Application of Probabilistically-Weighted Graphs to Image-Based Diagnosis of Alzheimer's Disease using Diffusion MRI

Syeda Maryam[a], Laura McCrackin[a], Mark Crowley[a], Yogesh Rathi[b], and Oleg Michailovich[a]

[a]University of Waterloo, Waterloo, Canada
[b]Harvard Medical School, Boston, United States

ABSTRACT

The world's aging population has given rise to an increasing awareness towards neurodegenerative disorders, including Alzheimers Disease (AD). Treatment options for AD are currently limited, but it is believed that future success depends on our ability to detect the onset of the disease in its early stages. The most frequently used tools for this include neuropsychological assessments, along with genetic, proteomic, and image-based diagnosis. Recently, the applicability of Diffusion Magnetic Resonance Imaging (dMRI) analysis for early diagnosis of AD has also been reported. The sensitivity of dMRI to the microstructural organization of cerebral tissue makes it particularly well-suited to detecting changes which are known to occur in the early stages of AD. Existing dMRI approaches can be divided into two broad categories: region-based and tract-based. In this work, we propose a new approach, which extends region-based approaches to the simultaneous characterization of multiple brain regions. Given a predefined set of features derived from dMRI data, we compute the probabilistic distances between different brain regions and treat the resulting connectivity pattern as an undirected, fully-connected graph. The characteristics of this graph are then used as markers to discriminate between AD subjects and normal controls (NC). Although in this preliminary work we omit subjects in the prodromal stage of AD, mild cognitive impairment (MCI), our method demonstrates perfect separability between AD and NC subject groups with substantial margin, and thus holds promise for fine-grained stratification of NC, MCI and AD populations.

Keywords: diffusion MRI, Alzheimer's Disease, early diagnosis, graph theory

1. INTRODUCTION

Alzheimer's Disease (AD) is a neurodegenerative disorder affecting nearly 44 million people worldwide, with an estimated 1 in 3 seniors developing AD or a related dementia during their lifetime. It is characterized by the atrophy of the cortical and subcortical grey matter, as well as by the Wallerian degeneration of white matter. It is believed that the first clinical symptoms start to appear over a decade after the onset of the disease,[1,2] at which point it is usually too late to reverse its effects. In particular, it is known that half the patients in the prodromal stages of the disease (amnestic mild cognitive impairment) will converge to advanced AD in the next 3-5 years.[3] Current efforts of the research community are directed towards characterizing the prognostic course of the disease in the hopes that it will facilitate the development of clinical solutions.

In addition to classical psychophysiological tests for AD, a wide range of diagnostic tools exists, including genetics, proteomics, and methods based on neuroimaging.[4] In this particular work, our focus is on diffusion MRI (dMRI), a medical imaging modality rooted in the dynamics of Brownian motion of water molecules. Unlike more conventional forms of MRI, dMRI is able to probe the microstructural integrity of the cerebral white matter, making it an unparalleled tool to explore the structural connectivity of the brain.

There are two main classes of approach for characterizing brain pathology using dMRI. The first is region-based techniques, a class of bottom-up methods where brain regions are considered independently. Generally, these begin with a hypothesis about the etiology of the disease, and lead to a prediction of how diffusion measures

Further author information: (Send correspondence to O.M.)
E-mail: olegm@uwaterloo.ca, Telephone: 1-519-888-4567 ext. 38247

behave in response to pathology. The second is tract-based methods, which analyze diffusion measures along specific neural pathways in the brain based on the results of fibre tractography. However, these can be limited by the inconsistency of results between different tractography algorithms, and often require expert input to extract meaningful information.

While region-based methods have successfully been used to quantify abnormalities in late-stage AD, it is well-known that AD affects multiple regions of the brain. Therefore, we seek to observe the correlations between brain regions rather than considering them independently, combining the strengths of both region-based and tract-based approaches.

In our approach, we first extract diffusion features from a number of brain regions which are known to be implicated in AD. We then quantify the probabilistic distances between these regions, and consider the resulting connectivity pattern as a fully-connected undirected graph. Finally, we use the characteristics of this graph as a basis for discriminating between different subject populations. The resulting degree of separability is substantially larger than those obtained in similar works. Although we only consider normal controls and AD subjects in this work, we believe our method shows strong promise for further stratification of various stages of the disease, including early and late mild cognitive impairment (MCI).

We begin with a brief introduction to dMRI, and in particular, the dMRI model we will be working with, Diffusion Tensor Imaging (DTI). We then describe the dMRI features used in this work, and our methods for achieving graph-based discrimination. Next, we briefly elaborate on our dataset and its preprocessing. Finally, we discuss the results of our method, and suggest areas for further work.

2. BACKGROUND

At its essence, diffusion MRI consists of the addition of a magnetic field gradient to conventional MRI pulse sequences, first to dephase the molecular spins and then reverse this dephasing, so that the resulting signal attenuation quantifies how much diffusion has taken place. Most dMRI acquisition protocols are based on the pulsed-gradient spin-echo sequence (PGSE) first introduced by Stejskal and Tanner in their seminal paper in 1956.[5]

A typical implementation of dMRI depends on a specific choice of user-defined parameters (n, b), with the vector n representing the direction of diffusion encoding and the scalar b, a function of the strength and timing of the magnetic field gradient, sensitizing the resulting measurements towards different diffusion regimes (e.g., fast vs. slow). Thus, for a fixed b, using several n allows exploring the anisotropic properties of water diffusion, while using different values of b makes it possible to tap into the microstructural characteristics of brain tissue. However, since each particular combination of (n, b) entails a full 3D scan, the total number of such combinations is usually limited in practice.

There are many different dMRI models, ranging from simple to more complex depending on the underlying model assumptions and acquisition requirements. The more sophisticated models include Diffusion Spectrum Imaging (DSI), which estimates the average diffusion propagator from diffusion data in a non-parametric fashion, allowing for a more versatile diffusion representation.[6] However, DSI requires an infeasibly large number of sampling points, which limits its practical value.

In contrast, Diffusion Tensor Imaging (DTI), popularized in the early 90s in the works of LeBihan[7] and Besser et al.,[8] is a simpler, parametric model for characterizing diffusion, making it the method of choice for clinical applications. DTI employs Gaussian distributions to model the displacement of water molecules, allowing diffusion to be represented using the covariance structure of the associated probability density. One of the most important advantages of DTI is its clinical practicality, since estimating the parameters of the model can be done with as few as 6 diffusion measurements, with an additional measurement for normalization. In practice, measurements are normally on the order of 25-30, which is still within the bounds of clinical requirements.[9]

While the use of sophisticated models is theoretically enticing, the acquisition time required by these protocols tends to increase in proportion to their complexity. For example, probing the diffusion regime at different scales (slow vs. fast) requires acquisition at multiple b-values, which leads to prohibitively long acquisition times. Longer scan times cannot be tolerated by many non-compliant patients, including children and the elderly, and this makes the data more likely to be compromised due to motion artifacts.

Furthermore, since DTI has been in use for the past decade, substantial longitudinal databases have already been acquired using the DTI acquisition protocol. For example, the Alzheimer's Disease Neuroimaging Initiative (ADNI) database consists of thousands of cases collected from both normal controls and AD subjects, showing various stages in the progression of the disease. While the development of new diffusion models and their associated acquisition protocols would render such massive databases obsolete, we still believe that these datasets can provide significant neurological insight.

While this work does not attempt to directly extract diffusion features which are attributed to more sophisticated models, we demonstrate how such features can be reduced to the chosen setting while still providing useful information.

3. METRICS

There are a number of different diffusion metrics that have previously been used for AD diagnosis.[10–12] For example, there has been much research on using classical dMRI features like Fractional Anisotropy (FA) and Mean Diffusivity (MD) for stratification of different subject groups. In addition to using these features, we also adapt metrics from more sophisticated acquisition schemes for use with our dataset. This particular choice of metrics is by no means exclusive; new features developed by the scientific community can also be added for our method. We now present an overview of the features we have used in this work.

3.1 DTI-based features

DTI-based features are among those most commonly used for AD research. In the DTI model, the diffusion signal S is described by a Gaussian model of the form

$$S(\mathbf{u}) = S_0\, e^{-b\mathbf{u}^T\mathbf{D}\mathbf{u}}, \tag{1}$$

where b is the b-value, $\mathbf{u}$ denotes the diffusion-encoding directions, S_0 is the reference signal acquired without diffusion-weighting (known as the $b0$-image) and $\mathbf{D}$ is the diffusion tensor. In practice, the diffusion signal is normalized by the $b0$-image to result in

$$E(\mathbf{u}) = S(\mathbf{u})/S_0. \tag{2}$$

The eigenvalues of $\mathbf{D}$ are rotation-invariant characteristics which can be used to describe the diffusion signal. In particular, FA and MD are defined according to

$$\mathrm{FA} = \sqrt{\frac{3}{2}}\frac{\sqrt{(\lambda_1 - \langle\lambda\rangle)^2 + (\lambda_2 - \langle\lambda\rangle)^2 + (\lambda_3 - \langle\lambda\rangle)^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}, \tag{3}$$

$$\mathrm{MD} = \langle\lambda\rangle = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3}. \tag{4}$$

FA is a measure of anisotropy scaled between 0 and 1, with 0 corresponding to isotropic diffusion (uniform in all directions) and 1 corresponding to highly anisotropic (directional) diffusion, while MD attempts to quantify the average diffusion magnitude.

We also consider Relative Anisotropy (RA) and Volume Ratio (VR), which are two other common DTI metrics:

$$\mathrm{RA} = \sqrt{\frac{1}{3}}\frac{\sqrt{(\lambda_1 - \langle\lambda\rangle)^2 + (\lambda_2 - \langle\lambda\rangle)^2 + (\lambda_3 - \langle\lambda\rangle)^2}}{\langle\lambda\rangle}, \tag{5}$$

$$\mathrm{VR} = \frac{\lambda_1\lambda_2\lambda_3}{\langle\lambda\rangle^3}. \tag{6}$$

An additional pair of metrics used in this work are those proposed by Westin et al.[13] Known as linearity (C_l) and planarity (C_p), these metrics characterize the shape of the diffusion tensor and are defined according to

$$\mathrm{C}_l = \frac{\lambda_1 - \lambda_2}{\lambda_1}, \tag{7}$$

$$\mathrm{C}_p = \frac{\lambda_2 - \lambda_3}{\lambda_1}. \tag{8}$$

3.2 Spherical Harmonic-based Features

Considering E to be a square-integrable spherical function allows it to be expressed in terms of spherical harmonics (SH). Specifically, E can be represented as

$$E(\mathbf{u}) = \sum_l \sum_m c_{lm} Y_{lm}(\mathbf{u}), \tag{9}$$

where l is the degree of the SH expansion, m is its order, and c_{lm} are the SH coefficients describing the spectral content of the signal. Although the coefficients themselves are not rotation-invariant, they can be used to compute rotation-invariant metrics. One example is rotation-invariant Fourier signatures,[14] which can be considered analogous to the power spectrum in classical Fourier analysis, and are defined according to

$$e_l = \frac{1}{2l+1} \sum_m |c_{lm}|^2, \tag{10}$$

where l assumes even values due to the antipodal symmetry of diffusion signals. The maximum value of l is determined by the bandwidth of diffusion signals. In the case of the ADNI dataset, a reasonable choice for the maximum value of l is 6. The SH coefficients, c_{lm}, are typically estimated through a fitting method. For the sake of simplicity, and without much loss of generality, we use a standard least squares fit.[15]

3.3 Non-Gaussianity Features

Özarslan et al. introduced measures to be used to quantify the deviation of observed diffusion signals from a Gaussian model.[16] However, calculating these measures requires the ensemble average propagator (EAP), which is not generally recoverable from single-shell diffusion measurements, like the ADNI dataset considered in this paper. Therefore, the method needs to be adapted before it can be applied in this setting. A particular way of doing so is explained below.

The normalized diffusion signal, E, is first expressed in terms of its SH coefficients, c_{lm}, which are found through a standard fitting algorithm. At the same time, E is also fitted to the classical DTI model in equation (1) to produce a Gaussian approximation of the signal, $\tilde{E}$, which is then also subjected to SH fitting to produce SH coefficients $\tilde{c}_{lm}$. Finally, we use an inner product between c_{lm} and $\tilde{c}_{lm}$ as a measure of similarity between them, which we define as

$$\langle c, \tilde{c} \rangle = \sum_{l,m} c_{lm} \tilde{c}_{lm}. \tag{11}$$

For convenience of notation, we use c and $\tilde{c}$ as the vector of SH coefficients in lexicographical order. We can then define non-Gaussianity (NG) as

$$\mathrm{NG} = \sqrt{1 - \left| \frac{\langle c, \tilde{c} \rangle}{||c||\,||\tilde{c}||} \right|^2}, \tag{12}$$

where $||\cdot||$ is the standard norm induced by the inner product. In this definition, NG can assume values between 0 and 1, with 0 corresponding to the case when E complies with the Gaussian model of equation (1). To improve the contrast of NG values, we subject them to a non-linear transformation defined by Özarslan et al.,[16] as given by

$$\sigma(t, \epsilon) = \frac{t^{3\epsilon}}{1 - 3t^{\epsilon} + 3t^{2\epsilon}}. \tag{13}$$

Here, t is the argument of the function and ϵ is a user-defined shape parameter which is adjusted to achieve the required range of emphasis; in our work, ϵ is equal to 0.02.

The measure of sphericity – or equivalently, isotropy (I)[16] – can be similarly adapted to our dataset. We use the SH representation of the original signal E to define this measure as

$$I = \sqrt{1 - \left| \frac{c_{0,0}}{||c||} \right|^2}. \tag{14}$$

I is then subjected to the same scaling procedure outlined in equation (13) to improve contrast.

3.4 Return-to-Origin Probability

The final feature we adapt in this work is Return-to-Origin Probability (RTOP).[16] This measure is believed to correlate with the average axon diameter, and is considered to be an important biomarker. The RTOP is formally defined as the value of the EAP $P(\mathbf{r})$ at $\mathbf{r} = 0$, i.e,

$$\text{RTOP} = P(\mathbf{r})|_{\mathbf{r}=0}, \tag{15}$$

where $\mathbf{r}$ is the spin displacement. As before, since we do not have access to $P(\mathbf{r})$, some adaptations of the concept are needed. In particular, in this work, we compute the EAP based on a mono-exponential model, which extrapolates the values of E observed on a spherical shell of radius q_0 to the whole $\mathbf{q}$-space according to

$$E(\mathbf{q}) = E(q_0\mathbf{u})^{\frac{q^2}{q_0^2}}, \tag{16}$$

where q is the length of the $\mathbf{q}$ vector.[17] Under this model, the RTOP can be shown to be given by

$$\text{RTOP} = \frac{\sqrt{\pi}}{4q_0^{\frac{3}{2}}} \int_0^{\pi}\int_0^{2\pi} ADC(\theta,\phi)^{-\frac{3}{2}} \sin(\theta) d\theta\, d\phi, \tag{17}$$

where $ADC(\theta,\phi)$ denotes the Apparent Diffusivity Coefficients evaluated in the standard spherical coordinate system (θ,ϕ). The above integration was performed numerically using Lebedev quadratures. The details of this derivation can be found in the thesis of Maryam.[18]

3.5 Additional Features

As previously mentioned, our choice of diffusion metrics is by no means exclusive; our methodology does not depend on a particular list of metrics, which can be either extended or modified to suit other applications or reflect new developments. It is worth noting, however, that it is prudent to check whether new additions are redundant first, as many popular features exhibit high correlation. A typical example is Generalized Fractional Anisotropy (GFA), which has been shown to be asymptotically equivalent to sphericity.[18] An empirical proof of this fact can be seen in Figure 1, which depicts a scatter plot of both metrics observed in the corpus collosum of a typical dataset.

4. METHODS

Despite careful selection, our diffusion metrics have been found to be considerably correlated. For instance, three of the chosen metrics (FA, RA and GFA) are variants of anisotropy measures. To ensure numerical stability and account for excessive correlations between the metrics, we employ Robust Principal Component Analysis (RPCA) to reduce our feature space before using it for further analysis.[19] To this end, diffusion metrics are calculated and aggregated into feature vectors on a voxel-by-voxel basis for all relevant brain regions of each subject (see Section 5). Subsequently, RPCA is applied to the feature set to select three principal components based on the convergence of its spectrum.

We now introduce some formal notation for the sake of convenience. Let Ω denote the domain of diffusion-encoded images and θ denote the reduced feature set. Within Ω, we define a number of smaller regions $\{\Omega_i\}_{i=1}^N$, which represent anatomically distinct and physiologically meaningful regions of the brain known to be implicated in AD. Hence, in this notation, $x \in \Omega_i$ will denote the spatial coordinates of a voxel within region i. Next, we assume the feature vector θ to have different probability distributions within each region. In particular, the features $\{\theta(x)\}_{x\in\Omega_i}$ observed within region Ω_i are assumed to be independent and identically distributed (i.i.d.) copies of a random variable distributed according to a distribution with the conditional density

$$p_i(\theta) = p(\theta(x) \mid x \in \Omega_i). \tag{18}$$

In this work, all p_is are considered to be multivariate Gaussian densities parametrized by their corresponding means μ_i and covariances Σ_i. While justifiable by the central limit theorem, in this work the assumption is used primarily for the sake of computational efficiency, and can be refined in future works.

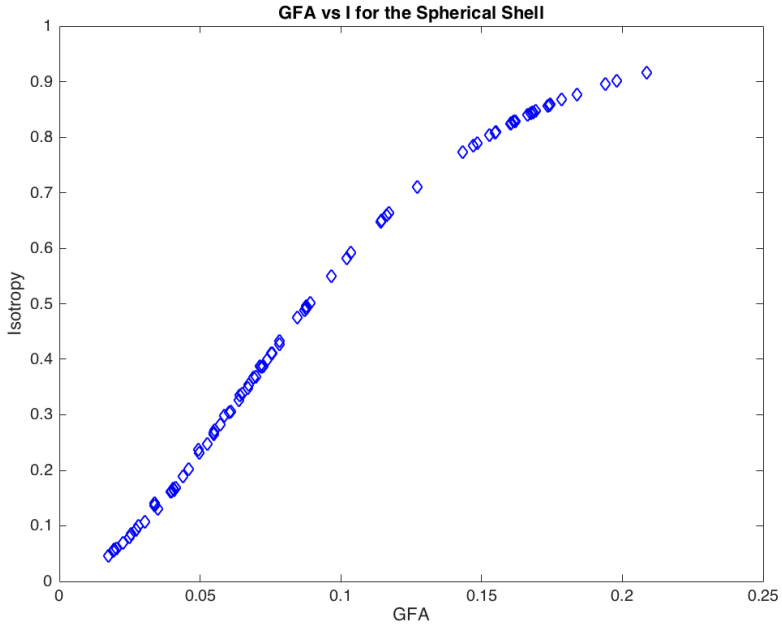

Figure 1: A strong correlation can be observed between the GFA and the sphericity values extracted from a region of the brain.

Next, we quantify the *regional* connectivity between each pair of Ω_i and Ω_j in terms of a probabilistic distance between their respective densities, p_i and p_j. There are a number of methods for evaluating probabilistic distances in the literature.[20–22] However, in our work, we use a standard choice rooted in the fundamental concepts of information theory. In particular, to quantify the information shared between regions, we use a symmetrized version of the Kullback-Leibler (KL) divergence[23,24] known as the J-divergence.[21]

Formally, the KL divergence is defined as

$$D_{KL}(p_i||p_j) = \int p_i(\theta) \log \frac{p_i(\theta)}{p_j(\theta)} d\theta, \tag{19}$$

where $p_i(\theta)$ and $p_j(\theta)$ are the conditional probability densities corresponding to regions i and j. The KL divergence has the additional advantage of having a closed form expression for multivariate Gaussian distributions, providing us with significant improvements in computational complexity. Namely,

$$D_{KL}(p_i||p_j) = \frac{1}{2}\left(\mathrm{Tr}(\Sigma_{p_j}^{-1}\Sigma_{p_i}) + (\mu_{p_j} - \mu_{p_i})^T \Sigma_{p_j}^{-1}(\mu_{p_j} - \mu_{p_i}) - k + \ln\left(\frac{\det \Sigma_{p_j}}{\det \Sigma_{p_i}}\right)\right). \tag{20}$$

Finally, the J-divergence can be expressed as

$$\mathcal{J}(p_i, p_j) = D_{KL}(p_i||p_j) + D_{KL}(p_j||p_i). \tag{21}$$

The aforementioned distances are then used to form an $N \times N$ matrix $\{d(i,j)\}_{i,j=1}^{N}$, which we consider to be the weight matrix for a fully-connected, undirected graph comprised of N nodes. Hence, each subject can be represented using a simple graph structure and its associated weights. Subsequently, the characteristics of this graph are employed as features for discriminating between different subject populations. To this end, we extract the graph distance (χ) and closeness centrality (η) metrics, which describe the overall connectivity of the graph. The metrics can be expressed as

$$\chi = \sum_i \sum_j d(i,j), \tag{22}$$

$$\eta = \frac{1}{N}\sum_{n=1}^{N}\left(\frac{N-1}{\sum_{j\neq i} d(i,j)}\right). \tag{23}$$

5. DATA AND PREPROCESSING

Data used in this study was acquired from the publicly-available Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. The ADNI project is a multi-site study aimed at combining clinical, neuroimaging, and genetic data to facilitate research towards improved diagnosis and treatment of Alzheimer's Disease. The ADNI database includes MRI and dMRI data which have been collected using standardized acquisition protocols from subjects belonging to the cognitively normal (NC), AD, or MCI subject groups[*].

A total of 40 age-matched subjects (mean age 72.6 $\pm$ 7.6) were selected for analysis, 20 cognitively normal and 20 with AD. Preprocessing steps for dMRI data involved image registrations for head motion and eddy current correction using tools available in the FSL library[†]. Segmentation of the brain was performed using the Freesurfer Software Suite[‡]. Finally, the segmented volumes extracted using Freesurfer were resampled to the diffusion space using the Advanced Normalization Tools software.[§]

The Freesurfer labels correpsonding to the following brain regions were selected for analysis: the corpus callosum, parahippocampal white matter, the posterior cingulate, the cerebral white matter (comprising the internal capsule among other regions), and the cerebellum white matter. These regions have been shown to be implicated in AD.[25–27]

6. RESULTS

Once the graph distance and closeness centrality have been calculated for each subject, their values can be plotted as shown in Figure 2(a). As one can see, these features are sufficient for perfect stratification of the AD and NC subject populations with substantial margin. For illustrative purposes, a separating line between the two classes is shown in black.

In this case, it can be seen that using closeness centrality alone would be sufficient to provide good separation. Graph distance does not improve separability, and appears to more or less uniformly distribute our data; it may, however, prove useful for finer-grained stratification in future experiments.

[*] More information can be found on the ADNI website at http://adni.loni.usc.edu/

[†] Specifically, FLIRT was used for image registration, with more details available at http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/

[‡] More information can be found at https://surfer.nmr.mgh.harvard.edu/

[§] The ANTs software is publicly available at http://stnava.github.io/ANTs/

Our resulting separation is particularly striking given that significant overlap is observed when any two of our candidate features are similarly plotted for any considered brain region. As an example, we consider the two most commonly used features in the literature, FA and MD 2(b), (c) and (d). As one can see, these features are unable to discriminate the two subject groups to the same degree observed with our method.

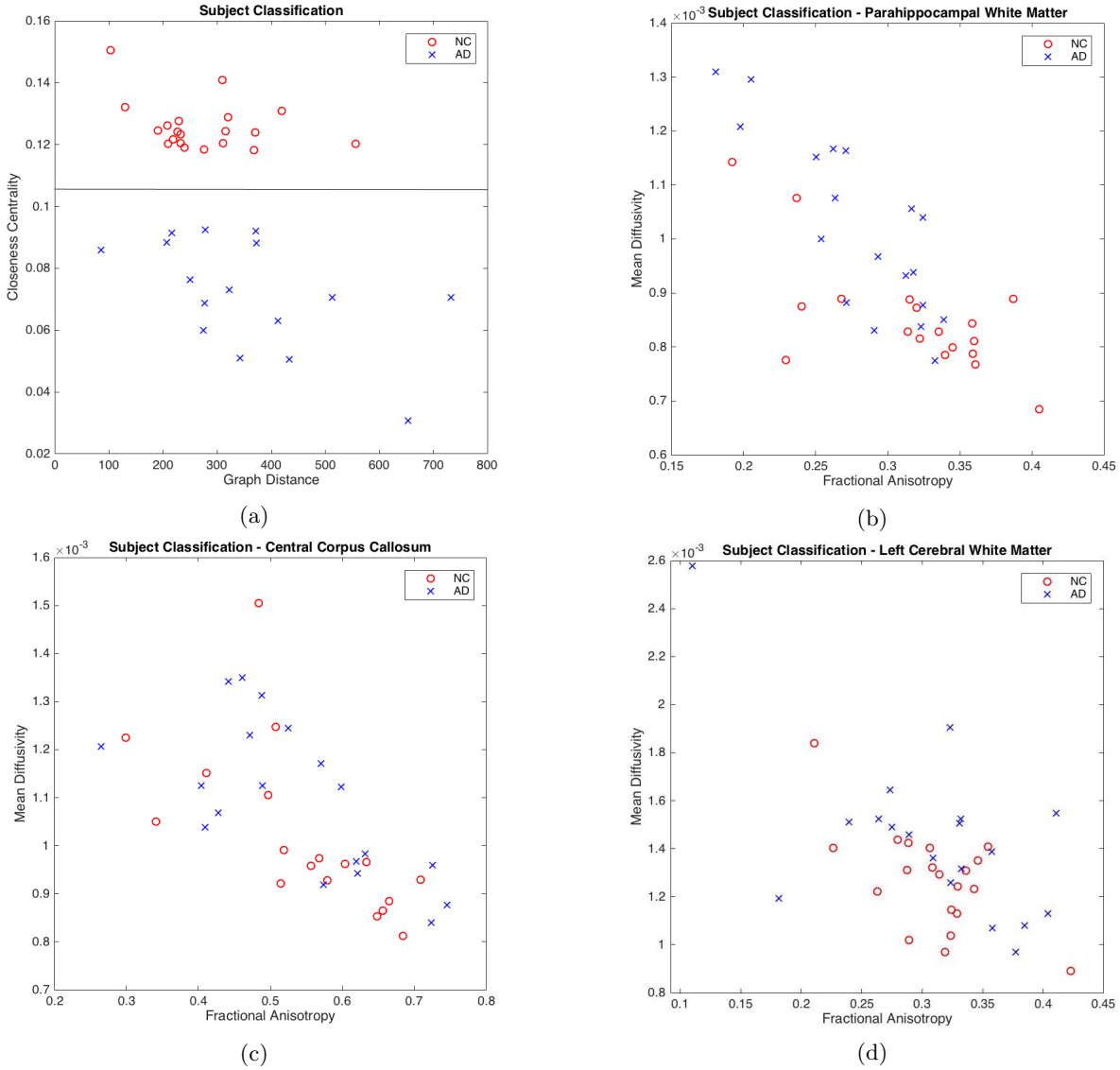

Figure 2: (a) Closeness centrality and graph distance corresponding to AD and NC populations. (b), (c) and (d) Fractional Anisotropy and Mean Diffusivity extracted from various regions of the brain for the same dataset. As can be observed, these features offer low discriminatory power between the two groups.

7. CONCLUSIONS

We have proposed in this paper a novel, hybrid approach to dMRI-based diagnosis, combining the robustness of region-based methods with the regional connectivity of the brain. Our results clearly demonstrate the utility of our method, and its practicality is further underscored by its algorithmic simplicity and computational efficiency: our method requires under two minutes on a standard desktop computer to run on our entire pre-processed dataset.

We believe this approach shows strong promise for further development. We hope to extend this method to perform finer-grained stratification by including subjects with MCI in addition to AD subjects and normal controls, as well as by considering longitudinal data for each subject group. Since our method does not depend on nominal metric values, but rather on differential characteristics such as regional similarity, it is likely to be less susceptible to inter-site variability – that is, differences between MRI scanners – than more traditional methods, which would be considerably advantageous when considering larger, multi-site study data. Finally, our method's flexibility to different selections of features allows for easy adaptation as further advances are made by the medical research community.

ACKNOWLEDGMENTS

This work was supported by the Canadian Institutes of Health Research (`http://www.cihr-irsc.gc.ca`). Computations were performed on the SOSCIP Consortiums Blue Gene/Q computing platform. SOSCIP is funded by the Federal Economic Development Agency of Southern Ontario, the Province of Ontario, IBM Canada Ltd., Ontario Centres of Excellence, Mitacs and 15 Ontario academic member institutions.

REFERENCES

[1] Dubois, B., Feldman, H. H., Jacova, C., Cummings, J. L., Dekosky, S. T., Barberger-Gateau, P., Barberger-Gateau, P., Delacourte, A., Frisoni, G., Fox, N. C., Galasko, D., Gauthier, S., Hampel, H., Jicha, G. A., Meguro, K., O'Brien, J., Pasquier, F., Robert, P., Rossor, M., Salloway, S., Sarazin, M., de Souza, L. C., Stern, Y., Visser, P. J., and Scheltens, P., "Revising the definition of Alzheimer's disease: A new lexicon," *Lancet Neurol* **9**, 1118 –1127 (2010).

[2] Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., Iwatsubo, T., Jack, C. R., Kaye, J., Montine, T. J., Park, D. C., Reiman, E. M., Rowe, C. C., Siemers, E., Stern, Y., Yaffe, K., Carrillo, M. C., Thies, B., Morrison-Bogorad, M., Wagster, M. V., and Phelps, C. H., "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 (2011).

[3] R. C. Petersen, J. Q. T., "Use of Alzheimer disease biomarkers: Potentially yes for clinical trials but not yet for clinical practice," *JAMA* **302**, 436–437 (2009).

[4] Weiner, M. W., Veitch, D. P., Aisen, P. S., Beckett, L. A., Cairns, N. J., Green, R. C., Harvey, D., Jack, C. R., Jagust, W., Liu, E., et al., "The alzheimer's disease neuroimaging initiative: a review of papers published since its inception," *Alzheimer's & Dementia* **9**(5), e111–e194 (2013).

[5] Stejskal, E. O. and Tanner, J. E., "Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient," *The journal of chemical physics* **42**(1), 288–292 (1965).

[6] Wedeen, V., Reese, T., Tuch, D., Weigel, M., Dou, J., Weiskoff, R., and Chessler, D., "Mapping fiber orientation spectra in cerebral white matter with fourier-transform diffusion mri," in [*Proceedings of the 8th Annual Meeting of ISMRM, Denver*], 82 (2000).

[7] Le Bihan, D., Mangin, J.-F., Poupon, C., Clark, C. A., Pappata, S., Molko, N., and Chabriat, H., "Diffusion tensor imaging: concepts and applications," *Journal of magnetic resonance imaging* **13**(4), 534–546 (2001).

[8] Pierpaoli, C., Jezzard, P., Basser, P. J., Barnett, A., and Di Chiro, G., "Diffusion tensor mr imaging of the human brain.," *Radiology* **201**(3), 637–648 (1996).

[9] Jones, D. K., "The effect of gradient sampling schemes on measures derived from diffusion tensor mri: a monte carlo study," *Magnetic Resonance in Medicine* **51**(4), 807–815 (2004).

[10] Zhang, Y., Schuff, N., Jahng, G.-H., Bayne, W., Mori, S., Schad, L., Mueller, S., Du, A.-T., Kramer, J., Yaffe, K., et al., "Diffusion tensor imaging of cingulum fibers in mild cognitive impairment and alzheimer disease," *Neurology* **68**(1), 13–19 (2007).

[11] Bozzali, M., Franceschi, M., Falini, A., Pontesilli, S., Cercignani, M., Magnani, G., Scotti, G., Comi, G., and Filippi, M., "Quantification of tissue damage in ad using diffusion tensor and magnetization transfer mri," *Neurology* **57**(6), 1135–1137 (2001).

[12] Medina, D., Urresta, F., Gabrieli, J. D., Moseley, M., Fleischman, D., Bennett, D. A., Leurgans, S., Turner, D. A., Stebbins, G. T., et al., "White matter changes in mild cognitive impairment and ad: a diffusion tensor imaging study," *Neurobiology of aging* **27**(5), 663–672 (2006).

[13] Westin, C.-F., Peled, S., Gudbjartsson, H., Kikinis, R., Jolesz, F. A., et al., "Geometrical diffusion measures for mri from tensor basis analysis," in [*Proceedings of ISMRM*], **97**, 1742 (1997).

[14] Frank, L. R., "Characterization of anisotropy in high angular resolution diffusion-weighted mri," *Magnetic Resonance in Medicine* **47**(6), 1083–1099 (2002).

[15] Descoteaux, M., Angelino, E., Fitzgibbons, S., and Deriche, R., "Regularized, fast, and robust analytical q-ball imaging," *Magnetic Resonance in Medicine* **58**(3), 497–510 (2007).

[16] Özarslan, E., Koay, C. G., Shepherd, T. M., Komlosh, M. E., İrfanoğlu, M. O., Pierpaoli, C., and Basser, P. J., "Mean apparent propagator (map) mri: a novel diffusion imaging method for mapping tissue microstructure," *NeuroImage* **78**, 16–32 (2013).

[17] Aganj, I., Lenglet, C., Sapiro, G., Yacoub, E., Ugurbil, K., and Harel, N., "Reconstruction of the orientation distribution function in single-and multiple-shell q-ball imaging within constant solid angle," *Magnetic Resonance in Medicine* **64**(2), 554–566 (2010).

[18] Maryam, S., "Predictive diagnosis of alzheimer's disease using diffusion mri," (2016).

[19] Rousseeuw, P. J. and Driessen, K. V., "A fast algorithm for the minimum covariance determinant estimator," *Technometrics* **41**(3), 212–223 (1999).

[20] Kailath, T., "The divergence and bhattacharyya distance measures in signal selection," *IEEE transactions on communication technology* **15**(1), 52–60 (1967).

[21] Lin, J., "Divergence measures based on the shannon entropy," *IEEE Transactions on Information theory* **37**(1), 145–151 (1991).

[22] Kantorovich, L. V. and Rubinstein, G. S., "On a space of completely additive functions," *Vestnik Leningrad. Univ* **13**(7), 52–59 (1958).

[23] Kullback, S. and Leibler, R. A., "On information and sufficiency," *The annals of mathematical statistics* **22**(1), 79–86 (1951).

[24] Kullback, S., [*Information theory and statistics*], Courier Corporation (1997).

[25] Biegon, A., Eberling, J., Richardson, B., Roos, M., Wong, S., Reed, B. R., and Jagust, W., "Human corpus callosum in aging and alzheimer's disease: a magnetic resonance imaging study," *Neurobiology of aging* **15**(4), 393–397 (1994).

[26] Salat, D., Tuch, D., Van der Kouwe, A., Greve, D., Pappu, V., Lee, S., Hevelone, N., Zaleta, A., Growdon, J., Corkin, S., et al., "White matter pathology isolates the hippocampal formation in alzheimer's disease," *Neurobiology of aging* **31**(2), 244–256 (2010).

[27] Zhang, Y., Schuff, N., Du, A.-T., Rosen, H. J., Kramer, J. H., Gorno-Tempini, M. L., Miller, B. L., and Weiner, M. W., "White matter damage in frontotemporal dementia and alzheimer's disease measured by diffusion mri," *Brain* **132**(9), 2579–2592 (2009).