Title
Thickness network features for prognostic applications in dementia

Journal
Neurobiology of Aging, 36(Suppl 1)

ISSN
0197-4580

Authors
Raamana, Pradeep Reddy
Weiner, Michael W
Wang, Lei
et al.

Publication Date
2015

DOI
10.1016/j.neurobiolaging.2014.05.040

Peer reviewed

HHS Public Access

Author manuscript

Neurobiol Aging. Author manuscript; available in PMC 2018 March 13.

Published in final edited form as:

Neurobiol Aging. 2015 January ; 36(Suppl 1): S91–S102. doi:10.1016/j.neurobiolaging.2014.05.040.

Thickness network features for prognostic applications in dementia

Pradeep Reddy Raamana[a], **Michael W. Weiner**[b], **Lei Wang**[c,1], **Mirza Faisal Beg**[a,*,1], and **for the Alzheimer's Disease Neuroimaging Initiative**[2]

[a]Department of Engineering Science, School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada

[b]Department of Radiology, Center for Imaging of Neurodegenerative Diseases, University of California, San Francisco, CA, USA

[c]Department of Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

Abstract

Regional analysis of cortical thickness has been studied extensively in building imaging biomarkers for early detection of Alzheimer's disease but not its interregional covariation of thickness. We present novel features based on the inter-regional covariation of cortical thickness. Initially, the cortical labels of each subject are partitioned into small patches (graph nodes) by spatial k-means clustering. A graph is then constructed by establishing a link between 2 nodes if the difference in thickness between the nodes is below a certain threshold. From this binary graph, a thickness network is computed using nodal degree, betweenness, and clustering coefficient measures. Fusing them with multiple kernel learning, it is observed that thickness network features discriminate mild cognitive impairment (MCI) converters from controls (CN) with an area under curve (AUC) of 0.83, 74% sensitivity and 76% specificity on a large subset obtained from the Alzheimer's Disease Neuroimaging Initiative data set. A comparison of predictive utility in Alzheimer's disease and/or CN classification (AUC of 0.92, 80% sensitivity [SENS] and 90% specificity [SPEC]), in discriminating CN from MCI (converters and nonconverters combined; AUC of 0.75, SENS and SPEC of 64% and 73%, respectively) and in discriminating between MCI nonconverters and MCI converters (AUC of 0.68, SENS and SPEC of 65% and 64%) is also presented. ThickNet features as defined here are novel, can be derived from a single magnetic resonance imaging scan, and demonstrate the potential for the computer-aided prognostic applications.

[*]Corresponding author at: Department of Engineering Science, School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada. Tel.: +1 778 782 5696; fax: +1 778 782 4951. mfbeg@sfu.ca (M.F. Beg).

[1]These authors contributed equally to this work.

[2]Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

Disclosure statement

All the authors declare that they have no conflicts of interest.

Keywords

Cortical thickness; Network properties; Fusion; Multiple kernel learning; Early detection; Mild cognitive impairment; Alzheimer

1. Introduction

Today, Alzheimer's disease (AD) is the most common type of dementia, accounting for 60%–80% of the cases (Alzheimer's Association, 2012), for which definitive diagnosis can only be made with the histopathologic confirmation of amyloid plaques and neurofibrillary tangles. Recent reports suggest that the Alzheimer pathology begins decades before any clinical symptoms appear (Amieva et al., 2008; Braak and Braak, 1991; Braak and Del Tredici, 2011), which highlights the importance and challenge in early detection of AD.

Structural magnetic resonance imaging (sMRI, T1-weighted) offers a noninvasive way to image and analyze the structure of the brain (high tissue-contrast) at 1 mm^3resolution and is routinely used in clinical practice. Structural alterations associated with AD can be detected before the onset of clinical symptoms (Jack et al., 2010), supporting the use of structural imaging features for the early detection of AD. A large body of research on neuroimaging techniques exists for the detection of AD and for the prediction of conversion in mild cognitive impairment (MCI), including analysis of gray matter densities (Duchesne et al., 2010; Koikkalainen et al., 2011; Misra et al., 2009), hippocampal shape (Beg et al., 2013; Coupé et al., 2012; Wang et al., 2007), amyloid deposition (La Joie et al., 2012; Tosun et al., 2011; Villain et al., 2012), functional magnetic resonance imaging (MRI) (Bullmore and Sporns, 2009; Wee et al., 2012), and also employing multimodality approaches (Walhovd et al., 2010; Westman et al., 2012). sMRI features have also been successfully applied for the differential diagnosis of AD and frontotemporal disease (Du et al., 2007; Raamana et al., 2012; Woodward et al., 2010).

The early-stage neurodegeneration observed in AD is subtle and spatially distributed over the brain, which makes cortical thickness features an ideal imaging-biomarker for AD. Cortical thickness has been the focus of numerous studies for the detection of probable AD (Desikan et al., 2009; Dickerson et al., 2009; Eskildsen et al., 2013; Mcevoy et al., 2011; Wolz et al., 2011). These studies show that early-stage cortical thickness by itself using only baseline MRI scans proved to be useful for the early detection of AD, but with limited utility, as was shown in Cuingnet et al. (2011). Cuingnet et al. (2011) performed an objective comparison of the predictive performance of published image processing methods, on a common data set, to predict conversion to AD in MCI patients. They observed that the performance of various method based solely on baseline cortical thickness has been limited at the best in accurately predicting conversion to AD in MCI subjects.

There has been a plethora of research in ROI-based analysis of cortical thickness (Cuingnet et al., 2011), but only few studies analyzed the covariation of thickness in different regions of the brain. We would like to capture the nature of the pairwise changes to characterize the topographic covariation in cortical thickness as associated with the progression of AD. Establishing links (akin to edges in a graph) using cortical thickness or gray matter density

extracted from sMRI allows for such a study, and these approaches are only beginning to be explored. Here, we briefly summarize the different studies published so far and refer the reader to the following publications for a comprehensive review of studies on anatomic covariance (Alexander-Bloch et al., 2013a; Evans, 2013; Iturria-Medina, 2013; Wen et al., 2011). These studies can broadly be divided into 2 categories based on the type of anatomic features they use, that is, whether they use cortical morphometric features (He et al., 2008) versus gray matter density and/or volume (Mechelli et al., 2005; Tijms et al., 2012; Yao et al., 2010) to establish links and whether they utilize only the cortex (Chen et al., 2008) or the brain volume entirely (Tijms et al., 2012; Wee et al., 2012) or specific volume of interest (Mechelli et al., 2005; Seeley et al., 2009). Initial studies on structural covariance were pioneered by the Alan Evans group based on the analysis of vertex-wise correlations in thickness Lerch et al. (2006). This article was followed by (Chen et al., 2008; Gong et al., 2012; He et al., 2008, 2009a; Khundrakpam et al., 2013), which revealed insights into the aberrant network properties. For example, (He et al., 2007, 2008, 2009a, 2009b) used graph-theory analysis to study the group differences in AD relative to controls (CN) and revealed an abnormal small-world architecture, significantly reduced nodal centrality, increased local efficiency (local clustering), and decreased global efficiency (increased mean path length).

Gong et al. (2012) presented a comparison of patterns of covariance in cortical thickness and that of diffusion based fiber connections. This comparison suggested that positive correlations in cortical thickness might be mediated by a fiber pathway, and that cortical thickness correlations present exclusive information, that is not offered by fiber connections. Chen et al. (2008) demonstrated the modularity of the human cortical network (based on a network of correlations in cortical thickness) and its organization into different topologic modules overlapping closely with known functional domains.

Alexander-Bloch et al. (2013b) analyzed a longitudinal data set of healthy young people and constructed structural and maturational networks based on the rate of change in thickness over time. They studied the link between maturational networks and structural networks to demonstrate the similarity in their topologic properties (global and nodal). Khundrakpam et al. (2013) studied developmental changes in structural network properties of cortical thickness and revealed a significant reduction in local efficiency, modularity, and increased global efficiency in late childhood.

Bassett et al. (2008) analyzed group differences between schizophrenia and healthy controls to show that schizophrenic patients exhibited reduced loss of frontal hubs and emergence of nonfrontal hubs. Raj et al. (2010) analyzed the covariance networks of thickness and curvature to localize seizures in temporal lobe epilepsy and present an interesting graph-level statistical analysis. Seeley et al. (2009) studied the relationship between neurodegeneration, anatomic covariance and intrinsic connectivity networks in 5 neurodegenerative syndromes. They showed that the patterns in syndrome-specific atrophy mirror that of structural and functional covariance (using the maximal atrophic region as the seed region).

These studies presented so far assert the utility of covariance properties in studying disease-related changes. However, these studies were mostly limited to either studying the existence,

or lack thereof, of small world properties, for example He et al. (2008), or any group differences in covariance properties that exist between patient and CN groups, for example Bassett et al. (2008). These studies have not, to date, constructed any features from such inter-regional covariation or performed evaluation of its diagnostic utility. We propose to utilize covariation patterns in cortical thickness as an imaging biomarker for AD. The progression of AD generally follows a stereotypical spatial pattern and hence pairwise covariation between cortical surface patches will likely complement existing features for early detection based on cortical thickness. We construct novel features based on the network properties of inter-regional links in the brain defined using cortical thickness. Furthermore, we fuse these thickness network (ThickNet) features using probabilistic multiple kernel learning approach and investigate their predictive utility in the detection of prodromal AD (MCI converters) on a large cohort from Alzheimer's Disease Neuroimaging Initiative (ADNI) data set. Further, we compare its performance in discriminating between CN and AD, CN and MCI converters (MCIc), MCIc and MCI nonconverters (MCInc) as well as between CN and MCI (MCIc + MCInc, combined).

2. Methods

2.1. Data set

Data used in the preparation of this article were obtained from the ADNI database (adni.loni.ucla.edu). The ADNI was launched in 2003 by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, the Food and Drug Administration, private pharmaceutical companies and nonprofit organizations, as a $60 million, 5-year public-private partnership. For up-to-date information, see www.adni-info.org.

Cuingnet et al. (2011) compared the performance of various published classification methods on fixed training and testing sets resulting in a comparable set of performance metrics. To enable comparison with a large set of similar methods, we utilized the same subset of 509 participants as studied in Cuingnet et al. (2011), except for a few exclusions whose cortical parcellation did not meet our quality control metrics, see Appendix for further details. We refer the reader to Cuingnet et al. (2011) for the complete description of the participants and demographics for the study cohort. Briefly, our study consists of 481 T1-weighted MR scans acquired at 1.5 T. MRI scans from the baseline visit were used when available (and from the screening visit otherwise). This gave MR images from 159 CN subjects, 56 MCIc subjects (who had converted to AD within 18 months), 130 MCI nonconverters (MCInc) subjects, and 136 AD subjects. In this article, we use the term prodromal AD to denote MCI converters (MCIc), and the 2 terms are used interchangeably.

2.2. Thickness measurement and processing

Initial cortical reconstruction and volumetric segmentation of the whole brain were performed with the Freesurfer image analysis suite (Fischl et al., 2002) to obtain pial and WM and/or GM surfaces. The resulting cortical parcellations were quality controlled. Errors were corrected whenever possible (they were excluded otherwise). In the space between the GM and/or WM and pial surfaces, a discrete approximation of Laplace equation was solved

(Gibson et al., 2009; Yezzi and Prince, 2003) using the tools developed by our group. Streamlines of this harmonic function define corresponding points on the surfaces, and the Euclidean distance between these points defines the cortical thickness. To perform group analysis, we registered the surface of each subject in the study to the surface of a common atlas (derived from averaging over 80 healthy subjects) using the tools from Fischl et al. (2002), see Appendix for further details. This establishes vertex-wise correspondence and enables group-wise analysis. Finally, cortical thickness was smoothed with a 10-mm full width at half height gaussian kernel to improve the signal-to-noise ratio and statistical power for subsequent analysis (Lerch and Evans, 2005).

2.3. Novel dimensionality reduction

Each cortex surface contained 327,684 vertices in the whole brain, and we have a limited number of subjects. To avoid the curse of dimensionality, we partitioned each cortical label (such as posterior cingulate and so forth from the 68 Freesurfer-derived cortical labels) containing thousands of vertices into a small number (say 10) of partitions by clustering vertices, within each label, using k-means clustering of vertex coordinates. The thickness feature for each subpartition is defined as the average thickness across vertices in that partition. This simple novel approach not only reduces the dimensionality of the features but also does it in an anatomically meaningful way, as opposed to other dimensionality reduction methods (such as PCA) which transform the features to an entirely different space which may lack physical meaning and anatomic relevance. Note that, clustering is done within each Freesurfer label, which prohibits linking vertices across different adjacent labels. Moreover, the vertex density of Freesurfer parcellation is sufficiently even and high to satisfy the k-means assumptions (Lee et al., 2006), and visual verification of partitioning confirmed the resulting label subdivisions were appropriate. Visualization of this subdivision of the cortex into 680 partitions is shown in Fig. 2B and 2A. As they are all registered to a common atlas, this subdivision of the cortex is propagated into the cortical surface of each subject to establish correspondence for further analysis.

It is worth noting that certain trade-offs exist in deciding the total number of partitions (TNP) for this method. When we choose to average across the entire freesurfer label (which can be quite large covering many gyri and sulci), we may lose the discriminatory signal. In contrary, when the TNP is excessively large, for example over 5000, we risk the curse of dimensionality as well as making the method overly sensitive to noise. Hence, we study the performance of this method for different values of TNP = 340, 680, 1020, 1360, and 1700, to avoid making an arbitrary choice.

The aforementioned parcellation method we used is very similar to those previously reported. For example, Hagmann et al. (2008) utilized region-growing method for the subdivision of each of the Freesurfer labels into a large number of patches, which is very similar to our method of clustering neighbouring vertices using k-means clustering method. The 2 methods are very similar, except for the difference in the stopping criteria of clustering. We restrict the number of patches for each FS label (say n = 10), whereas Hagmann et al. (2008) restrict the maximum size of patches to be 1.5 cm^2. In fact, the average patch size resulting from our method using 10 patches per Freesurfer label is 1.59

cm^2, which is very close. As noted previously, we comprehensively study the performance of our method for different number of patches (n = 5 to 25 in steps of 5), to avoid making an arbitrary choice as opposed to a fixed choice of 1.5 cm^2 in Hagmann et al. (2008).

2.4. ThickNet features

Once the pial surface is partitioned into large number of small sub-partitions (thought of as nodes), a network (graph) is constructed by establishing a link between 2 nodes if the absolute difference in thickness is below a specified threshold. The term network is used here in the abstract sense to mean a mathematical graph and not a functional and/or structural network connected by physical fiber tracts or connections. From this binary undirected graph, we compute thickness network measures, we term them ThickNet features, such as nodal degree, betweenness centrality, and clustering coefficient to represent each individual brain. ThickNet measures are intrinsic to each subject and offer insight into regional correlations in cortical thinning. The extraction of ThickNet features in illustrated in the form of a flowchart in Fig. 1.

Suppose, N is the set of all nodes in the network (the number of nodes n = NPP × 68, NPP = number of partitions per freesurfer label in each of the 68 freesurfer labels), and L is the set of all links in the network (l = number of links). Note, N equals TNP which is the total number of partitions in each subject's cortical surface. Let (i; j) be a link between nodes i and j (i; $j \in N$) and a_{ij} is the link status between i and j: a_{ij} = 1 when link (i; j) exists; a_{ij} = 0 otherwise. A link is defined between i and j, if $|MT_i - MT_j| <= \alpha$, where MT_x represents the mean thickness in the node x; $x \in N$. Here, α is the threshold to establish a link. A lenient threshold (α > 0.5 mm) allows large number of links in the cortex, whereas a stringent threshold (α ≥ 0.5 mm) allows relatively fewer links. It is important to note that spatial distance or spatial adjacency is not a criteria, as the method searches all possible pairwise links between all cortical subpartitions.

We chose to utilize nodal degree (measure of how connected each node is), betweenness centrality (measure of centrality) and clustering coefficient (measure of segregation) from the binary graph as properties to describe the network (Rubinov and Sporns, 2010). In brief, for a given node i, these are defined as:

$$\text{Nodal degree, } k_i = \sum_{j \in N} a_{ij} \quad (1)$$

$$\text{Betweenness centrality, } b_i = \frac{1}{(n-1)(n-2)} \sum_{h,j \in N, h \neq j, h \neq i, j \neq i} \frac{\rho_{hj}(i)}{\rho_{hj}} \quad (2)$$

$$\text{Clustering coefficient, } C_i = \frac{1}{n} \sum_{i \in N} \frac{2t_i}{k_i(k_i - 1)} \quad (3)$$

where $t_i = \frac{1}{2} \sum_{j,h \in N} a_{ij} a_{ih} a_{jh}$ is the number of triangles around node i; ρ_{hj} is the number of shortest paths between h and j and $\rho_{hj}(i)$ is the number of shortest paths between h and j that pass through i. Please note, k_i in Equation (3) is the nodal degree defined in Equation (1).

Intuitively, the degree of an individual node is equal to the number of links connected to that node, which therefore reflects the level of interaction of that node in the network. It is hypothesized that there are central nodes which participate in many short paths in the brain network. Betweenness centrality measures the fraction of all shortest paths in the network that pass through a given node. It is also known that human brain segregates specialized processing into interconnected groups of brain regions (clusters), clustering coefficient measures the clustering connectivity around a given node.

The ThickNet features for the CN and MCIc classes, in the form of group-differences, are visualized in Figs. 2 and 3. This may be helpful in obtaining better insight into the actual features in different classes. Fig. 3 shows that nodal degree exhibits the largest differences followed by betweenness and then clustering coefficient. Although the nodal degree features exhibit the largest differences, the differences are located in separated clusters (such as the 2 clusters in the post-central region) compared with other ThickNet features, which needs to be investigated further in future. These patterns of differences in nodal degree are similar to those seen in CN versus AD (not shown here for lack of space), which are just more pronounced and wider spread.

It is to be noted that the thickness network (as defined in this article) is an abstract construct that links 2 nodes based on similarity of thickness values between those nodes. This approach for network construction in the context of AD is similar to the extensive body of work by Evans et al. and other groups showing the construction of covariance structural networks based on cortical thickness or gray matter density in general. We are working on relating these networks to a biological explanation that relates these abstract constructs to the changes seen in the course of AD, and also developing appropriate visualizations that can aid in intuitive understanding of these ThickNet networks.

3. Evaluation of predictive utility

The ThickNet features reveal different properties of the regional links in thickness in the human brain. To maximize their utility for the early detection of AD, these features can be fused to form a composite set of features. Multiple kernel learning (MKL) is a natural choice for such a fusion of different features for the classification task. The procedure to evaluate the predictive utility is described in the following sections and also presented as a flowchart in Fig. 4.

3.1. Probabilistic MKL

One such method is the Variational Bayes probabilistic MKL (VBpMKL) which has been successfully applied to protein fold recognition (Damoulas and Girolami, 2008). This method combines multiple feature spaces, allowing a different kernel (e.g., gaussian and polynomial) for each feature space to embed them in a high-dimensional similarity space,

using a variational Bayes approximation to form a composite kernel. This composite kernel is fed to a multi-class model which applies Bayes theorem to learn the significance of each feature, as well as the kernel weights and kernel parameters automatically, without resorting to ad hoc parameter tuning. VBpMKL outputs probability estimates of membership to each class for each test subject, from which we can compute performance metrics (such as accuracy) as well as construct the receiver operating characteristic (ROC) curve.

3.2. Feature selection

Before fusion, further feature selection is done (within each feature set separately), by ranking each partition by its 2-sample t-statistic computed from the training set alone. All the partitions are ranked by their t-statistic, and the top K partitions are selected for training the classifier. We computed t-statistic with the alternative hypothesis that means are not equal assuming the variances are not equal, using the following formulae:

$$t = \frac{\overline{X_1} - \overline{X_2}}{S_{\overline{X_1} - \overline{X_2}}}, \; S_{\overline{X_1} - \overline{X_2}} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \quad (4)$$

Here, s^2 is an unbiased estimator of the variance of the 2 samples X_1 and X_2. $\overline{X_i}$ and n_i are the mean and the number of participants in each sample $i = 1, 2$, respectively.

3.3. Largest reduced dimensionality to avoid over fitting

There is an empirical relationship between the number of features (K) used to train the classifier and the minimum size of the training sample needed to avoid the curse of dimensionality, which is that for K number of features and small probability of error $p(e)$, the minimum sample size required $N_{\min} \geq \frac{K}{2p(e)}$ (Fitzpatrick and Sonka, 2000). If one would like to keep $p(e)$ below 5% with K features, we need at least $N_{\min} = K/(2 \times 0.05) = K \times 10$ subjects for training. We use this relation to determine the maximum number of features that can be used to train the classifier with an N_{train} number of samples in the training set, that is, $K_{\max} = \downarrow (N_{\text{train}}/10)$. This would give a $K_{\max} = 12$, 5, 15 and 5 for the experiments AD and/or CN, MCIc and/or CN, and MCI (MCIc + MCInc)/CN and MCIc/MCInc, respectively, when using the evaluation method to be described below in Section 3.4. We propose novel application of this approach derived from analytical results to set the largest dimensionality to avoid the possibility of over fitting.

After the selection of K features from each ThickNet measure, the 4 feature subsets are fed to the MKL classifier separately for each feature set. Depending on the combination strategy and parameters selected, this can sometimes result in simple concatenation of the subset (K) of features to arrive at a bigger set of features (4*K) to train the classifier. But this is only specific to some choices and is not always the case.

3.4. Repeated hold-out stratified Training set

Using this combination, that is, t-stat feature selection followed by VBpMKL as the classification system, we evaluate its predictive utility using a novel form of repeated

holdout method to handle class-imbalance. We eliminate class imbalance in the training set by first selecting a fixed percentage of subjects from the smallest class, and then selecting the same number from all the classes in the data set. We denote it as the repeated hold-out, stratified training set (RHsT) evaluation method. It is stratified in the sense that each class has an equal number of subjects in the training set to eliminate any class imbalance that may arise for typical uses of popular cross-validation methods. In each repetition, we hold out N_{train} subjects from each class for training and reserve the rest for testing the classifier. Here, N_{train} is determined by 95% of the smallest class in the experiment. For example, in the CN (n = 159) versus AD (n = 136) experiment, training set would consist of $N_{\text{train}} = \downarrow(0.95 \times 136) = 129$ subjects from both CN and AD classes and the testing set would have 30 CN and 7 AD subjects. In each repetition, we compute the accuracy, sensitivity, and specificity as well as area under curve (AUC) by constructing an ROC, from the predictions generated for the corresponding unseen test set. This method is repeated 100 times, each time creating random training and/or test sets, to avoid the bias that can arise from a single training and/or test sets (Cuingnet et al., 2011). The mean performance metrics, and their standard deviations, from the 100 repetitions are reported.

3.5. Comparison of performance improvement

To analyze performance improvement contributed by the proposed ThickNet (TN) features, we compared its performance relative to mean thickness (MT) features alone. The classification power of mean thickness features alone, in place of the ThickNet features, is evaluated while keeping the rest of the evaluation procedure (RHsT) the same. The best performance of the MT features (highest AUC over different values of TNP) is compared with that of TN method, and the results are shown in Table 2.

We tested this using 2 measures of AUC: (1) covering the entire ROC curve, which tests whether one method is better than the other for the entire range of sensitivity (blanket test); and (2) covering only a high-specificity portion of ROC curve which focuses on clinically relevant portion of the ROC. Accordingly, we define partial AUC (pAUC) as the AUC under partial ROC curve bounded by specificity over 85%. This upper threshold of 85% specificity is commonly accepted for a clinically relevant biomarker and/or test (McClish, 1989).

4. Results

The evaluation method as described in Section 3, and graphically summarized in Fig. 4, is applied to the fusion of the following 4 feature sets: mean thickness, nodal degree, betweenness centrality, and clustering coefficient at each partition. From preliminary trials in AD and/or CN classification, we observed the best performance from VBpMKL using a polynomial kernel (third degree) for each feature set and thereby fixing it as the kernel of choice for this study. The performance of the fusion method is evaluated in the following 4 binary classification experiments: CN versus AD, CN versus MCIc, CN versus MCI, and MCInc versus MCIc, to compare their predictive utility under different levels of separability.

For each such experiment, there are 2 parameters that change the feature extraction (of the mean thickness and the 3 network features): TNP and the link threshold a. Choice of a larger or smaller TNP is equivalent to selecting a different parcellation scheme (coarse to

finer resolution), and selecting different values for α can be interpreted as selecting different types of features, that is, weak connections with lenient alpha (large tolerance for similarity), and strong connections with stringent alpha (relatively low tolerance for similarity). To avoid making an arbitrary choice for these parameter values, we have studied the performance of our method for different combinations of TNP and α, with TNP = 340, 680,1020,1360, and 1700, and α was varied from 0.1 mm to 1.5 mm, in steps of 0.1 mm. The AUC for all the combinations are visualized in Fig. 5. These results show that AUC for ThickNet features is high for CN versus AD, and it is robust for a wide range of values for TNP and α. The results in the other three experiments show that (1) AUC decreases with decreasing separability across the 4 experiments; and (2) AUC is robust with varying values for TNP and α in all the experiments. These figures display the same trends that AUC exhibited in all the experiments. The results in 3 metrics combined assert the promising performance of ThickNet classifier, as well as its robustness to the parameters of feature extraction.

The best performance (highest AUC) of the ThickNet fusion classifier for different experiments are summarized in Table 1, with different performance metrics and the optimal ThickNet parameters TNP and α. Table 1 shows that ThickNet classifier exhibits not only high AUC but also a balanced performance in terms of similar sensitivity and specificity. Although there is a class imbalance in the data set, neither specificity nor sensitivity is over-learned by our classifier. Corresponding ROCs are visualized in Fig. 6, which are constructed by averaging the 100 ROCs obtained from the 100 repetitions of the RHST, using the vertical averaging method as described in Fawcett (2006).

The results from the comparison of performance improvement, as described in Section 3.5, are shown in Table 2 for all the classification experiments presented in Table 1. The corresponding ROCs for ThickNet and MT are compared in Fig. 6. Further summary of the comparison between MT and ThickNet features are presented in the form of descriptive statistics in Table 3.

The average weights from the 100 repetitions of RHsT for the ThickNet features, indicating their individual significance, are visualized in Fig. 7. The visualizations for sensitivity and specificity estimates of the ThickNet fusion classifier, in the lines of Fig. 5, are presented in Figs. 8 and 9, respectively.

5. Discussion

Our results confirm a general trend that the classification performance is the highest in CN versus AD experiment and it gradually decreases as the problem becomes increasingly challenging from CN versus MCIc, CN versus MCI, to MCIc versus MCInc. From Table 1, we observe that our method discriminates AD from CN with an AUC of 0.92. We can also observe that the performance of our method decreases as the problem becomes increasingly challenging, that is, in AUC for CN/MCI classification is 0.75 with 64% sensitivity and 73% specificity; in MCInc/MCIc classification, AUC is 0.68 with 65% sensitivity and 64% specificity, which is to be expected.

The most promising result in Table 1 is the performance in discriminating MCIc from CN with an AUC of 0.83 (74% sensitivity and 76% specificity). For a similar experiment, the mean thickness in each freesurfer label (or ROI, many times larger than our partition), noted as the thickness-ROI method in Cuingnet et al. (2011) resulted in 65% sensitivity and 94% specificity. This is not directly comparable with our result, as we fuse multiple ThickNet features whereas other study uses mean thickness alone; however, it demonstrates the improvement achieved by using the sophisticated method.

Another study (Eskildsen et al., 2013) that utilized mean thickness features for classification between progressive MCI (conversion in 24 months, results for 18 months are not reported) and stable MCI, reports an AUC of 0.67, 59% sensitivity and 70% specificity. For a similar experiment using mean thickness on classification between MCInc (stable MCI) and MCIc (conversion in 18 months), we obtain AUC of 0.68 with 65% sensitivity and 64% specificity. Our performance is slightly better than (Eskildsen et al., 2013), although it is to be noted the subsets being studied are different, and the 2 studies utilize different types of mean thickness. There have been many reports on MCInc and/or MCIc classification, as summarized in (Cuingnet et al., 2011), mostly with low classification performance. This may be because of MCInc being a very heterogeneous and unstable diagnostic group.

Another interesting point to note from Table 1 is that when the separability is higher (CN vs. AD and CN vs. MCIc), the best performance was obtained with a coarse partitioning (TNP = 340, relatively large patches) and a stringent threshold ($\alpha = 0.3$ mm that results in only few links). In contrast, in challenging problems with lower separability (CN vs. MCI and MCIc vs. MCInc), our method needed a very intricate network (TNP > 1000 resulting in large number of small areas and a lenient $\alpha = 1.3$ mm resulting in many links). A lenient α results in higher nodal degree (each node is connected to larger number of nodes) and smaller centrality (discourages local clustering) and clustering coefficient. This makes sense clinically and conforms to our understanding of the disease stages.

Figs. 8 and 9 presenting sensitivity and specificity, show that the performance of our method is robust across a wide range of values of the parameters TNP and α. Moreover, we notice that these metrics follow the same trends as AUC, when the performance of ThickNet fusion method is compared across the 4 different experiments.

The results presented thus far demonstrate the diagnostic utility of ThickNet features. An objective comparison with mean thickness (MT) features, as described in Section 3.5, is performed. The results from the comparison are presented in Fig. 6, Tables 2 and 3. Fig. 6 shows that ThickNet outperformed MT (encompassing the ROC), over a wide range of specificities, in all the experiments except for CN versus MCI. As most real features perform better than a random classifier, the ROC curves tend to be very similar in the low-specificity range (e.g., false positive range >15%). When we consider the clinically relevant pAUC (the second measure of ROC area), the results in Table 2 show that ThickNet method exhibits higher pAUC in all the experiments. Overall, these results comparing the classification performance of ThickNet features directly with mean thickness indicate that ThickNet features do add value by providing equivalent or better performance than mean thickness

alone. However, we should investigate further as to the clinical significance of the performance offered by the ThickNet features to understand its prognostic implications.

5.1. Individual significance of ThickNet features

For each run of RHsT, we obtain not only the prediction of the test set subjects but also the significance of each feature set in the fused classifier, estimated by VBpMKL. This allows to gain further insight into the contribution of different feature sets. The average weights from the 100 repetitions of RHsT for the ThickNet features are visualized in Fig. 7, for the 4 classification experiments.

Results from Fig. 7 show that all the ThickNet features are contributing to the classifier (non-zero weights), although in varying proportions. Note that, the contribution of mean thickness features is significant in all the classification problems. Whereas the contribution of ThickNet features increased with increasing difficulty of the classification problem, such as CN versus MCIc and MCIc versus MCInc. This further asserts their utility for the prognostic applications for early detection of AD.

6. Conclusions and future work

We present novel ThickNet features that can be extracted from a single time-point MRI scan and demonstrate their potential for individual patient diagnosis. As these features are not specific to a disease, they can be easily applied to other prognostic problems in neuroimaging. The diagnostic utility of ThickNet features is demonstrated by applying probabilistic MKL (preceded by t-statistic feature selection) to the challenging problem of detection of prodromal AD, that is, MCIc based on baseline MRI scan alone. We report an AUC of 0.83 for CN/MCIc classification problem with 74% sensitivity and 76% specificity, which is promising. Furthermore, we present a detailed comparison of the classification performance of the proposed ThickNet fusion method in AD/CN, MCIc/MCInc and MCI (MCIc + MCInc)/CN classification experiments.

We would like to note that there is likely significant room for improvement, for example, by computing more ThickNet features using additional measures of centrality, segregation, hub-likeness, and integration, as well as constructing weighted graphs from the regional links in cortical thickness as opposed to current choice of binary and undirected graphs in this study. Moreover, we could apply different (or multiple) kernels for each feature as well as tuning the kernel parameters, as opposed to the current results obtained with a fixed kernel (polynomial kernel, degree = 3). This framework is easily extensible, for example, in including features from other modalities such as positron emission tomography, diffusion-tensor imaging, as well as other morphologic and neuropsychological features. Each new feature can be tuned with an additional kernel, which can be easily fused with existing features. Moreover, as the classifier used is by design multi-class, this method can be readily applied to differential diagnosis, for example, discriminating AD from other types of dementia such as frontotemporal disease and vascular dementia and so forth.

Acknowledgments

We gratefully acknowledge funding support from Alzheimer Society of Canada Research Program (ASRP), Natural Sciences and Engineering Research Council of Canada (NSERC), Canadian Institutes of Health Research (CIHR) and Michael Smith Foundation for Health Research (MSFHR). They sincerely thank Dr Michael W. Weiner and the Freesurfer team at University of California, San Francisco for the computation and quality control of Freesurfer processing for ADNI data set.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (from National Institutes of Health grants U01 AG024904, P30 AG010129, and K01 AG030514), for which complete listing of acknowledgements can be found at http://adni.loni.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Manuscript_Citations.pdf. Data collection and sharing for this project was funded by the ADNI (National Institutes of Health grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimer Drug Discovery Foundation; BioClinica, Inc; Biogen Idec Inc; Bristol-Myers Squibb Foundation; Eisai; Elan Pharmaceuticals, Inc; Eli Lilly and Company; F. Hoffmann-La Roche Ltd, and its affiliated company Genentech, Inc; GE Healthcare; Innogenetics, N.V.; IXICO Ltd; Janssen Alzheimer Immunotherapy Research & Development, LLC; Johnson & Johnson Pharmaceutical Research & Development LLC; Medpace, Inc; Merck & Co, Inc; Meso Scale Diagnostics, LLC; NeuroRx Research; Novartis Pharmaceuticals Corporation; Pfizer; Piramal Imaging; Servier; Synarc Inc; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles. This research was also supported by National Institutes of Health grants P30 AG010129 and K01 AG030514.

Appendix

Excluded subjects from quality control

Here, we list the set of subjects that were excluded from analysis in Table 4.

Estimation of the common atlas

After the extraction of cortical thickness from each subject, we registered the surface of each subject to that of a common atlas. This atlas is derived from averaging 80 healthy controls using tools from Freesurfer. With the help of Talairach transform computed for each subject, Talairach (MNI305) coordinates for each vertex are computed. These coordinates (from the 80 subjects) are averaged after mapping them to the common surface (which overlays well on the average MNI305 volume). In the following, we list all the subjects that were part of this averaging in Table 5.

References

Alexander-Bloch A, Giedd JN, Bullmore E. Imaging structural co-variance between human brain regions. Nat Rev Neurosci. 2013a; 14:322–336. [PubMed: 23531697]

Alexander-Bloch A, Raznahan A, Bullmore E, Giedd J. The convergence of maturational change and structural covariance in human cortical networks. J Neurosci. 2013b; 33:2889–2899. [PubMed: 23407947]

Alzheimer's Association. 2012 Alzheimer's disease facts and figures. Alzheimer's Dement. 2012; 8:131–168. [PubMed: 22404854]

Amieva H, Le Goff M, Millet X, Orgogozo JM, Pérès K, Barberger Gateau P, Jacqmin Gadda H, Dartigues JF. Prodromal Alzheimer's disease: successive emergence of the clinical symptoms. Ann Neurol. 2008; 64:492–498. [PubMed: 19067364]

Bassett DS, Bullmore E, Verchinski BA, Mattay VS, Weinberger DR, Meyer-Lindenberg A. Hierarchical organization of human cortical networks in health and schizophrenia. J Neurosci. 2008; 28:9239–9248. [PubMed: 18784304]

Beg MF, Raamana PR, Barbieri S, Wang L. Comparison of four shape features for detecting hippocampal shape changes in early Alzheimer's. Stat Methods Med Res. 2013; 22:439–462. [PubMed: 22653846]

Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. Acta Neuropathologica. 1991; 82:239–259. [PubMed: 1759558]

Braak H, Del Tredici K. The pathological process underlying Alzheimer's disease in individuals under thirty. Acta Neuropathologica. 2011; 121:171–181. [PubMed: 21170538]

Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci. 2009; 10:186–198. [PubMed: 19190637]

Chen ZJ, He Y, Rosa-Neto P, Germann J, Evans A. Revealing modular architecture of human brain structural networks by using cortical thickness from MRI. Cereb Cortex. 2008; 18:2374–2381. [PubMed: 18267952]

Coupé P, Eskildsen SF, Manjón JV, Fonov VS, Pruessner JC, Allard M, Collins DL, Alzheimer's Disease Neuroimaging Initiative. Scoring by nonlocal image patch estimator for early detection of Alzheimer's disease. Neuroimage Clin. 2012; 1:141–152. [PubMed: 24179747]

Cuingnet R, Gerardin E, Tessieras J, Auzias G, Lehericy S, Habert MO, Chupin M, Benali H, Colliot O, Initiative ADN. Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. Neuroimage. 2011; 56:766–781. [PubMed: 20542124]

Damoulas T, Girolami MA. Probabilistic multi-class multi-kernel learning: on protein fold recognition and remote homology detection. Bioinformatics. 2008; 24:1264–1270. [PubMed: 18378524]

Desikan RS, Cabral HJ, Hess CP, Dillon WP, Glastonbury CM, Weiner MW, Schmansky NJ, Greve DN, Salat DH, Buckner RL, Fischl B, Alzheimer's Disease Neuroimaging Initiative. Automated MRI measures identify individuals with mild cognitive impairment and Alzheimer's disease. Brain. 2009; 132(Pt 8):2048–2057. [PubMed: 19460794]

Dickerson BC, Bakkour A, Salat D, Feczko E, Pacheco J, Greve D, Grodstein F, Wright C, Blacker D, Rosas H. The cortical signature of Alzheimer's disease: regionally specific cortical thinning relates to symptom severity in very mild to mild AD dementia and is detectable in asymptomatic amyloid-positive individuals. Cereb Cortex. 2009; 19:497–510. [PubMed: 18632739]

Du AT, Schuff N, Kramer JH, Rosen HJ, Gorno-Tempini ML, Rankin K, Miller BL, Weiner MW. Different regional patterns of cortical thinning in Alzheimer's disease and frontotemporal dementia. Brain. 2007; 130:1159–1166. [PubMed: 17353226]

Duchesne S, Bocti C, De Sousa K, Frisoni G, Chertkow H, Collins DL. Amnestic MCI future clinical status prediction using baseline MRI features. Neurobiol Aging. 2010; 31:1606–1617. [PubMed: 18947902]

Eskildsen SF, Coupé P, García-Lorenzo D, Fonov V, Pruessner JC, Collins DL. Prediction of Alzheimer's disease in subjects with mild cognitive impairment from the ADNI cohort using patterns of cortical thinning. Neuroimage. 2013; 65:511–521. [PubMed: 23036450]

Evans AC. Networks of anatomical covariance. Neuroimage. 2013; 80:489–504. [PubMed: 23711536]

Fawcett T. An introduction to ROC analysis. Pattern Recognition Lett. 2006; 27:861–874.

Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron. 2002; 33:341–355. [PubMed: 11832223]

Fitzpatrick, M., Sonka, M. Handbook of Medical Imaging Vol. 2: Medical Image Processing & Analysis (PM80) SPIE. International Society for Optical Engineering; Seattle, WA: 2000.

Gibson E, Wang L, Beg MF. Cortical thickness measurement using eulerian PDEs and surface-based global topological information. Org Human Brain Mapping, 15th Annual Meeting. 2009

Gong G, He Y, Chen ZJ, Evans AC. Convergence and divergence of thickness correlations with diffusion connections across the human cerebral cortex. Neuroimage. 2012; 59:1239–1248. [PubMed: 21884805]

Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, Sporns O. Mapping the structural core of human cerebral cortex. PLoS Biol. 2008; 6:e159. [PubMed: 18597554]

He Y, Chen ZJ, Evans AC. Small-world anatomical networks in the human brain revealed by cortical thickness from mri. Cereb Cortex. 2007; 17:2407–2419. [PubMed: 17204824]

He Y, Chen Z, Evans A. Structural insights into aberrant topological patterns of large-scale cortical networks in Alzheimer's disease. J Neurosci. 2008; 28:4756–4766. [PubMed: 18448652]

He Y, Chen Z, Gong G, Evans A. Neuronal networks in Alzheimer's disease. Neuroscientist. 2009a; 15:333–350. [PubMed: 19458383]

He Y, Dagher A, Chen Z, Charil A, Zijdenbos A, Worsley K, Evans A. Impaired small-world efficiency in structural cortical networks in multiple sclerosis associated with white matter lesion load. Brain. 2009b; 132:3366–3379. [PubMed: 19439423]

Iturria-Medina Y. Anatomical brain networks on the prediction of abnormal brain states. Brain Connect. 2013; 3:1–21. [PubMed: 23249224]

Jack CR Jr, Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, Petersen RC, Trojanowski JQ. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. Lancet Neurol. 2010; 9:119–128. [PubMed: 20083042]

Khundrakpam BS, Reid A, Brauer J, Carbonell F, Lewis J, Ameis S, Karama S, Lee J, Chen Z, Das S, Evans AC, Brain Development Cooperative Group. Ball WS, Byars AW, Schapiro M, Bommer W, Carr A, German A, Dunn S, Rivkin MJ, Waber D, Mulkern R, Vajapeyam S, Chiverton A, Davis P, Koo J, Marmor J, Mrakotsky C, Robertson R, McAnulty G, Brandt ME, Fletcher JM, Kramer LA, Yang G, McCormack C, Hebert KM, Volero H, Botteron K, McKinstry RC, Warren W, Nishino T, Robert Almli C, Todd R, Constantino J, McCracken JT, Levitt J, Alger J, O'Neil J, Toga A, Asarnow R, Fadale D, Heinichen L, Ireland C, Wang DJ, Moss E, Zimmerman RA, Bintliff B, Bradford R, Newman J, Evans AC, Arnaoutelis R, Bruce Pike G, Louis Collins D, Leonard G, Paus T, Zijdenbos A, Das S, Fonov V, Fu L, Harlap J, Leppert I, Milovan D, Vins D, Zeffiro T, Van Meter J, Lange N, Froimowitz MP, Botteron K, Robert Almli C, Rainey C, Henderson S, Nishino T, Warren W, Edwards JL, Dubois D, Smith K, Singer T, Wilber AA, Pierpaoli C, Basser PJ, Chang LC, Koay CG, Walker L, Freund L, Rumsey J, Baskir L, Stanford L, Sirocco K, Gwinn-Hardy K, Spinella G, McCracken JT, Alger JR, Levitt J, O'Neill J. Developmental changes in organization of structural brain networks. Cereb Cortex. 2013; 23:2072–2085. [PubMed: 22784607]

Koikkalainen J, Lötjönen J, Thurfjell L, Rueckert D, Waldemar G, Soininen H. Multi-template tensor-based morphometry: application to analysis of Alzheimer's disease. Neuroimage. 2011; 56:1134–1144. [PubMed: 21419228]

La Joie R, Perrotin A, Barré L, Hommet C, Mézenge F, Ibazizene M, Camus V, Abbas A, Landeau B, Guilloteau D, de La Sayette V, Eustache F, Desgranges B, Chételat G. Region-specific hierarchy between atrophy, hypometabolism, and β-amyloid (aβ) load in Alzheimer's disease dementia. J Neurosci. 2012; 32:16265–16273. [PubMed: 23152610]

Lee JK, Lee JM, Kim JS, Kim IY, Evans AC, Kim SI. A novel quantitative cross-validation of different cortical surface reconstruction algorithms using MRI phantom. Neuroimage. 2006; 31:572–584. [PubMed: 16503170]

Lerch JP, Evans AC. Cortical thickness analysis examined through power analysis and a population simulation. Neuroimage. 2005; 24:163–173. [PubMed: 15588607]

Lerch JP, Worsley K, Shaw WP, Greenstein DK, Lenroot RK, Giedd J, Evans AC. Mapping anatomical correlations across cerebral cortex (macacc) using cortical thickness from MRI. Neuroimage. 2006; 31:993–1003. [PubMed: 16624590]

McClish DK. Analyzing a portion of the ROC curve. Med Decis Making. 1989; 9:190–195. [PubMed: 2668680]

Mcevoy LK, Holland D, Hagler DJ, Fennema-Notestine C, Brewer JB, Dale AM, Alzheimer's Disease Neuroimaging Initiative. Mild cognitive impairment: baseline and longitudinal structural MR imaging measures improve predictive prognosis. Radiology. 2011; 259:834–843. [PubMed: 21471273]

Mechelli A, Friston KJ, Frackowiak RS, Price CJ. Structural covariance in the human cortex. J Neurosci. 2005; 25:8303–8310. [PubMed: 16148238]

Misra C, Fan Y, Davatzikos C. Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: results from ADNI. Neuroimage. 2009; 44:1415–1422. [PubMed: 19027862]

Raamana P, Wang L, Rosen H, Miller B, Weiner M, Beg MF. Differential diagnosis among Alzheimer's disease, frontotemporal disease and healthy aging: Comparative study using subcortical features. Alzheimer's Dement. 2012; 8:P163–P164.

Raj A, Mueller SG, Young K, Laxer KD, Weiner M. Network-level analysis of cortical thickness of the epileptic brain. Neuroimage. 2010; 52:1302–1313. [PubMed: 20553893]

Rubinov M, Sporns O. Complex network measures of brain connectivity: uses and interpretations. Neuroimage. 2010; 52:1059–1069. [PubMed: 19819337]

Seeley WW, Crawford RK, Zhou J, Miller BL, Greicius MD. Neurodegenerative diseases target large-scale human brain networks. Neuron. 2009; 62:42–52. [PubMed: 19376066]

Tijms BM, Seriès P, Willshaw DJ, Lawrie SM. Similarity-based extraction of individual networks from gray matter MRI scans. Cereb Cortex. 2012; 22:1530–1541. [PubMed: 21878484]

Tosun D, Schuff N, Mathis C, Jagust W, Weiner MW. Spatial patterns of brain amyloid-β burden and atrophy rate associations in mild cognitive impairment. Brain. 2011; 134:1077–1088. [PubMed: 21429865]

Villain N, Chételat G, Grassiot B, Bourgeat P, Jones G, Ellis KA, Ames D, Martins RN, Eustache F, Salvado O, Masters CL, Rowe CC, Villemagne VL, AIBL Research Group. Regional dynamics of amyloid- deposition in healthy elderly, mild cognitive impairment and Alzheimer's disease: a voxel-wise PiB-PET longitudinal study. Brain. 2012; 135:2126–2139. [PubMed: 22628162]

Walhovd K, Fjell A, Brewer J, McEvoy L, Fennema-Notestine C, Hagler D, Jennings R, Karow D, Dale A, Alzheimer's Disease Neuroimaging Initiative. Combining MR imaging, positron-emission tomography, and CSF biomarkers in the diagnosis and prognosis of Alzheimer disease. AJNR Am J Neuroradiol. 2010; 31:347–354. [PubMed: 20075088]

Wang L, Beg MF, Ratnanather JT, Ceritoglu C, Younes L, Morris JC, Csernansky JG, Miller MI. Large deformation diffeomorphism and momentum based hippocampal shape discrimination in dementia of the Alzheimer type. IEEE Trans Med Imaging. 2007; 26:462–470. [PubMed: 17427733]

Wee CY, Yap PT, Zhang D, Denny K, Browndyke JN, Potter GG, Welsh-Bohmer KA, Wang L, Shen D. Identification of MCI individuals using structural and functional connectivity networks. Neuroimage. 2012; 59:2045–2056. [PubMed: 22019883]

Wen W, He Y, Sachdev P. Structural brain networks and neuropsychiatric disorders. Curr Opin Psychiatry. 2011; 24:219–225. [PubMed: 21430538]

Westman E, Muehlboeck JS, Simmons A. Combining MRI and CSF measures for classification of Alzheimer's disease and prediction of mild cognitive impairment conversion. Neuroimage. 2012; 62:229–238. [PubMed: 22580170]

Wolz R, Julkunen V, Koikkalainen J, Niskanen E, Zhang DP, Rueckert D, Soininen H, Lötjönen J, Alzheimer's Disease Neuroimaging Initiative. Multi-method analysis of MRI images in early diagnostics of Alzheimer's disease. PLoS One. 2011; 6:e25446. [PubMed: 22022397]

Woodward M, Mackenzie IRA, Hsiung GYR, Jacova C, Feldman H. Multiple brain pathologies in dementia are common. Eur Geriatr Medecine. 2010; 1:259–265.

Yao Z, Zhang Y, Lin L, Zhou Y, Xu C, Jiang T, Alzheimer's Disease Neuroimaging Initiative. Abnormal cortical networks in mild cognitive impairment and Alzheimer's disease. PLoS Comput Biol. 2010; 6:e1001006. [PubMed: 21124954]

Yezzi AJ, Prince JL. An eulerian PDE approach for computing tissue thickness. IEEE Trans Med Imaging. 2003; 22:1332–1339. [PubMed: 14552586]

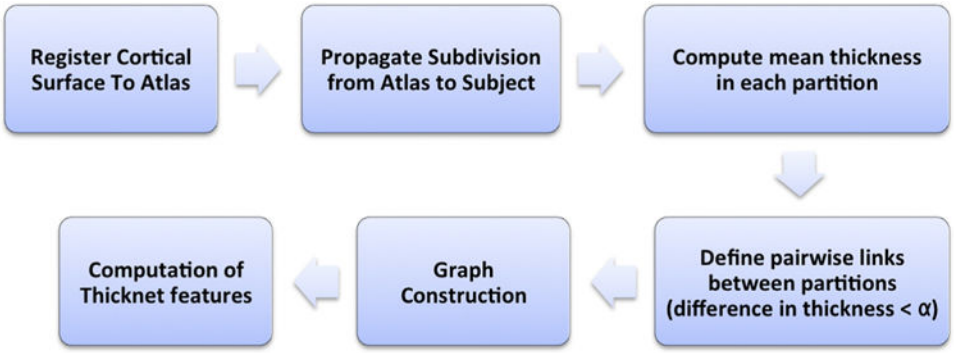

Fig. 1.

Flow chart describing the steps involved in the extraction of ThickNet features. Once the pial surfaces from all the subjects are registered to a common atlas, we subdivide the cortex of the atlas surface into a fixed number of partitions (or patches). This subdivision is propagated into cortical surface of each subject and mean thickness within each partition is computed for all the patches in every subject. Based on the similarity in thickness, links are defined between various pairs of partitions with difference in mean thickness below a certain threshold. The Boolean link status between all the pairwise connections forms the adjacency matrix of the graph. From this graph, we compute various ThickNet features. Please refer to Section 2.4 for a detailed description. (For interpretation of the references to color in this Figure, the reader is referred to the web version of this article.)

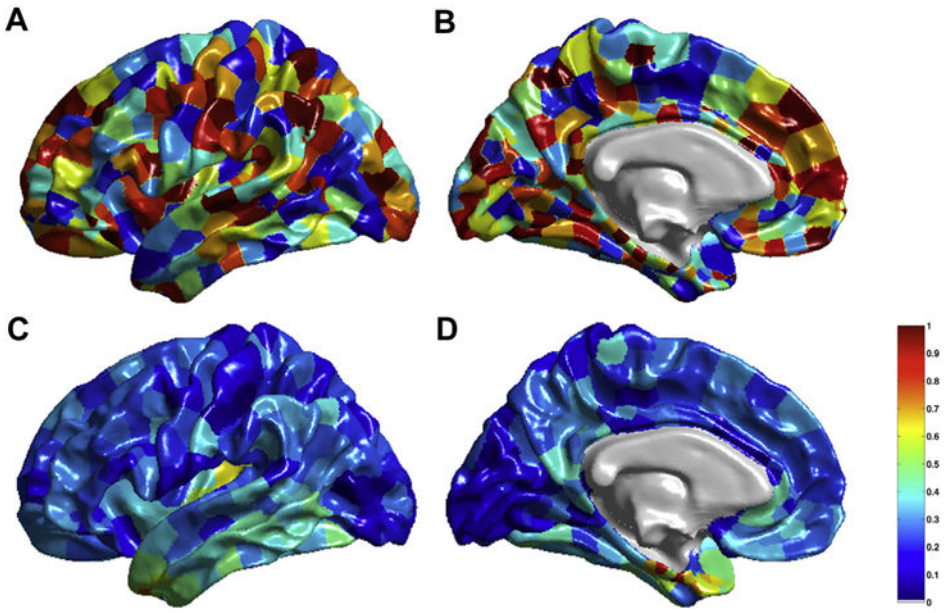

Fig. 2.

Visualization of the partitions on the atlas surface in the medial view (A) and lateral view (B), when TNP = 680. Also visualized here in C and D are the group differences in the mean thickness, that is, mean(CN)-mean(MCIc) at each partition, rescaled to [0,1] to enable comparison with other ThickNet features show in Fig. 3. Abbreviations: CN, controls; MCIc, mild cognitive impairment converters; TNP, total number of partitions. (For interpretation of the references to color in this Figure, the reader is referred to the web version of this article.)

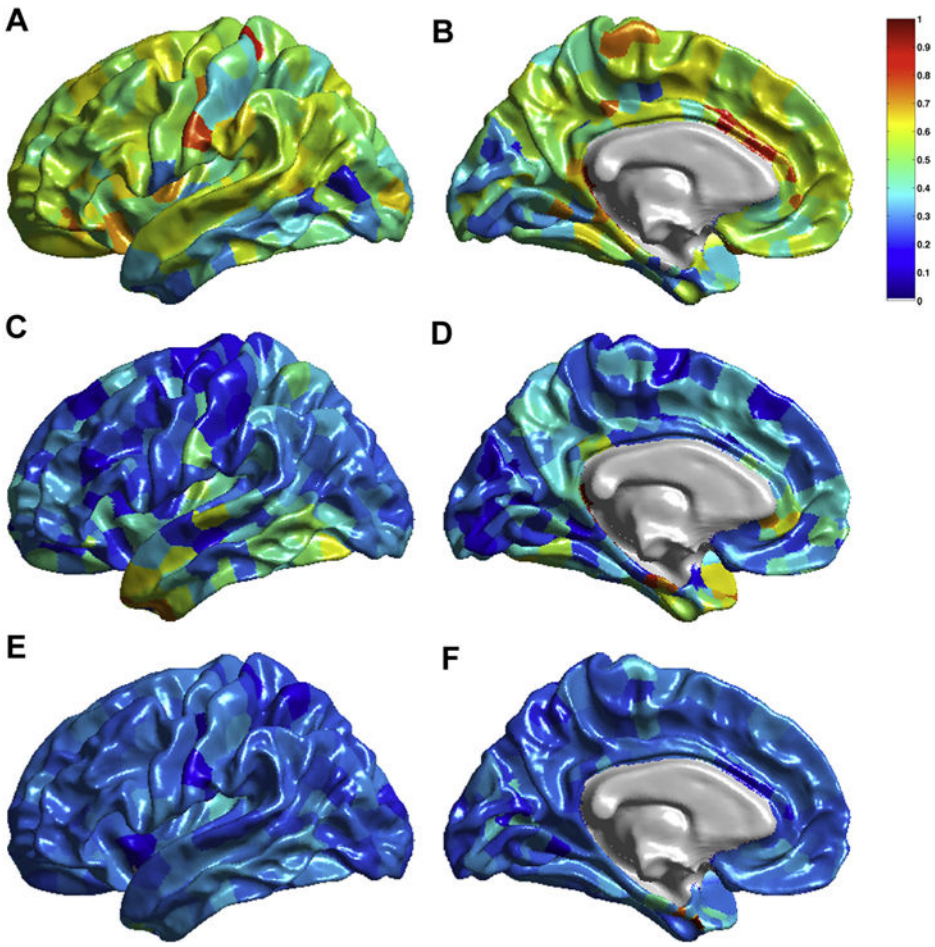

Fig. 3.

Visualization of the differences in group means, that is, mean(CN)-mean(MCIc) at each partition, of the ThickNet features when TNP = 680 and α = 0.30. Left column presents the medial view and the right column presents the lateral view of the group differences in each feature. The values of each feature in A–F are normalized to [0,1] to enable comparison across features. These values do not have any applicable units. Abbreviations: CN, controls; MCIc, mild cognitive impairment converters; TNP, total number of partitions. (For interpretation of the references to color in this Figure, the reader is referred to the web version of this article.)

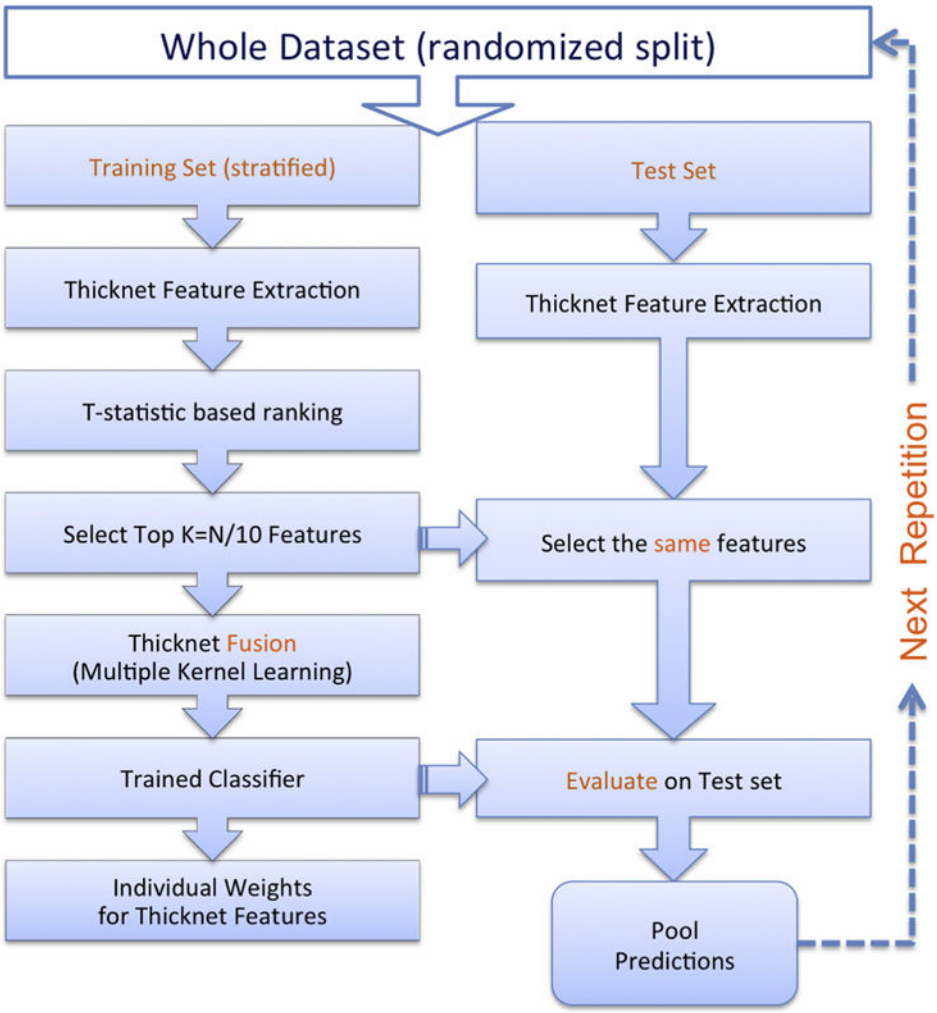

Fig. 4.

Flowchart illustrating the performance evaluation procedure utilized in this study. The training set is stratified in the sense that there is no class-imbalance (all the classes are equal in size) to limit any bias toward 1 particular class. Please note, this procedure is repeated 100 times. In each repetition, the performance metrics are computed based on the predictions from the corresponding test set only. In other words, we do not pool predictions across different repetitions, which may invalidate the computation of AUC. That would be invalid because the prediction scores in different repetitions are obtained from different classifiers, which may not be comparable or calibrated. Abbreviation: AUC, area under curve. (For interpretation of the references to color in this Figure, the reader is referred to the web version of this article.)

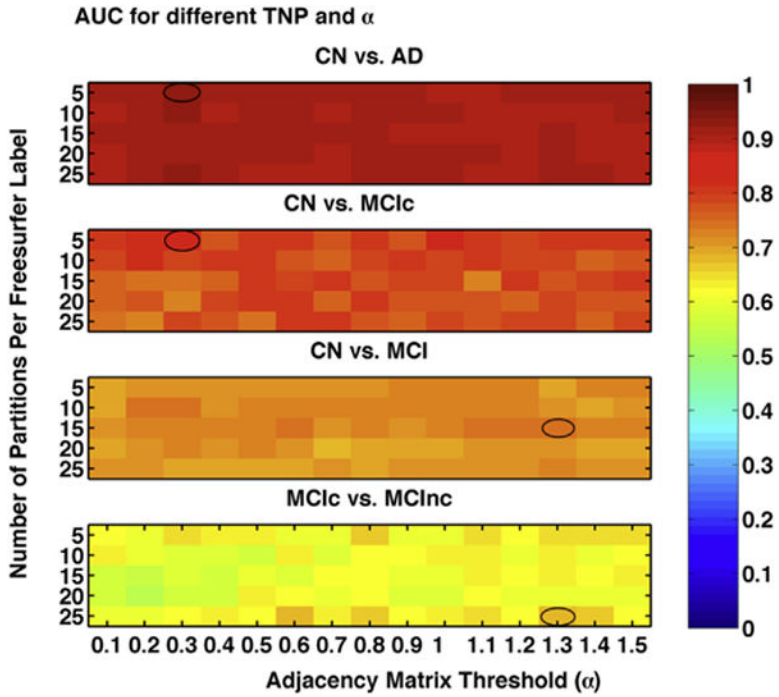

Fig. 5.

Comparison of AUC obtained from RHsT method for each combination of NPP and α. The combination with the best performance in each experiment is highlighted with a black oval. Abbreviations: AD, Alzheimer's disease; AUC, area under curve; MCI, mild cognitive impairment; MCIc, mild cognitive impairment converters; MCInc, mild cognitive impairment nonconverters; NPP, number of partitions per freesurfer label; RHsT, repeated hold-out, stratified training set; TNP, total number of partitions. (For interpretation of the references to color in this Figure, the reader is referred to the web version of this article.)

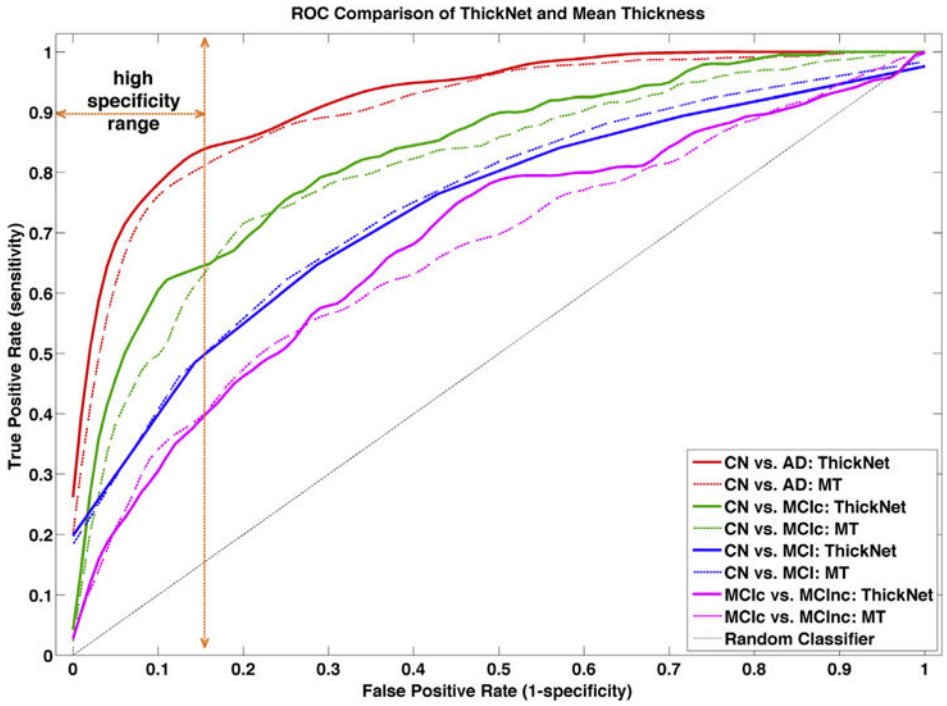

Fig. 6.

Comparison of ROC curves corresponding with the best performance of ThickNet fusion method in each experiment, displayed using solid lines. We also compare these ROC curves with those of mean thickness (MT) features. This comparison shows that ThickNet features outperform the MT features (dashed lines) in all the experiments except CN versus MCI. Abbreviations: CN, controls; MCI, mild cognitive impairment; ROC, receiver operating characteristic. (For interpretation of the references to color in this Figure, the reader is referred to the web version of this article.)

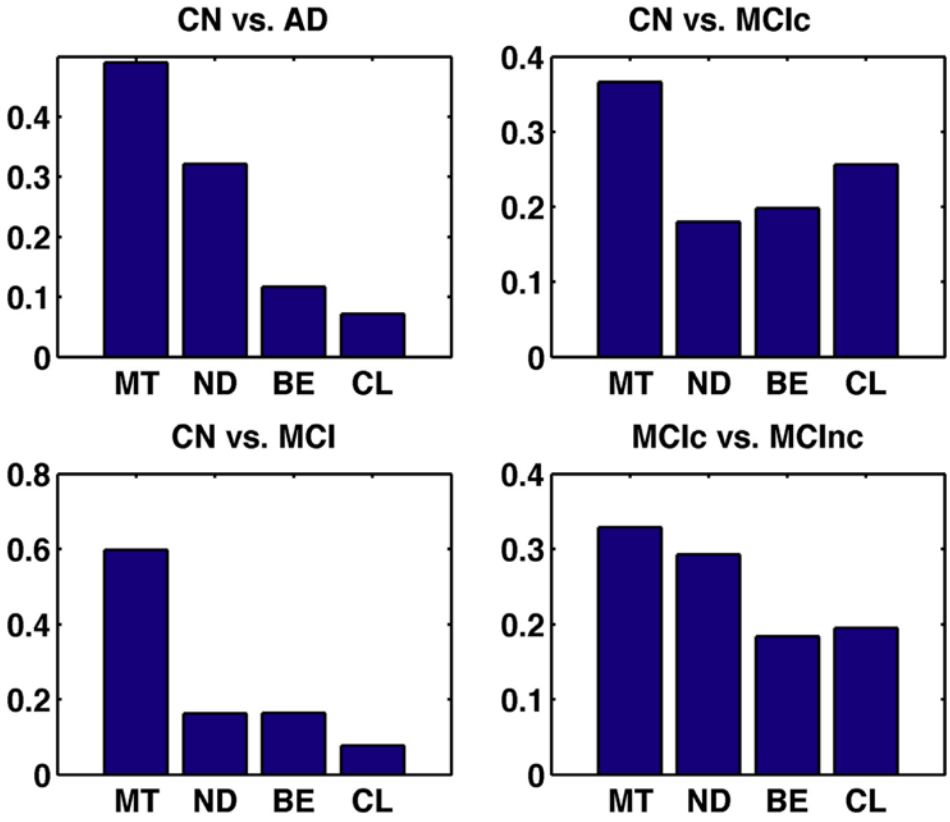

Fig. 7.

Individual contribution of ThickNet features toward classification in the probabilistic multiple kernel learning framework. These results show that all the ThickNet features contributed to discrimination, although in varying proportions. Note that, mean thickness contributed in all the classification problems, whereas the contribution of ThickNet features increased with increasing difficulty of the problem such as CN versus MCIc and MCIc versus MCInc. This only asserts their utility for the prognostic applications. Abbreviations: BE, betweenness centrality; CL, clustering coefficient; MT, mean thickness; ND, nodal degree. (For interpretation of the references to color in this Figure, the reader is referred to the web version of this article.)

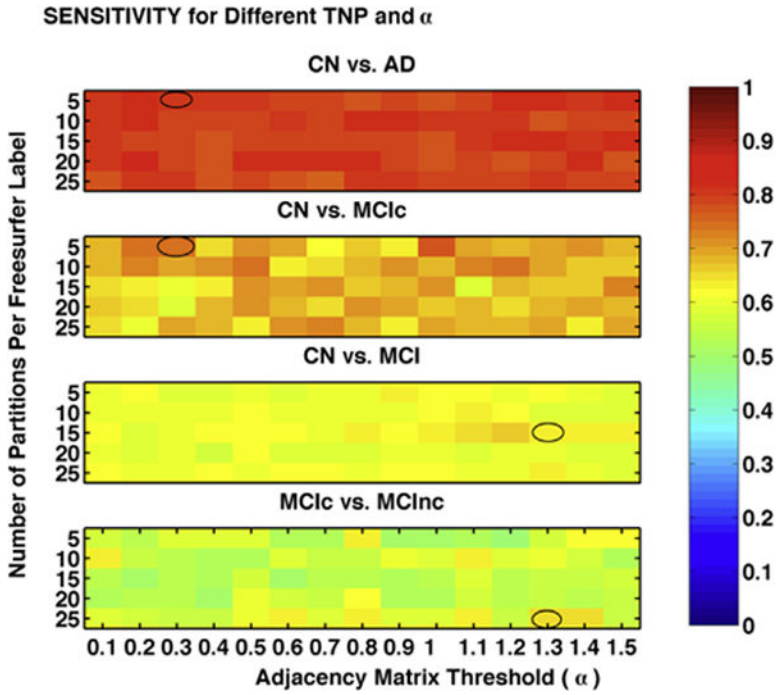

Fig. 8.

Comparison of sensitivity, as a function of TNP and α, obtained from RHsT method. The combination with the best performance (highest AUC) in each experiment is highlighted with a black oval. Abbreviations: AD, Alzheimer's disease; AUC, area under curve; MCI, mild cognitive impairment; MCIc, mild cognitive impairment converters; MCInc, mild cognitive impairment nonconverters; RHsT, repeated hold-out, stratified training set; TNP, total number of partitions. (For interpretation of the references to color in this Figure, the reader is referred to the web version of this article.)

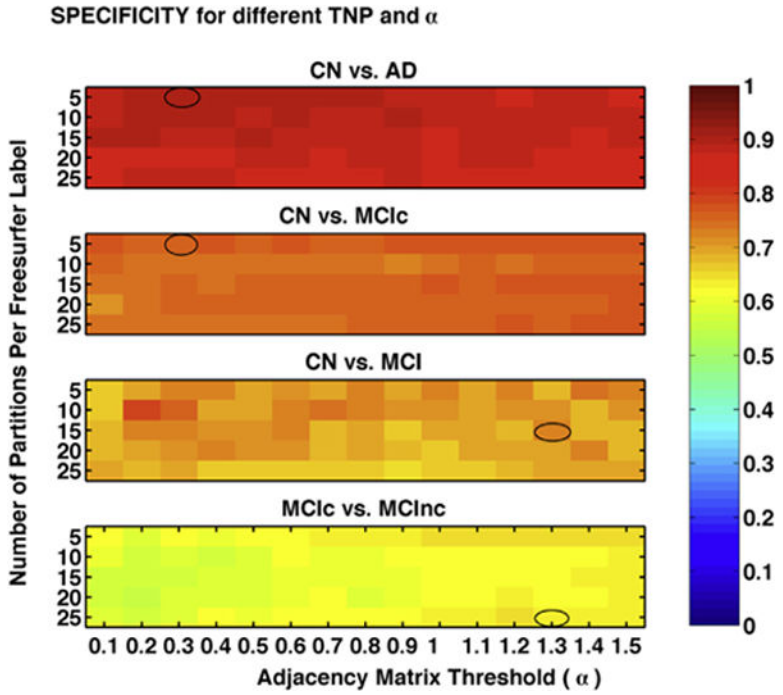

Fig. 9.

Comparison of specificity, as a function of TNP and α, obtained from RHsT method. The combination with the best performance (highest AUC) in each experiment is highlighted with a black oval. Abbreviations: AD, Alzheimer's disease; AUC, area under curve; MCI, mild cognitive impairment; MCIc, mild cognitive impairment converters; MCInc, mild cognitive impairment nonconverters; RHsT, repeated hold-out, stratified training set; TNP, total number of partitions. (For interpretation of the references to color in this Figure, the reader is referred to the web version of this article.)

Table 1

Best performance (highest AUC) of the ThickNet fusion method for each experiment, in various classification metrics (with their standard deviation [SD] from the 100 repetitions of RHsT) describing the performance

Experiment	AUC (SD)	ACC (SD)	SENS (SD)	SPEC (SD)	TNP	α
CN versus AD	0.92 (0.06)	0.89 (0.06)	0.80 (0.16)	0.90 (0.06)	340	0.30
CN versus MCIc	0.83 (0.15)	0.76 (0.04)	0.74 (0.32)	0.76 (0.04)	340	0.30
CN versus MCI	0.75 (0.10)	0.65 (0.07)	0.64 (0.08)	0.73 (0.17)	1020	1.30
MCIc versus MCInc	0.68 (0.21)	0.64 (0.06)	0.65 (0.36)	0.64 (0.06)	1700	1.30

The optimal TNP and threshold (α) is noted for each experiment.

Key: AD, Alzheimer's disease; AUC, area under curve; CN, controls; MCI, mild cognitive impairment; MCIc, mild cognitive impairment converters; MCInc, mild cognitive impairment nonconverters; RHsT, repeated hold-out, stratified training set; SD, standard deviation; SENS, sensitivity; SPEC, specificity; TNP, total number of partitions.

Table 2

Comparison of the performance of our novel ThickNet features (shortened as TN in this table for lack of space) and mean thickness (MT) features

Experiment	Best TNP and α		AUC		Partial AUC	
	MT	TN	MT	TN	MT	TN
CN versus AD	680	340, 0.30	0.916 (0.08)	0.924 (0.06)	0.090 (0.024)	0.097 (0.027)
CN versus MCIc	340	340, 0.30	0.807 (0.15)	0.832 (0.15)	0.057 (0.036)	0.068 (0.043)
CN versus MCI	1020	1020, 1.30	0.758 (0.08)	0.748 (0.10)	0.046 (0.025)	0.057 (0.029)
MCIc versus MCInc	680	1700, 1.30	0.670 (0.16)	0.683 (0.21)	0.029 (0.025)	0.035 (0.038)

Performance is reported in the form of mean (SD).

Key: AD, Alzheimer's disease; AUC, area under curve; CN, controls; MCI, mild cognitive impairment; MCIc, mild cognitive impairment converters; MCInc, mild cognitive impairment nonconverters; SD, standard deviation; TNP, total number of partitions.

Neurobiol Aging. Author manuscript; available in PMC 2018 March 13.

Table 3

Descriptive statistics from the comparison of the performance of ThickNet features and mean thickness features. Here, MT and TN are arrays of size 100 × 1 with AUCs (full or partial) for the 100 repetitions of holdout method

Experiment	Full ROC		Partial ROC	
	Mean(MT-TN)	SD(MT-TN)	Mean(MT-TN)	SD(MT-TN)
CN versus AD	−0.0082	0.0889	0.0002	0.0327
CN versus MCIc	−0.0248	0.2060	−0.0115	0.0568
CN versus MCI	0.0095	0.1165	−0.0107	0.0322
MCIc versus MCInc	−0.0134	0.2729	−0.0063	0.0456

Key: AD, Alzheimer's disease; AUC, area under curve; CN, controls; MCI, mild cognitive impairment; MCIc, mild cognitive impairment converters; MCInc, mild cognitive impairment nonconverters; MT, mean thickness; ROC, receiver operating characteristic; SD, standard deviation.

Table 4

List of identifiers of the subjects excluded from the analysis in this study

1070	118	1215	1218	1247	128	1394	222	240	344	388	41	42
461	555	604	695	723	725	750	843	855	856	860	887	954

Note, these are baseline subjects from ADNI 1. Subjects were excluded owing to their failure in either freesurfer cortical parcellation or estimation of cortical thickness from our laplacian streamlines method. Our quality control consisted of checking for permanent failure in Freesurfer automatic parcellation, visually examining for the presence of holes or handles in the pial or white surfaces (left or right hemisphere), or when the cortical surfaces have gross errors in following the structural boundaries. Furthermore, even with acceptable Freesurfer parcellation, some subjects were excluded if our thickness computation method based on Laplacian streamlines fails to estimate thickness in either left or the right hemisphere.

Key: ADNI, Alzheimer's Disease Neuroimaging Initiative.

Table 5

List of identifiers of the subjects used for the estimation of average atlas

295	1261	1280	907	981	602	484	498	681	68	1222	2	5	16	21	22	23	502	575	1035
519	520	558	883	899	1288	14	130	61	963	985	1063	403	824	843	920	1098	48	672	813
1023	1002	643	779	934	19	312	386	363	640	489	711	171	896	533	1014	173	601	405	506
605	680	260	622	863	232	969	1200	123	319	441	433	488	86	196	1194	1195	1197	1202	1203

Note these are baseline subjects from ADNI-1.

Key: ADNI, Alzheimer's Disease Neuroimaging Initiative.