Gaussian Discriminant Analysis for Optimal Delineation of Mild Cognitive Impairment in Alzheimer's Disease

Article *in* International Journal of Neural Systems · April 2018

DOI: 10.1142/S012906571850017X

International Journal of Neural Systems

GAUSSIAN DISCRIMINANT ANALYSIS FOR OPTIMAL DELINEATION OF MILD COGNITIVE IMPAIRMENT IN ALZHEIMER'S DISEASE

CHEN FANG[1], CHUNFEI LI[1], MERCEDES CABRERIZO[1], ARMANDO BARRETO[1], JEAN ANDRIAN[1], NAPHTALI RISHE[1], DAVID LOEWENSTEIN[2, 3, 4], RANJAN DUARA[2, 4, 5], MALEK ADJOUADI[1, 4, 5, *], FOR THE ALZHEIMER'S DISEASE NEUROIMAGING INITIATIVE[†]

[1]Center for Advanced Technology and Education (CATE), Florida International University, 10555 W Flagler St Miami, Florida 33174, USA
[2]Wien Center for Alzheimer's Disease & Memory Disorders, Mount Sinai Medical Center Miami Beach, Florida 33140, USA
[3] Department of Psychiatry & Behavioral Sciences, Miller School of Medicine, University of Miami Miami, Florida 33136, USA
[4]Florida Alzheimer's Disease Research Center (ADRC), University of Florida Gainesville, Florida 32610, USA
[5]Herbert Wertheim College of Medicine, Florida International University Miami, Florida 33174, USA
*E-mail: cfang002@fiu.edu, cli029@fiu.edu, mcabreriz@fiu.edu, barretoa@fiu.edu, andrianj@fiu.edu, nrishe@fiu.edu, DLoewenstein@med.miami.edu, ranjan.duara@msmc.com and *adjouadi@fiu.edu*

Over the past few years, several approaches have been proposed to assist in the early diagnosis of Alzheimer's disease (AD) and its prodromal stage of mild cognitive impairment (MCI). Using multimodal biomarkers for this high-dimensional classification problem, the widely used algorithms include Support Vector Machines (SVM), Sparse Representation-based classification (SRC), Deep Belief Networks (DBN), and Random Forest (RF). These widely used algorithms continue to yield unsatisfactory performance for delineating the MCI participants from the cognitively normal control (CN) group. A novel Gaussian discriminant analysis-based algorithm is thus introduced to achieve a more effective and accurate classification performance than the aforementioned state-of-the-art algorithms. This study makes use of magnetic resonance imaging (MRI) data uniquely as input to two separate high-dimensional decision spaces that reflect the structural measures of the two brain hemispheres. The data used include 190 CN, 305 MCI, and 133 AD subjects as part of the AD Big Data DREAM Challenge #1. Using 80% data for a tenfold cross validation, the proposed algorithm achieved an average F1 score of 95.89%, and an accuracy of 96.54% for discriminating AD from CN; and more importantly, an average F1 score of 92.08% and an accuracy of 90.26% for discriminating MCI from CN. Then a true test was implemented on the remaining 20% held-out test data. For discriminating MCI from CN, an accuracy of 80.61%, a sensitivity of 81.97%, and a specificity of 78.38% were obtained. These results show significant improvement over existing algorithms for discriminating the subtle differences between MCI participants and the CN group.

Keywords: Gaussian discriminant analysis, mild cognitive impairment, Alzheimer's disease, machine learning, classification algorithms, and computer aided diagnosis.

[*]Corresponding author

[†]Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.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.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

1. Introduction

According to the National Institute on Aging (NIA), before memory loss and other cognitive impairments can be observed as evidence for Alzheimer's Disease (AD), subtle changes to the brain have already started for a decade or more.[1] Although there still is no known cure for the disease, alleviation of specific symptoms is possible through treatment for some patients in the early or middle stages of AD.[2] Thus, accurate diagnosis of its prodromal stage, mild cognitive impairment (MCI), with a high risk to convert to AD, is essentially important as means to facilitate planning for early intervention and treatment.[3]

Multiple modalities of biomarkers have been found to be significantly sensitive in assessing the progression of AD. These include structural magnetic resonance imaging (MRI),[4–10] positron emission tomography (PET),[9–13] cerebrospinal fluid (CSF),[6,13,14] electroencephalographic (EEG) rhythms,[15–26] and Magnetoencephalography (MEG).[27,28] Using these modalities of biomarkers and taking advantage of advances made in the development of machine learning and deep learning algorithms over the past few years, several approaches have been proposed to assist in the early diagnosis of MCI.[7–9,11–13,29–31] Since no matter which modality or modalities of biomarkers are used, there will always be multiple variables for predicting the progression of the disease, which ultimately can be generalized as a high-dimensional classification problem.

Currently, many machine learning and deep learning algorithms capable of dealing with high-dimensional data have been applied to classification and regression analysis in the context of disease diagnosis and transition predictions. The more notable of these types of algorithms are Random Forest (RF), Support Vector Machines (SVM), Sparse Representation-based classification (SRC), and Deep Belief Networks (DBN).[4,7–9,11–13,29–31] Among these state-of-the-art algorithms, SVM continues to be one of the most widely used for the classification of AD and its prodromal stages. But SVM still faces serious challenges, especially in the selection of the kernel function parameters for nonlinear problems, even under the so-called kernel trick, which remain essentially difficult to overcome in view of the high variance in the main features that define the disease. In particular, for discriminating MCI from elderly cognitively normal control group (CN), the classification performance of SVM remains insufficient, ranging between 79% and 83% in accuracy, and the sensitivity is substantially lower than that for AD vs. CN (the easiest two groups to separate) and even not significantly better than chance.[4,8,9,29] Although many of the state-of-the-art strategies and techniques continue to advance our understanding of AD, there remain many challenges in the different experimental stages at determining more conclusive evidence for the accurate diagnosis and classification of AD, as expressed in studies.[30,31]

As a way to overcome such challenges, this study develops a machine learning classification algorithm based on the Gaussian discriminant analysis (GDA), introducing the use of dual decisional spaces, one for each hemisphere. Among those modalities, structural MRI is currently widely used for analyzing the gradual progression of atrophy patterns in key brain regions,[32] therefore, this study makes use of structural MRI as the unique input. To the best of our knowledge, this study is the first to apply GDA to the diagnosis of CN vs. MCI, with the CN vs. AD classification results included here only for comparative purposes.

2. Methodology

Several software pipelines are used to preprocess the raw MRI data as a first step. After the pre-processing step, morphometric (shape) data are derived from the images, including shape measures of all 25 labeled cortical regions. Then a noise detection procedure and a feature selection method based on the analysis of variance (ANOVA) are deployed to determine the statistical significance of each variable in the classification outcome. Then, a GDA-based classifier applied on the dual decision spaces is proposed for solving the boundaries between any two different groups of subjects (i.e., CN vs. MCI, CN vs. AD, and MCI vs. AD). The general framework of the proposed algorithm is illustrated in Fig. 1.

2.1. *Subjects*

The data used in preparation of this study were obtained from the Alzheimer’s disease Neuroimaging Initiative (ADNI) database, as part of the ADNI1: Complete 1Yr 1.5T collection and their assessments at baseline, which includes 628 individuals (190 CN, 305 MCI, and 133 AD).[33] The ADNI was launched in 2003 as a public-

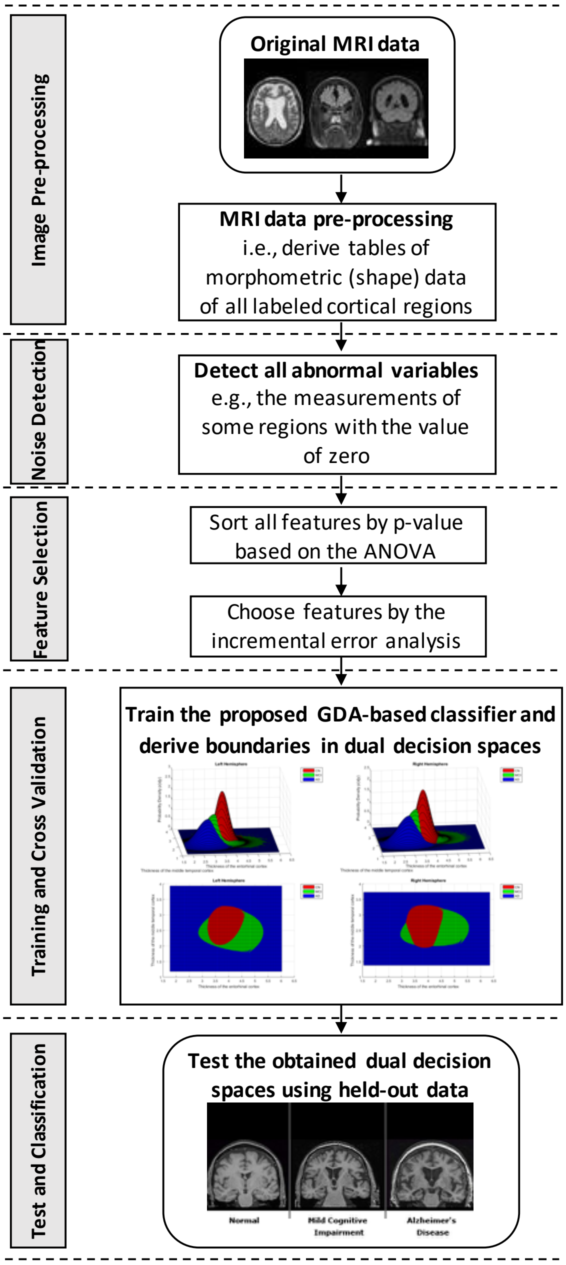


Fig. 1. General framework of the GDA-based dual high-dimensional decision spaces.

private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. The primary phenotype is diagnostic group and MMSE. All source imaging data consisted of 1.5 Tesla T1-weighted MRI volumes in the NIfTI (.nii.gz) format. Summary statistics and patient counts are listed in Table 1.

Table 1. Summary statistics of subjects.

Group	Patient Numbers	Mean ± SD			Male %	Female %
		MMSE	*Age*	*Years of Edu.*		
CN	190	29.1 ± 1.0	75.9 ± 5.1	16.1 ± 2.7	51.6	48.4
MCI	305	27.0 ± 1.8	74.9 ± 7.1	15.7 ± 3.0	64.9	35.1
AD	133	23.5 ± 1.9	74.8 ± 7.6	14.7 ± 3.1	51.1	48.9
Total	628	26.9 ± 2.6	75.2 ± 6.7	15.6 ± 3.0	58.0	42.0

CN: cognitively normal control, MCI: mild cognitive impairment, AD: Alzheimer's disease, MMSE: Mini–Mental State Examination, SD: standard deviation

2.2. *MRI data pre-processing*

Using three neuroimaging software pipelines: FreeSurfer,[34] Advanced Normalization Tools (ANTs),[35] and Mindboggle,[36] the original MRI data were preprocessed following the instruction provided by Alzheimer's Disease Big Data DREAM Challenge #1.[37] Tables of morphometric data were derived from the images using the following seven shape measures for all 25 FreeSurfer labeled cortical regions for both left and right hemispheres of the brain: 1) surface area; 2) travel depth; 3) geodesic depth; 4) mean curvature; 5) convexity; 6) thickness; and 7) volume. FreeSurfer pipeline (version 5.3) was applied to all T1-weighted images to generate labeled cortical surfaces, and labeled cortical and non-cortical volumes. Templates and atlases used by ANTs and Mindboggle can be found on the Mindboggle website.[38]

2.3. *Noise detection*

The aforementioned preprocessed MRI data of the 25 labeled cortical regions were used to generate two 175-variable (7×25) vector discriminators, for each subject (i.e., one 175-variable vector per hemisphere). This study reveals that separating the variables for each hemisphere of the brain yields a better classification performance than processing all features together, with details in support of this assertion provided in the Results Section. As for the few subjects whose vector discriminator involved atypical variables, for example, some regions having measurements of some areas to be zero, these subjects were removed from further investigation.

2.4. *Feature selection*

By the final stage of AD, brain tissue has atrophied significantly, so all shape measures mentioned above could have changed as well. Some of the subtle changes initially appear to take place in some specific areas of the brain, so determination of the key changed regions of interest (ROIs) can help to discriminate more specifically MCI from CN.

2.4.1. *ANOVA ranking*

An ANOVA was carried out on each of the 175 variables of the two vectors between any two groups (i.e., CN vs. MCI, CN vs. AD, and MCI vs. AD) to determine the significance of each variable in terms of classification outcome, and all variables were thereafter ranked according to their p-values. It should be noted that in the feature selection procedure, the Shapiro-Wilk test was employed for testing the normality of the shape measures and the average *p-value* is 0.28, which indicates that the data are from a normally distributed population.[39] Furthermore, equal weights are assigned to each of the shape measures so as to eliminate any bias.

2.4.2. *Incremental error analysis*

In order to maintain only few key variables and still ensure good classification performance, an incremental error analysis was performed to determine how many of the top-ranked variables ought to be included in the classifier.[3] In the initial phase, the proposed GDA-based classifier only uses the first-ranked variable. The error analysis was employed whereby introducing the next top-ranked variable in the classifier at each subsequent phase, and recording the corresponding classification statistics (i.e., F1 score, accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV)), which then would be compared with the previous phase. When the performance in terms of its classification statistics can no longer be improved, the optimal set of variables would have been obtained.

2.5. *GDA-based classifier*

Since there may be as many as 175 variables to be taken into consideration, the classifier must be able to resolve this high-dimensional classification problem. For this reason, and by using GDA, an important supervised machine learning algorithm for such classification problems, the proposed classifier is able to solve the boundaries between any two groups (i.e., CN vs. MCI, CN vs. AD, and MCI vs. AD). The proposed classification problem can then be formalized by having the machine learn to distinguish among CN ($y=0$), MCI ($y=1$), and AD ($y=2$), based on the selected features $\vec{x}\in\mathbf{R}^n$. Then, given a training set, the proposed algorithm can model $p(\vec{x}\mid y)$, the condition distribution of the n-dimensional vector $\vec{x}$ given $y\in\{0,1,2\}$, assumed to be distributed according to a multivariate Gaussian distribution (or multivariate normal distribution), whose density function is given by (1) as below:

$$p(\vec{x};\vec{\mu},\Sigma)=\frac{1}{\sqrt{(2\pi)^n\,|\Sigma|}}e^{-\frac{1}{2}(\vec{x}-\vec{\mu})^T\Sigma^{-1}(\vec{x}-\vec{\mu})} \tag{1}$$

where $\vec{\mu}\in\mathbf{R}^n$ is the mean vector, $\Sigma\in\mathbf{R}^{n\times n}$ is the covariance matrix, the same as the one used in other regression analysis methods (e.g., the principal component analysis), and $|\Sigma|$ and Σ^{-1} denote the determinant and inverse matrix of Σ , respectively. Note that n is the dimension of vector $\vec{x}$, i.e., the number of features included in the classifier. After modeling $p(\vec{x}\mid y)$, the proposed algorithm uses Bayes rule to derive subsequent distribution on y given $\vec{x}$ as follows:

$$p(y\mid\vec{x})=\frac{p(\vec{x}\mid y)\,p(y)}{p(\vec{x})} \tag{2}$$

Here, $p(y)$ is the class prior distribution, which could not be determined when given a certain subject, so it is assumed to be absolutely random (i.e., for all $i\neq j$, $p(y=i)=p(y=j)$). Furthermore, in order to make a prediction, it is not necessary to calculate the denominator $p(\vec{x})$, since

$$\begin{aligned}\arg\max_y p(y\mid\vec{x}) &= \arg\max_y \frac{p(\vec{x}\mid y)\,p(y)}{p(\vec{x})}\\ &= \arg\max_y p(\vec{x}\mid y)\,p(y)\end{aligned} \tag{3}$$

Therefore, for the purpose of classification, it only needs

$$\arg\max_y p(y\mid\vec{x})=\arg\max_y p(\vec{x}\mid y) \tag{4}$$

The classifier was applied to each hemisphere of the brain (using the two 175-variable vectors), and if either one of the two sides had been classified to be positive, the corresponding subject should be positive as well.

The performance of the proposed classifier was measured using the F1 score, accuracy, sensitivity, specificity, PPV, and NPV based on a tenfold cross validation process. For selecting the optimal set of variables, 80% of the noise-free detected subjects' data was used as the training set in a tenfold cross validation process, which were randomly assigned to ten subsets

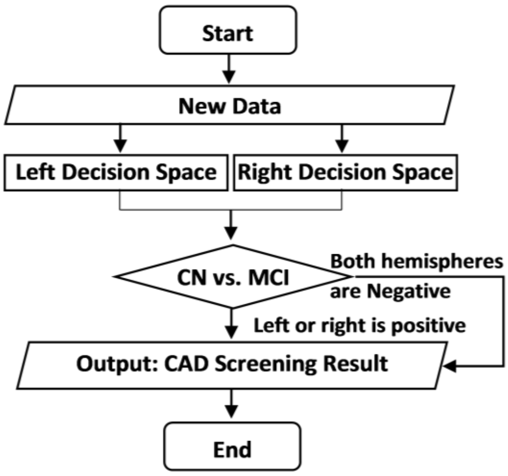


Fig. 2. Flowchart of the GDA-based dual decision space classification.

d_0, d_1, …, d_9, so that all subsets were of equal size. Then one of each of the ten sets was retained as the validation dataset, while the remaining nine datasets were used as training data; thus, every data point was used for both training and validation on each fold. Once the optimal set of variables was generated, the classification performance was evaluated by using the remaining 20% of the noise-free detected subject data points as the held-out test set.

Table 2. Number of significant variables selected for each comparison.

Groups	CN vs. MCI	CN vs. AD	MCI vs. AD
Side of brain	Number of significant variables (p-value < 0.01)		
Left	50	79	51
Right	44	68	41

CN: cognitively normal control, MCI: mild cognitive impairment, AD: Alzheimer's disease

As demonstrated in Fig. 2, data of the left and right hemispheres of the brain were processed separately, which means, for the final classification, each hemisphere had its own decision space, and as long as one decision space produces the positive result (i.e., MCI in CN vs. MCI, AD in CN vs. AD, and AD in MCI vs. AD), the tested subject is classified as such. This innovative process resulted in a significant improvement of the classification performance as demonstrated in the Results section, especially for the most challenging classification of CN vs. MCI.

3. Results

In this section, the experimental results of the feature selection process reveal the significance of different ROIs in patients for the three classification types: 1) CN vs. MCI; 2) CN vs. AD; and 3) MCI vs. AD. Evolution in the statistics during the incremental error analysis and

Table 3. Top-10 significant variables for each comparison.

Groups		CN vs. MCI		CN vs. AD		MCI vs. AD	
Side of brain	Rank	Measurements	*p-value*	Measurements	*p-value*	Measurements	*p-value*
Left	1	Thickness of entorhinal	$< 10^{-17}$	Thickness of entorhinal	$< 10^{-36}$	Thickness of inferior parietal	$< 10^{-8}$
	2	Curvature of entorhinal	$< 10^{-13}$	Curvature of entorhinal	$< 10^{-26}$	Thickness of entorhinal	$< 10^{-7}$
	3	Thickness of middle temporal	$< 10^{-12}$	Thickness of middle temporal	$< 10^{-24}$	Thickness of middle temporal	$< 10^{-6}$
	4	Thickness of inferior temporal	$< 10^{-10}$	Thickness of inferior temporal	$< 10^{-22}$	Thickness of inferior temporal	$< 10^{-6}$
	5	Curvature of middle temporal	$< 10^{-9}$	Curvature of middle temporal	$< 10^{-21}$	Volume of inferior parietal	$< 10^{-6}$
	6	Thickness of fusiform	$< 10^{-8}$	Thickness of inferior parietal	$< 10^{-18}$	Curvature of middle temporal	$< 10^{-6}$
	7	Curvature of insula	$< 10^{-8}$	Curvature of inferior temporal	$< 10^{-16}$	Curvature of inferior parietal	$< 10^{-6}$
	8	Curvature of parahippocampal	$< 10^{-8}$	Thickness of fusiform	$< 10^{-16}$	Volume of inferior temporal	$< 10^{-6}$
	9	Curvature of inferior temporal	$< 10^{-8}$	Curvature of parahippocampal	$< 10^{-15}$	Curvature of entorhinal	$< 10^{-5}$
	10	Thickness of superior temporal	$< 10^{-6}$	Curvature of inferior parietal	$< 10^{-14}$	Volume of middle temporal	$< 10^{-5}$
Right	1	Thickness of entorhinal	$< 10^{-13}$	Thickness of entorhinal	$< 10^{-31}$	Curvature of middle temporal	$< 10^{-8}$
	2	Thickness of middle temporal	$< 10^{-10}$	Thickness of middle temporal	$< 10^{-25}$	Thickness of entorhinal	$< 10^{-8}$
	3	Thickness of fusiform	$< 10^{-10}$	Curvature of middle temporal	$< 10^{-22}$	Thickness of middle temporal	$< 10^{-7}$
	4	Curvature of middle temporal	$< 10^{-8}$	Curvature of entorhinal	$< 10^{-21}$	Volume of inferior parietal	$< 10^{-6}$
	5	Curvature of superior temporal	$< 10^{-8}$	Thickness of inferior parietal	$< 10^{-18}$	Curvature of inferior temporal	$< 10^{-6}$
	6	Curvature of entorhinal	$< 10^{-8}$	Curvature of inferior temporal	$< 10^{-18}$	Curvature of entorhinal	$< 10^{-6}$
	7	Curvature of fusiform	$< 10^{-8}$	Thickness of inferior temporal	$< 10^{-17}$	Thickness of inferior parietal	$< 10^{-6}$
	8	Thickness of superior frontal	$< 10^{-7}$	Thickness of fusiform	$< 10^{-17}$	Volume of middle temporal	$< 10^{-5}$
	9	Thickness of inferior temporal	$< 10^{-7}$	Curvature of inferior parietal	$< 10^{-15}$	Volume of inferior temporal	$< 10^{-5}$
	10	Curvature of superior frontal	$< 10^{-7}$	Curvature of fusiform	$< 10^{-15}$	Thickness of inferior temporal	$< 10^{-5}$

CN: cognitively normal control, MCI: mild cognitive impairment, AD: Alzheimer's disease the classification performance of the GDA-based algorithm using the proposed dual decisional spaces are provided.

3.1. *Ranking of the variables*

After the noise detection process was applied, 9 subjects were removed because of the noisy data, which included measurements with zero values, so the final data used in the classification experiment included 619 individuals, among them, 187 CN, 301 MCI, and 131 AD. As mentioned earlier, ANOVA was performed for CN vs. MCI, CN vs. AD, and MCI vs. AD using two 175-variable vectors corresponding to the left and right hemispheres of the brain. For each group, all variables found at 0.01 level of significance (LOS) out of all 175 variables for each side of the brain (i.e., those variables with p-values less than 0.01) were used for the classification as shown in Table 2.

The top 10 ranked variables and their corresponding measurements are given in Table 3, where it can be observed that the entorhinal cortex is the most significant (first-ranked) cortical region for discriminating either MCI or AD from CN. This observation is consistent with recent studies indicating that indeed the entorhinal cortex is the first area to be implicated in AD,[40–42] providing credence to the validity of our feature selection method. Moreover, the entorhinal cortex has been proven to be a major source of projections to the hippocampus,[43] which plays an important role in converting short-term memory (also known as working memory) to long-term memory. Interestingly, for discriminating MCI from AD, the entorhinal cortex is relegated to the second top-ranked region.

Although the hippocampus area does not appear to be of higher significance than the entorhinal cortex in the feature selection process, it could still serve as an explanation for the symptom of AD in that short-term memory loss occurs earlier than long-term memory loss. Since, at the very beginning, direct connections to the hippocampus seem to have been affected. The second top-ranked cortical region, the middle temporal, is also

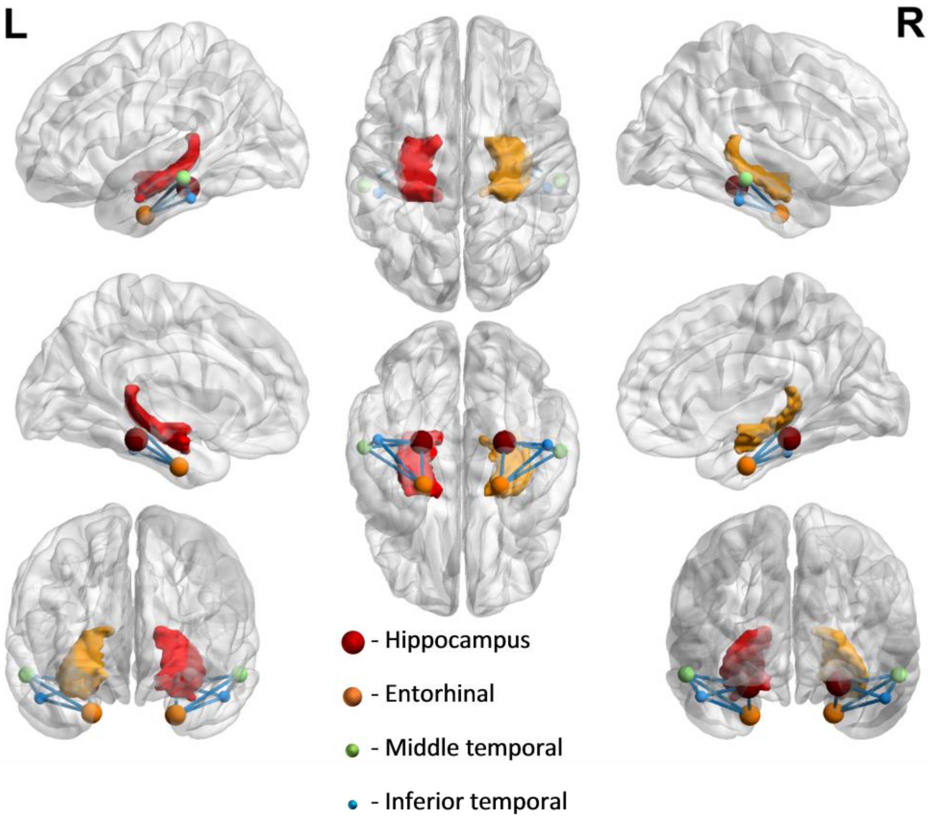


Fig. 3. Relative location of hippocampus and the top-three-ranked cortical regions (visualized with the BrainNet Viewer[45])

critical for long-term memory, to which the disrupted hippocampal connectivity has been found in the early stages of AD.[44] Moreover, in the human brain, all top three-ranked cortical regions, including the entorhinal, the middle temporal, and the inferior temporal are very close to the hippocampus as shown in Fig. 3. From Table 3, it also can be observed that for discriminating between MCI and AD, the significant variables are now much different to others. Hence, for the tenfold cross validation and the incremental error analysis, all three classifications were trained and validated separately in order to achieve the best performance.

3.2. *Optimal sets of variables*

To generate the optimal set of variables, in the tenfold cross validation, the aforementioned 80% of the noise-free detected data points included 500 individuals (150 CN, 240 MCI, and 110 AD), where all numbers were rounded to the nearest number divisible by 10 for the tenfold cross validation of the noise-free detected subjects included in this study (i.e., 619 = 187 CN + 301 MCI + 131 AD).

The purpose of applying the incremental error analysis was to obtain the best classification performance with the optimal number of variables (i.e., the number of dimensions in the decisional spaces). For each hemisphere, some classification statistics are illustrated in Fig. 4, where the horizontal axis indicates the number of significant variables included in each iteration.

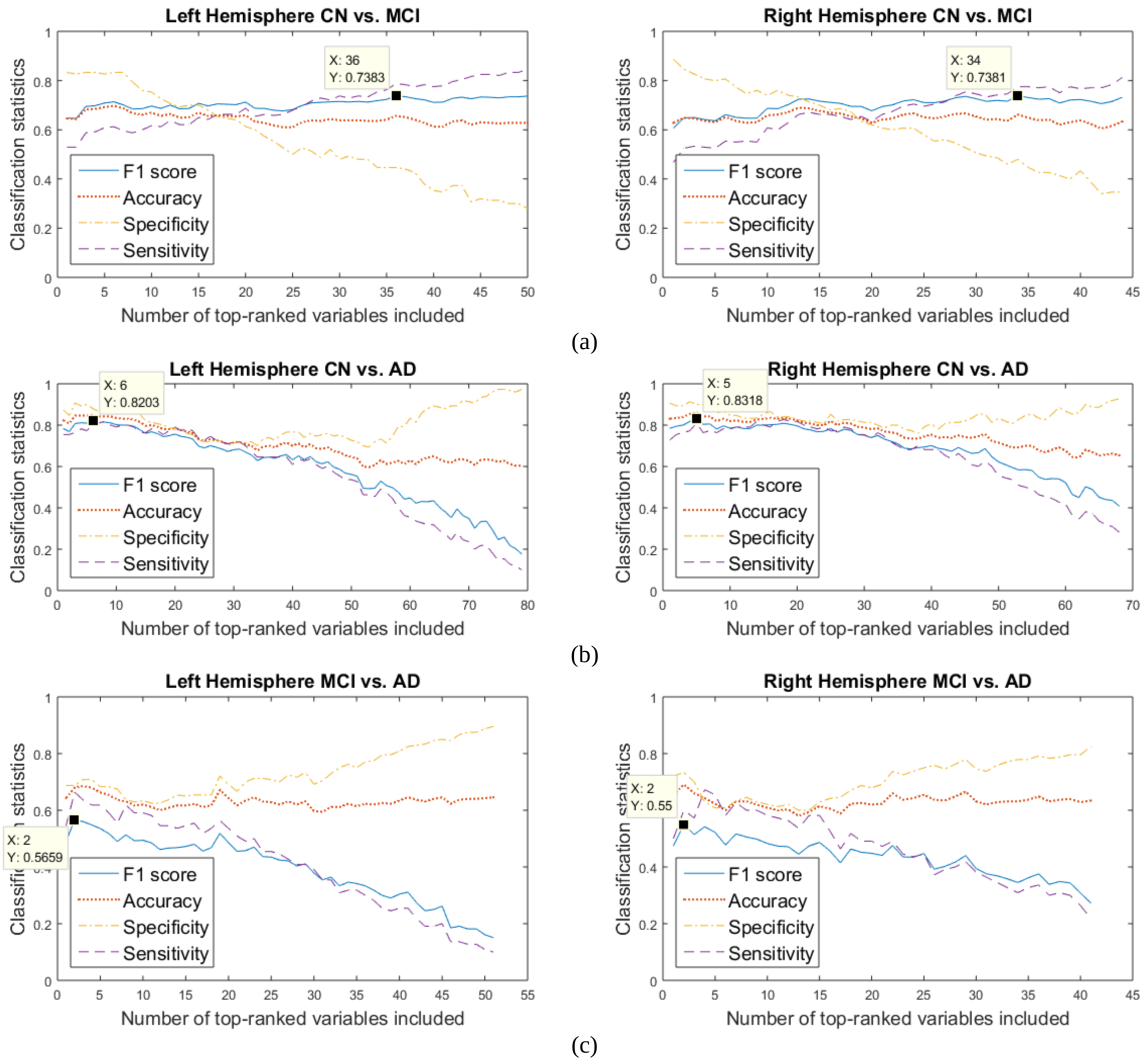
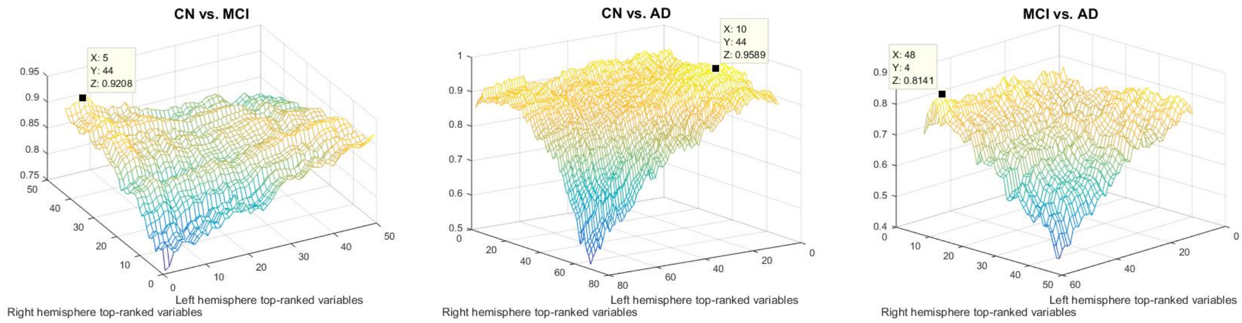


Fig. 4. Incremental error analysis performance of classification statistics (a) CN vs. MCI, (b) CN vs. AD, and (c) MCI vs. AD.



(a) (b) (c)

Fig. 5. Incremental error analysis performance of F1 score (a) CN vs. MCI, (b) CN vs. AD, and (c) MCI vs. AD.

In the tenfold cross validation and in the subsequent true test, four important parameters were computed, including the number of True Positives (TP) (i.e., the correctly classified positive subjects), the number of True Negatives (TN) (i.e., the correctly classified negative subjects), the number of False Positive (FP) (i.e., the negative subjects incorrectly classified as positive), and the number of False Negative (FN) (i.e., the positive subjects incorrectly classified as negative). For evaluating the classification performance, the following commonly used measures are computed for determining accuracy (5), sensitivity (6), specificity (7), positive predictive value - PPV (8), and negative predictive value - NPV (9):

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{5}$$

$$Sensitivity = \frac{TP}{TP+FN} \tag{6}$$

$$Specificity = \frac{TN}{TN+FP} \tag{7}$$

$$PPV = \frac{TP}{TP+FP} \tag{8}$$

$$NPV = \frac{TN}{TN+FN} \tag{9}$$

But due to the effect of imbalanced data, and as a clinical application, the classification performance was not only measured by the accuracy, which actually relies more on the sensitivity or recall and the PPV or precision, but also by using the F1 score, as expressed below, in order to select the optimal sets of variables,

$$F1 = \frac{2 \times Sensitivity \times PPV}{Sensitivity + PPV} = \frac{2TP}{2TP+FP+FN} \tag{10}$$

As the harmonic mean of sensitivity and PPV, the F1 score or balanced F-score is the widely used measure of performance in statistical analysis of binary classification. For the incremental error analysis (IEA), the set with the highest F1 score was selected, when several sets had the same F1 score, the one with the highest accuracy was chosen, then if still multiple choices were found, the one having the minimum size was finally selected.

As demonstrated in Fig. 4(a), for either one of the two hemispheres of the brain, the classification performance for CN vs. MCI yielded better than the average results obtained from other studies reported in Ref. 4, where the sensitivity is 78.75% and 77.50% for each decision space (i.e., each hemisphere), respectively. After combining the results of the two decisional spaces together and implementing the incremental error analysis again, the evolution of the F1 score is as illustrated in Fig. 5.

It can be observed that the final optimal sets are different from the ones obtained for each hemisphere before combining the two decision spaces together. For all comparisons, the performance was improved significantly as shown in Table 4. Moreover, for the most difficult two groups to delineate, CN vs. MCI, significant enhancements in classification statistics were achieved, including an increase in F1 score average from 73.82% to 92.08% and increments of 24.37% for accuracy, 13.96% for sensitivity, 41.00% for specificity, 22.11% for PPV, and 30.37% for NPV, respectively. Comparatively to the more recently reported cross validation performances of some of the state-of-the-art-

Table 4. Summary of tenfold cross validation performance improved after combining the dual decision spaces.

Groups	CN vs. MCI			CN vs. AD			MCI vs. AD		
Decision Space	Left	Right	Comb.	Left	Right	Comb.	Left	Right	Comb.
F1 %	73.83	73.81	92.08	82.03	83.18	95.89	56.59	55.00	81.41
ACC %	65.64	66.15	90.26	85.00	86.15	96.54	68.00	69.14	89.43
SEN %	78.75	77.50	92.08	80.91	80.91	95.45	66.36	60.00	73.64
SPE %	44.67	48.00	87.33	88.00	90.00	97.33	68.75	73.33	96.67
PPV %	69.49	70.45	92.08	83.18	85.58	96.33	49.32	50.77	91.01
NPV %	56.78	57.14	87.33	86.27	86.54	96.69	81.68	80.00	88.89
Number of the optimal variables	36	34	L: 5 R: 44	6	5	L: 10 R: 44	2	2	L: 48 R: 4

Table 5. Comparison of cross validation performance with some recent studies.

Groups Reference	Modalities	Classifier	Source of Data (CN+MCI+AD)	CN vs. MCI			CN vs. AD			MCI vs. AD		
				ACC %	SEN %	SPE %	ACC %	SEN %	SPE %	ACC %	SEN %	SPE %
M. Liu et al.,[7] 2012	MRI	SRC	ADNI (229 + 225 + 198)	87.85	85.26	**90.40**	90.80	86.32	94.76	-	-	-
L. Khedher et al.,[8] 2015	MRI	SVM	ADNI (229 + 401 + 188)	81.89	82.61	81.62	88.49	91.27	85.11	85.41	**87.03**	83.78
T. Ye et al.,[9] 2016	MRI + PET	SVM	ADNI (52 + 99 + 51)	82.13	87.68	71.54	95.92	94.71	97.12	-	-	-
T. Tong et al.,[13] 2017	MRI + PET + CSF + Genetic	RF	ADNI (35+75+37)	79.50	85.10	67.10	91.80	88.90	94.70			
L. Khedher et al.,[29] 2017	MRI	SVM	ADNI (-)	79.00	82.00	76.00	89.00	92.00	86.00	85.00	85.00	86.00
A. Ortiz et al.,[46] 2017	MRI	DBN	ADNI (-)	83.00	-	-	90.00	-	-	84.00	-	-
Proposed Study	MRI	GDA	ADNI (190 + 305 + 133)	**90.26**	**92.08**	87.33	**96.54**	**95.45**	**97.33**	**89.64**	84.29	**90.67**

CN: cognitively normal control, MCI: mild cognitive impairment, AD: Alzheimer's disease, Comb.: combining left and right, ACC: accuracy, SEN: sensitivity, SPE: specificity, PPV: positive predictive value, NPV: negative predictive value

approaches,[7–9,12,13,29,46] the proposed study achieves remarkable improvements in performance, especially in delineating MCI from CN, even when MRI is the only modality used for this study. As shown in Table 5, except for the specificity of the CN vs. MCI classification and the sensitivity of the MCI vs. AD classification, the proposed method yielded the best cross validation performance in a comparative assessment to all other methods.

3.3. Classification performance

In order to obtain a reliable measure of the classification performance, the remaining 20% of the noise-free detected data points were used as the held-out test data (37 CN, 61 MCI, and 21 AD) using the obtained optimal sets of variables. The results are presented in Table 6. Although the classification performance was not as good as that obtained in the tenfold cross validation, the results are still better than state-of-the-art-algorithm reviewed in Ref. 4 and the recently proposed state-of-the-art approach in Ref. 12, as shown in Table 7. Since not all studies implemented the held-out data true test, only the results from Ref. 4 and Ref. 12 were considered for comparison to our proposed method, which also used ADNI data. For discriminating AD from CN, the proposed GDA-based algorithm achieved an accuracy of 93.10%, sensitivity of 90.48%, specificity of 94.59%, PPV of 90.48%, and NPV of 94.59%; these results for these two groups were expected. But more importantly, an accuracy of 80.61%, sensitivity of 81.97%, specificity of 78.38%, PPV of 86.21%, and NPV of 72.50% were obtained for discriminating MCI from CN; results that are considered as the best classification performance obtained so far using the GDA method.

4. Discussion

The merits of the proposed GDA-based dual decision space algorithm are not only reflected through the good classification performance it achieved, but also in the strategic way it looked at the two hemispheres of the brain separately. The classification was performed using

Table 6. Summary of the proposed GDA-based dual high-dimensional decision spaces classification performance.

Groups	CN vs. MCI			CN vs. AD			MCI vs. AD		
Decision Space	Left	Right	Comb.	Left	Right	Comb.	Left	Right	Comb.
ACC %	55.10	52.04	80.61	75.86	70.69	93.10	65.85	67.07	85.37
SEN %	73.77	65.57	81.97	71.43	66.67	90.48	42.86	38.10	52.38
SPE %	24.32	29.73	78.38	78.38	72.94	94.59	73.77	77.05	96.72
PPV %	61.64	60.61	86.21	65.22	58.33	90.48	36.00	36.36	84.62
NPV %	36.00	34.38	72.50	82.86	79.41	94.59	78.95	78.33	85.51

Table 7. Comparison of classification performance with other studies using true test with held-out data.

Groups Reference	Modalities	Classifier	CN vs. MCI					CN vs. AD				
			ACC %	SEN %	SPE %	PPV %	NPV %	ACC %	SEN %	SPE %	PPV %	NPV %
R. Cuingnet et al.,[4] 2011	MRI	SVM	-	73.00	74.00	56.00	**86.00**	-	82.00	89.00	86.00	86.00
H. Aidos et al.,[12] 2017	PET	SVM	61.90	54.70	69.20	-	-	84.40	76.90	91.90	-	-
Proposed Study	MRI	GDA	**80.61**	**81.97**	**78.38**	**86.21**	72.50	**93.10**	**90.48**	**94.59**	**90.48**	**94.59**

CN: cognitively normal control, MCI: mild cognitive impairment, AD: Alzheimer's disease, Comb.: combining left and right, ACC: accuracy, SEN: sensitivity, SPE: specificity, PPV: positive predictive value, NPV: negative predictive value

two decision spaces (i.e., the left and right hemispheres of the brain), respectively, then as long as one of them produces a positive result (MCI or AD), the given subject is classified as a positive one. Since the boundaries have been obtained, it would be very effective to classify a subject. A normality test was conducted, which proved that the original data were normally distributed; therefore, GDA was the method of choice as a more efficient way to address the anticipated nonlinear boundaries between the different groups (CN, MCI and AD). Empirical evaluations demonstrated that the proposed GDA-based algorithm, as illustrated in Fig. 6, proved to be easier for implementation and provided better results than logistic regression and SVM with Gaussian or RBF kernel. And taking advantage of the covariance matrix, the correlation of different variables is taken into account by the proposed GDA-based classifier, which is deemed essentially important and often ignored in some probabilistic classification algorithms like Naive Bayes.

It ought to be noted that in this study, the classification performance has been improved significantly by using only structural MRI data. Evidently, there are many other sensitive biomarkers including PET, CSF, EEG, among others, and some cognitive markers like failure to recover from proactive interference (frPSI),[47] that could be integrated in the proposed analysis that made use of only MRI measurements. In a multimodal neuroimaging approach, diagnosis, prediction and classification of AD are all processes that would be greatly enhanced, with a focus placed on the early detection of the MCI stage and hence timely planning of therapeutic interventions and treatment.[31]

So far, most of the current investigations assumed only binary or two-way classification, where validation experiments were based on two-group comparisons, i.e., CN vs. MCI, CN vs. AD, and MCI vs. AD. Such binary classifications limit the clinical diagnosis for a given patient, which could belong in any of the three groups. In those three-way classification studies, the performance is still not insufficient, which can achieve the overall accuracy around 60%.[13] The proposed algorithm is not able to implement three-way classification yet, therefore, more efforts and further investigations need to be concentrated on the multimodal multi-class classification of different stages of AD for our future work.

5. Conclusion

This study proposed GDA-based dual high-dimensional decision spaces for the diagnosis of MCI in AD using structural MRI data as the unique input. The feature selection in this study demonstrates that the entorhinal cortex is the most significant cortical region for distinguishing CN from MCI and more evidently for AD, which is consistent with recent studies that concluded that the entorhinal cortex, deep in the brain, is the first area to be implicated in AD. As a clinical application, when selecting the optimal sets of variables, the classification performance is measured by the F1 score instead of the accuracy in consideration of the imbalanced data. Another major contribution of this study is that by performing the feature selection and

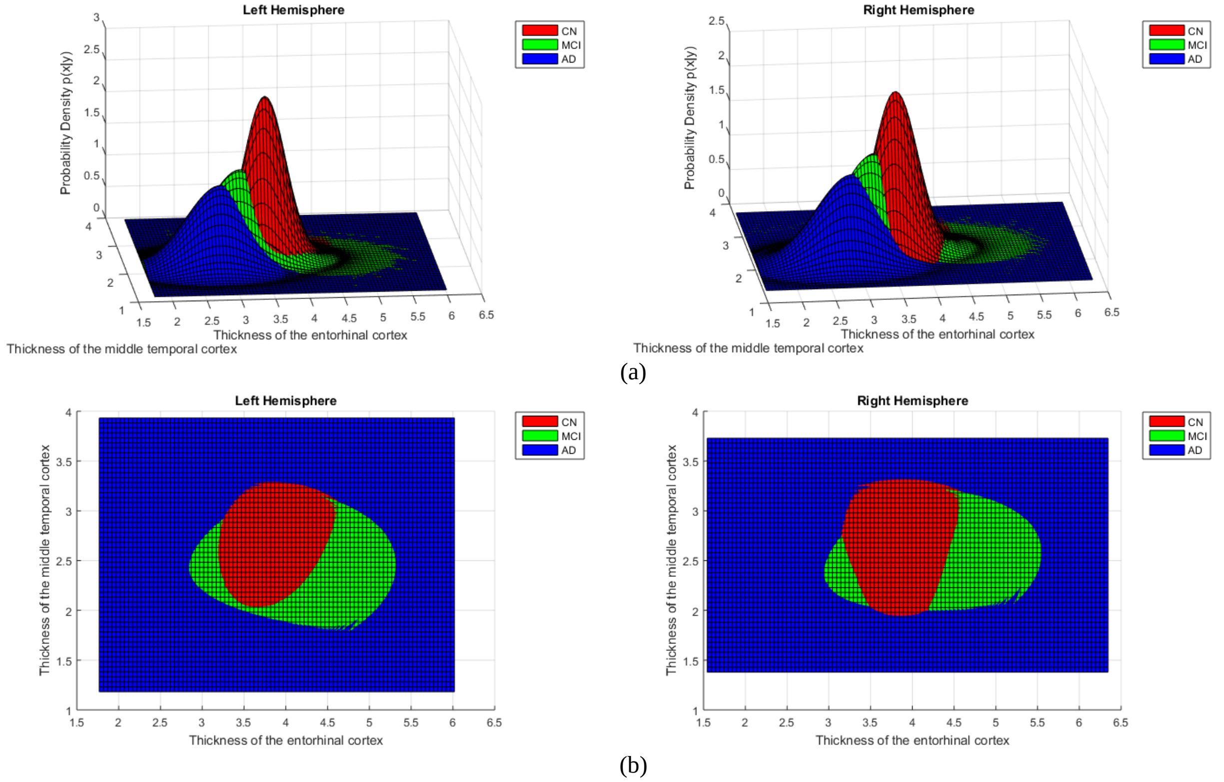


Fig. 6. The GDA-based dual high-dimensional decision spaces for CN, MCI, and AD with top-two ranked features (a) Multivariate Gaussian distribution, and (b) Classification boundaries of GDA-based dual decision spaces.

training process to both left and right hemispheres of the brain separately, then generating dual decision spaces instead of typically using only one decision space, the classification performance is shown to improve significantly.

Availability of data and material

Data are available to researchers by applying to the respective organization, ADNI. Application is required to protect participant confidentiality. The ADNI data are available at (http://adni.loni.usc.edu/).

Acknowledgements

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. 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 Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

We acknowledge the critical support provided by the National Science Foundation under grants: CNS-1532061, CNS-1551221, CNS-1338922, and CNS-1429345. The generous support of the Ware Foundation is greatly appreciated.

We appreciate the support provided by the Florida-based Alzheimer's Disease Research Center (1P50AG047266-01A1), R01 AG047649-01A1, the Florida Department of Health, Ed and Ethel Moore Alzheimer's Disease Research Program, and the Wien Center for Alzheimer's Disease & Memory Disorders, Mount Sinai Medical Center, Miami, FL.

References

1. NIA, *Basics of Alzheimer's Disease and Dementia*. Available:https://www.nia.nih.gov/alzheimers/topics/alzmeimers-basics. Accessed: 24-Jan-2017.
2. NIA, How Is Alzheimer's Disease Treated? Available: https://www.nia.nih.gov/health/how-alzheimers-disease-treated. Accessed: 8-Dec-2017.
3. Q. Zhou, M. Goryawala, M. Cabrerizo, J. Wang, W. Barker, D. A. Loewenstein, R. Duara, and M. Adjouadi, An Optimal Decisional Space for the Classification of Alzheimer's Disease and Mild Cognitive Impairment, *IEEE Trans. Biomed. Eng.* **61**(8) (2014) 2245–2253.
4. R. C. A. Cuingnet, E. Gerardin, J. Tessieras, G. Auzias, S. Lehéricy, M.-O. Habert, M. Chupin, H. Benali, and O. Colliot, Automatic classification of patients with Alzheimers disease from structural MRI: A comparison of ten methods using the ADNI database, *Neuroimage* **56**(2) (2011) 766–781.
5. F. J. Martinez-Murcia, J. M. Górriz, J. Ramírez, and A. Ortiz, A Structural Parametrization of the Brain Using Hidden Markov Models-Based Paths in Alzheimer's Disease, *Int. J. Neural Syst.* **26**(7) (2016) 1650024.
6. D. Zhang, Y. Wang, L. Zhou, H. Yuan, D. Shen, and Alzheimer's Disease Neuroimaging Initiative, Multimodal classification of Alzheimer's disease and mild cognitive impairment, *Neuroimage* **55**(3) (2011) 856–867.
7. M. Liu, D. Zhang, D. Shen, and Alzheimer's Disease Neuroimaging Initiative, Ensemble sparse classification of Alzheimer's disease, *Neuroimage* **60**(2) (2012) 1106–1016.
8. L. Khedher, J. Ramírez, J. M. Górriz, A. Brahim, F. Segovia, and Alzheimer's Disease Neuroimaging Initiative, Early diagnosis of Alzheimer's disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images, *Neurocomputing* **151**(1) (2015) 139–150.
9. T. Ye, C. Zu, B. Jie, D. Shen, D. Zhang, and Alzheimer's Disease Neuroimaging Initiative, Discriminative multi-task feature selection for multi-modality classification of Alzheimer's disease, *Brain Imaging Behav.***10**(3) (2016) 739–749.
10. R. Romero-Garcia, M. Atienza, and J. L. Cantero, Different scales of cortical organization are selectively targeted during the progression to Alzheimer's disease, *Int. J. Neural Syst.* **26**(2) (2016), 1650003.
11. I. Illán, J. Górriz, J. Ramírez, D. Salas-Gonzalez, M. López, F. Segovia, R. Chaves, M. Gómez-Rio, and C. Puntonet, 18F-FDG PET imaging analysis for computer aided Alzheimer's diagnosis, *Info. Sci.* **181**(4) (2011) 903–916.
12. H. Aidos, and A. Fred, Discrimination of Alzheimer's Disease using longitudinal information, *Data Mining and Knowledge Discovery*, **31**(4) (2017) 1006–1030.
13. T. Tong, K. Gray, Q. Gao, L. Chen, and D. Rueckert, Multi-modal classification of Alzheimers disease using nonlinear graph fusion, *Pattern Recognition*, **63** (2017) 171–181.
14. P. Scheltens, CSF biomarkers in relation to neuroimaging in mild cognitive impairment and Alzheimers disease, *Alzheimers Dement.* **6**(4) (2010).
15. F. C. Morabito, M. Campolo, D. Labate, G. Morabito, L. Bonanno, A. Bramanti, S. D. Salvo, A. Marra, and P. Bramanti, A longitudinal EEG study of Alzheimer's disease progression based on a complex network approach, *Int. J. Neural Syst.* **25**(2) (2015) 1550005.
16. N. Mammone, L. Bonanno, S. D. Salvo, S. Marino, P. Bramanti, A. Bramanti, and F. C. Morabito, Permutation disalignment index as an indirect, EEG-based, measure of brain connectivity in MCI and AD patients, *Int. J. Neural Syst.* **27**(5) (2017) 1750020.
17. N. Mammone, S. D. Salvo, C. Ieracitano, S. Marino, A. Marra, F. Corallo, and F. C. Morabito, A permutation disalignment index-based complex network approach to evaluate longitudinal changes in brain-electrical connectivity, *Entropy* **19**(10) 548.
18. N. Houmani, G. Dreyfus, and F. B. Vialatte, Epoch-based entropy for early screening of Alzheimer's disease, Int. *J. Neural Syst.* **25**(8) (2015) 1550032.
19. J. Dauwels, F. Vialatte, T. Musha, and A. Cichocki, A comparative study of synchrony measures for the early diagnosis of Alzheimers disease based on EEG, *Neuroimage* **49**(1) (2010) 668–693.
20. H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, Alzheimers Disease: Models of Computation and Analysis of EEGs, *Clinical EEG and Neurosci.* **36**(3) (2005) 131–140.
21. H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, A Spatio-temporal Wavelet-Chaos Methodology for EEG-based Diagnosis of Alzheimer's Disease, *Neurosci. Letters* **444**(2) (2008) 190-194.

22. M. Ahmadlou, H. Adeli, and A. Adeli, New diagnostic EEG markers of the Alzheimer's disease using visibility graph, *J. Neural Transmission* **117**(9) (2010) 1099–1109.
23. M. Ahmadlou, H. Adeli, and A. Adeli, Fractality and a Wavelet-chaos-Methodology for EEG-based Diagnosis of Alzheimer Disease, *Alzheimer Disease & Associated Disorders* **25**(1) (2011) 85–92.
24. Z. Sankari and H. Adeli, Probabilistic neural networks for diagnosis of Alzheimers disease using conventional and wavelet coherence, *J. Neurosci. Methods* **197**(1) (2011) 165–170.
25. Z. Sankari, H. Adeli, and A. Adeli, Intrahemispheric, interhemispheric, and distal EEG coherence in Alzheimer's disease, *Clinical Neurophysiol.* **122**(5) (2011) 897–906.
26. M. Ahmadlou, A. Adeli, R. Bajo, and H. Adeli, Complexity of functional connectivity networks in mild cognitive impairment subjects during a working memory task, *Clinical Neurophysiol.* **125**(4) (2014) 694–702.
27. J. P. Amezquita-Sanchez, A. Adeli, and H. Adeli, A new methodology for automated diagnosis of mild cognitive impairment (MCI) using magnetoencephalography (MEG), *Behavioural Brain Res.* **305** (2016) 174–180.
28. D. López-Sanz, P. Garcés, B. Álvarez, M.L. Delgado, R. López, and F. Maestú, Network disruption in the preclinical stages of Alzheimer´s Disease: from Subjective Cognitive Decline to Mild Cognitive Impairment, *Int. J. Neural Sys.* **27**(8) (2017), 1750041.
29. L. Khedher, I. A. Illán, J. M. Górriz, J. Ramírez, A. Brahim, and A. Meyer-Baese, Independent Component Analysis-Support Vector Machine-Based Computer-Aided Diagnosis System for Alzheimer's with Visual Support, *Int. J. Neural Syst.* **27**(3) (2017) 1650050.
30. H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, Alzheimers disease and models of computation: Imaging, classification, and neural models, *J. Alzheimers Disease* **7**(3) (2005) 187–199.
31. G. Mirzaei, A. Adeli, and H. Adeli, Imaging and machine learning techniques for diagnosis of Alzheimer's disease, *Reviews in the Neurosci.* **27**(8) (2016) 857–870.
32. C. Li, C. Fang, M. Adjouadi, M. Cabrerizo, A. Barreto, J. Andrian, R. Duara, and D. Loewenstein, A Neuroimaging Feature Extraction Model for Imaging Genetics with Application to Alzheimer's Disease, *Proc. 17th IEEE BIBE* (2017) 15–20.
33. University of South California, *Alzheimer's Disease Neuroimaging Initiative (ADNI) database*, Available: http://adni.loni.usc.edu/methods/mri-analysis/adni-standardized-data/. Accessed 13 Jun 2016.
34. M. Reuter, N. J. Schmansky, H. D. Rosas, and B. Fischl, Within-Subject Template Estimation for Unbiased Longitudinal Image Analysis, *Neuroimage*. **61**(4) (2012) 1402–1418.
35. B. B. Avants, N. Tustison, and G. Song, *Advanced Normalization Tools (ANTS),* Available: ftp://ftp.ie.freshrpms.net/pub/sourceforge/a/project/ad/advants/ANTS/ANTS_1_9_x/ants.pdf. Accessed 24 Jan 2017.
36. A. Klein, S. S. Ghosh, F. S. Bao, J. Giard, Y. Häme, E. Stavsky, N. Lee, B. Rossa, M. Reuter, E. C. Neto, and A. Keshavan, Mindboggling morphometry of human brains, *PLOS Computational Biology* **13**(2) (2017) e1005350.
37. Sage Bionetworks, *Alzheimer's Disease Big Data DREAM Challenge #1, Available:* https://www.synapse.org/#!Synapse:syn2290704/wiki/64710. Accessed 24 Jan 2017.
38. Mindboggle-101, *Mindboggle Data, Available:* http://www.mindboggle.info/data.html. Accessed 24 Jan 2017.
39. S. S. Shapiro, and M. B. Wilk, An analysis of variance test for normality (complete sample). *Biometrika.* **52**(3–4) (1965) 591–611.
40. H. Braak, and K. D. Tredici, Alzheimer's disease: pathogenesis and prevention, *Alzheimers Dement.* **8**(3) (2012) 227–233.
41. T. Gómez-Isla, J. L. Price, D. W. McKeel, J. C. Morris, J. H. Growdon, and B. T. Hyman, Profound loss of layer II entorhinal cortex neurons occurs in very mild Alzheimer's disease, *J. Neurosci.* **16**(14) (1996) 4491–4500.
42. U. A. Khan, L. Liu, F. A. Provenzano, D. E. Berman, C. P. Profaci, R. Sloan, R. Mayeux, K. E. Duff, S. A. Small, Molecular drivers and cortical spread of lateral entorhinal cortex dysfunction in preclinical Alzheimer's disease, *Nat. Neurosci.* **17**(2) (2014) 304–311.
43. L. R. Squire, and S. Zola-Morgan, The medial temporal lobe memory system, *Science* **253**(5026) (1991) 1380–1386.
44. L. Wang, Y. Zang, Y. He, M. Liang, X Zhang, L. Tian, T. Wu, T. Jian, and K. Li, Changes in hippocampal connectivity in the early stages of Alzheimer's disease: Evidence from resting state fMRI, *Neuroimage* **31**(2) (2006) 496–504.
45. M. Xia, J. Wang, and Y. He, BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics, *PLoS ONE* **8**(7) (2013) e68910.
46. A. Ortiz, J. Munilla, J. M. Górriz, and J. Ramírez, Ensembles of deep learning architectures for the early diagnosis of the Alzheimer's disease, *Int. J. Neural Syst.* **26**(7) (2016) 1650025.
47. D. A. Loewenstein, R. E. Curiel, S. Dekosky, M. Rosselli, R. Bauer, M. Grieg-Custo, A. Penate, C. Li, G. Lizagarra, T. Golde, M. Adjouadi, and R. Duara, Recovery from Proactive Semantic Interference and MRI Volume: A Replication and Extension Study, *J. Alzheimers Disease* **59**(1) (2017) 131–139.