has been incorporated in the diagnostic criteria for AD (McKhann et al., 2011). It is now accepted that the neurodegenerative cascade in AD begins in the brain years, decades even, before the clinical and radiological manifestations of the illness. The dementia is preceded by a prodromal phase of mild cognitive impairment (Albert et al., 2011), and this, in turn, by a pre-clinical phase (Sperling et al., 2011) of variable duration. Understanding the biological changes, occurring in these early phases, is of paramount importance, as it would open a window

of opportunity for future disease-modifying treatments. While it is clear that neurodegeneration in AD occurs in a rather stereotyped fashion in the majority of cases (West et al., 1994; Perl, 2010; Landin-Romero et al., 2017), it is not known exactly what drives the propagation of the disease within an individual, and what is behind the variations in the patterns of atrophy between individuals. To which extent neurodegeneration propagates through anatomical contiguity is yet to be clarified.

MRI can provide significant information on topological organization of the brain (Yao et al., 2010; Bullmore and Bassett, 2011; Alexander-Bloch et al., 2012; Tijms et al., 2013b), thus graph theory has been widely used to study AD which is known to involve both a structural and a functional disruption of brain connectivity (He et al., 2008; Stam et al., 2009; Ciftçi, 2011; de Haan et al., 2012). These studies reported altered local and global graph properties, supporting the clinical relevance of brain networks, especially within group-wise association studies (Crossley et al., 2014; Daianu et al., 2015).

Up to now, graph models of the brain have been based on two distinct approaches (Suk et al., 2014): (i) voxel-wise and (ii) region of interest analyses. We propose here a novel approach based on parceling MRI brain scans in rectangular boxes, that we call "patches," of fixed dimensions representing the nodes of a network. Then, we measure pairwise similarity measurements between the nodes to define network connections. Therefore, our approach does not inherit the intrinsic computational burden and lack of statistical power affecting voxel wise descriptions (Davatzikos, 2004). Besides, as it is based on unsupervised segmentations of the brain, it avoids *a priori* assumptions about localization of disease effects and typical bias deriving from segmentation errors (Amoroso et al., 2015). In addition, as brain disease has often a diffuse effect, affecting multiple voxels, but not necessarily corresponding to entire anatomical structures, the proposed approach has the potential to better suit the description of pathological changes in the brain, reflecting biological variability.

Specifically for network science, recent studies have investigated the limitations of traditional approaches to describe real systems (Mucha et al., 2010; Lee et al., 2012; Boccaletti et al., 2014) and have pointed out that context information plays a fundamental role. Analogously, we introduce here the novel perspective of multiplex networks (from now onward also multiplexes). Multiplexes are multi-layer systems with a fixed number of nodes that can be linked in different interacting layers, to investigate inter-subject characterization, rather than group-wise differences. In this study, multiplex-based measures are investigated to detect subtle brain atrophy effects, taking into account inter-subject variability; then, proper measures are used to feed random forest classifiers and reveal the emergence of statistically significant AD-related patterns altering the topological organization of the brain.

2. MATERIALS AND METHODS

2.1. Subjects

In this study we used a training set $\mathcal{D}_{train}$ composed of 67 T1 MRI scans. The sample, described in Boccardi et al. (2015), includes 29

normal controls (NC) and 38 AD subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI). We also employed an independent test set of 148 subjects $\mathcal{D}_{test}$, composed by 52 NC, 48 AD and 48 subjects with mild cognitive impairment converting to AD (cMCI). Conversions to AD occurred in a range of [30, 108] months following the baseline diagnosis. $\mathcal{D}_{test}$ subjects were randomly chosen within the whole ADNI in order to match the demographic characteristics of training subjects. The training sample (67) and the test sample (148) are of sufficient size for the construction of robust classification models (Mukherjee et al., 2003; Beleites et al., 2013). All 215 participants underwent whole-brain MRI at 34 different sites. Both 1.5 and 3.0 T scans were included in $\mathcal{D}_{train}$ and $\mathcal{D}_{test}$. Indeed, 1.5 and 3 T scans do not significantly differ in their power to detect neurodegenerative changes as shown in Ho et al. (2010)

ADNI images consisted of MPRAGE MRI brain scans, which were normalized with the MNI152 brain template of size of $197 \times 233 \times 189$ mm^3 and resolution of $1 \times 1 \times 1$ mm^3; as a consequence in the following paragraphs voxels and mm^3 will be interchangeably used. Clinical and demographic information, including the Mini Mental State Examination (MMSE) score, age, years of education and gender for the $\mathcal{D}_{train}$ and $\mathcal{D}_{test}$ is detailed in **Table 1**. Except for MMSE scores, there were no significant differences among the three groups.

The study encompassed three principal phases: image processing, multiplex network analysis and information content assessment. The first phase is devoted to data normalization, it consists of processing steps which mitigate data heterogeneity; secondly, a network model is assigned to each subject and the comprehensive multiplex model describing the whole cohort is built; finally, quantitative measures are extracted from the model and are used to train a classifier. The overall processing pipeline is schematically represented in **Figure 1** and will be explained in detail in the following sections.

2.2. Image Processing

The nodes of the networks describing each subject should share the same anatomical content in order to be compared. Thus, the proposed approach requires that the same anatomical regions should roughly overlap in order to be robust to subtle local differences, due for example to subject morphological variability, or small registration failures.
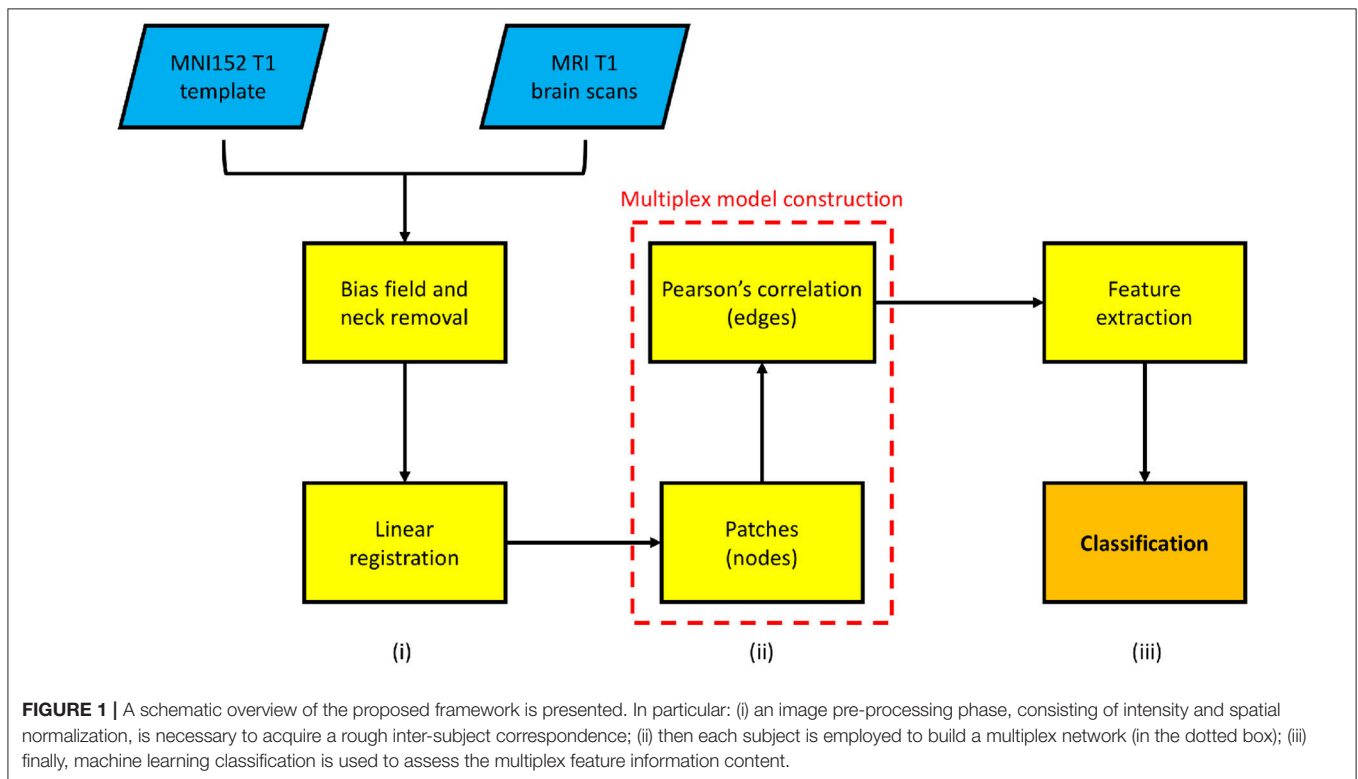
Accordingly, intra-cranial regions were extracted and MRI scan intensity differences, yielded by bias field, were normalized with the Oxford FMRIB library FSL (Jenkinson et al., 2012). Then, spatial normalization was performed to co-register the different images into the common coordinate space provided by the MNI152 template. An affine registration was performed with the FSL Linear Registration Tool (FLIRT) with a standard parameter configuration.

Finally, we divided the brain of each subject into the two hemispheres by the medial longitudinal fissure. Starting from this sagittal plane, it was possible to uniformly cover each hemisphere with an equal number of rectangular ($l_1 \times l_2 \times l_3$) boxes, from now onward referred to as "patches," covering the whole brain, see **Figure 2**. It is worth noting that, once the MRI scans and the template had been co-registered, they shared the same reference

TABLE 1 | Group size and gender information are reported for each class.

	$\mathcal{D}_{\textbf{train}}$		$\mathcal{D}_{\textbf{test}}$			**Total**
Disease status	AD (38)	NC (29)	AD (48)	NC (52)	cMCI (48)	215
Female/male	18/20	13/16	22/26	25/27	21/27	99/116
Age (years)	73.55 ± 8.00	74.97 ± 6.30	78.41 ± 6.02	74.77 ± 6.01	76.12 ± 5.89	75.56 ± 6.44
Education (years)	15.29 ± 3.43	16.93 ± 2.78	15.42 ± 3.13	16.37 ± 2.98	15.33 ± 3.03	15.87 ± 3.07
MMSE	22.68 ± 2.27	29.07 ± 0.92	24.32 ± 1.89	28.97 ± 0.81	26.78 ± 1.77	26.36 ± 1.53
n sites	23	19	26	29	18	34

The table also provides age, years of education and MMSE (mean and standard deviation). The disease status reported is referred at the baseline visit. According to t-test statistics, MMSE scores resulted significantly different with a p-value $p < 0.01$ for all comparisons except those between $\mathcal{D}_{train}$ and $\mathcal{D}_{test}$ normal controls and between $\mathcal{D}_{train}$ and $\mathcal{D}_{test}$ AD patients.

FIGURE 1 | A schematic overview of the proposed framework is presented. In particular: (i) an image pre-processing phase, consisting of intensity and spatial normalization, is necessary to acquire a rough inter-subject correspondence; (ii) then each subject is employed to build a multiplex network (in the dotted box); (iii) finally, machine learning classification is used to assess the multiplex feature information content.

space and therefore the anatomical content of each patch was almost the same.

The size D of the patches was chosen considering that too small patches could be considerably affected by registration noise, while a size too large, may make it impossible to distinguish subtle disease effects, often diffused to different parts of a region, due to natural inter-subject variability. To investigate how the size of patches affected the quality of the analysis, the overall patch volume D was varied from a minimum of 1,000 to a maximum of 4,000 voxels. The l_1, l_2, and l_3 values were chosen in order to obtain patches whose dimensions were divisor of the image size and divided regularly the image. Then, only the patches whose voxels overlapped the template brain mask more than 10% were considered.

The patches were considered nodes of a network whose connections represented the grade of similarity between them.

We, therefore, used different similarity metrics and a multiplex network framework, in order to extract inter- and intra-subject characteristics.

2.3. Multiplex Network Construction

Graph theory provides tools to concisely quantify the properties of complex networks that describe interrelationships (represented by edges) between the objects of interest (represented by nodes). In this work, for each image and, thus, for each subject, we built an N node undirected weighted network with nodes defined by brain MRI patches and edges defined by pairwise Pearson's correlation among them. Therefore, multiplex network $\mathcal{G} = \{\mathcal{G}_1, \mathcal{G}_2, ..., \mathcal{G}_\alpha, ..., \mathcal{G}_M\}$ was, in this case, a collection of single subject weighted networks $\mathcal{G}_\alpha = (N, \mathcal{E}_\alpha, \mathcal{W}_\alpha)$ (see **Figure 3** for a pictorial representation) sharing a common number of nodes N, while the set of links $\mathcal{E}_\alpha$
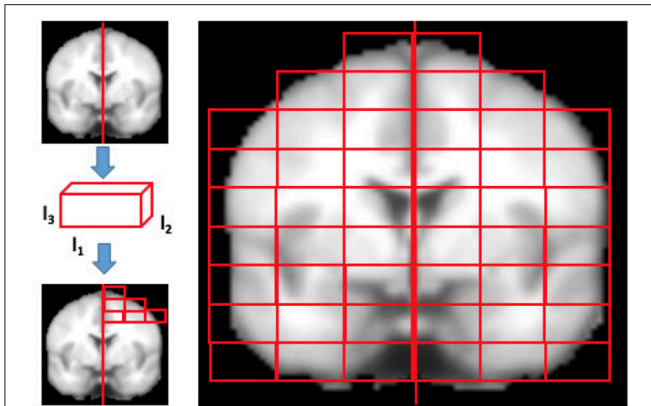
FIGURE 2 | The figure qualitatively shows how MRI brain scans are segmented in rectangular patches of dimensions $l_1 \times l_2 \times l_3$. Firstly, the brains normalized to MNI152 template are divided in left and right hemispheres using the medial longitudinal fissure, then the patch dimensions are set and finally the brain is segmented. Only patches overlapping the brain for at least the 10% of their content are kept, others are discarded.

FIGURE 3 | At the top: the multiplex network with M layers and N nodes. At the bottom: the representation of multi-links for the different pairs of network nodes. Within each layer different nodes can be connected with a link and a specific weight. This context information is then used to detect different patterns.

given patches s_i and s_j of dimension D, the Pearson's correlation coefficient r_{ij} is defined with $i, j = (1, ..., N)$:

$$r_{i,j} = \frac{\sum_{k=1}^{D}(s_i^k - \bar{s}_i)(s_j^k - \bar{s}_j)}{\sqrt{\sum_{k=1}^{D}(s_i^k - \bar{s}_i)^2}\sqrt{\sum_{k=1}^{D}(s_j^k - \bar{s}_j)^2}} \qquad (1)$$

The numerator is the sum over the product of the voxels intensities s_i^k and s_j^k at each voxel position k after subtraction of the patch average values, and the denominator is the product of the standard deviations of s_i and s_j gray-level distributions.

Pearson's correlation was chosen to model the effects of atrophy, as it is fast to implement and compute, simple to understand and interpret, and it does not require any scaling or centering of the patches as it is intrinsically normalized. In addition, correlation is a similarity criterion that associates corresponding voxels within patches, therefore taking into account spatial relationships between voxels. To investigate the importance of preserving spatial voxel correspondence when building the multiplex, a preliminary study about similarity metrics had been previously performed (see **Supplementary Material**), which demonstrated that Pearson's correlation was the optimal choice.

Pearson's correlation admits negative values, thus in principle it could be adopted for a directed weighted network description. In the case discussed here, it is worth noting that negative correlations can be found, for example, between patches in which gray matter and white matter undergo a left-right inversion. As a result, distinguishing positive and negative correlations would have included in the multiplex model a left-right bias. As asymmetry is a common characteristic of atrophy in AD, it was decided to consider undirected networks (see **Figure 4**).

Network edges can be weighted or unweighted. Unweighted network topology is easier to study and interpret, and has computational advantages. On the one hand, even if in several cases the decision to binarize a weighted network with a suitable threshold could be appropriate, this would seem a forced decision in our case, with the patch similarity being an intrinsically continuous measure. On the other hand, weighted networks can include weak relationships that might be spurious and introduce noise into the graph. Therefore, we decided to threshold the networks by setting to 0 all connections whose absolute correlation was less than moderate ($|r| < 0.3$), in order to exclude noisy interrelationships in the model, and reducing as much as possible the loss of important links. For higher correlations, weights were kept in the model, thus resulting in a weighted undirected network representation for each subject:
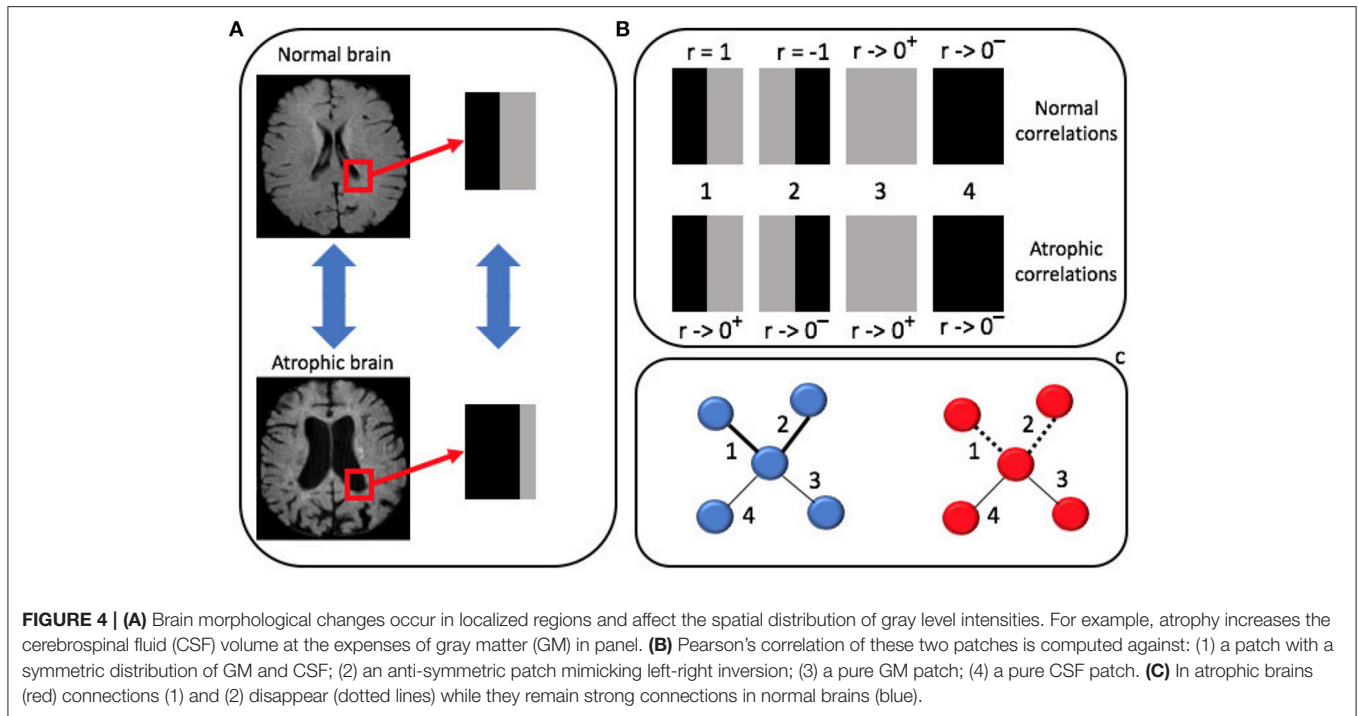
$$w_{ij} = \begin{cases} 0, & \text{if } |r_{ij}| \leq 0.3 \\ r_{ij}, & \text{otherwise} \end{cases} \qquad (2)$$

An investigation on how the threshold affects the multiplex network ability to detect diseased patterns is reported in the following section 3.1.

In a multiplex it is possible to introduce several topological characteristics that are usually adopted to describe a complex network (Menichetti et al., 2014; Amoroso et al., 2018). In our
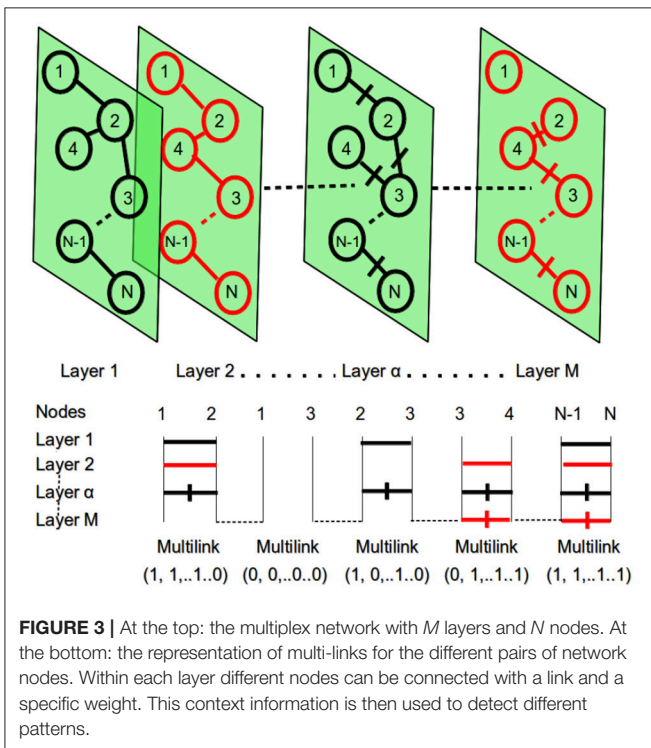
changed depending on the layer (subject) α. Each network $\mathcal{G}_\alpha$ can also be represented by the corresponding adjacency matrix $A_\alpha = a_{ij}^\alpha$, a useful notation to investigate the network properties. Hence, the proposed model is a multiplex composed of $M = 67$ weighted undirected networks: each representing an MRI brain scan, and including N nodes or patches. For each layer, interrelationships were described by $\mathcal{W}_\alpha = \{w_{ij}^\alpha\}$ in which w_{ij} were given in terms of Pearson's correlation. In particular,

FIGURE 4 | (A) Brain morphological changes occur in localized regions and affect the spatial distribution of gray level intensities. For example, atrophy increases the cerebrospinal fluid (CSF) volume at the expenses of gray matter (GM) in panel. **(B)** Pearson's correlation of these two patches is computed against: (1) a patch with a symmetric distribution of GM and CSF; (2) an anti-symmetric patch mimicking left-right inversion; (3) a pure GM patch; (4) a pure CSF patch. **(C)** In atrophic brains (red) connections (1) and (2) disappear (dotted lines) while they remain strong connections in normal brains (blue).

approach we employed the following indicators: the strength s_i^α and the inverse participation ratio Y_i^α of a node i in layer α:

$$s_i^\alpha = \sum_{j=1}^{N} w_{ij}^\alpha \qquad (3)$$

$$Y_i^\alpha = \sum_{j=1}^{N} \left(\frac{w_{ij}^\alpha}{s_i^\alpha} \right)^2 \qquad (4)$$

Strength measurements denote which nodes are more relevant within the network describing a single layer (i.e., a subject) of the multiplex. Inverse participation ratio attains the heterogeneity of the weight distribution within each layer.

Along with these two measurements we also evaluated the conditional means of strength $s(k)^\alpha$ and inverse participation $Y(k)^\alpha$ against the nodes with degree k:

$$s(k)^\alpha = \frac{1}{N_k} \sum_{i=1}^{N} s_i^\alpha \delta(k_i^\alpha, k) \qquad (5)$$

$$Y(k)^\alpha = \frac{1}{N_k} \sum_{i=1}^{N} Y_i^\alpha \delta(k_i^\alpha, k) \qquad (6)$$

Summation is extended over the N_k nodes having degree k; as summation includes a Kronecker δ function, the only non-null terms, for both strength and inverse participation, are referred to nodes i of the layer α whose degree is k. These quantities help to understand how weights are distributed within each layer, thus, for example, distinguishing whether, on average, the weights of central nodes and less connected nodes are identically distributed
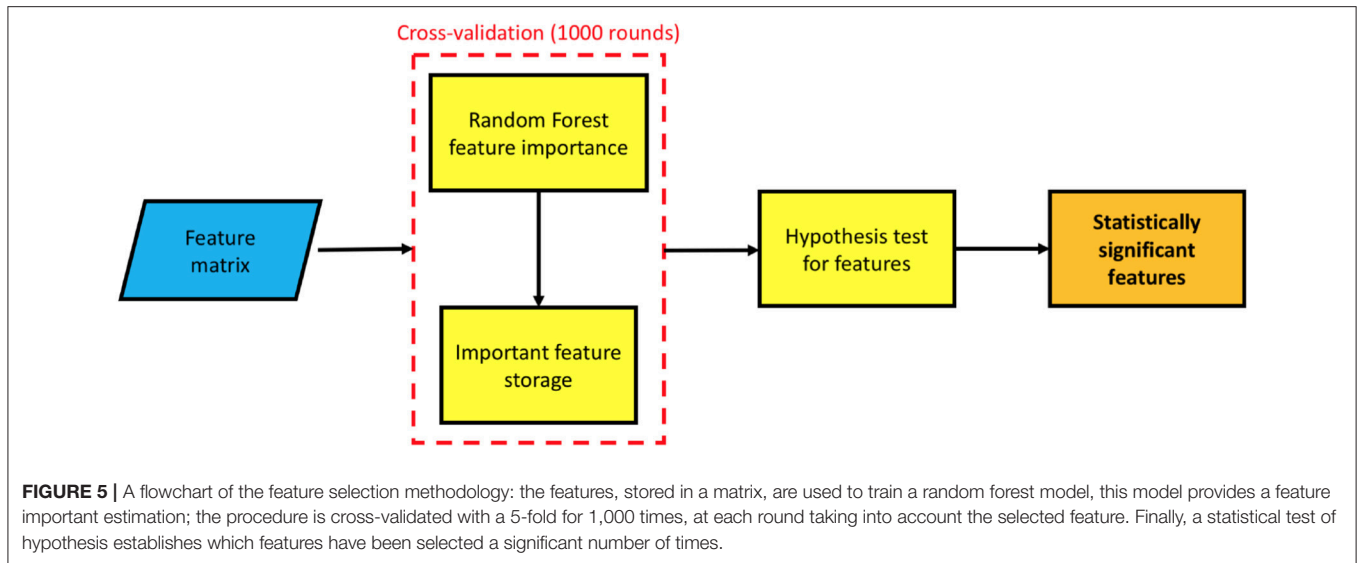
or not. Several studies have already pointed out, especially with group-wise single layer approaches (Tijms et al., 2013a), how these features can describe significant differences among healthy and diseased subjects.

However, it is reasonable to assume that further evidence of significant differences between subjects, can arise from the context information provided by the multiplex framework. Accordingly, this information content was exploited by considering the aggregate adjacency matrix $A^{multi} = a_{ij}^{multi}$ where:

$$a_{ij}^{multi} = \{1 \ if \ \exists \alpha | w_{ij}^\alpha > 0 \quad \wedge \quad 0 \quad otherwise\} \qquad (7)$$

The matrix A^{multi} naturally allowed us to re-introduce the previous measurements within a global perspective. In fact, it was possible to compute for each node an aggregated degree and then use it to weight the previously defined strength and inverse participation. Analogously, we used A^{multi} to define the aggregate degree for each node and then re-computing the conditional means. In this way we introduced in the description of each node the information produced by the whole multiplex.

In conclusion each network was described by 8N features (4N single layer and 4N multiplex features), resulting in a $M \times 8N$ feature representation which from now on we will call $\mathcal{F}_{train}$. It is worthwhile to note that this characterization was independent from the clinical status of the subjects as the multiplex had been built blindly to diagnosis. This base of knowledge was then investigated with supervised machine learning models to extract specific disease effect patterns.

FIGURE 5 | A flowchart of the feature selection methodology: the features, stored in a matrix, are used to train a random forest model, this model provides a feature important estimation; the procedure is cross-validated with a 5-fold for 1,000 times, at each round taking into account the selected feature. Finally, a statistical test of hypothesis establishes which features have been selected a significant number of times.

2.4. Assessment and Validation

The multiplex characterization of the images yielded a simple matrix representation, which could be used to feed machine learning models, and unveil discriminating anatomical patterns.

The number of features f, involved in this approach, could easily reach values ranging from $\sim 10^3$ to $\sim 10^4$ outnumbering the number of the available training samples. Thus, to prevent over-training issues, arising from the curse of dimensionality and assess the multiplex framework, a feature selection was necessary. A flowchart of the whole feature selection method is represented in **Figure 5**.

A 5-fold cross-validation feature importance selection was performed within a wrapper-based strategy. We randomly divided 1,000 times $\mathcal{F}_{train}$ in a training and a validation test. For each cross-validation round, we built a multiplex model on training subjects, then we computed the important features. In particular, we measured the total decrease in node impurities, in terms of Gini index, from splitting on the variable, averaged over all trees. The selected features were stored for later use and used to train a second random forest classifier which was used to predict the diagnosis of the validation subjects. An evaluation of the informative content of this representation is presented in section 3.2. In both cases random forests were grown with 500 trees, a number large enough for the out-of-bag error to reach the typical training plateau. At each split $\sqrt{f}$ features were randomly sampled.

As previously mentioned, for each cross-validation round different features were selected, thus a quantitative criterion was necessary to determine the most important features, independently from training set. This problem was solved by taking into account the overall occurrence rate of each feature and interpreting it as a success rate. As a consequence a binomial distribution was observed and an experimental p-value could be computed to test the randomness hypothesis. We tested it with a $p < 0.01$ to select a more exiguous number of features, then we established which ones had shown a significant probability

of occurrence. Once the best features had been selected, we used them to train a new ensemble model on $\mathcal{D}_{train}$ and tested it on $\mathcal{D}_{test}$ to assess the method robustness and evaluate the informative content carried by multiplex features.
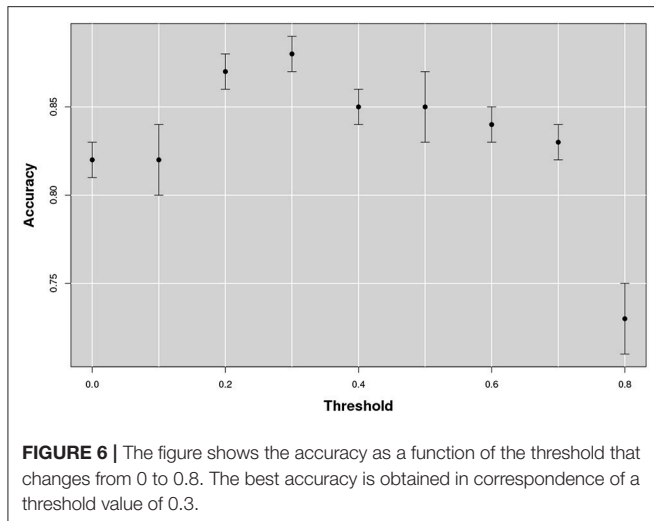
For test subjects, single layer features were straightforwardly computed. Features accounting the whole multiplex structure were in turn computed adding the test subject to the training multiplex but keeping fixed $\mathcal{F}_{train}$. The reason for this choice can be justified considering the perturbation induced by the addition of one layer is small.

It is worth noting that features like strength and inverse participation have a direct interpretation, being directly related to a single patch of the brain network whilst conditional means, by definition, are related to several nodes sharing a common degree k. For classification purposes this is not an issue, being based on computed features; on the contrary this is relevant in order to provide an anatomical interpretation and a diagnostic value of the features selected.

2.5. Anatomical Interpretation

Since the identification of the nodes is based on a purely mathematical approach, it seemed important to investigate the relationship between network features and anatomical areas of interest for the disease.

Nodes, whose features were significantly related to AD, were localized on the reference template and the corresponding atlas. We adopted Harvard-Oxford cortical and sub-cortical structural atlases (Desikan et al., 2006). For conditional mean features, which intrinsically encode the information contained in different nodes, we identified nodes significantly related to AD. Next, for each one, we recorded subject by subject the patches having the degree k used to compute that specific conditional mean feature. Then, we computed an occurrence rate taking into account how many times a patch had been used to compute that conditional mean. At this point, patches significantly correlated to AD were identified by interpreting the occurrence as a success rate, and

FIGURE 6 | The figure shows the accuracy as a function of the threshold that changes from 0 to 0.8. The best accuracy is obtained in correspondence of a threshold value of 0.3.

FIGURE 7 | The figure represents the accuracy for the NC-AD classification as a function of the patch size. The existence of a robust plateau, in correspondence of $[2, 250, 3, 200]$ voxels, is highlighted in the circle. These results suggest the existence of an optimal dimensional scale for multiplex describing AD atrophy patterns.

testing the hypothesis of randomness according to a binomial distribution with $p < 0.01$. This methodology allowed us to detect a restricted number of anatomical districts associated to AD, as shown in section 3.3.

3. RESULTS

3.1. Threshold Assessment

Since this approach could in principle heavily depend on the threshold value adopted to discard negligible correlations, the threshold values ranging from 0 to 0.8 were explored with a 0.1 step. Then, for each threshold value a different multiplex was constructed. The patch dimension adopted was 3,000 mm^3. The training classification performance was measured in terms of accuracy, see **Figure 6**.

The classification accuracy reached its maximum value with a 0.3 threshold value and it remained stable over 0.85 for a large range of correlations $[0.2, 0.5]$. With lower or higher threshold performances showed a significant decrease, especially above the 0.8 threshold; in which case more of the 50% of the networks resulted empty.

3.2. Scale Selection and Informative Content

Firstly, we investigated on training the optimal number of nodes N to be adopted and, secondly, whether the features thus arising could be used to distinguish NC and AD subjects on the available datasets. This is because the number of nodes N of the multiplex, as well as the correlation measure among the different patches, depends on the patch size. As there was no *a priori* reason to choose the patch size, we examined to which extent the size of the patch affected the classification accuracy in discriminating healthy controls and AD subjects from the training data subset (see **Figure 7**).

From this analysis we found that the optimal size for the patch was of $10 \times 15 \times 20$ mm^3 equal to an overall volume of 3,000 mm^3. Accuracy increased with the patch size until the range $[2250, 3200]$ mm^3 was reached. At this scale, discarding
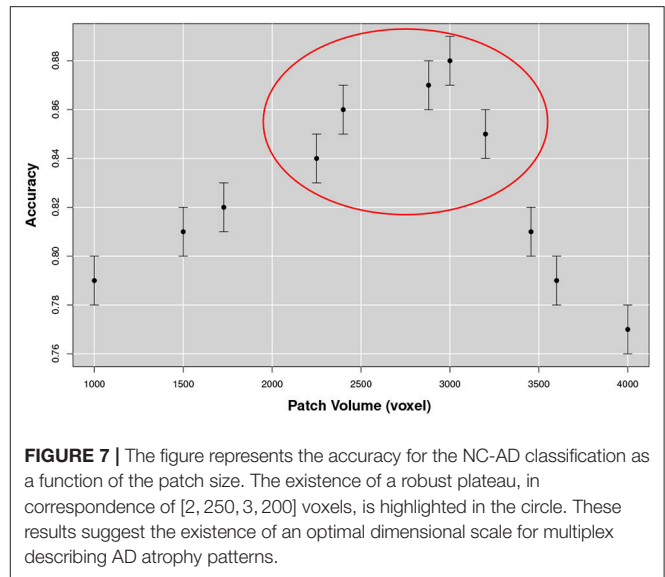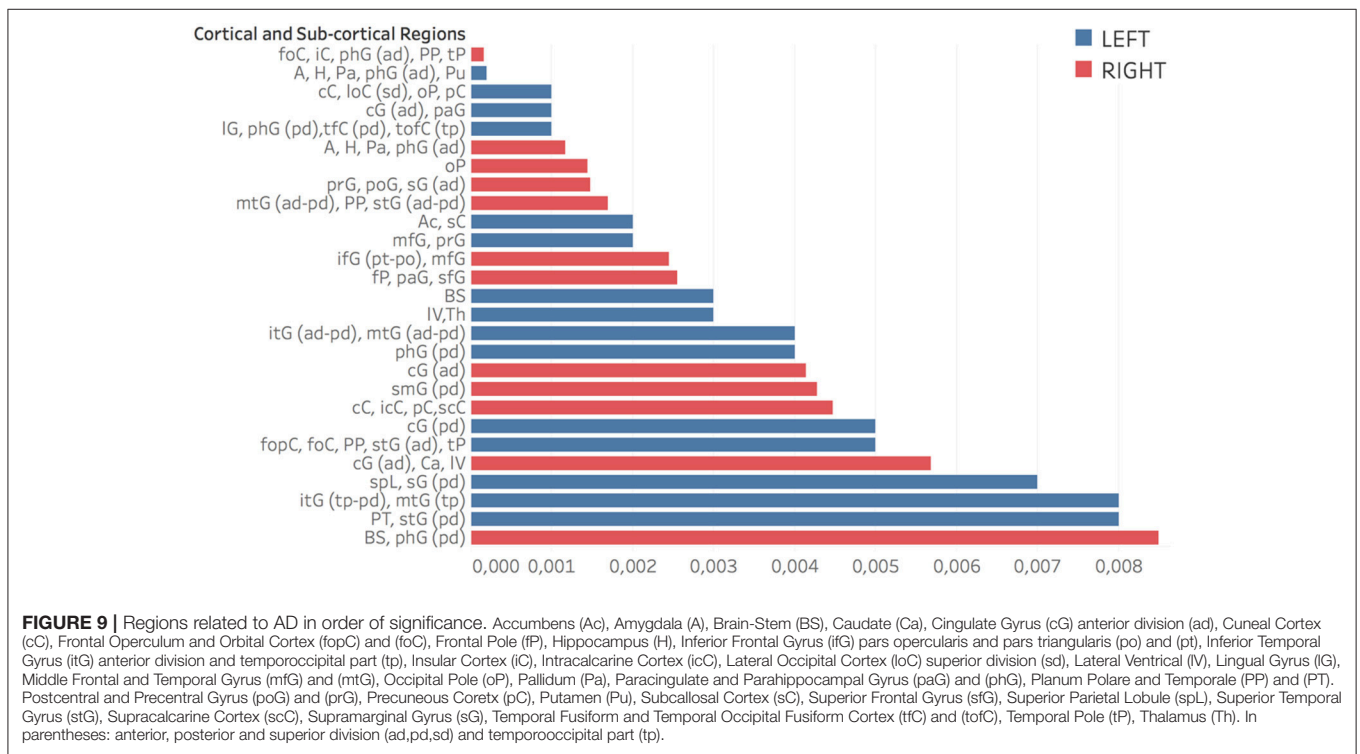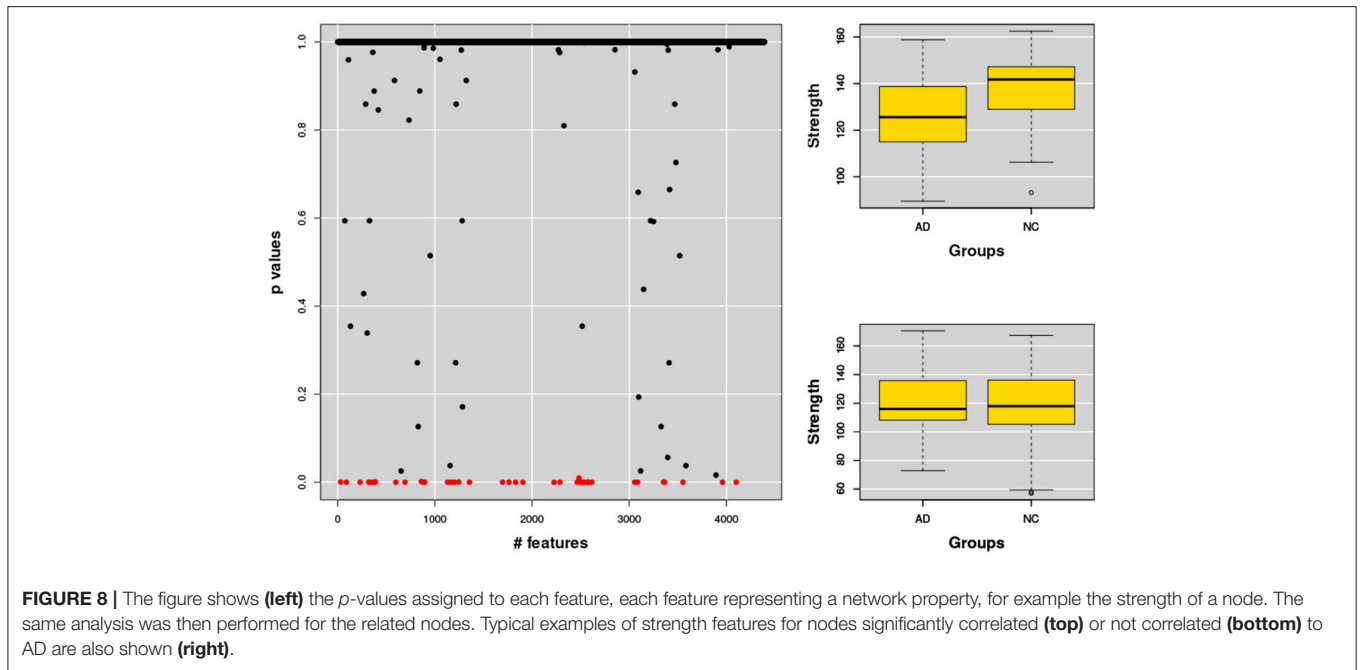
the patches overlapping the template brain with less than 10% of voxels, 549 patches were obtained for each image. The corresponding accuracy value was on average 0.88 with a 0.01 standard error and a sensitivity and a specificity respectively of 0.90 ± 0.01 and 0.88 ± 0.02. We compared this performance using 180 structural morphological features, obtained by FreeSurfer (6.0 version) (Fischl, 2012), with the same classification strategy, including a first random forest wrapper for feature selection and a second random forest classifier for prediction. In this case classification performance was on average significantly lower 0.83 ± 0.01 confirming the effectiveness of the multiplex characterization.

3.3. Anatomical Characterization

Once the optimal dimension of multiplex network had been fixed we selected the most representative features according to their relative importance. As explained in section 2.4 we selected those features whose contribution to the classification was considerably distant from the null hypothesis of a random behavior, see **Figure 8** for a typical example.

The whole base of knowledge consisted of 32 significant patches, 18 ($\sim$ 56%) in the left hemisphere and 14 in the right, including 27 different cortical and sub-cortical regions listed in the following **Figure 9** in order of significance. As a region can be included in different patches (provided at least one of its voxels belongs to the considered patch), only most significant p-value entries are reported.

In **Figure 10** some representative brain axial planes are shown, as well as the Harvard-Oxford atlas we used for this assessment. In the left hemisphere, patches corresponding to amygdala, hippocampus, para-hippocampal gyrus, pallidum and putamen showed the strongest association to AD ($p = 0.0001$). For cingulate and para-cingulate giri, pre-cuneus, cuneus, and occipital cortex $p = 0.001$. Other significant patches ($p = 0.002$) were located in middle frontal gyrus and pre-frontal gyrus, nucleus accumbens, brain stem and thalamus.
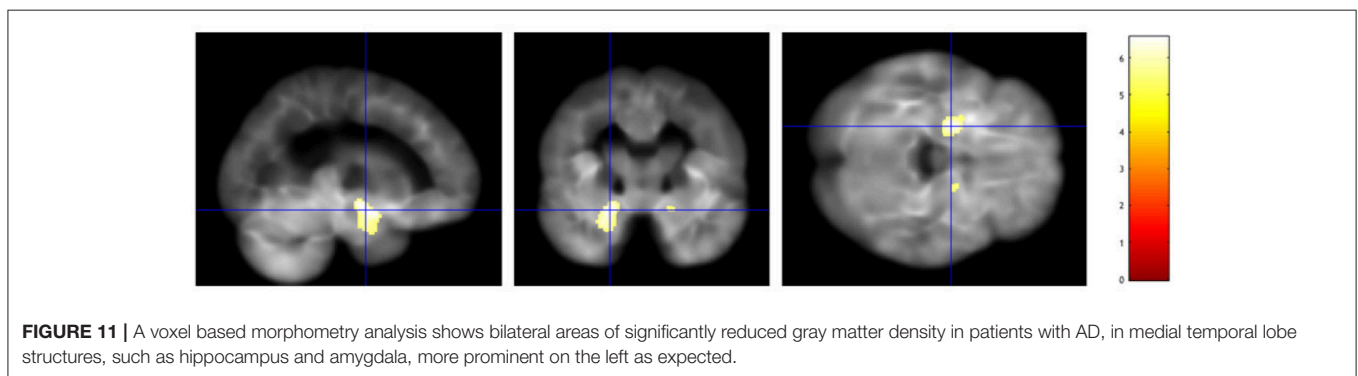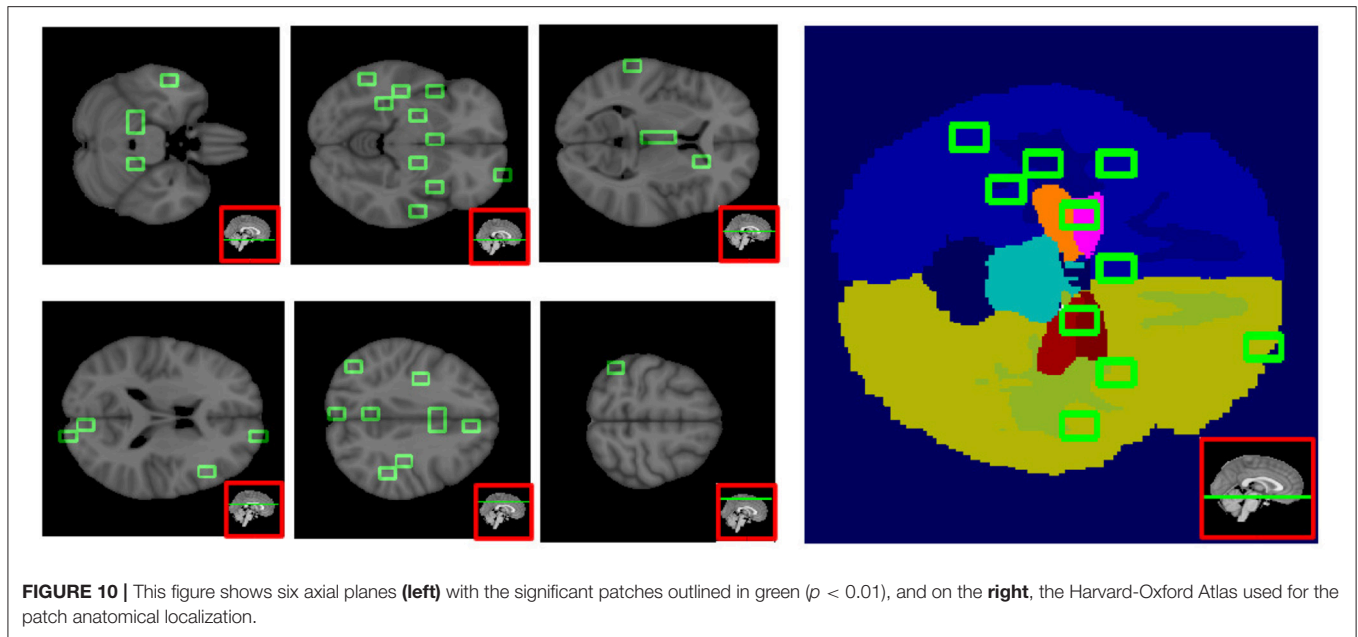
FIGURE 8 | The figure shows **(left)** the p-values assigned to each feature, each feature representing a network property, for example the strength of a node. The same analysis was then performed for the related nodes. Typical examples of strength features for nodes significantly correlated **(top)** or not correlated **(bottom)** to AD are also shown **(right)**.

FIGURE 9 | Regions related to AD in order of significance. Accumbens (Ac), Amygdala (A), Brain-Stem (BS), Caudate (Ca), Cingulate Gyrus (cG) anterior division (ad), Cuneal Cortex (cC), Frontal Operculum and Orbital Cortex (fopC) and (foC), Frontal Pole (fP), Hippocampus (H), Inferior Frontal Gyrus (ifG) pars opercularis and pars triangularis (po) and (pt), Inferior Temporal Gyrus (itG) anterior division and temporooccipital part (tp), Insular Cortex (iC), Intracalcarine Cortex (icC), Lateral Occipital Cortex (loC) superior division (sd), Lateral Ventrical (lV), Lingual Gyrus (lG), Middle Frontal and Temporal Gyrus (mfG) and (mtG), Occipital Pole (oP), Pallidum (Pa), Paracingulate and Parahippocampal Gyrus (paG) and (phG), Planum Polare and Temporale (PP) and (PT). Postcentral and Precentral Gyrus (poG) and (prG), Precuneous Coretx (pC), Putamen (Pu), Subcallosal Cortex (sC), Superior Frontal Gyrus (sfG), Superior Parietal Lobule (spL), Superior Temporal Gyrus (stG), Supracalcarine Cortex (scC), Supramarginal Gyrus (sG), Temporal Fusiform and Temporal Occipital Fusiform Cortex (tfC) and (tofC), Temporal Pole (tP), Thalamus (Th). In parentheses: anterior, posterior and superior division (ad,pd,sd) and temporooccipital part (tp).

On the right, $p = 0.0001$ for orbito-frontal cortex, insular cortex, prarahippocampal gyrus, planum polare and planum temporale; $p = 0.001$ for the parahippocampal-amygdalar complex, occipital pole, pre- and post-central gyri, supramarginal gyrus, middle and superior temporal gyri; $p = 0.002$ for inferior, middle and superior frontal gyri, frontal pole, and paracingulate gyrus.

It is interesting to note that frontal lobe involvement was more prominent on the right.

3.4. Multiplex Networks vs Voxel Based Morphometry

In order to establish if this new approach may offer any advantages over existing widely used methods, we analyzed

FIGURE 10 | This figure shows six axial planes **(left)** with the significant patches outlined in green ($p < 0.01$), and on the **right**, the Harvard-Oxford Atlas used for the patch anatomical localization.

FIGURE 11 | A voxel based morphometry analysis shows bilateral areas of significantly reduced gray matter density in patients with AD, in medial temporal lobe structures, such as hippocampus and amygdala, more prominent on the left as expected.

the same data set with Voxel Based Morphometry (VBM) (Ashburner and Friston, 2000).

We followed the standard prescription for VBM with the publicly available SPM 12 suite[1]. Firstly, a segmentation of brain tissues was performed, followed by non-linear normalization with the SPM tool DARTEL to create a study specific template. Secondly, we performed a smoothing with an isotropic Gaussian filter with a full width at half maximum of 8 mm. Lastly, a two-sample analysis was performed with a t statistics to investigate significant group-wise differences in atrophy between NC and AD on training subjects. Significant voxels, with 5% family-wise correction, are represented below in **Figure 11**.

The VBM analysis showed significant reduction in gray matter density in bilateral peri-hippocampal regions, more prominent of the left.

3.5. Left/right Characterization
Since the VBM analysis confirmed that left-sided changes were more prominent, two dedicated tests were carried out to further explore the lateralization. Firstly, we used the $\mathcal{D}_{train}$ to compute

the multiplex features, then we selected only those inherent to the left (right) hemisphere and trained the classification models. The feature selection and the cross-validation procedures described in section 2.4 were perfectly replicated as the goal of this test was to quantify the information content of features related to left (right) hemisphere regions. We found that left patches were able to discriminate NC from AD patients with an accuracy of 0.87 ± 0.01 while right hemisphere features were able to reach the accuracy value 0.85 ± 0.01. Left hemisphere remained responsible for a greater part of the overall information of the multiplex framework, which was 0.88 ± 0.01.

It must be taken into account that each patch, summarizes a network of interrelationships with other patches independently from its spatial collocation. As an example, the strength of a node denotes the sum of its connections, the fact that a node of the left hemisphere is significantly related to AD does not prevent its strength to be the result of its correlation with the right hemisphere.

As a consequence, a second test was performed. We considered the multiplexes of left and right hemispheres separately. This was done dividing each brain scan in two
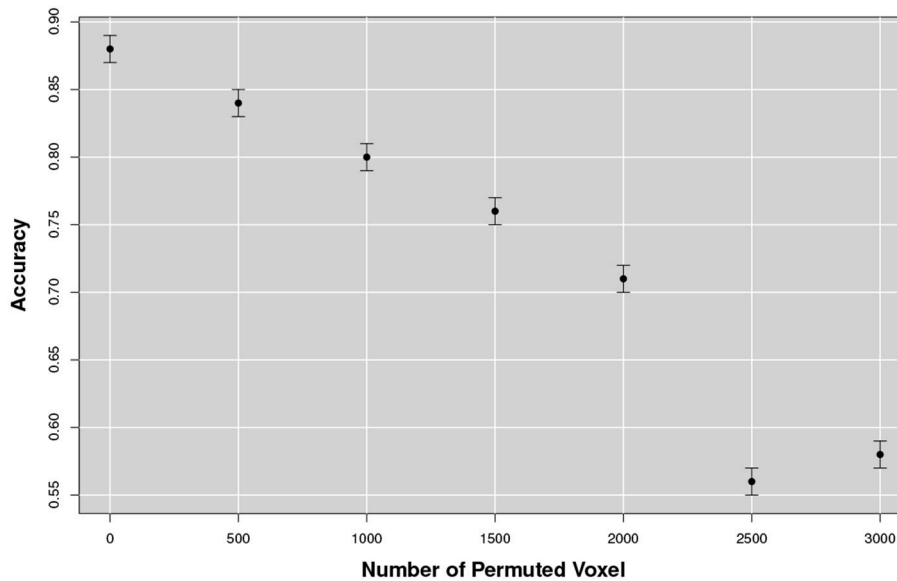
FIGURE 12 | Accuracy varying with the number of permuted voxel within a patch. Classification performance decreased as the number of shuffled voxels was increased. Noticeably, a drastic drop was observed when the shuffle reached values of about $2,500 \sim 3,000$ voxels.

different images containing the two hemispheres and then using only one half to build the multiplex. Accordingly, the multiplex features computed in this case could be genuinely considered as related to only one hemisphere. Even in this case we performed feature-selection and cross-validation analyses reproducing the whole brain procedure. Classification accuracy for NC-AD when using left multiplex was 0.83 ± 0.01, for right we found 0.81 ± 0.01, thus confirming the greater involvement of the left hemisphere but also signaling a definite deterioration of the information content if compared with the whole brain multiplex.

3.6. Robustness and Generalization

To investigate if classification performance was related to the random permutation of voxels inside a patch, we firstly shuffled a varying number of voxel within each patch, while keeping the patch decomposition stable, thus affecting the Pearson's correlation pairwise measurement. Then we measured the classification accuracy. The training results are presented in **Figure 12**.

The test was repeated 100 times increasing the size of the shuffle by 500 voxels at the time. It could be noticed that for small variations, under 1,000 voxels, performance did not suffer a significant deterioration; but with $2,500$ voxel permutation a drastic drop of the performance was observed, a value comparable with the dimensional scale determined in section 3.2.

To further assess the method robustness we also performed a classical non-parametric statistical permutation test. This consisted in the permutation of the clinical labels of each subject belonging to $\mathcal{D}_{train}$. We performed 1,000 random permutation and observed (see **Figure 13**) a consistent decrease of the

classification performance suggesting that the selected features do characterize the disease.

Training set randomization effectively established that the multiplex framework was able to model a significant structure in the $\mathcal{D}_{train}$ data between the multiplex features and the clinical label. Moreover, given the normality of the performance distribution obtained by permuting the labels, it was possible to assign a p-value to the performance obtained without permutations. The result showed that the multiplex model was able to identify a significant ($p < 0.001$) class structure within the $\mathcal{D}_{train}$ data. Otherwise, it would not have been possible to reject the null hypothesis underlying this test, i.e., that labels and features were independent, so that in fact no difference really existed between the classes.

As a further assessment we performed a binary classification on the $\mathcal{D}_{test}$ for the NC-AD and NC-cMCI classes. The analysis was repeated using 100 bootstrapped $\mathcal{D}_{test}$ sets to provide a measurement of the performance uncertainty. We found in terms of accuracy, respectively 0.86 ± 0.01 and 0.84 ± 0.01. The respective specificity were 0.74 ± 0.01 and 0.72 ± 0.01, while sensitivity reached higher values for both cases: 0.96 ± 0.01 and 0.94 ± 0.01. Remarkably, the NC-cMCI classification performance compared well with NC-AD classification confirming the method reliability and its informative content.

The small, but significant, performance deterioration (training accuracy was 0.88 ± 0.01, see section 3.2) could be expected, mainly because even if the test perturbation of the training multiplex was considered small, it should not be completely neglected. The implementation of larger training sets could in principle mitigate this effect. A summary of the classification performances obtained for the different groups are shown in **Table 2**.
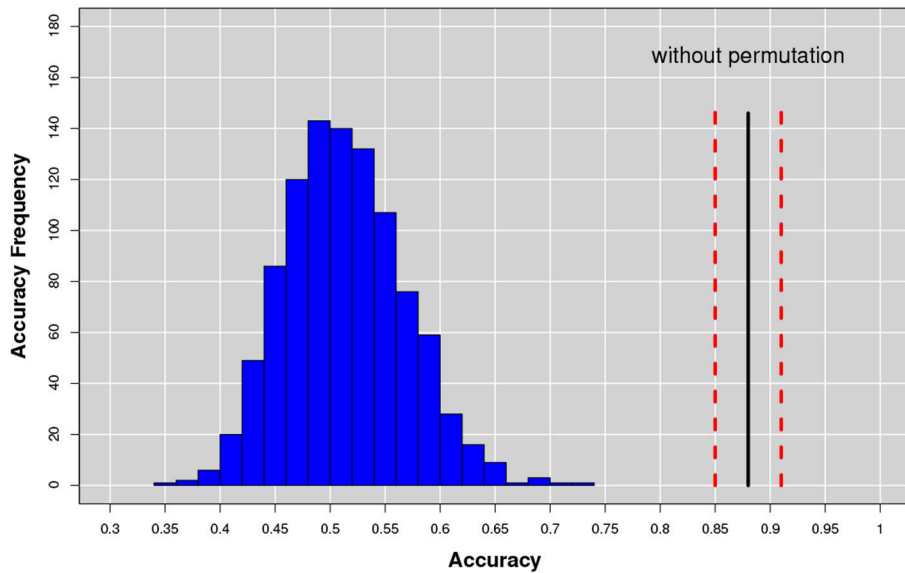
FIGURE 13 | The accuracy distribution for the binary problem NC-AD on the $\mathcal{D}_{train}$ with a random permutation of the subject labels. The average value (continuous line) and the relative uncertainty (dotted lines) of best training performances obtained without permutation are also represented for direct comparison.

TABLE 2 | Summary of the classification performances in terms of accuracy, sensitivity specificity and relative standard errors for the different groups: NC-AD used for the training, NC-AD and NC-cMCI considered for the validation.

Groups	Accuracy	Sensitivity	Specificity
$\mathcal{D}_{train}(NC - AD)$	0.88 ± 0.01	0.90 ± 0.01	0.88 ± 0.02
$\mathcal{D}_{test}(NC - AD)$	0.86 ± 0.01	0.96 ± 0.01	0.74 ± 0.01
$\mathcal{D}_{test}(NC - cMCI)$	0.84 ± 0.01	0.94 ± 0.01	0.72 ± 0.01

It is worth noting that these performances were obtained using a subset of 70 features including both single-layer and multiplex features.

4. DISCUSSION

The proposed approach aims at modeling brain atrophy in AD through inter-subject multiplex networks whose nodes are represented by brain patches and edges by pairwise Pearson's correlations. Metrics preserving the spatial information as Pearson's correlation and Mutual Information yield accurate results, with the first to be preferred for interpretability and performance consideration. To discard negligible correlations and improve the method sensitivity (as a result of a higher signal to noise ratio) we removed edges with weight below a threshold value of 0.3. Applying this threshold the method appeared robust and the classification performance remained stable over a broad range of correlations ([0.2, 0.5]). Outside this range a performance drop was observed. This is because lower threshold values introduced noisy correlations within the model, thus concealing the effective network information, whilst greater threshold values were too penalizing as informative links were neglected. However, in this study and other similar works (De Vico Fallani et al., 2017), determining the optimal threshold remains an open issue and somehow it limits the robustness of the results.

The method proved to have high sensitivity and high discriminatory power, being therefore suitable both for descriptive and classificatory purposes. As to sensitivity, an optimal volume size for the detection of AD effects, maximizing the informative content of the multiplex, was identified as ranging from 2, 250 to 3, 200 mm^3. This range can be easily interpreted considering that brain differences may be missed on smaller scales, due for example to misregistration errors; dimensional scales too large, on the contrary, may not capture subtle differences affecting small portions of the brain.

The high sensitivity of the method in the detection of illness related brain changes was demonstrated by the number of regions that were identified as significantly associated with AD. The detected regions comprised hippocampus and para-hippocampal-amygdalar complex, pallidum and putamen, cingulate and paracingulate giri, pre-cuneus, cuneus, and occipital cortex, middle frontal gyrus, pre-central gyrus, accumbens, sub-callosal cortex and brain stem.

While the prominent role in AD pathology of medial temporal lobe structures is widely recognized, the involvement of several other cortical and subcortical areas may be less obvious.

The cingulate cortex is a key component of the default mode network (Buckner et al., 2008), and its early involvement in AD pathology, has been amply demonstrated by functional and structural studies (Minoshima et al., 1997; Yokoi et al., 2018). The same is true for posterior areas, such as cuneus and pre-cuneus, also known to be affected by the illness in early stages

(Baron et al., 2001; Bailly et al., 2015). As to the involvement of subcortical gray matter in AD, this has also been recognized, and shown to correlate with cognitive impairment (de Jong et al., 2008). Volume loss of the nucleus accumbens was found to increase the risk of progression from MCI to AD (Yi et al., 2015).

The brain stem is a key area in the early pathophysiology of Parkinson's disease, another common neurodegenerative disorder, and alterations of the brain stem in AD have been shown both *in vivo* (Braun and Van Eldik, 2018), and post-mortem (Simic et al., 2009).

It was striking how VBM on the same data set was able to detect only atrophy of the perihippocampal regions. The method here described seems more sensitive than standard VBM (Good et al., 2002), while studies adopting advanced VBM methodologies have also shown better results (Karas et al., 2003).

The whole base of knowledge consisted of 32 regions significant patches, but only 22 concerned single-layer measures; the multiplex model thus allowed a consistent increment (+46%) in the detection of significant brain regions.

The results also confirmed asymmetry in the spatial distribution of significant patches, mostly located in the left hemisphere, in keeping with several other studies (Fennema-Notestine et al., 2009; Derflinger et al., 2011; Long et al., 2018). This asymmetry has a direct effect on the informative content.

As to the application of this methodology to disease classification studies, the method is based on the assumption that the introduction of a test subject in the multiplex is not able to significantly perturb the multiplex itself, so that trained models can be easily used for prediction. In fact, on $\mathcal{D}_{test}$ there is not a great deterioration of the classification performance and the reliability of the framework remains optimal for classification purposes. The framework is robust and accurate, its informative content does not show extreme variations with random shuffling of the voxels inside the patches.

Classification performances are accurate and comparable with recent classification-focused studies (Bron et al., 2015; Moradi et al., 2015; Salvatore et al., 2015; Feng et al., 2018). Even though providing a diagnosis support system is not the main goal of this work, results are encouraging in this sense. Indeed, multiplex model features are able to efficiently capture inter-subject variability underlining disease pattern. An even more refined classification could have been achieved including, as suggested by our previous works, structural features (Amoroso et al., 2014) or longitudinal information (Chincarini et al., 2016).

The method was robust and able to provide a sensitive and informative base of knowledge. This was in particular true when the results were compared with the classification performance using FreeSurfer features. While the present study has been focused of the application of multiplex to disease classification, the method has great versatility and lends itself to a variety of purposes, including the identification of "disease signature" for more anatomically heterogeneous forms of neurodegenerative disorder, such as tauophathies or synucleinopathies, where the model could be enriched with additional clinical or genetic data.

5. CONCLUSION

In this paper we propose a novel approach based on multiplex networks to characterize brain structural variations related to AD. We investigated the information content provided by multiplex networks and showed that they produce an accurate modeling of the disease.

We demonstrated how this framework is able to provide a robust method for AD characterization: (i) it shows the existence of an optimal scale for the description of disease effects of [2,250, 3,200] voxels. (ii) Starting from a robust unsupervised brain parcellation, it correctly identifies cerebral region significantly related to AD. It also confirms that AD pathology is more prominent in the left hemisphere. (iii) Multiplex networks are a robust and effective method to describe disease patterns. In fact, after a training phase that gives in cross-validation an accuracy of 0.88 ± 0.01, the multiplex base of knowledge, on the independent dataset $\mathcal{D}_{test}$, is able to accurately distinguish between NC and AD subjects with an accuracy of 0.86 ± 0.01 and can be suitably employed also for NC and cMCI classification with an accuracy of 0.84 ± 0.01.

The information content provided by multiplex characterization was able to efficiently detect disease patterns. Also the method is very suitable to application to longitudinal studies, ideally in association with functional imaging, to improve our understanding of the different patterns of neurodegeneration in different diseases. The impact of variables such as the degree of atrophy, disease duration, site or scanner type could also be investigated in further studies.

ETHICS STATEMENT

AUTHOR CONTRIBUTIONS

NA and ML conceived and conducted the analyses, SB gave clinical support. All authors NA, RB, SB, ML, TM, AM, and ST analyzed the results and reviewed the manuscript.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2018.00365/full#supplementary-material

REFERENCES

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., et al. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 7, 270–279. doi: 10.1016/j.jalz.2011.03.008

Alexander-Bloch, A. F., Vértes, P. E., Stidd, R., Lalonde, F., Clasen, L., Rapoport, J., et al. (2012). The anatomical distance of functional connections predicts brain network topology in health and schizophrenia. *Cereb. Cortex* 23, 127–138. doi: 10.1093/cercor/bhr388

Alzheimer's Association (2018). 2018 Alzheimer's disease facts and figures. *Alzheimers Dement.* 14, 367–429. doi: 10.1016/j.jalz.2018.02.001

Amoroso, N., Errico, R., and Bellotti, R. (2014). "PRISMA-CAD : fully automated method for Computer-Aided Diagnosis of Dementia based on structural MRI data," in *Proceedings of the Computer-Aided Diagnosis of Dementia Based on Structural MRI Data, MICCAI 2014* (Boston, MA), 16–24.

Amoroso, N., Errico, R., Bruno, S., Chincarini, A., Garuccio, E., Sensi, F., et al. (2015). Hippocampal unified multi-atlas network (HUMAN): protocol and scale validation of a novel segmentation tool. *Phys. Med. Biol.* 60, 8851. doi: 10.1088/0031-9155/60/22/8851

Amoroso, N., La Rocca, M., Monaco, A., Bellotti, R., and Tangaro, S. (2018). Complex networks reveal early MRI markers of Parkinson's disease. *Med. Image Anal.* 48, 12–24. doi: 10.1016/j.media.2018.05.004

Ashburner, J., and Friston, K. J. (2000). Voxel-based morphometry—the methods. *Neuroimage* 11, 805–821. doi: 10.1006/nimg.2000.0582

Bailly, M., Destrieux, C., Hommet, C., Mondon, K., Cottier, J.-P., Beaufils, E., et al. (2015). Precuneus and cingulate cortex atrophy and hypometabolism in patients with Alzheimer's disease and mild cognitive impairment: MRI and 18F-FDG PET quantitative analysis using freesurfer. *BioMed Res. Int.* 2015:583931. doi: 10.1155/2015/583931

Baron, J. C., Chételat, G., Desgranges, B., Perchey, G., Landeau, B., de la Sayette, V., et al. (2001). *In vivo* mapping of gray matter loss with voxel-based morphometry in mild Alzheimer's disease. *Neuroimage* 14, 298–309. doi: 10.1006/nimg.2001.0848

Beleites, C., Neugebauer, U., Bocklitz, T., Krafft, C., and Popp, J. (2013). Sample size planning for classification models. *Anal. Chim. Acta* 760, 25–33. doi: 10.1016/j.aca.2012.11.007

Boccaletti, S., Bianconi, G., Criado, R., Del Genio, C. I., Gómez-Gardeñes, J., Romance, M., et al. (2014). The structure and dynamics of multilayer networks. *Phys. Rep.* 544, 1–122. doi: 10.1016/j.physrep.2014.07.001

Boccardi, M., Bocchetta, M., Morency, F. C., Collins, D. L., Nishikawa, M., Ganzola, R., et al. (2015). Training labels for hippocampal segmentation based on the EADC-ADNI harmonized hippocampal protocol. *Alzheimers Dement.* 11, 175–183. doi: 10.1016/j.jalz.2014.12.002

Braun, D. J., and Van Eldik, L. J. (2018). *In vivo* brainstem imaging in Alzheimer's disease: potential for biomarker development. *Front. Aging Neurosci.* 10:266. doi: 10.3389/fnagi.2018.00266

Bron, E. E., Smits, M., van der Flier, W. M., Vrenken, H., Barkhof, F., Scheltens, P., et al. (2015). Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: the CADDementia challenge. *NeuroImage* 111, 562–579. doi: 10.1016/j.neuroimage.2015.01.048

Buckner, R. L., Andrews-Hanna, J. R., and Schacter, D. L. (2008). The brain's default network. *Ann. N. Y. Acad. Sci.* 1124, 1–38. doi: 10.1196/annals.1440.011

Bullmore, E. T., and Bassett, D. S. (2011). Brain graphs: graphical models of the human brain connectome. *Ann. Rev. Clin. Psychol.* 7, 113–140. doi: 10.1146/annurev-clinpsy-040510-143934

Chincarini, A., Sensi, F., Rei, L., Gemme, G., Squarcia, S., Longo, R., et al. (2016). Integrating longitudinal information in hippocampal volume measurements for the early detection of Alzheimer's disease. *NeuroImage* 125, 834–847. doi: 10.1016/j.neuroimage.2015.10.065

Ciftçi, K. (2011). Minimum spanning tree reflects the alterations of the default mode network during Alzheimer's disease. *Ann. Biomed. Eng.* 39, 1493–1504. doi: 10.1007/s10439-011-0258-9

Crossley, N. A., Mechelli, A., Scott, J., Carletti, F., Fox, P. T., McGuire, P., et al. (2014). The hubs of the human connectome are generally implicated in the anatomy of brain disorders. *Brain* 137, 2382–2395. doi: 10.1093/brain/awu132

Daianu, M., Jahanshad, N., Nir, T. M., Jack, C. R., Weiner, M. W., Bernstein, M. A., et al. (2015). Rich club analysis in the Alzheimer's disease connectome reveals a relatively undisturbed structural core network. *Human Brain Mapp.* 36, 3087–3103. doi: 10.1002/hbm.22830

Davatzikos, C. (2004). Why voxel-based morphometric analysis should be used with great caution when characterizing group differences. *Neuroimage* 23, 17–20. doi: 10.1016/j.neuroimage.2004.05.010

de Haan, W., van der Flier, W. M., Koene, T., Smits, L. L., Scheltens, P., and Stam, C. J. (2012). Disrupted modular brain dynamics reflect cognitive dysfunction in Alzheimer's disease. *Neuroimage* 59, 3085–3093. doi: 10.1016/j.neuroimage.2011.11.055

de Jong, L., van der Hiele, K., Veer, I. M., Houwing, J. J., Westendorp, R. G., Bollen, E. L., et al. (2008). Strongly reduced volumes of putamen and thalamus in Alzheimer's disease: an MRI study. *Brain* 131, 3277–3285. doi: 10.1093/brain/awn278

De Vico Fallani, F., Latora, V., and Chavez, M. (2017). A topological criterion for filtering information in complex brain networks. *PLoS Comput. Biol.* 13:e1005305. doi: 10.1371/journal.pcbi.1005305

Derflinger, S., Sorg, C., Gaser, C., Myers, N., Arsic, M., Kurz, A., et al. (2011). Grey-matter atrophy in Alzheimer's disease is asymmetric but not lateralized. *J. Alzheimers Dis.* 25, 347. doi: 10.3233/JAD-2011-110041

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31, 968–980. doi: 10.1016/j.neuroimage.2006.01.021

Feng, X., Yang, J., Laine, A. F., and Angelini, E. D. (2018). "Alzheimer's disease diagnosis based on anatomically stratified texture analysis of the hippocampus in structural MRI," in *15th International Symposium on Biomedical Imaging (ISBI 2018), 2018 IEEE* (Washington, DC: IEEE), 1546–1549.

Fennema-Notestine, C., Hagler, D. J., McEvoy, L. K., Fleisher, A. S., Wu, E. H., Karow, D. S., et al. (2009). Structural MRI biomarkers for preclinical and mild Alzheimer's disease. *Human Brain Mapp.* 30, 3238–3253. doi: 10.1002/hbm.20744

Fischl, B. (2012). FreeSurfer. *Neuroimage* 62, 774–781. doi: 10.1016/j.neuroimage.2012.01.021

Good, C. D., Johnsrude, I. S., Ashburner, J., Henson, R. N., Fristen, K., and Frackowiak, R. S. (2002). "A voxel-based morphometric study of ageing in 465 normal adult human brains," in *Biomedical Imaging, 2002. 5th IEEE EMBS International Summer School on* (IEEE), 16.

He, Y., Chen, Z., and Evans, A. (2008). Structural insights into aberrant topological patterns of large-scale cortical networks in Alzheimer's disease. *J. Neurosci.* 28, 4756–4766. doi: 10.1523/JNEUROSCI.0141-08.2008

Ho, A. J., Hua, X., Lee, S., Leow, A. D., Yanovsky, I., Gutman, B., et al. (2010). Comparing 3 t and 1.5 t mri for tracking alzheimer's disease progression with tensor-based morphometry. *Human Brain Mapp.* 31, 499–514. doi: 10.1002/hbm.20882

Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., and Smith, S. M. (2012). Fsl. *Neuroimage* 62, 782–790. doi: 10.1016/j.neuroimage.2011.09.015

Karas, G. B., Burton, E. J., Rombouts, S. A., van Schijndel, R. A., O'Brien, J. T., Scheltens, Ph., et al. (2003). A comprehensive study of gray matter loss in patients with Alzheimer's disease using optimized voxel-based morphometry. *Neuroimage* 18, 895–907. doi: 10.1016/S1053-8119(03)00041-7

Landin-Romero, R., Kumfor, F., Leyton, C. E., Irish, M., Hodges, J. R., and Piguet, O. (2017). Disease-specific patterns of cortical and subcortical degeneration in a longitudinal study of Alzheimer's disease and behavioural-variant frontotemporal dementia. *Neuroimage* 151, 72–80. doi: 10.1016/j.neuroimage.2016.03.032

Lebedeva, A. K., Westman, E., Borza, T., Beyer, M. K., Engedal, K., Aarsland, D., et al. (2017). Mri-based classification models in prediction of mild cognitive impairment and dementia in late-life depression. *Front. Aging Neurosci.* 9:13. doi: 10.3389/fnagi.2017.00013

Lee, K.-M., Kim, J. Y., Cho, W.-k., Goh, K., and Kim, I. (2012). Correlated multiplexity and connectivity of multiplex random networks. *New J. Phys.* 14:033027. doi: 10.1088/1367-2630/14/3/033027

Long, X., Jiang, C., and Zhang, L. (2018). Morphological Biomarker Differentiating MCI Converters from Nonconverters: longitudinal Evidence Based on Hemispheric Asymmetry. *Behav. Neurol.* 2018:3954101. doi: 10.1155/2018/3954101

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R. Jr., Kawas, C. H., et al. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 7, 263–269. doi: 10.1016/j.jalz.2011.03.005

Menichetti, G., Remondini, D., Panzarasa, P., Mondragón, R. J., and Bianconi, G. (2014). Weighted multiplex networks. *PLoS ONE* 9:e97857. doi: 10.1371/journal.pone.0097857

Minoshima, S., Giordani, B., Berent, S., Frey, K. A., Foster, N. L., and Kuhl, D. E. (1997). Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. *Ann. Neurol.* 42, 85–94.

Moradi, E., Pepe, A., Gaser, C., Huttunen, H., and Tohka, J. (2015). Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. *NeuroImage* 104, 398–412. doi: 10.1016/j.neuroimage.2014.10.002

Mucha, P. J., Richardson, T., Macon, K., Porter, M. A., and Onnela, J. P. (2010). Community structure in time-dependent, multiscale, and multiplex networks. *Science* 328, 876–878. doi: 10.1126/science.1184819

Mukherjee, S., Tamayo, P., Rogers, S., Rifkin, R., Engle, A., Campbell, C., et al. (2003). Estimating dataset size requirements for classifying DNA microarray data. *J. Comput. Biol.* 10, 119–142. doi: 10.1089/106652703321825928

Perl, D. P. (2010). Neuropathology of Alzheimer's disease. *Mount Sinai J.* 77, 32–42. doi: 10.1002/msj.20157

Salvatore, C., Cerasa, A., Battista, P., Gilardi, M. C., Quattrone, A., Castiglioni, I., et al. (2015). Magnetic resonance imaging biomarkers for the early diagnosis of Alzheimer's disease: a machine learning approach. *Front. Neurosci.* 9:307. doi: 10.3389/fnins.2015.00307

Simic, G., Stanic, G., Mladinov, M., Jovanov-Milosevic, N., Kostovic, I., and Hof, P. R. (2009). Does Alzheimer's disease begin in the brainstem? *Neuropathol. Appl. Neurobiol.* 35, 532–554. doi: 10.1111/j.1365-2990.2009.01038.x

Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., et al. (2011). Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 7, 280–292. doi: 10.1016/j.jalz.2011.03.003

Stam, C. J., de Haan, W., Daffertshofer, A., Jones, B. F., Manshanden, I., van Cappellen van Walsum, A. M., et al. (2009). Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer's disease. *Brain* 132, 213–224. doi: 10.1093/brain/awn262

Suk, H. I., Lee, S. W., Shen, D., and the Alzheimers Disease Neuroimaging Initiative (2014). Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *NeuroImage* 101, 569–582. doi: 10.1016/j.neuroimage.2014.06.077

Tijms, B. M., Möller, C., Vrenken, H., Wink, A. M., de Haan, W., van der Flier, W. M., et al. (2013a). Single-subject grey matter graphs in Alzheimer's disease. *PLoS ONE* 8:e58921. doi: 10.1371/journal.pone.0058921

Tijms, B. M., Wink, A. M., de Haan, W., van der Flier, W. M., Stam, C. J., Scheltens, P., et al. (2013b). Alzheimer's disease: connecting findings from graph theoretical studies of brain networks. *Neurobiol. Aging* 34, 2023–2036. doi: 10.1016/j.neurobiolaging.2013.02.020

Wei, R., Li, C., Fogelson, N., and Li, L. (2016). Prediction of conversion from mild cognitive impairment to alzheimer's disease using mri and structural network features. *Front. Aging Neurosci.* 8:76. doi: 10.3389/fnagi.2016.00076

West, M. J., Coleman, P. D., Flood, D. G., and Troncoso, J. C. (1994). Differences in the pattern of hippocampal neuronal loss in normal ageing and Alzheimer's disease. *Lancet* 344, 769–772.

Yao, Z., Zhang, Y., Lin, L., Zhou, Y., Xu, C., Jiang, T., et al. (2010). Abnormal cortical networks in mild cognitive impairment and Alzheimer's disease. *PLoS Comput. Biol.* 6:e1001006. doi: 10.1371/journal.pcbi.1001006

Yi, H. A., Möller, C., Dieleman, N., Bouwman, F. H., Barkhof, F., Scheltens, P., et al. (2015). Relation between subcortical grey matter atrophy and conversion from mild cognitive impairment to Alzheimer's disease. *J. Neurol. Neurosurg. Psychiatry* 87, 425–432. doi: 10.1136/jnnp-2014-309105

Yokoi, T., Watanabe, H., Yamaguchi, H., Bagarinao, E., Masuda, M., Imai, K., et al. (2018). Involvement of the precuneus/posterior cingulate cortex is significant for the development of Alzheimer's disease: a PET (THK5351, PiB) and resting fMRI study. *Front. Aging Neurosci.* 10:304. doi: 10.3389/fnagi.2018.00304

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.