3-26-2020

Development of Gaussian Learning Algorithms for Early Detection of Alzheimer's Disease

Chen Fang
Florida International University, cfang002@fiu.edu

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

DEVELOPMENT OF GAUSSIAN LEARNING ALGORITHMS FOR EARLY DETECTION OF ALZHEIMER'S DISEASE

A dissertation submitted in partial fulfillment of

the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

ELECTRICAL AND COMPUTER ENGINEERING

by

Chen Fang

2020

To: Dean John L. Volakis
College of Engineering and Computing

This dissertation, written by Chen Fang, and entitled Development of Gaussian Learning Algorithms for Early Detection of Alzheimer's Disease, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

_____
Mercedes Cabrerizo

_____
Armando Barreto

_____
Jean Andrian

_____
Naphtali Rishe

_____
Malek Adjouadi, Major Professor

Date of Defense: March 26, 2020

The dissertation of Chen Fang is approved.

_____
Dean John L. Volakis
College of Engineering and Computing

_____
Andrés G. Gil
Vice President for Research and Economic Development
and Dean of the University Graduate School

Florida International University, 2020

DEDICATION

This dissertation is dedicated to my parents and my wife, without their patience, understanding, support, and most of all love, the completion of this work would not have been possible.

ACKNOWLEDGMENTS

I would like to express my most sincere appreciation to my major advisor, Dr. Malek Adjouadi, for his patience, encouragement and guidance in my research and for his kindness and consideration throughout the past six years that I pursued my Master and Ph.D. degrees at FIU. This work would not have been possible without his continuous support and invaluable academic advice.

I am grateful to all of those with whom I have had the pleasure to work during this wonderful journey. Many thanks to my Dissertation Committee members, Dr. Mercedes Cabrerizo, Dr. Armando Barreto, Dr. Jean Andrian, and Dr. Naphtali Rishe for their professional guidance about both scientific research and life in general.

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.

I am grateful for the continued support from the National Science Foundation (NSF) under NSF grants CNS-1920182, CNS-1532061, CNS-1338922, and CNS-1551221. I also greatly appreciate the support of NIH-NIA grants (1R01AG055638-01A1, 1R01AG061106-01), State of Florida grant (8AZ23), the 1Florida Alzheimer's Disease Research Center (ADRC) (NIA 1P50AG047266-01A1) and the Ware Foundation.

ABSTRACT OF THE DISSERTATION

DEVELOPMENT OF GAUSSIAN LEARNING ALGORITHMS FOR EARLY DETECTION OF ALZHEIMER'S DISEASE

by

Chen Fang

Florida International University, 2020

Miami, Florida

Professor Malek Adjouadi, Major Professor

Alzheimer's disease (AD) is the most common form of dementia affecting 10% of the population over the age of 65 and the growing costs in managing AD are estimated to be $259 billion, according to data reported in the 2017 by the Alzheimer's Association. Moreover, with cognitive decline, daily life of the affected persons and their families are severely impacted. Taking advantage of the diagnosis of AD and its prodromal stage of mild cognitive impairment (MCI), an early treatment may help patients preserve the quality of life and slow the progression of the disease, even though the underlying disease cannot be reversed or stopped. This research aims to develop Gaussian learning algorithms, natural language processing (NLP) techniques, and mathematical models to effectively delineate the MCI participants from the cognitively normal (CN) group, and identify the most significant brain regions and patterns of changes associated with the progression of AD. The focus will be placed on the earliest manifestations of the disease (early MCI or EMCI) to plan for effective curative/therapeutic interventions and protocols.

Multiple modalities of biomarkers have been found to be significantly sensitive in assessing the progression of AD. In this work, several novel multimodal classification

frameworks based on proposed Gaussian Learning algorithms are created and applied to neuroimaging data. Classification based on the combination of structural magnetic resonance imaging (MRI), positron emission tomography (PET), and cerebrospinal fluid (CSF) biomarkers is seen as the most reliable approach for high-accuracy classification.

Additionally, changes in linguistic complexity may provide complementary information for the diagnosis and prognosis of AD. For this research endeavor, an NLP-oriented neuropsychological assessment is developed to automatically analyze the distinguishing characteristics of text data in MCI group versus those in CN group. Early findings suggest significant linguistic differences between CN and MCI subjects in terms of word usage, vocabulary, recall, fragmented sentences.

In summary, the results obtained indicate a high potential of the neuroimaging-based classification and NLP-oriented assessment to be utilized as a practically computer aided diagnosis system for classification and prediction of AD and its prodromal stages. Future work will ultimately focus on early signs of AD that could help in the planning of curative and therapeutic intervention to slow the progression of the disease.

TABLE OF CONTENTS

LIST OF TABLES

TABLE PAGE

LIST OF FIGURES

FIGURE PAGE

CHAPTER I. 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, and moreover, Alzheimer's disease (AD) may be estimated as the third leading cause of death for the older population in the United States, just behind heart disease and cancer [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. 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 [2]. Multiple modalities of biomarkers have been found to be significantly sensitive in assessing the progression of AD. These include structural magnetic resonance imaging (MRI) [3-11], positron emission tomography (PET) [5, 6, 10, 12-14], cerebrospinal fluid (CSF) [6, 13-15], electroencephalographic (EEG) rhythms [16-29], and Magnetoencephalography (MEG) [30, 31]. 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 [8-10, 12-14, 32-34]. 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) [3, 8-10, 12-14, 32-34]. 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, 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 [3, 9, 10, 32]. 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 [33, 34].

As a way to overcome such challenges, this study first develops a series of 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 [35], 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. The feature selection results of these proposed methods demonstrate 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 concluding that the entorhinal cortex, deep in the brain, is the first area to be implicated in AD [36-38].

In recent years, machine learning approaches have been applied in a growing number of studies to characterize patterns of structural, functional and metabolic difference discernible from multimodal neuroimaging data. The high-dimensionality nature of neuroimaging data often raises a necessity for dimensionality reduction and feature selection to obtain an optimal decision space. The results reported in some recent studies indicate that appropriate decision-making methods could improve the classification accuracy regardless of the sample size [39-42]. Voxel-based MRI studies have demonstrated that widely distributed cortical and subcortical brain regions show atrophy patterns in MCI, preceding the onset of AD [43-47]. A recent study has indicated the clinical utility of PET imaging for differential diagnosis in early onset dementia in support of clinical diagnosis of participants with AD and noncarrier APOE ε4 status who are older than 70 years [48]. Empirical evidence suggests that appropriate feature selection could preserve the complementary inter-modality information; therefore, the proposed dimensionality reduction model shows great potential for extracting relevant information from all modalities associated with the progression of AD.

Currently, the Principal Component Analysis (PCA) model remains the most widely used method in dimensionality reduction and feature selection tasks [49, 50]. However, for machine learning tasks like classification and regression analyses, PCA is

applied as an unsupervised method not considering the interclass information, such as data labels and target values; therefore, in many cases the consequently implemented feature selection methods may not be able to find the optimal decision spaces for the corresponding tasks. Moreover, the importance of PCA generated components is estimated by the variance, which are not often equivalent to the significance of those components in machine learning tasks.

This study also aims to introduce a supervised dimensionality reduction algorithm to characterize the important Gaussian discriminative components with respect to the structural, functional or metabolic measurements as observed in the MRI-PET combination associated with different stages of AD, focusing on the prodromal stage of MCI [51, 52]. The stage of MCI is subdivided into two stages, early MCI (EMCI) and late MCI (LMCI), as defined in the Alzheimer's disease Neuroimaging Initiative (ADNI) data. Since alleviation of specific symptoms is possible through therapeutic interventions for some patients in the early or middle stages of AD, effective diagnosis of EMCI from CN group is essentially important for the planning of early treatment. However, instead of utilizing PCA computed variances to determine the significances of different components, the proposed Gaussian discriminative component analysis (GDCA) makes use of GDA classifiers to reveal the discriminability of different components in terms of each component's performance obtained by a designate machine learning task. This process is shown to lead to stable, reliable and accurate dimensionality reduction in multimodal neuroimaging biomarkers for effective classification, enhanced diagnosis and the monitoring of disease progression.

Additionally, in order to yield more effective and accurate results, typically, neuropsychological tests are used to screen patients for the discrimination of different stages of AD, for example, the Mini-Mental State Examination (MMSE). Furthermore, the 1980s have seen the development of computerized neuropsychological testing systems that have contributed substantially in assessing cognitive function and memory decline in elderly patients in the early stages of Alzheimer's disease [53-59]. In these thorough reviews and studies, the general consensus is that although the psychometric measurements were beneficial in assessing memory decline and could actually augment well-known and common neuropsychological tests such as MMSE, the psychometric properties and measurements varied from one test to another, and that each battery of tests could be judged only on a case by case basis and on the type of application that is under consideration. This suggests the need for the development of a more generalized application that takes into consideration how well-conceived and easy to understand is the test.

As a well-known early symptom, some linguistic complexity changes of language have been associated with the progression of MCI, for instance, forgetting some frequently used terms [60]. Language troubles, occurred with executive function and memory troubles, are the first clinical cognitive signs of AD. Although the language performance depends on the education, age and some other factors for a given subject, word finding difficulties and semantic paraphasia are still the most obvious symptoms among early AD or MCI patients. During the early stage, affected by amnesic troubles, patients keep their language capacity almost intact. Tests of language capacity (spoken and written, production and perception) are usually performed during dedicated interviews. The most frequent tests are related to oral expression (chat on a given topic, description of pictures,

etc.) and the assessment of lexical production (naming of pictures, fluency, etc.), which allow performing qualitative and quantitative assessment of language capacity in controlled situations [53]. But, currently, the investigations of linguistic manifestations of AD are mostly relied on manual analysis, while as one of the language use characteristics sensitive to the effects of AD, linguistic complexity-based neuropsychological test cannot be easily and effectively operated and measured using manual approaches [61]. Hence, it is essentially important to build computerized neuropsychological tests that can automatically implement the linguistic analysis and be more widely used in the diagnosis of MCI.

The last part of this study aims at developing a natural language processing (NLP)-oriented computerized neuropsychological assessment taking advantage of the existing NLP techniques to find significant patterns of changes in linguistic complexity associated with the progression of AD. By applying NLP techniques, it is possible to precisely diagnose MCI based on the analysis of some conversation contents, such as, spoken features, lexical features and syntactic features [62]. Some case studies have indicated that the NLP technic is able to find clear patterns of decline in syntactic and grammatical complexity [60, 61]. The proposed system will extract key features from discourse transcripts, and evaluate on non-scripted news conferences from President Ronald Reagan (RR), who was diagnosed with AD in 1994, and some other presidents, who have no known diagnosis of AD, then indicate that over time, the patterns of the linguistic complexity changes are statistically significantly distinguishable between RR and CN people.

CHAPTER II. A GAUSSIAN DISCRIMINANT ANALYSIS-BASED GENERATIVE LEARNING ALGORITHM FOR THE EARLY DIAGNOISIS OF MILD COGNITIVE IMPAIRMENT IN ALZHEIMER'S DISEASE

2.1. Goal

This chapter aims to introduce a novel generative learning algorithm based on the GDA, which can achieve more effective and accurate classification performance than SVM. To the best of our knowledge, this study is the first that applies GDA to the classification of AD and MCI, with a focus placed on the classification of MCI vs. CN. The feature extraction 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 concluding that the entorhinal cortex, deep in the brain, is the first area to be implicated in AD [36-38].

2.2. Methodology

In this section, the GDA-based generative algorithm is presented for the classification of AD and MCI. First, several software pipelines are used to pre-process original MRI data. Second, after the pre-processing step, morphometric (shape) data could be derived from the images, including shape measures of all 25 labeled cortical regions. Then a noise detection procedure and a feature extraction method based on the analysis of variance (ANOVA) are employed to determine the statistical significance of each variable in the classification outcome. Third, a GDA-based classifier 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 presented in Fig. 1.

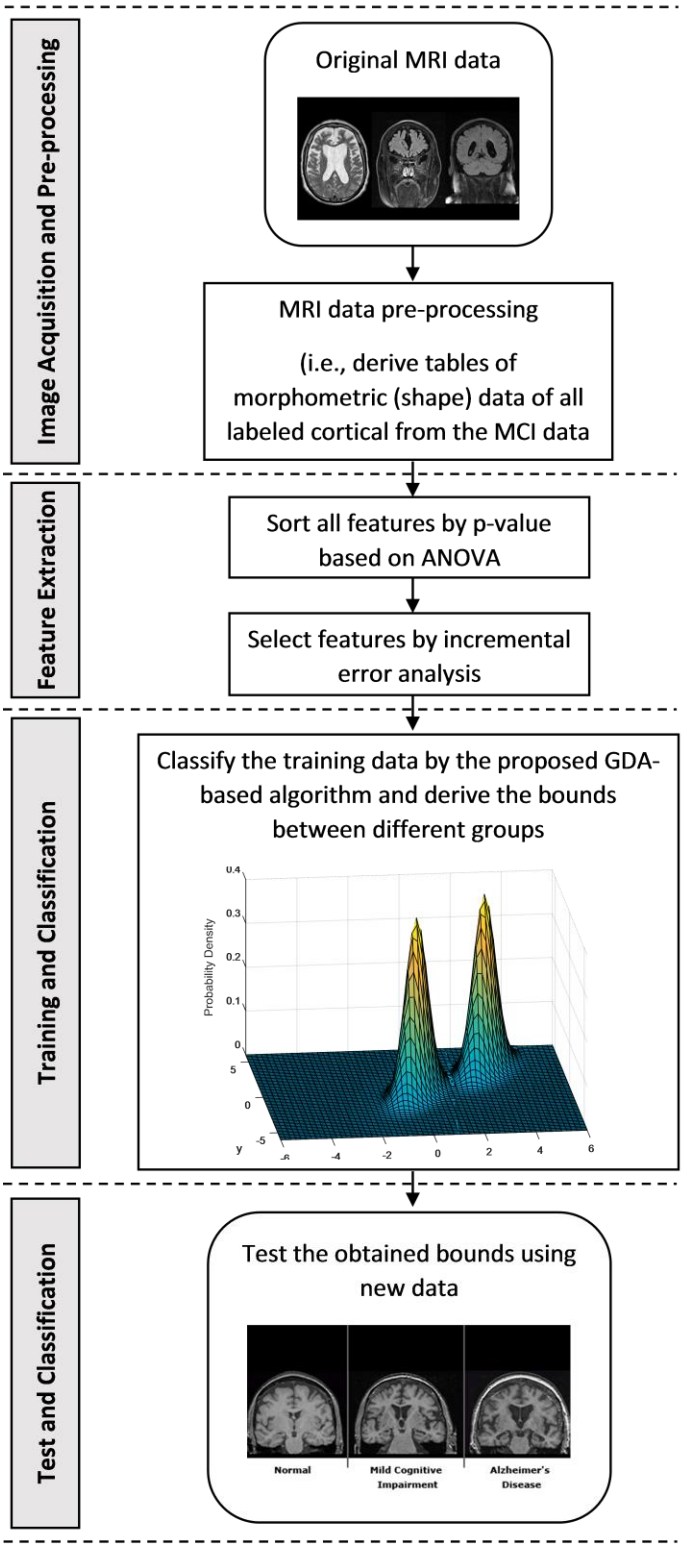


Fig.1. General Framework of the GDA-based Algorithm.

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

2.2.1. Subjects

The data used in the preparation of this study were obtained from the 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). The primary phenotype is diagnostic group and MMSE. All source imaging data consist of 1.5 Tesla T1-weighted MRI volumes in the NIfTI (.nii.gz) format from the ADNI1: Complete 1Yr 1.5T Data Collection. Summary statistics and patient counts are listed in Table 1.

2.2.2. MRI Data Pre-processing

Using three neuroimaging software pipelines: FreeSurfer [63], Advanced Normalization Tools (ANTs) [64], and Mindboggle [65], the original MRI data were pre-processed following the instruction provided by Alzheimer's Disease Big Data DREAM Challenge #1 [66]. 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; 7) volume. FreeSurfer pipeline (version 5.3) was applied to all T1-weighted images to generate labeled cortical surfaces, and labeled

cortical and noncortical volumes. Templates and atlases used by ANTs and Mindboggle could be found on the Mindboggle website [67].

2.2.3. Noise Detection

The aforementioned pre-processed MRI data of the 25 labeled cortical regions were used to generate two 175-variable (7×25) vector discriminator, for each subject (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 can be found in Section 2.4. As for subjects which involved abnormal variables, for example, some regions having measurements of some areas to be zero, these subjects should be regarded as noises and be deleted from the dataset for further investigation.

2.2.4. 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. The 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, equal weights were assigned to each of the shape measures so as to eliminate any bias.

2.2.5. Incremental Error Analysis

In order to maintain only few but 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 [2]. In the initial phase, the proposed GDA-based classifier only used 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, which then would be compared with the previous phase. Until the performance cannot be improved significantly anymore, the optimal set of variables would have been obtained.

2.2.6. GDA-based Classifier

Since there may be as many as 175 variables to be taken into consideration, the classifier must be capable of dealing with a high dimensional classification problem. For this reason, and by using GDA, an important generative 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 classifier was applied to each hemisphere of the brain (i.e., the two n-dimensional vectors), and if either one of the two sides was classified to be positive, the corresponding subject should be positive as well.

2.2.7. Classification Experiments

The performance of the proposed classifier was measured by using the accuracy, sensitivity, specificity, positive predictive value, and negative predictive value based on tenfold cross validation process. For selecting the optimal set of variables, 75% of the noise detected subject data points were used in the tenfold cross validation process. After the

optimal set of variables was generated, the classification performance was evaluated by using the remaining 25% of the noise detected subject data points.

2.3. Results

In this section, the experimental results of the feature selection process reveal the significance of different ROIs in patients for the three classifications: 1) CN vs. MCI; 2) CN vs. AD; 3) MCI vs. AD. The statistics evolution during the incremental error analysis and the classification performance of the proposed GDA-based classifier are also presented.

2.3.1. Top-ranked Variables

After the noise detection process, 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 were used for the classification as shown in Table 2.

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

Table 3. Top-3 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	$< 10^{-6}$
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	$< 10^{-7}$

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

The top 3 ranked variables and their corresponding measurements are given in Table 3. From this table, it can be observed that the entorhinal cortex is the most significant cortical region for discriminating either MCI or AD from CN. Even though for distinguishing MCI from AD, it is the second top-ranked region. This observation is consistent with recent studies indicating that indeed the entorhinal cortex is the first area to be implicated in AD [36-38], which comes in support of the validity of our feature selection method. Moreover, the entorhinal cortex has been proven as a major source of projections to the hippocampus [68], which plays an important role in converting short-term memory (also known as working memory) to long-term memory.

2.3.2. Optimal Sets of Variables

For the tenfold cross validation to generate the optimal set of variables, the aforementioned 75% of the noise detected data points included 470 individuals (140 CN, 230 MCI, and 100 AD, all divisible by 10 for the tenfold cross validation) of the noise free subjects included in this study (619 = 187 CN + 301 MCI + 131 AD). The purpose of applying the incremental error analysis is to achieve the best classification performance

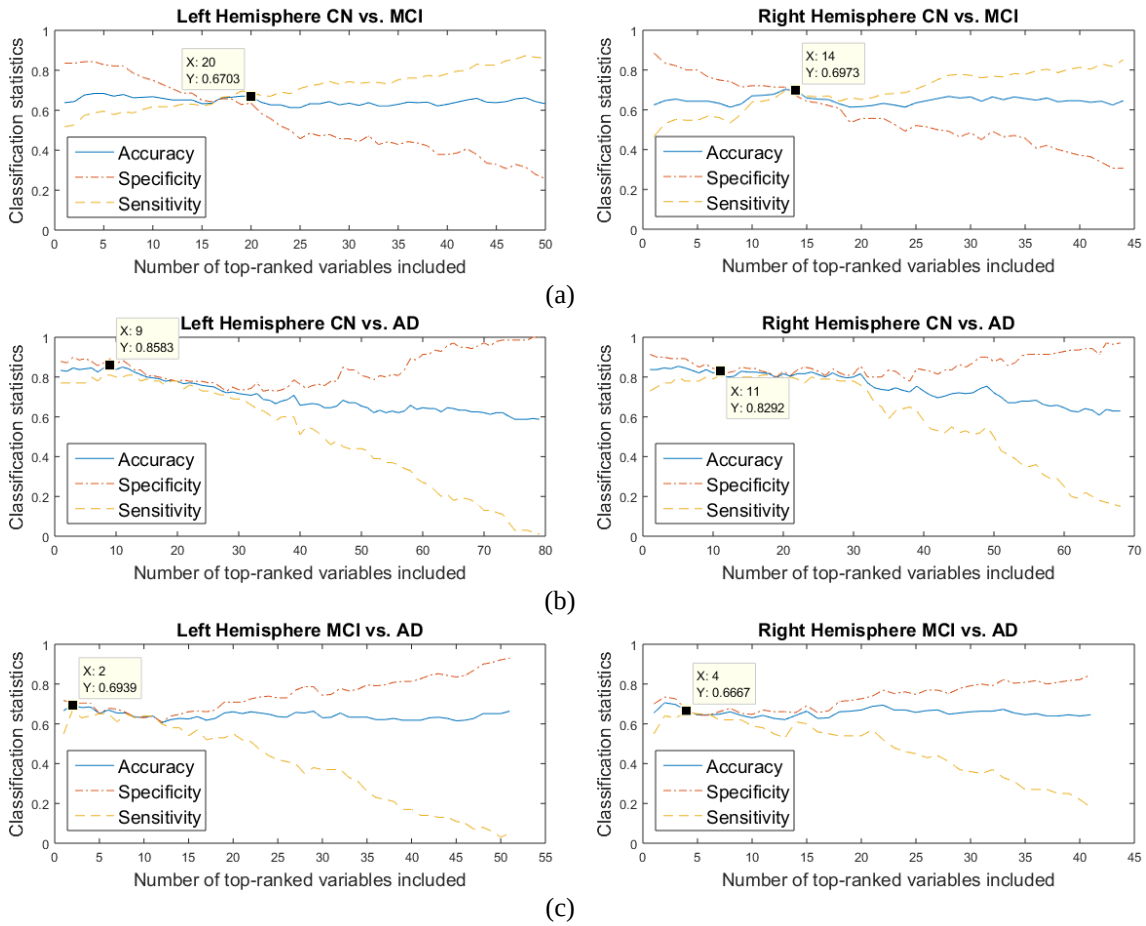


Fig. 2. Incremental Error Analysis Performance of Classification Statistics: (a) CN vs. MCI, (b) CN vs. AD, and (c) MCI. vs. AD.

with the minimum number of variables (i.e., the number of dimensions in the decision spaces). Some classification statistics of each hemisphere are illustrated in Fig. 2, where the horizontal axis indicates the number of significant variables included in each iteration.

It should be noted that as a clinical application, the classification performance was not only measured by the accuracy, which actually relies more on the sensitivity and specificity. For selecting the optimal set of variables, the set with the highest sensitivity was selected, when several sets had the same sensitivity value, the one with the highest specificity was chosen, then if still multiple choices are found, the one with the highest accuracy was finally taken. There was an exception for CN vs. MCI, the sensitivity

Table 4. Summary of Tenfold Cross Validation Performance Improved after Combining Two Decision Spaces

Groups	CN vs. MCI			CN vs. AD			MCI vs. AD		
Decision Space	*Left*	*Right*	*Comb.*	*Left*	*Right*	*Comb.*	*Left*	*Right*	*Comb.*
ACC %	67.03	69.73	84.86	85.83	82.92	94.17	69.39	66.67	81.82
SEN %	69.13	71.30	84.78	81.00	81.00	93.00	67.00	66.00	83.00
SPE %	63.57	67.14	85.00	89.29	84.29	95.00	70.43	66.96	81.30
PPV %	75.71	78.10	90.28	84.38	78.64	93.00	49.63	46.48	65.87
NPV %	55.63	58.75	77.27	86.81	86.13	95.00	83.08	81.91	91.67
Number of the optimal variables	Left: 20 Right: 14			Left: 9 Right: 11			Left: 2 Right: 4		

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

continued increasing while adding more variables, and from the clinical point of view, the higher sensitivity and the lower specificity would be preferred, therefore once the sensitivity was greater than the specificity, the set with the highest accuracy was then chosen.

As demonstrated in Fig. 2(a), for each side of the brain, the classification performance for CN vs. MCI yielded average results just as other studies reviewed in [3] have shown, where the sensitivity is 69.13% and 71.30% for each decision space (i.e., each hemisphere), respectively. Moreover, after combining the results of the two decision spaces together, the performance was significantly improved for all comparisons as shown in Table 4, which are very competitive in comparison to results reported in other recent studies [5, 6, 9] as indicated in Table 5.

Table 5. Comparison of Cross Validation Performance with Some Recent Studies

Groups Reference	Subjects (CN+MCI+AD)	Modalities	CN vs. MCI			CN vs. AD			MCI vs. AD		
			ACC %	SEN %	SPE %	ACC %	SEN %	SPE %	ACC %	SEN %	SPE %
L. Xu et al. [5], 2016	117+110+113	MRI+FDG +PET	74.50	66.40	82.10	94.80	95.60	94.00	-	-	-
D. Zhang et al. [6], 2011	52+99+51	MRI+PET +CSF	76.40	81.80	66.00	93.20	93.00	93.30	-	-	-
L. Khedher et al. [9], 2015	229+401+188	MRI	81.89	82.16	81.62	88.49	91.27	85.11	87.3	88.65	85.41
Proposed Study	190+305+133	MRI	84.86	84.78	85.00	94.17	93.00	95.00	81.82	83.00	81.30

CN: cognitively normal control, MCI: mild cognitive impairment, AD: Alzheimer's disease, ACC: accuracy, SEN: sensitivity, SPE: specificity

Table 6. Summary of the Proposed GDA-based Generative Learning Algorithm Classification Performance

Groups	CN vs. MCI			CN vs. AD			MCI vs. AD		
Decision Space	*Left*	*Right*	*Comb.*	*Left*	*Right*	*Comb.*	*Left*	*Right*	*Comb.*
ACC %	59.32	63.56	82.20	78.21	73.08	85.90	64.71	63.73	75.49
SEN %	63.38	64.79	83.10	70.97	70.97	83.87	41.94	61.29	64.52
SPE %	53.19	61.70	80.85	82.98	74.47	87.23	74.65	64.79	80.28
PPV %	67.16	71.88	86.76	73.33	64.71	81.25	41.94	43.18	58.82
NPV %	49.02	53.70	76.00	81.25	79.55	89.13	74.65	79.31	83.82

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

2.3.3. Classification Performance

By using the remaining 25% of the noise free data points as the held-out test data (47 CN, 71 MCI, and 31 AD), the classification performance of the trained GDA-based classifier (i.e., the obtained optimal set of variables) is presented in Table 6. Even though, the classification performance was not as good as that in the tenfold cross validation, the results obtained are still better than state-of-the-art-algorithms reviewed in [3], especially for the challenging CN vs. MCI classification.

2.4. Discussion

The merits of the proposed GDA-based machine learning algorithm are not only reflected through the good classification performance it achieved, but also in the way it looked at the two hemispheres of the brain separately. The classification was performed to 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. As demonstrated in Fig. 3, the boundaries are nonlinear, which could not be solved by other linear classification algorithms, such as, SVM and logistic regression.

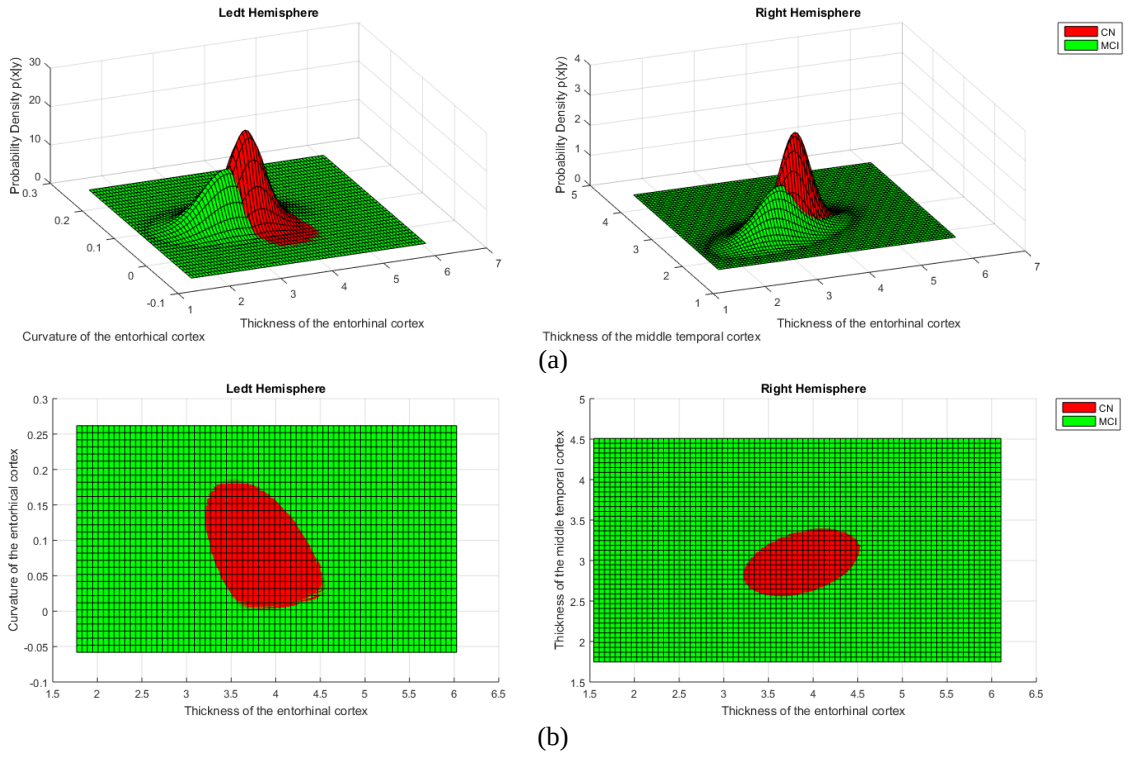


Fig. 3. The Boundaries for the Classification of CN vs. MCI Using Top-2 Significant Variables: (a) The multivariate Gaussian distribution of CN and MCI, and (b) The classification boundaries between CN and MCI.

And taking advantage of the covariance matrix, the correlation of different variables is taken into account by the proposed GDA-based algorithm, which are essentially important and often being ignored in some probabilistic classification algorithms like Naive Bayes.

It ought to be noted that in this study, the classification performance has been significantly improved by using only structural MRI data. But there are a lot of other sensitive biomarkers to the progression of AD and MCI, including PET, CSF, among others, to improve even further such classification results with a focus placed on CN vs. MCI. What is unique in this approach is the fact that instead of building only one decision space, the left and right hemispheres of the brain are separated into two decision spaces, respectively. The original intention of this approach was quite simple, and is based on the fact that every person has his or her dominant hand, and as such, right-handers use their

left brain more than the right one, and vice versa. Hence, regarding the progression of AD and MCI, the more frequently used side of the brain could be assumed to be affected more than the less used side.

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

3.1. Goal

Over the past few years, several approaches have been proposed to assist in the early diagnosis of AD and its prodromal stage of MCI. Using multimodal biomarkers for this high-dimensional classification problem, the widely used algorithms continue to yield unsatisfactory performance for delineating the MCI participants from the CN group. In this chapter, we develop a machine learning classification algorithm based on the 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 [35], 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 by training dual decisional spaces simultaneously, with the CN vs. AD classification results included here only for comparative purposes.

3.2. 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 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. 4.

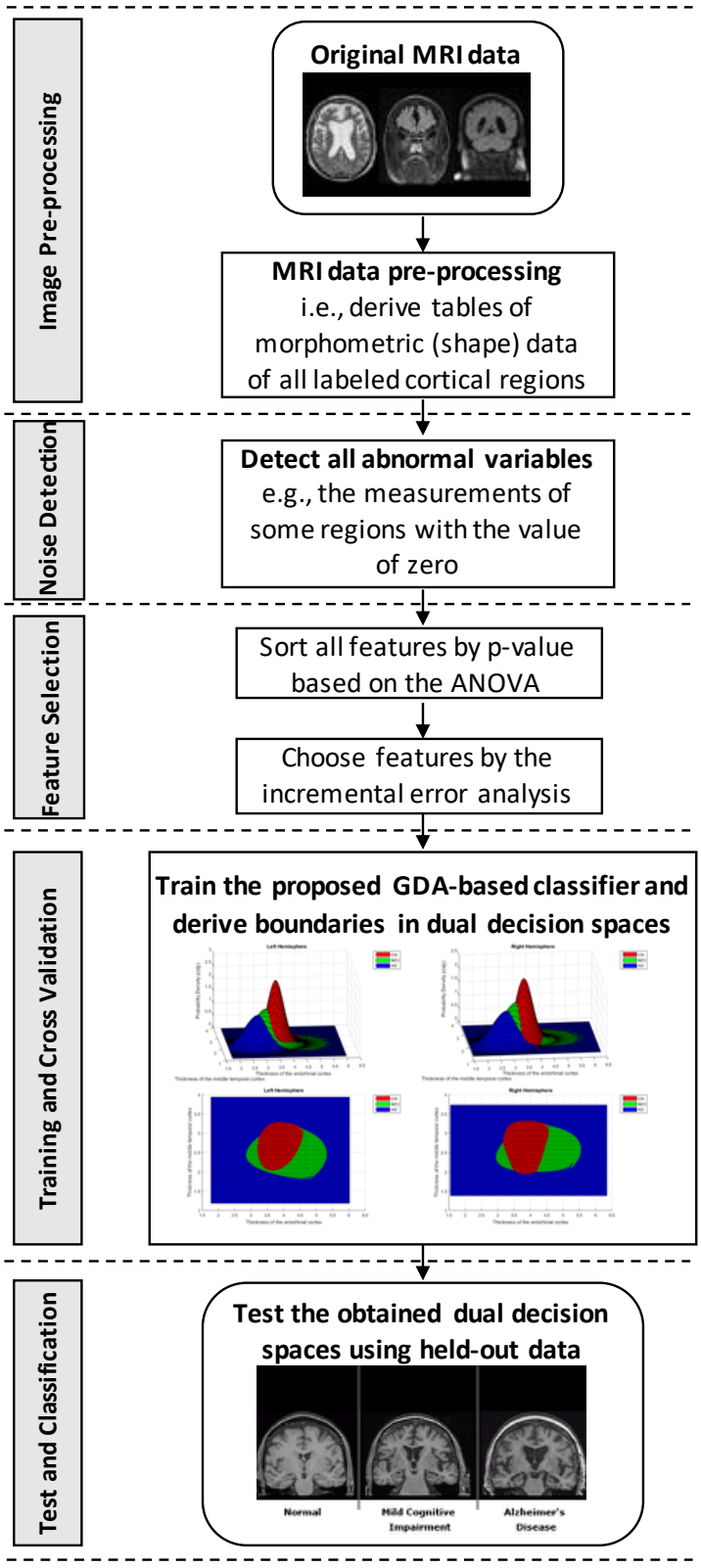


Fig. 4. General Framework of the GDA-based Dual High-dimensional Decision Spaces Algorithm.

3.2.1. 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). The ADNI was launched in 2003 as a public-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 as same as the one used in Chapter II.

3.2.2. MRI Data Pre-processing

Using three neuroimaging software pipelines: FreeSurfer [63], ANTs [64], and Mindboggle [65], the original MRI data were preprocessed following the instruction provided by Alzheimer's Disease Big Data DREAM Challenge #1 [66]. 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 [67].

3.2.3. Noise Detection

The 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 Section 3.3. 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.

3.2.4. 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 ROIs can help to discriminate more specifically MCI from CN.

3.2.4.1. 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 [69]. Furthermore, equal weights are assigned to each of the shape measures so as to eliminate any bias.

3.2.4.2. 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 [2]. 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.

3.2.5. 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 $x \in \mathbb{R}^n$. Then, given a training set, the proposed algorithm can model $p(x|y)$, the condition distribution of the n-dimensional vector 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:

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

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

$$p(y|x) = \frac{p(x|y)p(y)}{p(x)} \quad (2)$$

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 $p(x)$, since

$$\arg\max_y p(y|x) = \arg\max_y \frac{p(x|y)p(y)}{p(x)} = \arg\max_y p(x|y)p(y) \quad (3)$$

Therefore, for the purpose of classification, it only needs

$$\arg\max_y p(y|x) = \arg\max_y p(x|y) \quad (4)$$

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 $d_0, d_1, \ldots, d_9$, 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.

As demonstrated in Fig. 5, 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 Section 3.3, especially for the most challenging classification of CN vs. MCI.

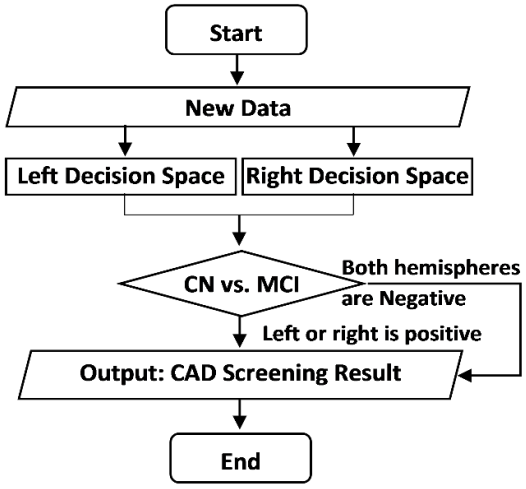


Fig. 5. Flowchart of the GDA-based Dual Decision Space Classification Process.

3.3. 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 the classification performance of the GDA-based algorithm using the proposed dual decisional spaces are provided.

3.3.1. 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 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 as same as those used in Chapter II.

The top 10 ranked variables and their corresponding measurements are given in Table 7, 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 [36-38], 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 [68], which plays an important role in converting short-term memory (also known as working memory) to long-term memory. Interestingly, for

Table 7. 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

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 critical for long-term memory, to which the disrupted hippocampal connectivity has been

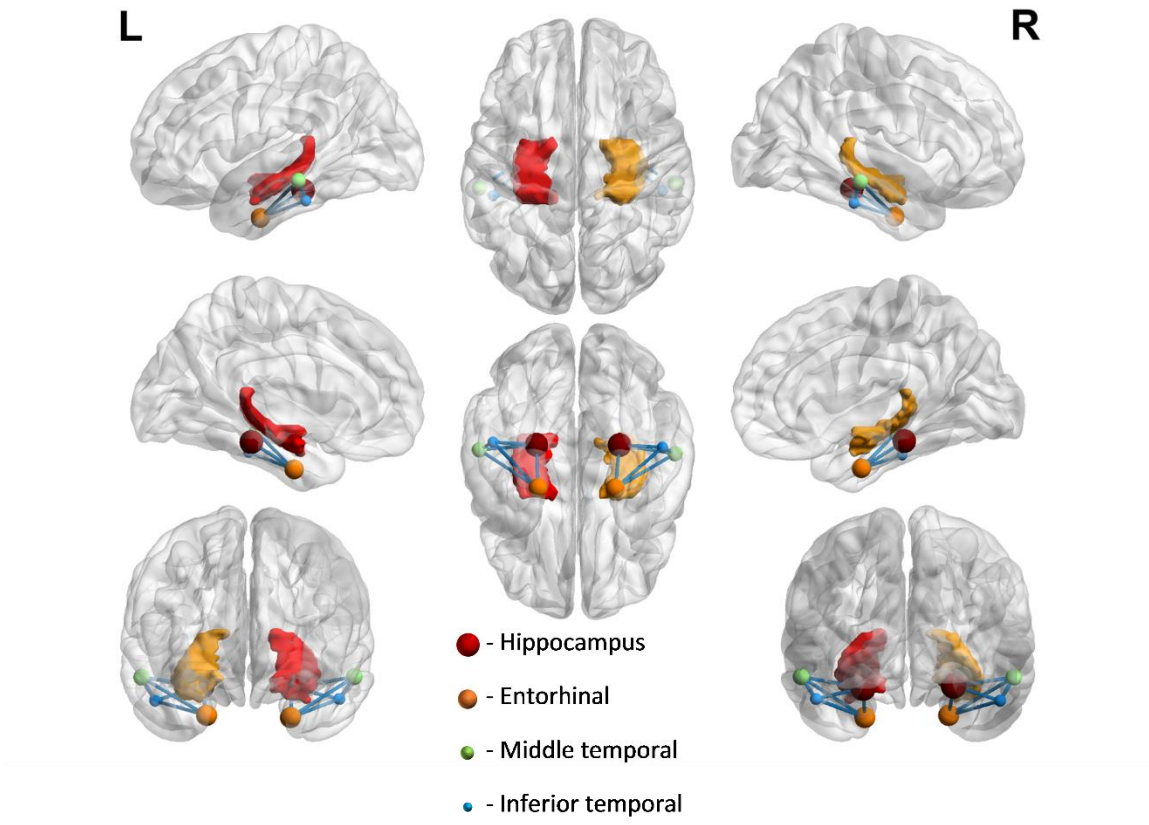


Fig. 6. Relative Location of Hippocampus and the Top-three-ranked Cortical Regions (visualized with the BrainNet Viewer [71]).

found in the early stages of AD [70]. 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. 6. From Table 7, 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.

3.3.2. Optimal Sets of Variables

To generate the optimal sets 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. 7, where the horizontal axis indicates the number of significant variables included in each iteration.

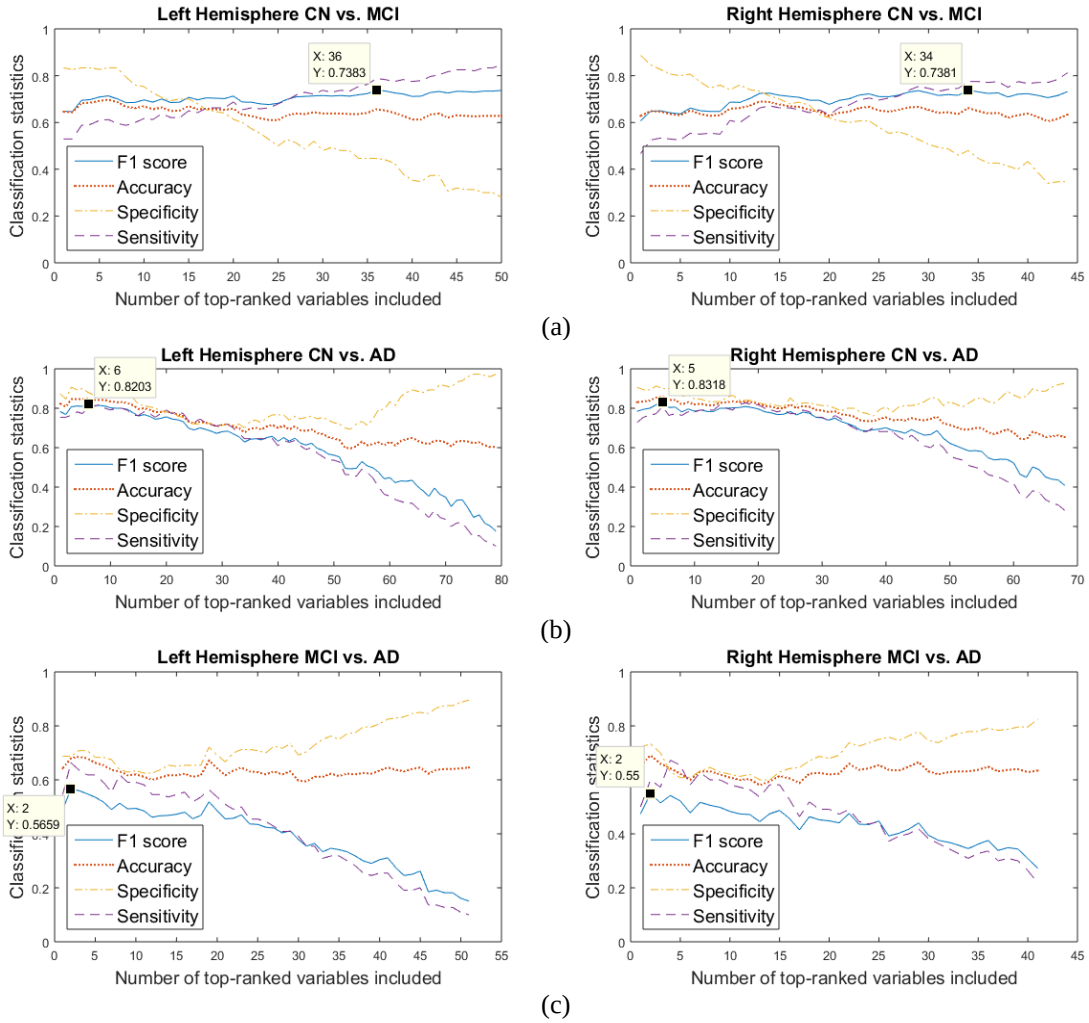


Fig. 7. Incremental Error Analysis Performance of Classification Statistics: (a) CN vs. MCI, (b) CN vs. AD, and (c) MCI vs. AD.

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):

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

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

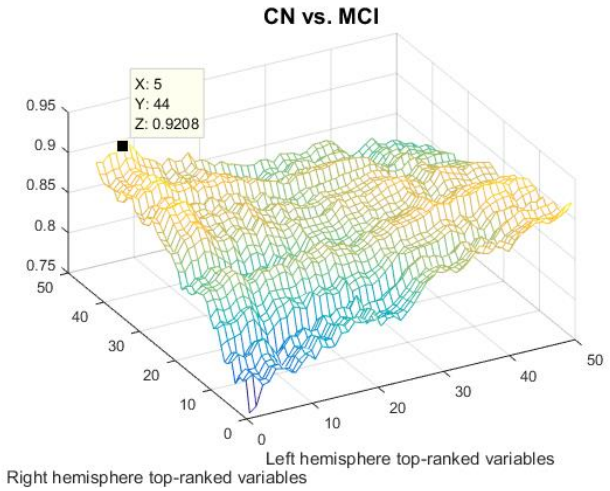
$$Specificity = \frac{TN}{TN + FP} \tag{7}$$

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

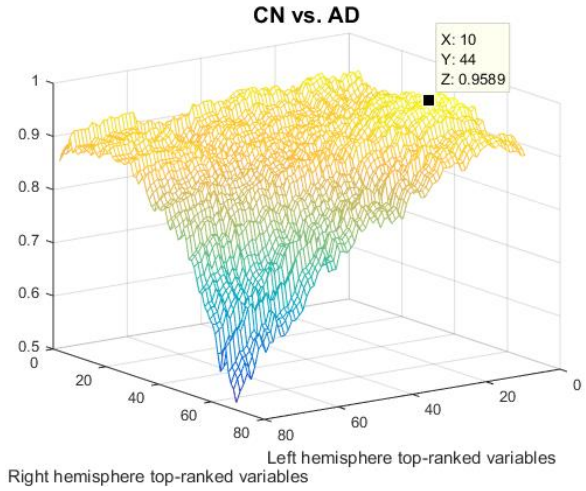
$$NPV = \frac{TN}{TN + FN} \tag{9}$$

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,

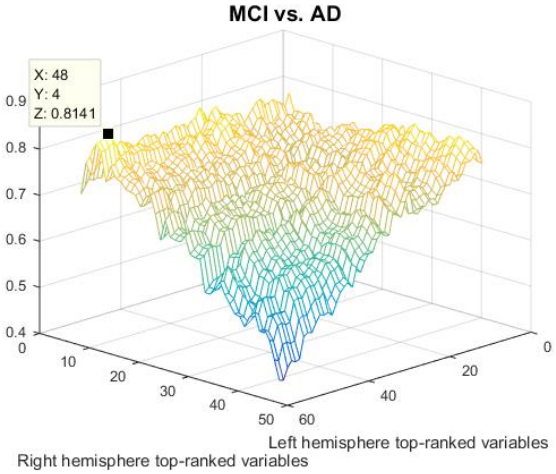
$$F1 = \frac{2 \times Sensitivity \times PPV}{Sensitivity + PPV} = \frac{2TP}{2TP + FP + FN} \tag{10}$$



(a)



(b)



(c)

Fig. 8. Simultaneous Incremental Error Analysis Performance of F1 Score: (a) CN vs. MCI, (b) CN vs. AD, and (c) MCI vs. AD.

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. 7(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 [3], 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 simultaneously implementing the incremental error analysis again, the evolution of the F1 score is as illustrated in Fig. 8.

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 8. 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-approaches [8-10, 12, 14, 32, 72], 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 9, except for the specificity of the CN vs. MCI classification and the sensitivity of the MCI vs. AD classification, the proposed

Table 8. 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

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

Table 9. 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. [8], 2012	MRI	SRC	ADNI (229 + 225 + 198)	87.85	85.26	**90.40**	90.80	86.32	94.76	-	-	-
L. Khedher et al. [9], 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. [10], 2016	MRI + PET	SVM	ADNI (52 + 99 + 51)	82.13	87.68	71.54	95.92	94.71	97.12	-	-	-
T. Tong et al. [12], 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. [14], 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. [72], 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

method yielded the best cross validation performance in a comparative assessment to all other methods.

3.3.3. 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 10. Although the classification performance was not as good as that

Table 10. Classification Performance of the Proposed GDA-based Dual High-dimensional Decision Space Algorithm

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 11. Comparison of Classification Performance Using Held-out Test 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. [3], 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

obtained in the tenfold cross validation, the results are still better than state-of-the-art-algorithm reviewed in [3] and the recently proposed state-of-the-art approach in [12], as shown in Table 11. Since not all studies implemented the held-out true test, only the results from [3] and [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.

3.4. 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 dual 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. 9, 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

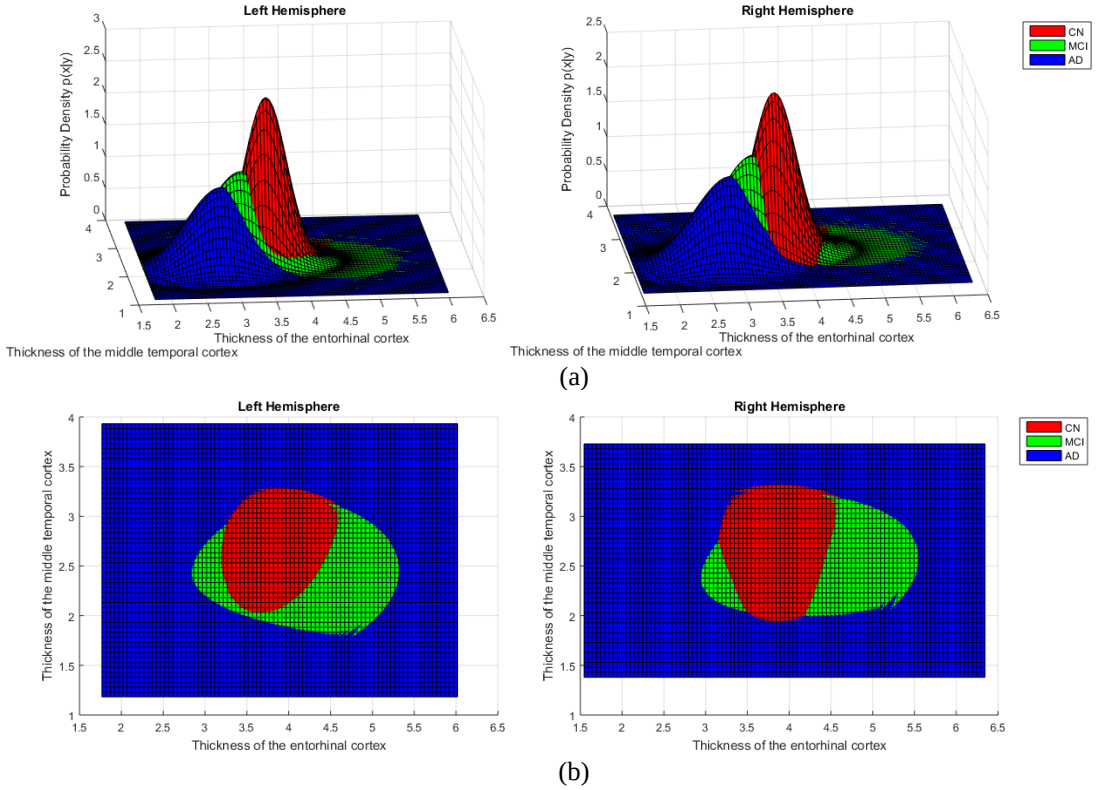


Fig. 9. The GDA-based Dual High-dimensional Decision Spaces for CN, MCI, and AD with Top-two Ranked Features: (a) Multivariate Gaussian distribution, and (b) Classification boundaries.

covariance matrix, the correlation of different variables is considered 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) [44], 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 [34].

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% [14]. 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.

CHAPTER IV. A NOVEL GAUSSIAN DISCRIMINANT ANALYSIS-BASED COMPUTER AIDED DIAGNOSIS SYSTEM FOR SCREENING DIFFERENT STAGES OF ALZHEIMER'S DISEASE

4.1. Goal

This chapter aims to introduce a novel classification algorithm that relies on a global feature selection, with the dual purpose of improved classification accuracy as well as enhanced prospects for its seamless integration into a computer aided diagnosis (CAD) system. Thus, instead of extracting different sets of features for those three comparisons (CN vs. MCI, CN vs. AD, and MCI vs. AD), respectively, in this study, only one optimal set of features will be generated. This proposed global feature selection indicates that the entorhinal cortex is the most significant cortical region associated with the progression of AD, which is consistent with recent studies concluding that the entorhinal cortex is the signature region to be implicated in AD [36-38, 73, 74]. Furthermore, by deploying left and right hemispheres of the brain into two distinct decision spaces, the classification performance is improved significantly.

4.2. Materials

The information of the subjects used in this study and the MRI data pre-processing procedure are presented in this section.

4.2.1. Subjects

The data used in the preparation of this study were obtained from the Alzheimer's disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-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. As part of the ADNI1: Complete 1Yr 1.5T collection and their assessments at baseline, the data include 628 individuals (190 CN, 305 MCI, and 133 AD). The primary phenotype is diagnostic group and Mini–Mental State Examination (MMSE). All source imaging data consist of 1.5 Tesla T1-weighted MRI volumes in the NIfTI (.nii.gz) format from the ADNI1: Complete 1Yr 1.5T Data Collection. Summary statistics and patient counts are listed in Table 1 as same as the one used in Chapter II and Chapter III.

4.2.2. MRI Data Pre-processing

Following the instruction provided by Alzheimer's Disease Big Data DREAM Challenge #1 [66], the original MRI data were pre-processed using three neuroimaging software pipelines: FreeSurfer [63], ANTs [64], and Mindboggle software [65]. Tables of morphometric data were derived from the images, including 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. All T1-weighted images were processed by FreeSurfer pipeline (version 5.3) in order to generate labeled cortical surfaces, and labeled cortical and noncortical volumes. Templates and atlases used by ANTs and Mindboggle could be found on the Mindboggle website [67].

4.2.3. Noise Detection

For each subject, the pre-processed MRI data of the 25 labeled cortical regions were used to obtain two 175-feature (7×25) vector discriminators corresponding to left and right hemisphere of the brain, respectively. Manipulating the features of each hemisphere

separately is shown to yield a better classification performance, of which details in support of this assertion can be found in the Results section. Any subjects with abnormal variables were eliminated from further investigation, for example, some of their cortical regions having measurements of some features to be zero.

4.3. Methods

In this section, the classification algorithm with global feature selection is presented for the early diagnosis of AD. The global feature selection method based on the ANOVA and the incremental error analysis are employed to determine the statistical significance of each feature in the classification outcome. A GDA-based classifier is then proposed for resolving the typical binary classification problem and introducing a CAD system for AD.

4.3.1. Global Feature Selection

Brain tissue has atrophied significantly at the final stage of AD, so all aforementioned shape measures could have changed as well. But in early onset of AD, the disease initially impacts only some specific regions of the brain, thus, determination of the key changes in ROIs can help achieve more specific diagnosis of different stages of AD.

4.3.1.1. Global Feature Ranking

Instead of performing the ANOVA on each of the 175 features of the two brain hemispheres between any two groups (i.e., CN vs. MCI, CN vs. AD, and MCI vs. AD), in the global feature selection scenario , the ANOVA was carried out on each of the 175 features of the two brain hemispheres between all three groups (CN, MCI and AD) to determine the significance of each feature in terms of classification outcome, and all features were thereafter ranked according to their p-values. Accordingly, only one set of

global features is obtained corresponding to each hemisphere of the brain. In order to eliminate any bias, equal weights were assigned to each of the shape measures.

4.3.1.2. Incremental Error Analysis

In order to maintain the few but essential features for achieving the best classification performance, an incremental error analysis was deployed for deciding how many of the top-ranked features ought to be involved in the final classifier. In the initial phase, the classifier only used the first-ranked feature, and the next top-ranked feature is thereafter introduced in the classifier at each subsequent phase, and recording the corresponding classification statistics, which then would be compared with the previous phase, until all essential features are determined. In this study, as a machine learning classification problem, the performance was estimated by the F1 score, and when the performance can no longer be improved, then the optimal set of essential features would have been obtained. The F1 score used for selecting the optimal sets of features is expressed as Equation (10) in Chapter III. It should be noted that as a clinical application, the classification performance may not only be measured by the accuracy, but actually relies more on the sensitivity and PPV.

4.3.2. GDA-based Classifier

Since the classifier may have as many as 175 features to be taken into consideration, it must be able to deal with high-dimensional data. Due to this reason, GDA, an important generative learning algorithm for high dimensional classification problems, has been used to establish the proposed classifier capable of recognizing the different patterns between any given groups. The detailed definition and formulas of the GDA classifier used in this study can be found in Chapter III, Section 3.2.5.

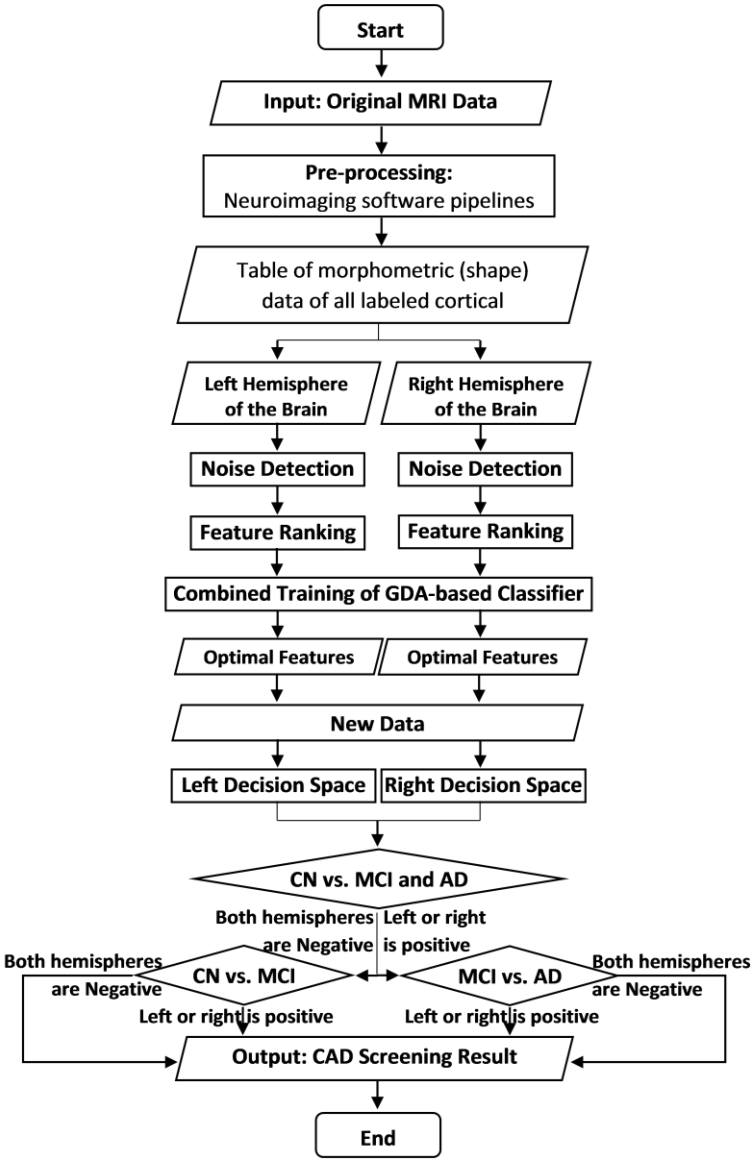


Fig.10.General Flowchart of the CAD System Based on the Proposed Classification Algorithm.

4.3.3. Classification Experiments

The performance of the proposed classification algorithm was measured by using the F1 score, accuracy, sensitivity, specificity, PPV, and NPV, based on a tenfold cross-validation process using 80% of the noise-free detected subject data points. All data points were randomly assigned to ten sets d_0, d_1, …, d_8 and d_9, so that all sets were of equal size. Then each one of the 10 sets was retained as the validation data while the remaining 9 sets

were used as training data. Therefore, on each fold, each data point was used for both training and validation. After the optimal sets of features were obtained, the remaining 20% of the noise-free detected data were utilized to estimate the classification performance.

The typical binary classification implemented in the past assumed two different groups of subjects at a time (i.e., CN vs. MCI, CN vs. AD and MCI vs. AD). However, for reasons explained earlier on the limited clinical use of such binary classifications, on the basis of a tenfold cross validation and a held-out test data set, this study implements a CAD system with a new classification process, where the CN group is delineated form the combined MCI and AD groups as a first classification phase, followed by the delineation of the MCI group in the second phase, using strategically the two hemispheres of the brain as distinct dimensional spaces in the decision process as shown in Fig. 10. It should be noted that the data of the left and right hemispheres of the brain were processed separately, except when we are generating the optimal set of features during the combined training of the GDA-based classifier. Furthermore, for the proposed classification algorithm, each hemisphere had its own decision space, and as long as one decision space produced a positive result, the tested subject should be classified as positive as well. This procedure could be considered as an innovative method, through which the classification performance is significantly improved as presented in the following section.

4.4. Results

The experimental results of the global feature selection reveal what are the most significant ROIs associated with the progression of AD. The evolution of all statistics in the incremental error analysis and the performance of the proposed classification algorithm are also presented.

4.4.1. Top-ranked Global Features

After pre-processing the original MRI images, 9 subjects were eliminated because of the noisy data which included measurements with zero values. Thus, 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 carried out on the two sets of 175-feature vectors corresponding to the left and right hemispheres of the brain between all three groups. All features found at 0.01 LOS out of all 175 measurements for each side of the brain were then used in the classification experiment as shown in Table 12.

The top-10 ranked features of each hemisphere of the brain are given in Table 13. Revealed from this table, the entorhinal cortex is the most significant cortical region associated with the progression of AD in both hemispheres of the brain. This observation

Table 12. Number of Significant Features Selected

Side of brain	Number of significant features (*p-value* < 0.01)
Left	71
Right	66

Table 13. Top-10 Ranked Features

Side of brain	Rank	Measurements	*p-value*
Left	1	Thickness of entorhinal	$< 10^{-33}$
	2	Thickness of middletemporal	$< 10^{-24}$
	3	Curvature of entorhinal	$< 10^{-24}$
	4	Thickness of inferior temporal	$< 10^{-21}$
	5	Curvature of middle temporal	$< 10^{-20}$
	6	Thickness of inferior parietal	$< 10^{-18}$
	7	Thickness of fusiform	$< 10^{-16}$
	8	Curvature of inferior temporal	$< 10^{-16}$
	9	Curvature of parahippocampal	$< 10^{-14}$
	10	Curvature of inferior parietal	$< 10^{-14}$
Right	1	Thickness of entorhinal	$< 10^{-29}$
	2	Thickness of middle temporal	$< 10^{-24}$
	3	Curvature of middle temporal	$< 10^{-22}$
	4	Curvature of entorhinal	$< 10^{-19}$
	5	Thickness of fusiform	$< 10^{-17}$
	6	Curvature of inferior temporal	$< 10^{-17}$
	7	Thickness of inferior parietal	$< 10^{-16}$
	8	Thickness of inferior temporal	$< 10^{-16}$
	9	Curvature of fusiform	$< 10^{-15}$
	10	Curvature of inferior parietal	$< 10^{-14}$

is consistent with recent studies indicating that indeed the entorhinal cortex is the first area to be implicated in AD [36-38]. Moreover, the entorhinal cortex has been proven as a major source of projections to the hippocampus [68], which plays an important role in converting short-term memory (also known as working memory) to long-term memory in the brain. The second top-ranked cortical region, the middle temporal, is also critical for long-term memory, to which the disrupted hippocampal connectivity has been found in the early stages of AD [70]. Involving the other three significant ROIs identified from the global feature ranking, the inferior temporal, fusiform, and parahippocampal, all top-5 ranked features are AD signature regions investigated previously [73], which provide credence to the validity of our global feature ranking method.

4.4.2. Optimal Feature Sets

The purpose of applying the incremental error analysis is to obtain the best classification performance with only few but essential features (i.e., the dimensional degree of the decision spaces). For generating the optimal feature sets, the tenfold cross validation retained 80% of the noise-free detected subjects included in this study (i.e., 619 = 187 CN + 301 MCI + 131 AD), involving 500 individuals (150 CN, 240 MCI, and 110 AD), all numbers were rounded to the nearest number divisible by 10 for the tenfold cross validation.

As mentioned in the Section 4.3, the performance of the proposed classification algorithm was estimated by the two-phase classification strategy that overcame the typical binary classification process. This facilitated the design of the CAD system shown in Fig. 10, and improved the classification results, especially as it pertains to the challenging delineation of MCI from the CN group. Therefore, for the tenfold cross validation, Fig. 11

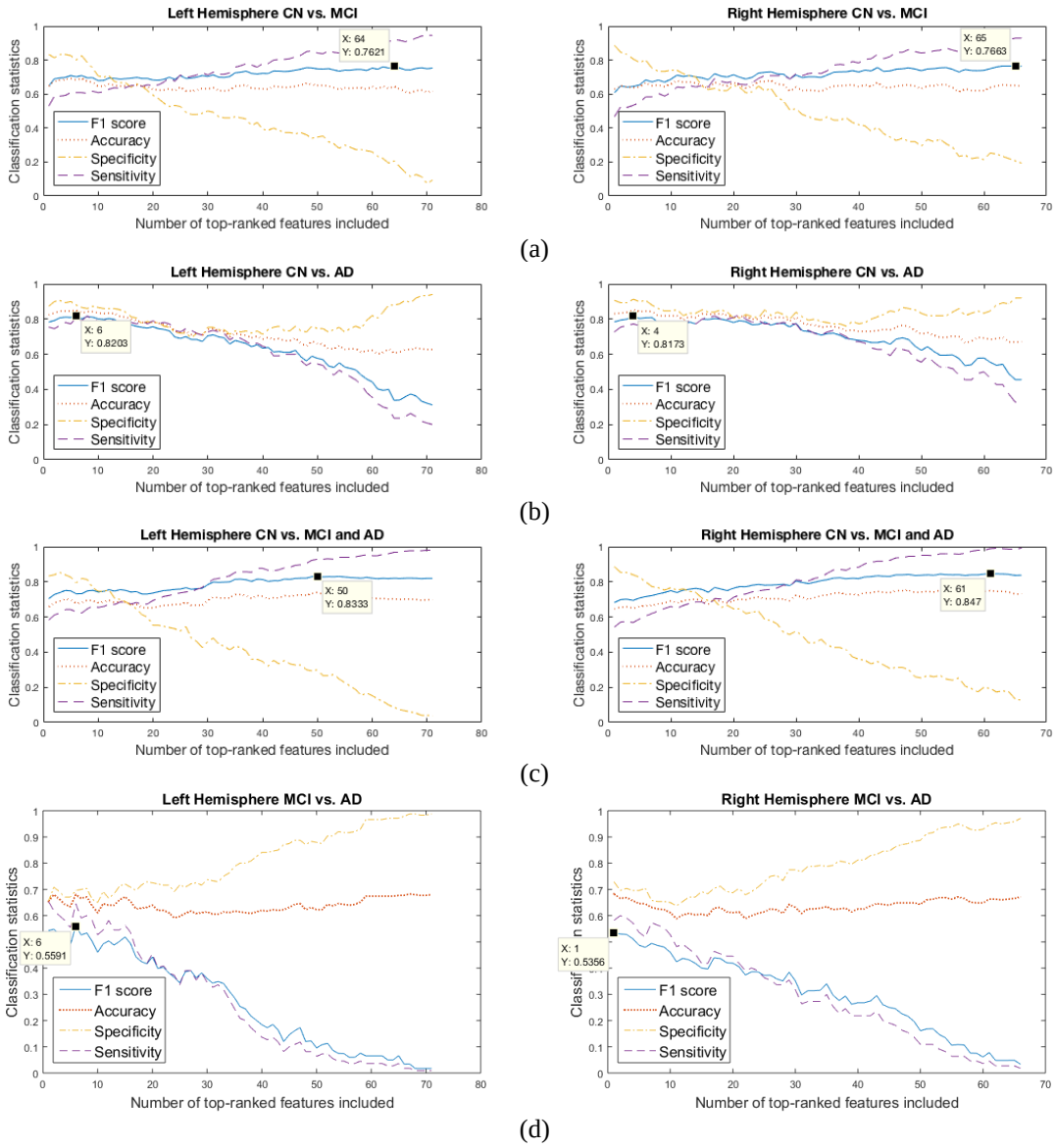
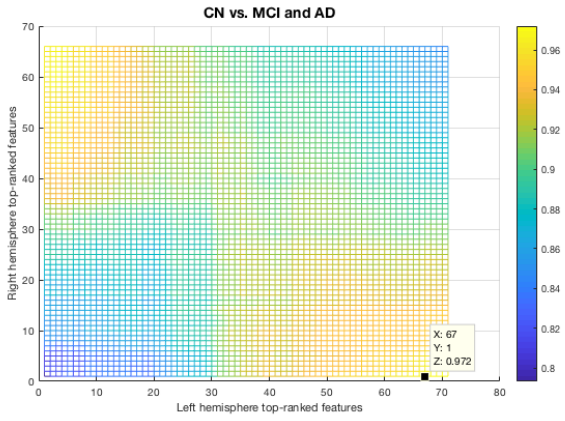


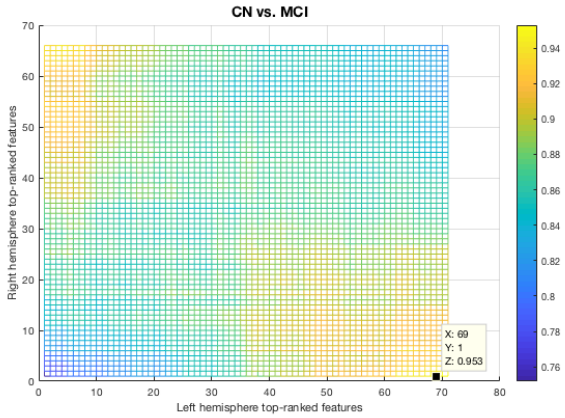
Fig. 11. Incremental Error Analysis Performance of Classification Statistics for Each Hemisphere: (a) CN vs. MCI, (b) CN vs. AD, (c) CN vs. MCI and AD, and (d) MCI vs. AD.

illustrates the evolution of some main classification statistics of each hemisphere, where the horizontal axis indicates the number of top-ranked significant features included in each iteration.

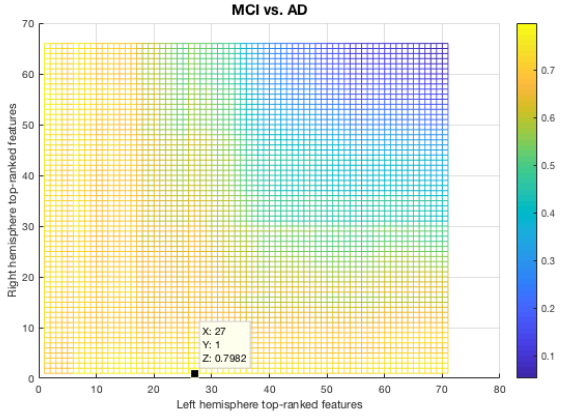
After combining the results of the two decision spaces together by the proposed CAD system, the evolution of F1 score of the three critical CAD classifications is thereafter demonstrated in Fig. 12. It can be observed that the final optimal feature sets are different



(a)



(b)



(c)

Fig. 12. Combining Incremental Error Analysis Performance of F1 Score: (a) CN vs. MCI and AD, (b) CN vs. MCI, and (c) MCI vs. AD.

from the ones obtained for each hemisphere before combining the two decision spaces. For each case, the classification performance was improved significantly as shown in Table 14. For delineating MCI from CN, significant enhancements in classification statistics were

Table 14. Summary of the Tenfold Cross Validation Performance

Classification	CN vs. MCI			CN vs. AD			CN vs. MCI and AD			MCI vs. AD		
Decision Space	*Left*	*Right*	*Comb.*	*Left*	*Right*	*Comb.*	*Left*	*Right*	*Comb.*	*Left*	*Right*	*Comb.*
F1 %	76.21	76.63	95.30	82.03	81.73	96.23	83.33	84.70	97.20	55.91	55.56	79.82
ACC %	64.62	65.13	94.10	85.00	85.38	96.92	74.00	75.00	96.00	68.00	68.29	87.43
SEN %	92.08	92.92	97.08	80.91	77.27	92.73	92.86	98.86	99.14	64.55	58.18	79.09
SPE %	20.67	20.67	89.33	88.00	91.33	100.0	30.00	19.33	88.67	69.58	72.92	91.25
PPV %	65.00	65.20	93.57	83.18	86.73	100.0	75.58	74.09	95.33	49.31	49.61	80.56
NPV %	62.00	64.58	95.04	86.27	84.57	94.94	64.29	87.88	97.79	81.07	79.19	90.50
Number of the optimal features	64	65	L: 69 R: 1	6	4	L: 4 R: 42	50	61	L: 67 R: 1	6	1	L: 27 R: 1

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

achieved, including an increase in F1 score average from 76.42% to 95.30% and increments of 29.23% for accuracy, 4.58% for sensitivity, 68.66% for specificity, 28.47% for PPV, and 31.75% for NPV, respectively. For discriminating MCI and AD together from CN group, an average increment of 24.03% for all classification statistics was obtained.

4.4.3. Classification Performance on Held-out Test Data

In order to obtain a reliable measure of the proposed classification algorithm performance, the remaining 20% of the noise-free detected data points were used as the held-out test data. Taking advantage of the obtained optimal sets of global features, the true test results are shown in Table 15. The classification performance was not as good as that achieved in the tenfold cross validation, but for the typically binary classification, the reported results are still very competitive in comparison to results reported in other recent studies also using ADNI database [12, 75], as indicated in Table 16. In comparison to such recent studies, the obtained results in delineating the MCI from CN are significantly

Table 15. Summary of the Classification Performance using the Held-out Test Data

Classification	ACC %	SEN %	SPE %	PPV %	NPV %
CN vs. MCI	88.78	91.08	83.78	90.32	86.11
CN vs. AD	93.10	90.48	94.59	90.48	94.59
CN vs. MCI and AD	93.28	98.78	81.08	92.05	96.77
MCI vs. AD	81.71	71.43	85.25	62.50	89.66

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

Table 16. Comparison of Binary Classification Performance with Some Recent Studies Using ADNI Data

Classification	CN vs. MCI			CN vs. AD		
Reference	*ACC %*	*SEN %*	*SPE %*	*ACC %*	*SEN %*	*SPE %*
H. Aidos et al., [12], 2017	61.90	54.70	69.20	84.40	76.90	91.90
J. Zhang et al., [75], 2017	79.02	90.46	59.90	88.30	79.61	93.69
Proposed Study	**88.78**	**91.08**	**83.78**	**93.10**	**90.48**	**94.59**

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

improved with an accuracy of 88.78%, a sensitivity of 91.08%, and a specificity of 83.78%, which are even more effective than the results achieved for discriminating AD from CN in those studies. For the CN vs. AD classification, performance improvement is also attained, especially for the sensitivity, where the proposed algorithm yielded an average increment of 12.23%. For delineating MCI and AD together from CN, an accuracy of 93.28%, a sensitivity of 98.78%, a specificity of 81.08%, PPV of 92.05%, and NPV of 96.77% were obtained. Then, an accuracy of 81.71%, a sensitivity of 71.43%, a specificity of 85.25%, PPV of 62.50%, and NPV of 89.96% were accomplished for discriminating AD from MCI. These results can be considered as the best CAD system performance achieved so far, which also suggest a high potential of the proposed CAD system to be applied as a practical clinical screening test for AD.

CHAPTER V. GAUSSIAN DISCRIMINATIVE COMPONENT ANALYSIS FOR EARLY DETECTION OF ALZHEIMER'S DISEASE: A SUPERVISED DIMENSIONALITY REDUCTION ALGORITHM

5.1. Goal

This chapter aims to introduce a novel supervised dimensionality reduction algorithm to characterize the important Gaussian discriminative components with respect to the structural, functional or metabolic measurements as observed in the MRI-PET combination associated with different stages of AD, focusing on the prodromal stage of MCI [51, 52]. The stage of MCI is subdivided into two stages, EMCI and LMCI, as defined in the ADNI database. Since alleviation of specific symptoms is possible through therapeutic interventions for some patients in the early or middle stages of AD, effective diagnosis of EMCI from CN group is essentially important for the planning of early treatment. However, instead of utilizing PCA computed variances to determine the significances of different components, the proposed Gaussian discriminative component analysis (GDCA) makes use of GDA classifiers to reveal the discriminability of different components in terms of each component's performance obtained by a designate machine learning task. The proposed method conducts a dimensionality reduction algorithm taking into consideration the interclass information to define an optimal eigenspace that maximizes the discriminability of selected eigenvectors. This process is shown to lead to stable, reliable and accurate dimensionality reduction in multimodal neuroimaging biomarkers for effective classification, enhanced diagnosis and the monitoring of disease progression. The proposed GDCA model has a high potential for deployment as a computer aided clinical diagnosis system for AD.

5.2. Materials

The information of the subjects used in this study, the MRI data pre-processing, and MRI/PET image registration procedure are presented in this section.

5.2.1. Participants and Clinical Data

The data used in conducting this study were collected from the ADNI database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-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 predict and gauge the progression of AD. A total of 906 subjects were considered for this study, which were categorized into groups of CN (251), EMCI (297), LMCI (196) and AD (162). All individuals underwent structural MRI and Florbetapir (F18-AV45) PET imaging, where the time gap between the two imaging modalities was less than 3 months. Details of MRI and AV45 PET data acquisition can be found on the ADNI website. Summary statistics and participants counts are listed in Table 17.

Table 17. Participant Demographic and Clinical Information

Symbol	CN (n=251)	EMCI (n=297)	LMCI (n=196)	AD (n=162)
F/M	128/123	132/165	85/111	68/94
Age_PET[b]	75.5 (6.5) [a]	71.5 (7.4)	73.8 (8.1)	74.9 (7.8)
Age_MRI[b]	75.3 (6.6)	71.3 (7.4)	73.6 (8.0)	74.7 (7.8)
Education	16.43 (2.6)	15.99 (2.7)	16.31 (2.7)	15.76 (2.7)
MMSE[cd]	29.04 (1.2)	28.32 (1.6)	27.61 (1.9)	22.77 (2.7)
RAVLT_immediate[cd]	45.3 (10.6)	39.5 (10.8)	33.2 (10.8)	22.3 (7.0)

[a] Values are represented as mean (standard deviation), except gender (F for female, M for male), which are frequencies instead
[b] Significant group differences (ANOVA for continuous and Chi-square test for categorical values, significance level is 0.05 by default)
[c] Significant group differences (ANCOVA adjusted for Age_PET)
[d] Significant differences for all between-group post-hoc tests (Tukey's HSD test)

5.2.2. Image Processing

5.2.2.1. MRI Data Pre-processing

The FreeSurfer (Version 5.3.0) was firstly performed under Linux system (centos4_x86_64) to transform the original MRI to the standard MNI 305 space, yielding the image referred to as T1.mgz, which is used as the reference image in the registration procedure, followed by skull-striping, segmenting, and delineating cortical and subcortical regions with the corresponding image result termed as aparc+aseg.mgz. Derived from the images, the following three shape measures were then calculated as morphological features on each of the 68 FreeSurfer labeled cortical regions for both hemispheres (34 per hemisphere): 1) cortical thickness, 2) surface area, and 3) cortical volume. Since version 5.3 of FreeSurfer was available, we tested the same data with FreeSurfer 6.0 and found minimal differences ranging from 1 to 5% and showing no statistical differences in terms of standardized uptake value ratio measurements (SUVRs).

5.2.2.2. MRI and PET Registration

With 12 degrees of freedom (DOF) onto the postprocessed T1 image, the AV45 PET was linearly registered (using trilinear interpolation), so that the regional amyloid deposition and gray matter atrophy are compared directly (i.e., thickness for cortical regions [76-79]), using the FMRIB Software Library (FSL) [80]. Moreover, in order to gain as much information as possible from PET images, which have relatively low resolution, the original AV45 PET with skull was utilized in this step. This registration process introduced in a recent study [81] guaranteed that AV45 PET image had the same segmentation and parcellation as the MRI image.

Combined with aparc+aseg.mgz images, the registered AV45 PET was inspected to obtain the mean standardized uptake values (SUV) for all 68 FreeSurfer labeled cortical regions. The SUV of the whole cerebellum, including 4 regions of interest (left/right cerebellum cortex and left/right white matter), was used as the reference region. Finally, regional SUVs of those 68 cortical regions were normalized by the SUV of the whole cerebellum to get the cortical-to-cerebellum SUVRs. Accordingly, overall there are 4 different types of neuroimaging features associated with each of the 68 cortical regions, yielding 272 (4×68) features for each subject in the dataset.

5.3. Methods

After obtaining all needed features derived from raw multimodal neuroimaging data, as aforementioned, a 272-dimensional feature vector was generated for each subject in the data set. In this section, the proposed GDCA algorithm is presented for the effective dimensionality reduction and early diagnosis of AD. The standard PCA is applied to the original data to find the principal components. Then, the discriminability of each component is estimated by a one-dimensional GDA classifier, and consequently, all components are sorted in order of the corresponding classification performance. Finally, the recursive feature elimination (RFE) is employed to determine the optimal dimensionality reduction of the Gaussian discriminant components in the classification outcome. Fig. 13 demonstrates the flowchart of the proposed GDCA model.

5.3.1. Gaussian Discriminative Component Analysis

5.3.1.1. Eigenvectors of the Covariance Matrix

The proposed classification problem can be formulated by having the machine learn to distinguish between CN ($y = 0$), EMCI ($y = 1$), LMCI ($y = 2$), and AD ($y = 3$), based

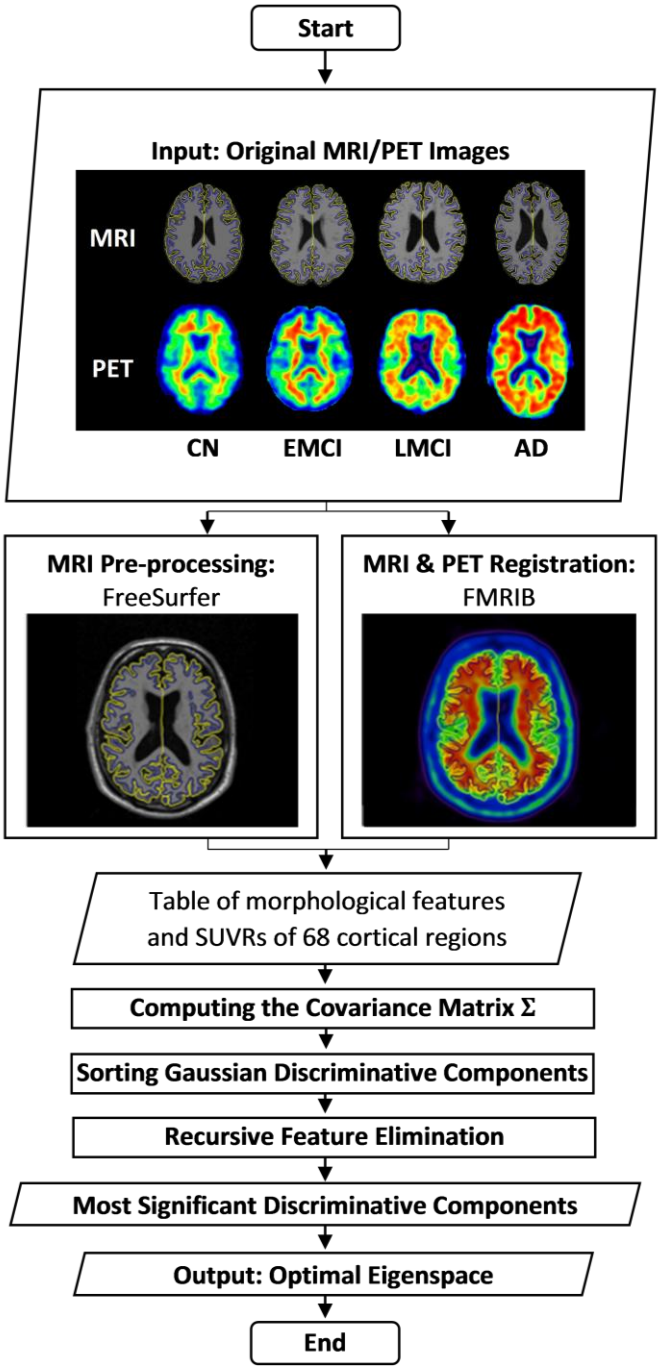


Fig. 13. General Flowchart of the Proposed GDCA Algorithm.

on the extracted features $x \in \mathbb{R}^n$. In order to determine the potential directions of Gaussian discriminative components of all features, the standard PCA method is carried out. Prior to running PCA, the data need to be normalized as follows:

$$\overline{x} = \frac{x - \mu}{\sigma} \tag{11}$$

where $\mu \in \mathbb{R}^n$ and $\sigma \in \mathbb{R}^n$ are the mean vector and standard deviation vector of all data, respectively. This process zeros out the mean of the data, and rescales each feature to have unit variance, which ensures different features to have the same scale. After normalization, the covariance matrix Σ can then be computed utilizing the normalized data by the formula below:

$$\Sigma = \frac{1}{m}\sum_{i=0}^{m} \overline{x}_i \cdot \overline{x}_i{}^T \tag{12}$$

where m is the total number of data points considered and $\overline{x}_i{}^T$ is the transpose of the normalized data point $\overline{x}_i$. Then to project the original data into a *k*-dimensional subspace ($k \leq n$), the eigenvector $u_j \in \mathbb{R}^n$ $(j \leq k)$ of the covariance matrix Σ can be computed to obtain the transformed features $x' \in \mathbb{R}^k$.

5.3.1.2. Supervised Dimensionality Reduction

As indicated earlier, the PCA model sorts the extracted eigenvectors (i.e., the direction of principal components) based on the variance represented by each eigenvector, without considering any information from the labels of data as an unsupervised algorithm. But, in general, only reducing the dimensionality to retain as much as possible of the variance cannot help in deciding the optimal subspace towards an optimal performance if a supervised machine learning scenario is contemplated. As a consequence, the proposed method capitalizes on a supervised dimensionality reduction model making use of a GDA-based classifier. Given the eigenvectors which were computed based on the covariance matrix Σ given in (12), the GDA model will be trained on each new feature in the

transformed space to determine the discriminability of each component according to the corresponding classification performance, subsequently sorting the extracted principal components in order of their discriminability.

GDA can model $p(x'|y)$, the distribution of the feature vector x' in the transformed feature space given $y \in \{0,1,2,3\}$, assumed to be distributed according to a Gaussian distribution, with the generalized density function given in (13):

$$p(x'; \mu', \Sigma') = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2}(x' - \mu')^T \Sigma'^{-1}(x' - \mu')\right) \tag{13}$$

where $\mu' \in \mathbb{R}^k$ is the mean vector in the new transformed feature space, Σ' is the new covariance matrix, $|\Sigma'|$ and Σ'^{-1} denote the determinant and inverse matrix of Σ', respectively. To determine the discriminability of each component, since $x' \in \mathbb{R}^1$, the μ' is the mean of the transformed feature, and Σ' is the variance of the transformed feature. After modelling $p(x'|y)$, Bayes rule is used to derive the posterior distribution on y given x' as:

$$p(y|x') = \frac{p(x'|y)p(y)}{p(x')} \tag{14}$$

Here, $p(y)$ denotes the class prior distribution, which cannot be determined when given a certain subject, so it is assumed to be absolutely random (for all $a \neq b, p(y = a) = p(y = b)$). Furthermore, to make a prediction, it is not necessary to calculate $p(x')$, since

$$\arg\max_y p(y|x') = \arg\max_y \frac{p(x'|y)p(y)}{p(x')} = \arg\max_y p(x'|y)p(y) \tag{15}$$

Therefore, for classification purposes, the following formula is used instead:

$$\arg\max_y p(y|x') = \arg\max_y p(x'|y) \tag{16}$$

5.3.1.3. Recursive Component Elimination

The GDA classifier is applied to each component, so that for each eigenvector, the transformed features can be ranked in terms of classification outcome using cross validation. In this study, the classification accuracy is used as the key metric to measure performance, which means the discriminability of each component is determined by its corresponding classification accuracy expressed as Equation (5) in Chapter III.

Setting the computed accuracies as assigned weights to discriminative components, the RFE is performed to select the optimal Gaussian discriminative components by recursively considering smaller and smaller sets of components. First, the entire set of components were applied to the classifier and estimated by the cross-validation performance. Then, the least important component is eliminated from current set of components. That procedure is recursively repeated on the pruned set until the desired set of Gaussian discriminative components is found with an optimal classification performance.

5.3.2. Classification Based on GDCA

From the proposed GDCA dimensionality reduction model, the optimal components are obtained in terms of the classification performance (i.e., the accuracy of that set of components selected), and would then be applied to other classification algorithms. Some other metrics are used as well, since, as a clinical application, the classification performance may not only be evaluated by the accuracy, but could also rely on precision, recall (or sensitivity) and specificity. The F1 score is also a widely used measure of performance in statistical analysis of binary classification, by which both

precision and recall are taken into consideration. The formulas used to calculate these four metrics are expressed as Equations (6)-(10) in Chapter III.

In order to assess the ability of the obtained transformed feature space in performance improvement, several widely used classification algorithms are applied on the original feature space as well as the dimensionality reduced new feature space, including linear SVM, multilayer perceptron (MLP), and gradient boosting (GB) classifiers. To demonstrate the advantage of the proposed GDCA over other widely used dimensionality reduction methods, PCA and univariate feature selection are carried out on the best performed classifier among the ones mentioned above.

5.4. Experiments and Results

The focus of this study is placed on demonstrating how the proposed dimensionality reduction model can determine the most discriminative components associated with the progression of MCI and improve the classification performance. Also, in order to predict the progression of AD, a multiclass classification was carried out on those three groups of AD patients (i.e., EMCI, LMCI, AD), therefore, we could further compare the proposed dimensionality reduction with other widely used methods based on the multiclass classification performance. Scikit-learn, free software machine learning library, was used to implement all classification algorithms with cross validation procedure and built-in experiment pipeline [82]. In the classification experiments, all subjects were randomly split into training, validation, test sets with 80% of the data used for training, 10% for cross validation, and 10% for the held-out true test. In order to demonstrate the advantage of the proposed GDCA method over other dimensionality reduction methods, 5-fold cross validation was carried out using PCA, univariate feature selection and the proposed GDCA

Table 18. The Classification Accuracy of Top-10 Gaussian Discriminative Components and the Corresponding PCA Rank

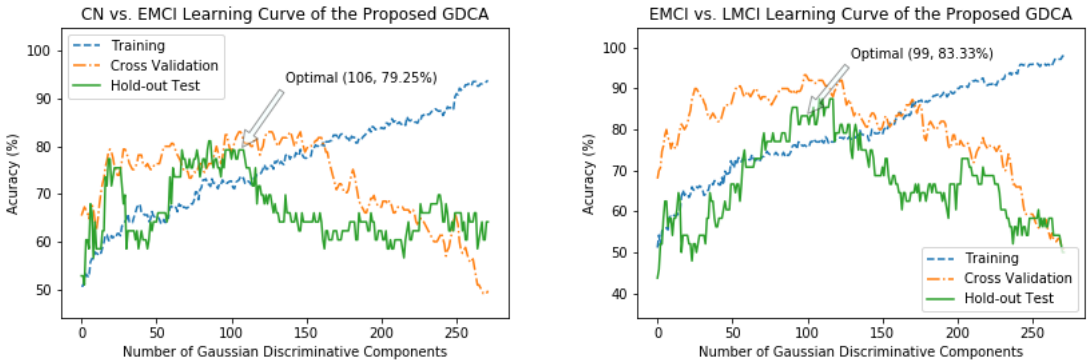
CN vs. EMCI			EMCI vs. LMCI		
GDCA Rank	Accuracy	PCA Rank	GDCA Rank	Accuracy	PCA Rank
1	65.45%	22	1	68.00%	204
2	65.45%	186	2	66.00%	9
3	63.64%	64	3	66.00%	35
4	63.64%	148	4	66.00%	105
5	63.64%	207	5	66.00%	132
6	63.64%	241	6	64.00%	64
7	63.64%	262	7	64.00%	170
8	63.64%	267	8	64.00%	239
9	61.82%	6	9	62.00%	3
10	61.82%	62	10	62.00%	74

methods. Finally, a CAD application for detecting different stages of AD was presented to reveal the potential of this GDCA model to be deployed as a CAD system.

5.4.1. Gaussian Discriminative Components

Given the eigenvectors of the covariance matrix calculated by the whole data, Table 18 shows the classification accuracy of top-10 Gaussian discriminative components based on the binary classification (i.e., CN vs. EMCI, EMCI vs. LMCI) on the validation data, and the PCA rank of these components are also provided to demonstrate the difference between GDCA and PCA. As shown in Table 18, the principal components with higher variance do not necessarily yield better performance in the classification task than those with lower variance, which may help in delineating the subtle changes associated with CN vs. EMCI and with EMCI vs. LMCI.

With the Gaussian discriminative components ranked, the RFE was applied on the validation data to find the optimal set of components that yielded the best validation performance in terms of overall classification accuracy. Consequently, these optimal discriminative components were used to evaluate the proposed GDCA on the held-out test data. Fig. 14(a) illustrates the CN vs. EMCI learning curves of the training, validation and testing when increasing the number of Gaussian discriminative components involved in the



(a) (b)

Fig. 14. The Learning Curves of the Training, Validation and Testing with Different Numbers of Gaussian Discriminative Components: (a) CN vs. EMCI classification; (b) EMCI vs. LMCI classification.

Table 19. The Benchmark CN vs. EMCI and EMCI vs. LMCI Classification Performance Based on the Proposed GDCA Model

Classification	CN vs. EMCI				EMCI vs. LMCI			
Performance	F1 Score	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall
Cross Validation	86.15%	83.64%	80.00%	93.33%	92.31%	94.00%	94.74%	90.00%
Hold-out Test	80.70%	79.25%	82.14%	79.31%	77.78%	83.33%	82.35%	73.68%

classifier. It can be observed that the proposed model was able to learn the generic discriminative components through the cross validation and performed similarly on the held-out test data. Based on the cross validation, the highest accuracy of 79.25% was obtained by using the first 106 Gaussian discriminative components. The GDCA results are shown in Table 19, which also sets a performance benchmark for further classification performance comparison using several different machine learning algorithms. Another challenging task of detecting different stages in AD is in distinguishing LMCI from EMCI, because LMCI may have higher risk in developing AD. Thus, EMCI vs. LMCI classification was carried out following the same procedure, and the results are illustrated in Table 19 and Fig. 14(b), where the best cross validation performance was attained by including the first 99 Gaussian discriminative components into the model with an accuracy of 83.33%.

Table 20. Binary Classification Performance Comparison of Original Features and GDCA-transformed Features

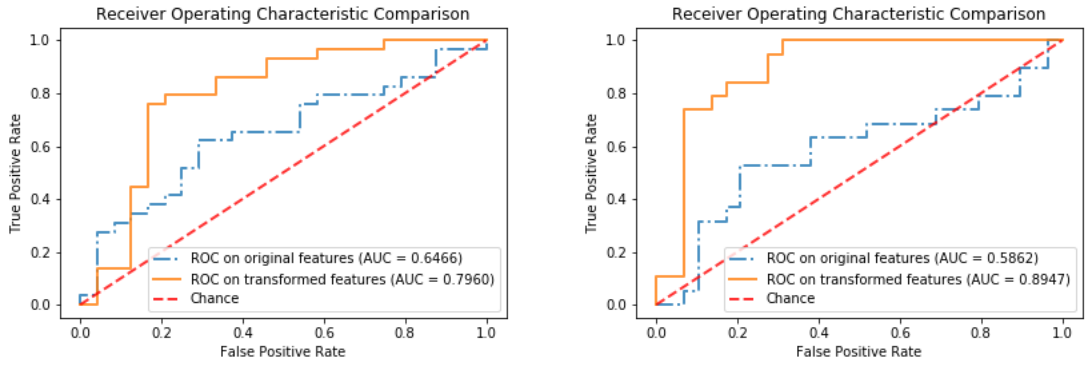
Task	Feature	Original features				Transformed features			
	Classifier	F1 Score	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall
CN vs. EMCI	SVM	72.73%	66.04%	64.86%	**82.76%**	78.69%	75.47%	75.00%	**82.76%**
	MLP	**75.41%**	**71.70%**	**71.88%**	79.31%	78.57%	77.36%	81.48%	75.86%
	GB	**75.41%**	**71.70%**	**71.88%**	79.31%	77.19%	75.47%	78.57%	75.86%
	GDA	66.67%	64.15%	67.86%	65.52%	**80.70%**	**79.25%**	**82.14%**	79.31%
EMCI vs. LMCI	SVM	54.05%	64.58%	55.56%	52.63%	65.00%	70.83%	61.90%	68.42%
	MLP	**59.46%**	**68.75%**	**61.11%**	57.89%	72.73%	75.00%	64.00%	**84.21%**
	GB	48.48%	64.58%	57.14%	42.11%	60.00%	75.00%	81.82%	47.37%
	GDA	52.00%	50.00%	41.94%	**68.42%**	**77.78%**	**83.33%**	**82.35%**	73.68%

5.4.2. Binary Classification Performance Comparison

By applying the relevant classifiers (i.e., SVM, MLP, and GB) to the original data and to the dimensionality-reduced data, the corresponding results are given in Table 20. Unlike the proposed GDCA, these algorithms may give us various results due to the random initialization. The classification experiments were run multiple times, and the overall best results were reported in Table 20. It can be observed that after introducing the proposed dimensionality reduction model, all the selected classifiers achieved better performance on the transformed feature space than obtained on the original features, which adds credence to the validity of the proposed GDCA model. Moreover, although state-of-the-art MLP and GB algorithms established better performance than the GDA algorithm on the original features as a result of the underlying feature selection process, for both CN vs. EMCI and EMCI vs. LMCI, they did not surpass the benchmark performance yielded by the proposed GDCA algorithm. However, because of the random initialization, classification algorithms like SVM, MPL, GB may not always achieve the global optimal solution, only the GDA classifier is applied here for the multiclass classification experiment.

As another widely used metric in choosing binary classification models, the receiver operating characteristic (ROC) curve and the area under the curve (AUC) were

used to measure the classification performance. The AUC score can reveal the discriminability of a classification model and to indicate if the false positive and true positive rates achieved by a model are significantly above random chance. The ROC curves and the corresponding AUC scores of held-out tests on original and transformed feature spaces are demonstrated in Fig. 15, and it can be observed that, after carrying out the proposed GDCA model, the AUC scores improved significantly by 0.15 for CN vs. EMCI classification and by 0.31 for EMCI vs. LMCI classification. In Table 21, the results obtained by the proposed GDCA model are compared with those obtained using most recent state-of-the-art methods based on ADNI data [83-89]. It should be noted that, as



(a) (b)

Fig. 15. ROC Curves and AUC Scores on Original Features and GDCA Transformed Features: (a) CN vs. EMCI classification; (b) EMCI vs. LMCI classification.

Table 21. CN vs. EMCI and EMCI vs. LMCI Classification Performance Comparison with Recent State-of-the-art Studies Based on ADNI Data

Classification	Subjects (CN/EMCI/LMCI)	CN vs. EMCI				EMCI vs. LMCI			
Performance		Accuracy	Recall	Specificity	AUC	Accuracy	Recall	Specificity	AUC
Pei et al. [83], 2018	-/18/18	-	-	-	-	70.00%	-	-	0.7088
Hett et al. [84], 2019	62/65/34	-	-	-	-	70.80%	-	-	0.6240
Jie et al. [85], 2018	50/56/43	-	-	-	-	74.80%	-	-	0.7200
Jie et al. [86], 2018	50/56/43	78.30%	74.00%	**82.10%**	0.7710	78.80%	82.10%	74.40%	0.7830
Wee et al. [87], 2019	300/314/208	53.00%	60.40%	55.00%	-	63.10%	61.30%	77.60%	-
Yang et al. [88], 2019	29/29/18	77.59%	59.09%	-	0.6849	76.60%	66.20%	-	0.7682
Kam et al. [89], 2019	48/49/-	76.07%	76.27%	75.87%	-	-	-	-	-
Proposed	251/297/196	**79.25%**	**79.31%**	79.17%	**0.7960**	**83.33%**	**82.35%**	**89.66%**	**0.8947**

shown in Table 21, although most of the studies used relatively small dataset, the proposed model still achieved overall best performance for both CN vs. EMCI classification and EMCI vs. LMCI classification; and for the only study having the same number of subjects [87], the proposed study obtained significantly better performance.

5.4.3. EMCI vs. LMCI vs. AD Multiclass Classification

The same pipeline was followed for the multiclass classification experiments, and since the F1 score, precision, and recall would no longer be available, the confusion matrix was used instead to evaluate the performance with each row corresponding to the true class. The diagonal elements of the confusion matrix represent the number of points for which the predicted label is equal to the true label, while off-diagonal elements are those that are misclassified by the classifier. Fig. 16 demonstrates the learning curve of the multiclass classification experiment using the proposed GDCA, where the best cross validation

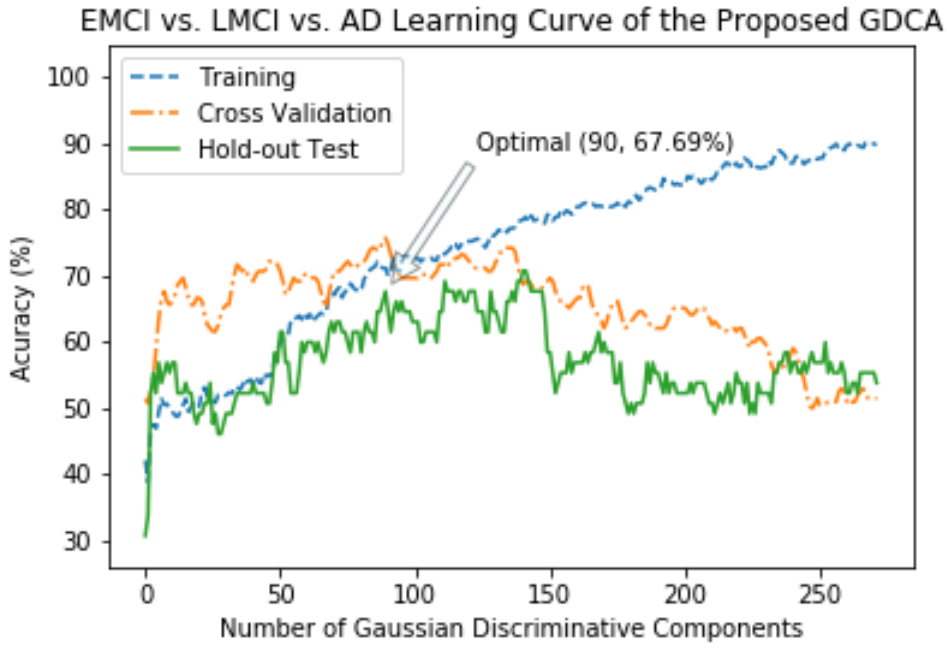
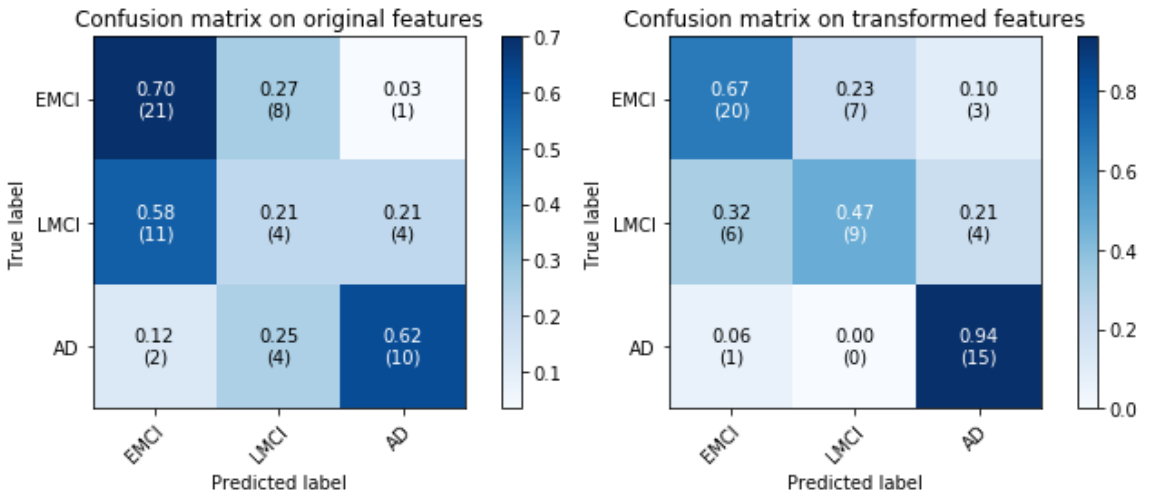


Fig. 16. The Learning Curves of the Training, Validation and Testing with Different Numbers of Gaussian Discriminative Components for EMCI vs. LMCI vs. AD Classification.



(a) (b)

Fig. 17. EMCI vs. LMCI vs. AD Classification Confusion Matrices: (a) All features were used; (b) GDCA-transformed features were used.

Table 22. MCI vs. AD Classification Performance by Converting the EMCI vs. LMCI vs. AD Classification Results

Features	F1 Score	Accuracy	Precision	Recall
Original	64.52%	83.08%	66.67%	62.50%
Transformed	78.95%	87.69%	68.18%	93.75%

performance was achieved by using the first 90 Gaussian discriminative component. It can be observed that the learning curve associated with the held-out test is closer to the learning curve of cross validation in comparison to the learning curve results shown in Fig. 14, since there were three classes instead of two classes, which enabled the model to learn more generic discriminative components across all three classes.

Fig. 17 shows the confusion matrices of the held-out test on the original features and GDCA transformed features. The overall classification accuracy using the transformed features was 67.69%, compared to 53.85% if all original features were utilized. As shown in Fig. 17, after applying the proposed GDCA model, the classifier could more precisely distinguish LMCI and AD from EMCI group, so that the overall classification performance was improved significantly. Additionally, in Table 22, the multiclass classification results

are converted to binary classification results of MCI vs. AD, showing that the proposed method could effectively discriminate AD from MCI with a 31.25% increase on recall.

5.4.4. Dimensionality Reduction Performance Comparison

Since the proposed GDCA method is capable of defining the most discriminative directions of all eigenvectors, noted improvements were obtained in the classification results. To demonstrate how this process differs from other widely used dimensionality reduction methods, 5-fold cross validation was implemented for the EMCI vs. LMCI vs. AD multiclass classification task by applying the PCA, univariate selection and proposed GDCA methods. The PCA method, as aforementioned, utilizes PCA computed variances to determine the significances of the principal components. For univariate selection method, the eigenvectors of the covariance matrix are not computed, and instead it selects the best features based on univariate statistical tests. In this study, the ANOVA was performed as the univariate statistical test to determine the significances of the different features.

Moreover, rather than adding one feature at a time, the different percentiles were used to illustrate the classification performance of these dimensionality reduction methods varying the percentile of features selected. The same GDA classifier was applied to all these three dimensionality reduction methods so as to eliminate any bias. The 5-fold cross validation results of these methods are shown in Fig. 18. As demonstrated in Fig. 18, for the univariate selection method, by selecting one percent of features, it achieved better performance than PCA and GDCA methods, but when adding more features, it was unable to improve the performance as significantly as did the PCA and GDCA methods. Overall, the proposed GDCA method obtained an average accuracy of 63.13% with 20 percent

dimensions of all features, which is significantly better than the best performance achieved by the PCA and univariate methods. The difference between these dimensionality

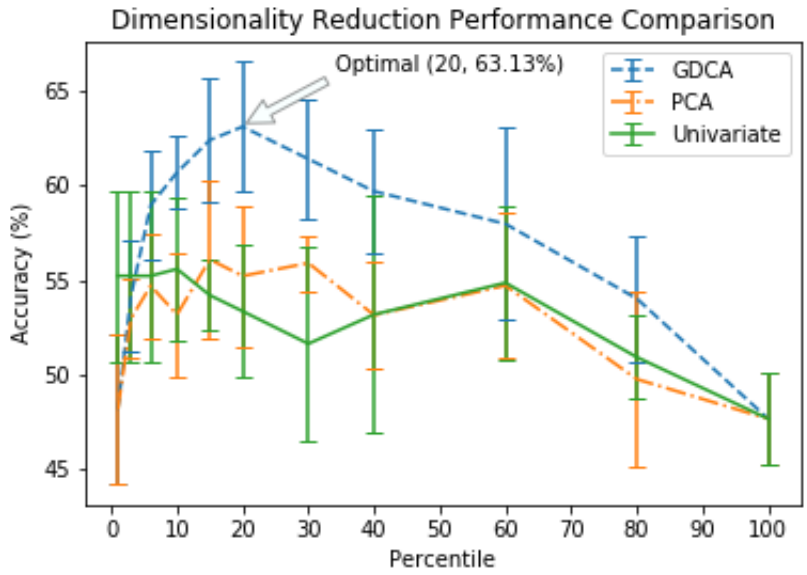


Fig. 18. EMCI vs. LMCI vs. AD Cross Validation Performance of Different Dimensionality Reduction Methods Varying the Percentile of Features Selected.

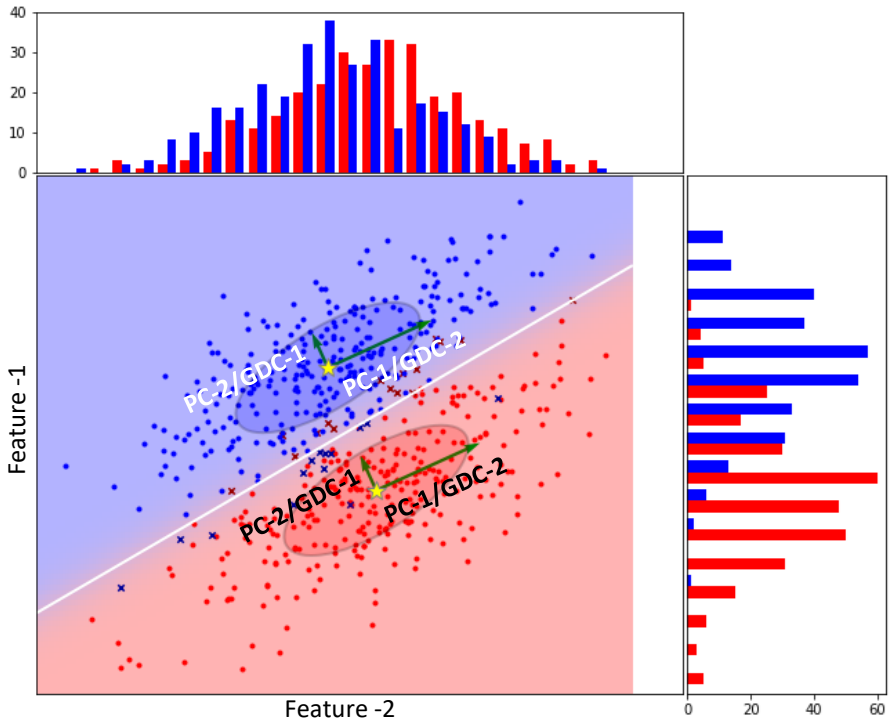


Fig. 19. A Visualized Example of Different Dimensionality Reduction Methods.

reduction approaches are also visualized in Fig. 19, in this example, there were two features associated with two classes, that required to reduce the two-dimensional space to its one-dimensional representation. It can be observed that the proposed GDCA approach would yield the best performance, since it selected the most discriminative component (i.e., GDC-1), and in contrast, the PCA method chose the second discriminative component (i.e., PC-1) and the univariate method picked the most discriminative feature (i.e., feature-1) from the original features.

5.4.5. Computer Aided Diagnosis Based on GDCA

The previous sections have indicated that the proposed GDCA model was able to identify the most discriminative components associated with different stages of AD as a multiclass classification problem. But, in order to apply the proposed model to a practical CAD system, the trained model should be able to include the CN group, allowing a given

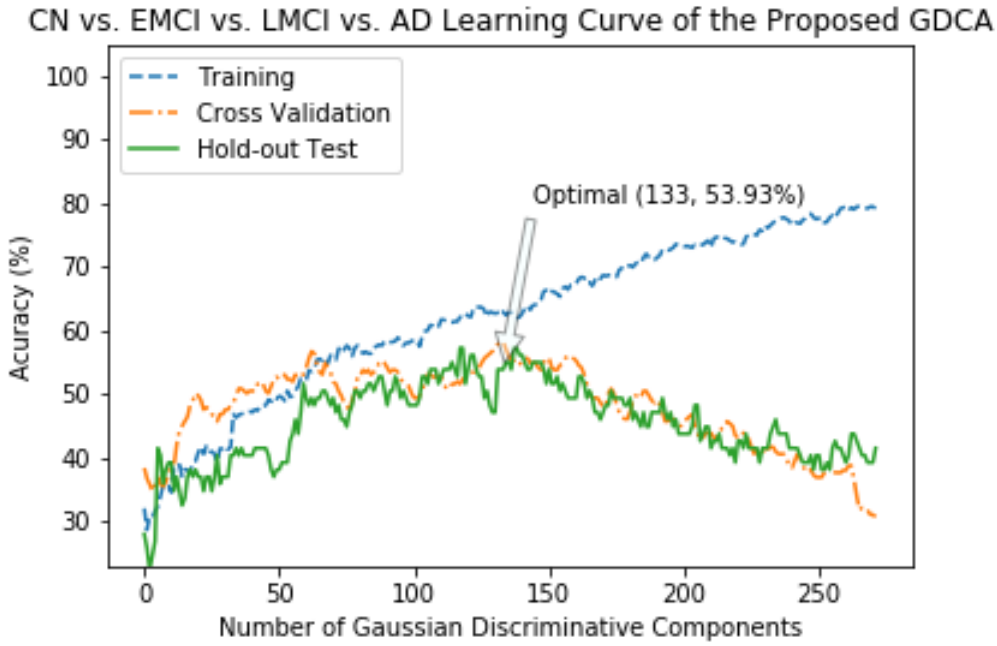
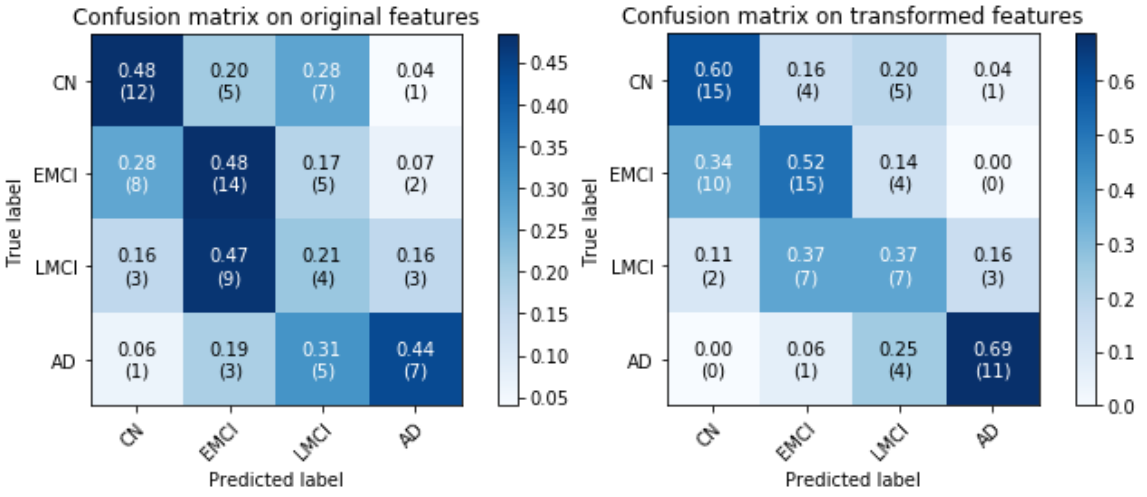


Fig. 20. The Learning Curves of the Training, Validation and Testing with Different Numbers of Gaussian Discriminative Components for the Proposed GDCA-based CAD Application.

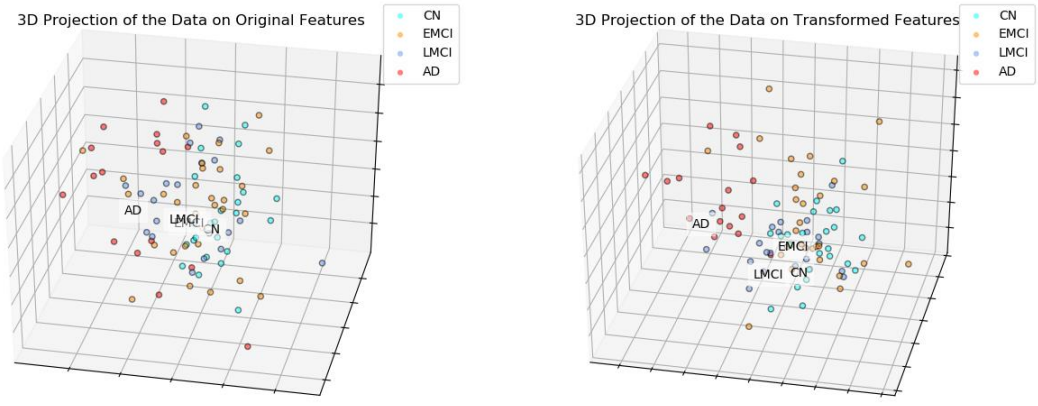
subject in the classification process to belong to any of the 4 groups: CN, EMCI, LMCI and AD. Therefore, in this section, a multimodal multiclass classification neuroimaging CAD application involving all four groups (CN, EMCI, LMCI and AD) is presented utilizing the proposed GDCA model.

The learning curve of the GDCA-based CAD application is shown in Fig. 20, where the best cross validation performance was obtained by using the first 133 Gaussian discriminative components. Now, since more interclass information was involved during



(a) (b)

Fig. 21. CN vs. EMCI vs. LMCI vs. AD Classification Confusion Matrices: (a) All features were used; (b) GDCA-transformed features were used.

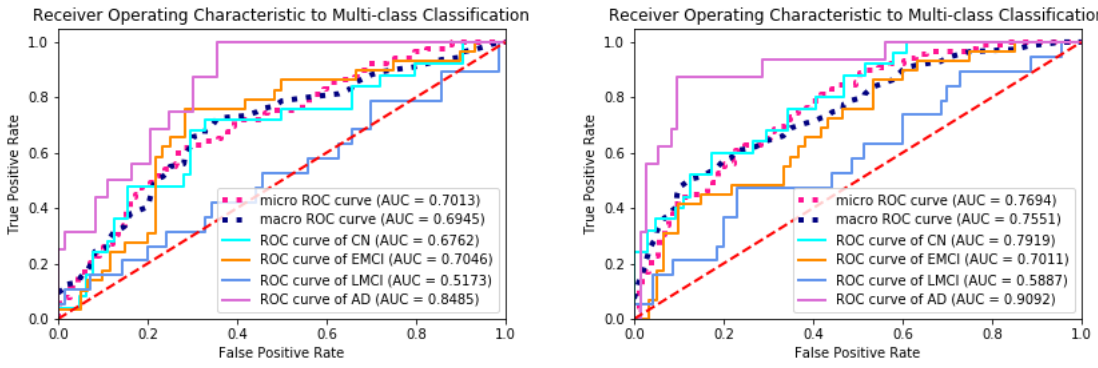


(a) (b)

Fig. 22. CN vs. EMCI vs. LMCI vs. AD 3-dimensional Visualization by Projecting the Data onto the Affine Subspace: (a) All features were used; (b) GDCA-transformed features were used.

the training, more generic discriminative components across all four classes were captured, which resulted in a small gap between the learning curves of the cross validation and the held-out test. Fig. 21 demonstrates the confusion matrices of the held-out true test on the original features and GDCA transformed features. As the most complicated task in AD classification, the accuracy of 53.93% was attained, which reached only 41.57% when all original features were used. Making use of GDA, Fig. 22 illustrates the 3-dimensional visualization by projecting the high dimensional data onto the affine subspace generated by the estimated class means of all classes. In Fig. 21 and Fig. 22, it can be observed that, after applying the proposed GDCA model, the classifier could detect the subtle difference between MCI group (i.e., EMCI and LMCI) and CN group as well as MCI group and AD group more effectively, in particular, more CN and AD subjects were correctly detected.

Furthermore, in order to illustrate the performance improvement of the GDCA-based CAD application, some extension of ROC to multiclass classification were carried out, including, one-against-rest ROC curve for each class, micro-averaging and macro-averaging ROC curves. Micro-averaging considers each element of the label indicator matrix as a binary prediction, while macro-averaging gives equal weight to the

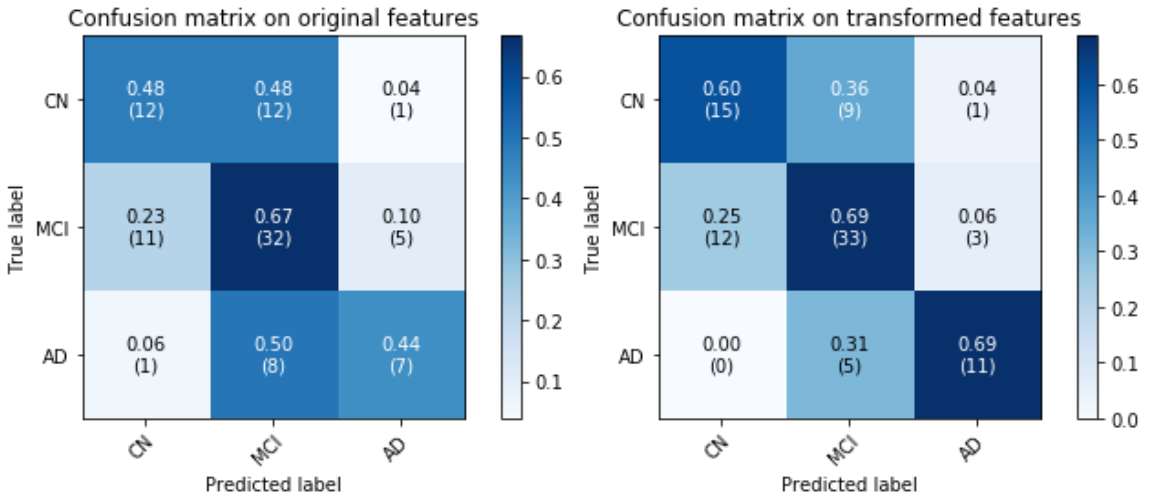


(a) (b)

Fig. 23. ROC Curves to Multiclass Classification and AUC Scores for the Proposed GDCA-based CAD Application: (a) All features were used; (b) GDCA-transformed features were used.

classification of each label. The ROC curves and the corresponding AUC scores are demonstrated in Fig. 23, and it can be observed that, after carrying out the proposed GDCA model, the micro-averaging and macro-averaging AUC scores were increased significantly by 9.71% and 8.73%, respectively. For AD vs. rest and CN vs. rest, the performances were also improved significantly, and AUC scores of 0.7919 and 0.9092, respectively were achieved.

As shown in Fig. 22, after applying the GDCA model, the classification improvement was attributed to more of the CN and AD subjects correctly distinguished from EMCI and LMCI groups. Therefore, in order to demonstrate the performance improvement on CN vs. MCI vs. AD classification, the CAD results of EMCI and LMCI were combined together as MCI. By combining those results, the confusion matrices on the original features and GDCA transformed features are shown in Fig. 24. After combining, the overall classification accuracy on original and transformed features was 57.30% and 66.29%, respectively. And more notably, if the MCI and AD results were



(a) (b)

Fig. 24. CN vs. MCI vs. AD Classification Confusion Matrices by Combining EMCI and LMCI: (a) All features were used; (b) GDCA-transformed features were used.

further combined as diseased group, it indicates that the proposed GDCA-based CAD application can effectively discriminate diseased subjects from the CN group with an accuracy of 75.28%, an F1 score of 82.51%, a precision of 83.87%, and a recall of 81.25%. These results show that the proposed GDCA model has a high potential for use as a clinical CAD system using multimodal neuroimaging data.

CHAPTER VI. COMPUTERIZED NEUROPSYCHOLOGICAL ASSESSMENT IN MILD COGNITIVE IMPAIRMENT BASED ON NATURAL LANGUAGE PROCESSING -ORIENTED FEATURE EXTRACTION

6.1. Goal

Typically, neuropsychological testing helps medical experts situate a given patient in continuum of the AD spectrum, especially in the continuum between CN and the prodromal stage of MCI. This chapter focuses on developing an NLP-oriented computerized neuropsychological assessment taking advantage of the existing NLP techniques to find significant patterns of changes in linguistic complexity associated with the progression of AD. By applying NLP techniques, it is possible to precisely diagnose MCI based on the analysis of some conversation contents, such as, spoken features, lexical features and syntactic features [62]. Some case studies have indicated that the NLP technic is able to find clear patterns of decline in syntactic and grammatical complexity [60, 61]. The proposed system will extract key features from discourse transcripts, and evaluate on non-scripted news conferences from President Ronald Reagan (RR), who was diagnosed with Alzheimer's disease in 1994, and some other presidents, who have no known diagnosis of AD, then indicate that over time, the patterns of the linguistic complexity changes are significantly distinguishable between RR and CN people.

6.2. Methodology

The innovative work for the proposed study is the development of a computerized neuropsychological test taking advantage of advances in state-of-the-art NLP technologies, so the test can automatically extract key features and discover the significant linguistic

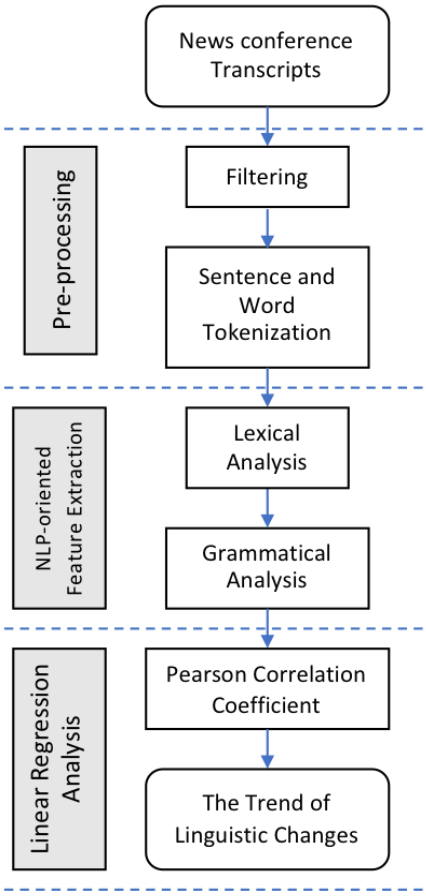


Fig. 25. General Framework of the Proposed Computerized Assessment.

changes associated with MCI. The general framework of the proposed NLP-oriented computerized neuropsychological assessment system is shown in Fig. 25.

6.2.1. Subjects

As one of the most well-known AD patients, RR was diagnosed with AD in 1994, six years after he left office; however, speculation of his cognitive decline while in office has been the subject of both academic scholarship and popular debate. Since RR's presidency was 6 years before his diagnosis of AD, he may have been at the MCI stage during his term(s) of office. The availability of official archives of presidential transcripts offers an opportunity to address questions regarding linguistic changes over time for MCI [90]. And for the control group, this study chooses President Dwight D. Eisenhower (DE),

and President George H. W. Bush (GB), who have no known diagnosis of AD, to avoid the influence for the linguistic changes caused by the age, where all three presidents were over 60 years old (RR was 69, DE was 62, and GB was 64 at the beginning of their presidency, respectively), that onset of AD first affects people in their mid-60s in most cases.

Using 46 transcripts of RR's press conferences (1981–1988), the extracted linguistic features would show change longitudinally with MCI. For maintaining equivalence of the data size, 46 transcripts of the press conferences of DE (1953–1961), and GB (1989–1993) were randomly selected for comparing trends in linguistic changes against those of RR. The data of the news conference transcripts were obtained from the American Presidency Project archive as a data source for this project [90]. The project is a comprehensive and organized searchable database of presidential documents, including transcripts of speeches, transcripts of news conferences, and other public documents. It ought to be noted that for the linguistic analysis in this study, the 46 transcripts of every president were grouped into 9 subsets in order of the date when they were made. In order to maintain the same number of transcripts and ensure enough document size in one subset, each subset included five consecutive transcripts, except the last one which contains the final six transcripts.

6.2.2. Pre-processing

There are typically two parts of Presidential press conferences: 1) a prepared statement or announcement read by the president and 2) a spontaneous question/answer session where the president can take questions from members of the media. It is often the case that the prepared statements are at least partially written by presidential speechwriters or staff members. As a result, for our analysis the focus was placed only on the latter since

the non-scripted nature of the discourse is cognitively more sensitive for our linguistic analysis.

6.2.2.1. Filtering

To generate the corpus for analysis, a transcript was downloaded, and then the prepared statement or announcement by the president and all discourses by other individuals (e.g., questions, statements and answers) were omitted. All annotations in brackets need to be filtered that refer to addendums to the transcripts by the editors to provide context (e.g., <Laughter>, Walt <Walter Rodgers, Associated Press Radio>, etc.).

6.2.2.2. Tokenization

Using the Natural Language Processing Tool Kit in Python (NLTK) [91], sentence tokenization and word tokenization were performed for each transcript. Tokenization is the process of splitting a stream of text into some smaller meaningful elements. After the sentence tokenization, every single sentence was derived from the original transcriptions. Then, implementing the word tokenization, each sentence would be broken up into words for further feature extraction processing.

6.2.3. Lexical Analysis

The first set of features extracted from the news conference transcripts was based on the lexical analysis, which generated 7 features corresponding to three types of words: 1) unique words, 2) non-specific words, and 3) specific words.

6.2.3.1. Unique Words

In order to obtain all the unique words in each transcript, each word was reduced to their stem, base, or root form. For example, the words item, items and itemization share a common root form, item, therefore, they were treated as the same unique word. The

stemming process was automatically implemented by using the Snowball Stemmer in NLTK [91], then the number of unique word was the first feature extracted from the transcripts. The number of unique words should remain the same for CN people, but for MCI patients, who have started forgetting some frequently used words, as a result, that number should decrease over the progression of MCI.

6.2.3.2. Non-specific Words

According to the aforementioned unique words, for each subset of transcripts, those words appearing in every transcript were considered as non-specific (NS) words, which means they are not that specific for each transcript but just depend on the word-using habit of a person. NS words should remain the same amount for CN people, but when some memory decline occurs due to MCI, the patients may have some difficulties to find different words to describe the similar thing or situation, i.e., the diversity of their discourse would also decline in the change of the use of NS terms. Therefore, the number of NS words, the frequency of NS words (i.e., the number of NS words/the number of unique words), and the rate of NS words (i.e., the total count of NS words/the number of NS words) are the three features related to NS words.

6.2.3.3. Specific Words

Unlike the NS words, specific (SPE) words are those words which are more meaningful and interesting for a given conversation. In this study, for a subset of transcripts, those words exactly appeared in only one transcript were regarded as being specific to that transcript. For CN people, SPE words should remain the same amount, but for MCI, the memory problem should cause a decrease of the usage of SPE terms. Then, there were three features related to SPE words, the number of SPE words, the frequency of

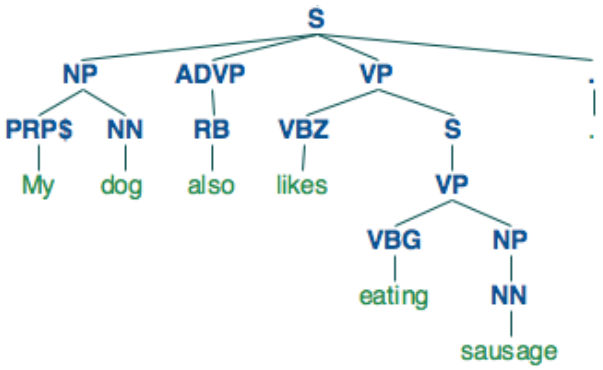


Fig. 26. The Parsing Tree of the Sentence 'My dog also likes eating sausage.'

SPE words (i.e., the number of SPE words/the number of unique words), and the rate of SPE words (i.e., the total count of SPE words/the number of SPE words).

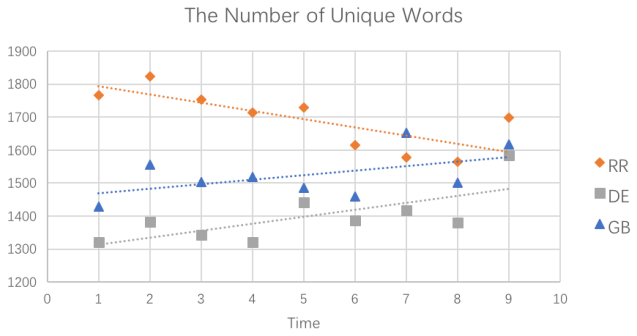
6.2.4. Grammatical Analysis

The second set of features extracted from the news conference transcripts took advantage of the grammatical analysis. The Stanford English PCFG parser was used to calculate the grammatical complexity of the transcripts [92]. For each sentence obtained from the sentence tokenization, the parser can produce a hierarchical tree representation of a given sentence as shown in Fig. 26. For the given sentence in Fig. 26, 'My dog also likes eating sausage.', three measures were introduced, the depth of the parsing tree, the width of the parsing tree, and the ratio of the depth to the width, in this example, which are 6, 7 and 0.86 (i.e., 6/7), respectively. Then, for each subset of the transcripts, the average of the three measures of all the sentence were used as the features of the grammatical complexity, which should be expected to indicate different patterns between CN and MCI.

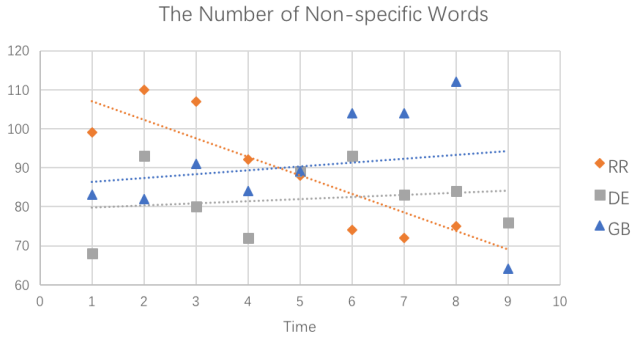
6.2.5. Linear Regression Analysis

For each of the 10 linguistic features, a linear regression analysis based on the Pearson correlation coefficient (PCC) was carried out to evaluate statistically significant changes in these variables over time for RR, DE, and GB. PCC is a measure of the linear

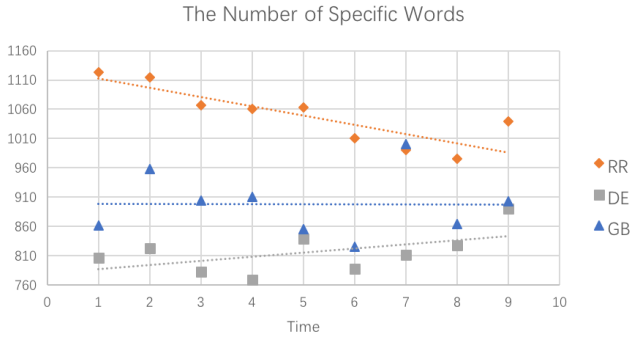
correlation between two variables. It has a value between +1 and −1, where 1 is total positive linear correlation, 0 is no linear correlation, and −1 is total negative linear correlation. For each obtained PPC, the linear regression analysis also computes the corresponding *p-value*, so that the statistical significance could be evaluated as well.



(a)



(b)



(c)

Fig. 27. The Trend of 3 Lexical Analysis Counting Features.

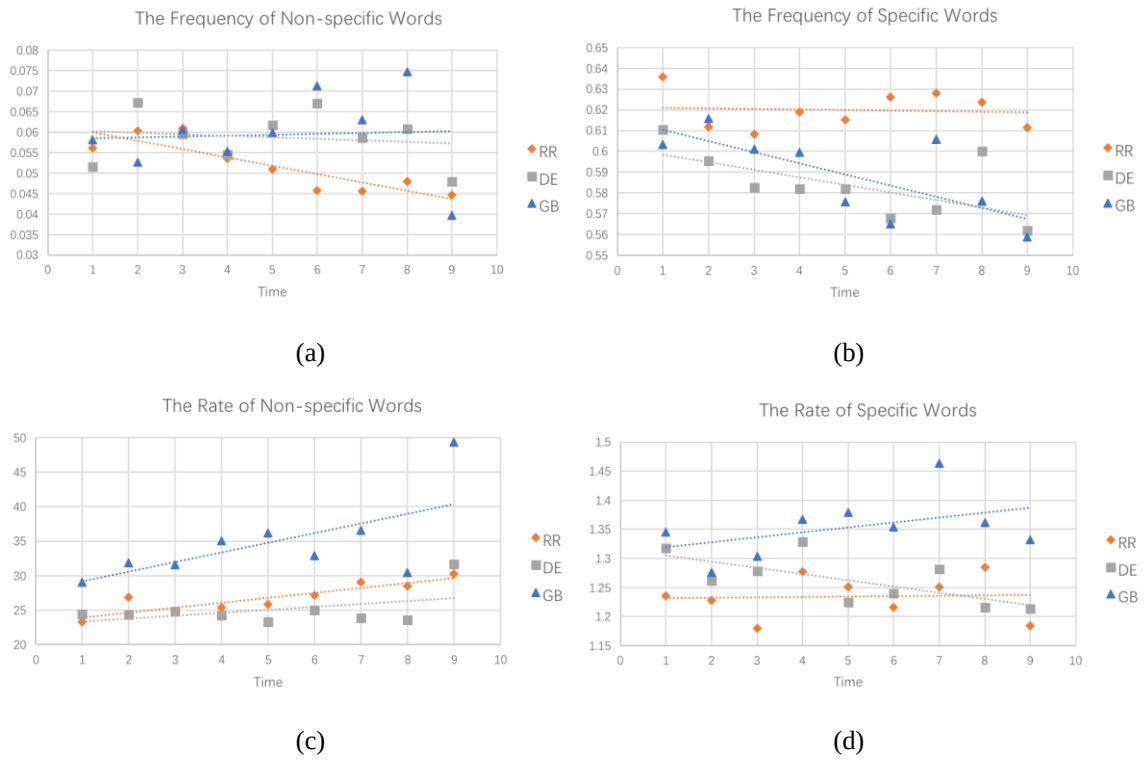


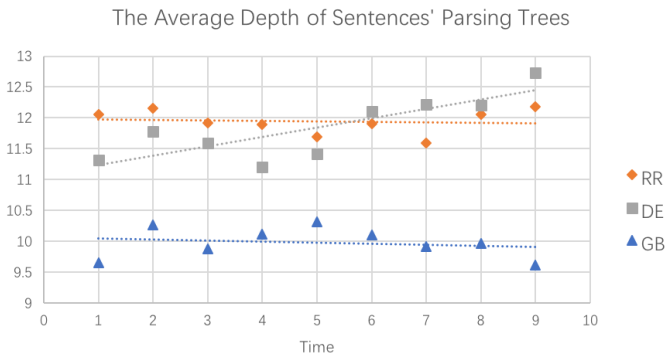
Fig. 28. The trend of 4 Lexical Analysis Ratio Features.

6.3. Results

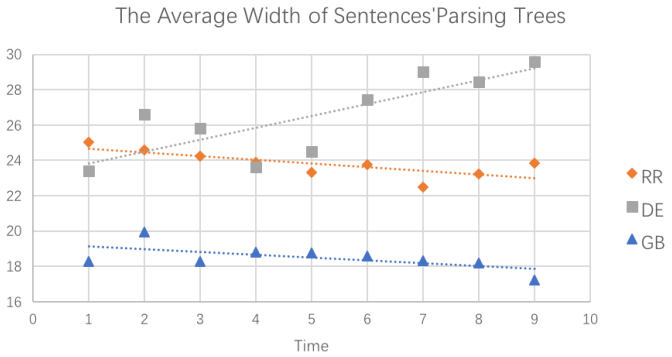
In this section, results of the proposed computerized assessment reveal significant different trend of linguistic complexity changes between MCI subject (i.e., RR) and CN subjects (i.e., DE and GB). The PCCs and statistical significances of all 10 features of each president are presented.

6.3.1. Linguistic Complexity Changes

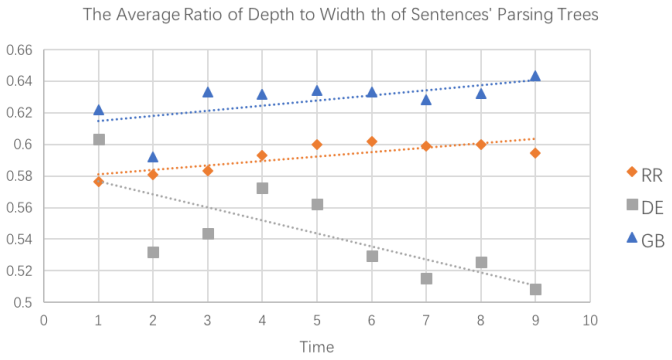
Fig. 27, Fig. 28 and Fig. 29 demonstrate the expected patterns of linguistic complexity changes over time. As shown in Fig. 27(a), (b), (c) for RR, the number of unique words, NS words, and SPE words, indicate a significant decreasing trend during his presidency, but the values of the same features for DE and GB slightly increased or remained at the same level over time. The same pattern can be observed in Fig. 28(a) for the frequency of NS words. From Fig. 28(b), the decline of the frequency of SPE words is



(a)



(b)



(c)

Fig. 29. The Trend of 3 Grammatical Analysis Features.

apparent for GB and DE, whereas the frequency of SPE words of RR indicates no change trend over time. For the grammatical complexity changes, there are no pattern found to be significantly distinguishable between RR and others as shown in Fig. 29.

6.3.2. Pearson Correlation Coefficient

The linear regression analysis supports the observations made in Fig. 27, Fig. 28, Fig. 29, and in Table 23. For RR, the number of unique words, NS words, and SPE words, and the frequency of NS words all decreased at 0.95 level of significance (LOS), but for DE and GB, these features all fairly increased or remained the same. Moreover, the frequency of SPE words for RR almost stayed at the same level with the PCC of -0.075, but for DE and GB, this feature indicated a significant reduction at 0.90 LOS with the PCC of -0.626 and -0.724, respectively. It ought to be noted that since GB held one term of office while both RR and DE completed two terms, some statistical values of GB may not reflect the real pattern associated with CN as compared to those obtained from DE. From Fig. 27, Fig. 28, Fig. 29, and Table 23, it can be observed that except for the rate of NS words, all other features delineate the opposite trend of linguistic complexity changes between RR and DE.

Table 23. Linear Regression Analysis of 10 Features

President	RR		DE		GB	
	Coeff.	*p-value*	*Coeff.*	*p-value*	*Coeff.*	*p-value*
Number of Unique words	-0.766	0.016	0.711	0.032	0.518	0.153
Number of NS words	-0.880	0.002	0.164	0.673	0.188	0.629
Number of SPE words	-0.856	0.003	0.537	0.136	-0.007	0.985
Frequency of NS words	-0.876	0.002	-0.159	0.683	0.057	0.884
Frequency of SPE words	-0.075	0.848	-0.626	0.071	-0.724	0.027
Rate of NS words	0.886	0.001	0.452	0.222	0.640	0.063
Rate of SPE words	0.054	0.891	-0.679	0.044	0.447	0.228
Depth of parsing trees	-0.106	0.787	0.818	0.007	-0.184	0.635
Width of parsing trees	-0.752	0.019	0.797	0.010	-0.606	0.083
Ratio of depth to width	0.819	0.007	-0.746	0.021	0.622	0.074

RR: Ronald Reagan, DE: Dwight D. Eisenhower, GB: George H. W. Bush, NS: non-specific, SPE: specific, Coeff.: coefficient

CHAPTER VII. COCLUSION

In Chapter II, the dissertation showed that, by using only structural MRI data, the proposed GDA-based generative learning algorithm achieved an accuracy of 85.90%, sensitivity of 83.87%, specificity of 87.23%, positive predictive value of 81.25%, and negative predictive value of 89.13% for discriminating AD from CN; these results for these two groups were expected. But more importantly, an accuracy of 82.20%, sensitivity of 83.10%, specificity of 80.85%, positive predictive value of 86.76%, and negative predictive value of 76.00% were obtained for discriminating MCI from CN that are very competitive. The proposed GDA-based classifier is capable of solving nonlinear boundaries for discriminating AD and MCI from CN in consideration of correlations among all variables. And most importantly, by separating left and right hemispheres of the brain into two decision spaces, and then combining results of these two spaces to determine the final outcomes, the classification performance is improved significantly.

Chapter III 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 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.

In Chapter IV, this dissertation presented a novel classification algorithm with global feature selection using structural MRI data as the unique input for the CAD in AD classification. The noise detection process was able to remove all abnormal subject datasets. Then based on ANOVA, the global feature ranking procedure sorted all features depending on their statistical significance associated with the progression of AD, then the IEA obtained the optimal set of global features taking advantage of the F1 score instead of the accuracy in consideration of the imbalanced data. The GDA-based classifier was effectively and accurately trained to discriminate different stages of AD using the innovative two high-dimensional decision spaces reflecting the left and right hemispheres. Furthermore, with the strategic two-phase classification process, the presented method achieved perhaps one of the best classification performance in delineating the most challenging groups of MCI and CN with an unprecedented accuracy of 88.78%, a sensitivity of 91.08%, and a specificity of 83.78%.

In Chapter V, a novel GDCA dimensionality reduction algorithm was proposed to characterize the optimal Gaussian discriminative components of the original high dimensional feature space, maximizing as a consequence the discriminability of selected eigenvectors. The CN vs. EMCI classification results indicated that the proposed supervised method was able to delineate the subtlest changes associated with the EMCI group. After transforming the original features to the optimal Gaussian discriminative components, a high accuracy of 79.25%, an F1 score of 80.70% and an AUC score of 0.7960 were obtained, which showed high potential of the proposed method for clinical

diagnosis of the early stage of AD. For EMCI vs. LMCI classification, the proposed model achieved a high accuracy of 83.33%, an F1 score of 77.78%, and an AUC score of 0.8947. These results of CN vs. EMCI classification and EMCI vs. LMCI classification are considered as the best classification performance obtained so far.

Chapter V also carried out a multiclass classification was also for the detection of the different stages in AD (i.e., EMCI, LMCI, and AD). An overall accuracy of 67.69% was achieved, and moreover, the proposed method was able to distinguish AD from MCI with an accuracy of 87.69% and a recall of 93.75%, respectively. The comparison with other widely used dimensionality reduction methods indicated that the proposed method could significantly reduce the dimensionality of the data and still accomplish an effective classification performance. A CAD application based on the proposed GDCA model was also presented, which attained an overall accuracy of 66.29% for CN vs. MCI vs. AD classification, and more notably, for distinguishing diseased subjects (i.e., MCI and AD) from CN group, with an accuracy of 75.28%. The future work will ultimately focus on taking advantage of the proposed GDCA algorithm to build a CAD system that could help in delineating the EMCI group in a multiclass classification process that could be helpful in the planning of early treatment and therapeutic interventions.

In Chapter VI, the dissertation indicated that, even though President Reagan was not diagnosed with AD until August of 1994, the results of the proposed computerized assessment suggest that changes in linguistic complexity were becoming detectable years prior to clinical diagnosis. In the news conferences of RR, the significant decline was found for the average length of sentences and the use of unique words, NS words and SPE words. An increasing trend of the ratio of depth to width of the sentences' parsing trees over time

was significant during President Reagan's presidency. Analysis of his transcripts revealed significant differences in variables known to be associated with the onset of MCI. To address the potential confound associated with changes resulting from healthy aging, by comparing RR's transcripts to those of DE and GB. At the start of their presidencies, DE was 62, RR was 69 years old, and GB was 64 (the year 1953, 1981 and 1989 respectively). Although the three age spans differ slightly, DE and GB provide the most comparable cases among the modern American presidents. Furthermore, the results are consistent with an early study indicating that RR's presidential debates contained the detectable patterns of AD as early as 1980 [93]. Instead of typically using simple statistic and fixed feature sets, by applying the state-of-the-art NLP technologies, this study indicated the significantly different patterns of the linguistic complexity changes between CN and MCI subjects, which can be applied to the diagnosis of MCI.

LIST OF REFERENCES

[1] NIA. "Basics of Alzheimer's Disease and Dementia." https://www.nia.nih.gov/alzheimers/topics/alzheimers-basics (accessed August 2, 2017).

[2] Q. Zhou *et al.*, "An optimal decisional space for the classification of Alzheimer's disease and mild cognitive impairment," (in eng), *IEEE Trans Biomed Eng*, vol. 61, no. 8, pp. 2245-53, Aug 2014, doi: 10.1109/TBME.2014.2310709.

[3] R. Cuingnet *et al.*, "Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database," (in eng), *Neuroimage*, vol. 56, no. 2, pp. 766-81, May 2011, doi: 10.1016/j.neuroimage.2010.06.013.

[4] Z. Long *et al.*, "A support vector machine-based method to identify mild cognitive impairment with multi-level characteristics of magnetic resonance imaging," (in eng), *Neuroscience*, vol. 331, pp. 169-76, 09 2016, doi: 10.1016/j.neuroscience.2016.06.025.

[5] L. Xu, X. Wu, K. Chen, and L. Yao, "Multi-modality sparse representation-based classification for Alzheimer's disease and mild cognitive impairment," (in eng), *Comput Methods Programs Biomed*, vol. 122, no. 2, pp. 182-90, Nov 2015, doi: 10.1016/j.cmpb.2015.08.004.

[6] D. Zhang, Y. Wang, L. Zhou, H. Yuan, D. Shen, and A. s. D. N. Initiative, "Multimodal classification of Alzheimer's disease and mild cognitive impairment," (in eng), *Neuroimage*, vol. 55, no. 3, pp. 856-67, Apr 2011, doi: 10.1016/j.neuroimage.2011.01.008.

[7] F. Martinez-Murcia, J. Gorriz, J. Ramirez, and A. Ortiz, "A Structural Parametrization of the Brain Using Hidden Markov Models-Based Paths in Alzheimer's Disease," (in English), *International Journal of Neural Systems*, Article vol. 26, no. 7, NOV 2016 2016, Art no. ARTN 1650024, doi: 10.1142/S0129065716500246.

[8] M. Liu, D. Zhang, D. Shen, and A. s. D. N. Initiative, "Ensemble sparse classification of Alzheimer's disease," (in eng), *Neuroimage*, vol. 60, no. 2, pp. 1106-16, Apr 2012, doi: 10.1016/j.neuroimage.2012.01.055.

[9] L. Khedher, J. Ramirez, J. Gorriz, A. Brahim, F. Segovia, and A. D. N. Initia, "Early diagnosis of Alzheimer's disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images," (in English), *Neurocomputing*, Article vol. 151, pp. 139-150, MAR 3 2015 2015, doi: 10.1016/j.neucom.2014.09.072.

[10] T. Ye, C. Zu, B. Jie, D. Shen, D. Zhang, and A. s. D. Neuroimaging, "Discriminative multi-task feature selection for multi-modality classification of Alzheimer's disease," (in English), *Brain Imaging and Behavior,* Article vol. 10, no. 3, pp. 739-749, SEP 2016 2016, doi: 10.1007/s11682-015-9437-x.

[11] R. Romero-Garcia, M. Atienza, and J. Cantero, "Different Scales of Cortical Organization are Selectively Targeted in the Progression to Alzheimer's Disease," (in English), *International Journal of Neural Systems,* Article vol. 26, no. 2, MAR 2016 2016, Art no. ARTN 1650003, doi: 10.1142/S0129065716500039.

[12] H. Aidos, A. Fred, and A. s. D. N. Initi, "Discrimination of Alzheimer's Disease using longitudinal information," (in English), *Data Mining and Knowledge Discovery,* Article vol. 31, no. 4, pp. 1006-1030, JUL 2017 2017, doi: 10.1007/s10618-017-0502-5.

[13] I. Illan *et al.*, "F-18-FDG PET imaging analysis for computer aided Alzheimer's diagnosis," (in English), *Information Sciences,* Article vol. 181, no. 4, pp. 903-916, FEB 15 2011 2011, doi: 10.1016/j.ins.2010.10.027.

[14] T. Tong, K. Gray, Q. Gao, L. Chen, D. Rueckert, and A. D. N. Initia, "Multi-modal classification of Alzheimer's disease using nonlinear graph fusion," (in English), *Pattern Recognition,* Article vol. 63, pp. 171-181, MAR 2017 2017, doi: 10.1016/j.patcog.2016.10.009.

[15] M. Ewers *et al.*, "Prediction of conversion from mild cognitive impairment to Alzheimer's disease dementia based upon biomarkers and neuropsychological test performance," (in English), *Neurobiology of Aging,* Article vol. 33, no. 7, pp. 1203-+, JUL 2012 2012, doi: 10.1016/j.neurobiolaging.2010.10.019.

[16] C. Huang, L. Wahlund, T. Dierks, P. Julin, B. Winblad, and V. Jelic, "Discrimination of Alzheimer's disease and mild cognitive impairment by equivalent EEG sources: a cross-sectional and longitudinal study," (in English), *Clinical Neurophysiology,* Article vol. 111, no. 11, pp. 1961-1967, NOV 2000 2000, doi: 10.1016/S1388-2457(00)00454-5.

[17] E. Gallego-Jutgla, J. Sole-Casals, F. Vialatte, M. Elgendi, A. Cichocki, and J. Dauwels, "A hybrid feature selection approach for the early diagnosis of Alzheimer's disease," (in English), *Journal of Neural Engineering,* Article vol. 12, no. 1, FEB 2015 2015, Art no. ARTN 016018, doi: 10.1088/1741-2560/12/1/016018.

[18] F. C. Morabito *et al.*, "A longitudinal EEG study of Alzheimer's disease progression based on a complex network approach," (in eng), *Int J Neural Syst,* vol. 25, no. 2, p. 1550005, Mar 2015, doi: 10.1142/S0129065715500057.

[19] N. Mammone *et al.*, "Permutation Disalignment Index as an Indirect, EEG-Based, Measure of Brain Connectivity in MCI and AD Patients," (in eng), *Int J Neural Syst,* vol. 27, no. 5, p. 1750020, Aug 2017, doi: 10.1142/S0129065717500204.

[20] N. Mammone *et al.*, "A Permutation Disalignment Index-Based Complex Network Approach to Evaluate Longitudinal Changes in Brain-Electrical Connectivity," (in English), *Entropy,* Article vol. 19, no. 10, OCT 2017 2017, Art no. ARTN 548, doi: 10.3390/e19100548.

[21] N. Houmani, G. Dreyfus, and F. Vialatte, "Epoch-based Entropy for Early Screening of Alzheimer's Disease," (in English), *International Journal of Neural Systems,* Article vol. 25, no. 8, DEC 2015 2015, Art no. ARTN 1550032, doi: 10.1142/S012906571550032X.

[22] J. Dauwels, F. Vialatte, T. Musha, and A. Cichocki, "A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG," (in English), *Neuroimage,* Article vol. 49, no. 1, pp. 668-693, JAN 1 2010 2010, doi: 10.1016/j.neuroimage.2009.06.056.

[23] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, "Alzheimer's disease: Models of computation and analysis of EEGs," (in English), *Clinical Eeg and Neuroscience,* Review vol. 36, no. 3, pp. 131-140, JUL 2005 2005, doi: 10.1177/155005940503600303.

[24] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, "A spatio-temporal wavelet-chaos methodology for EEG-based diagnosis of Alzheimer's disease," (in English), *Neuroscience Letters,* Article vol. 444, no. 2, pp. 190-194, OCT 24 2008 2008, doi: 10.1016/j.neulet.2008.08.008.

[25] M. Ahmadlou, H. Adeli, and A. Adeli, "New diagnostic EEG markers of the Alzheimer's disease using visibility graph," (in English), *Journal of Neural Transmission,* Article vol. 117, no. 9, pp. 1099-1109, SEP 2010 2010, doi: 10.1007/s00702-010-0450-3.

[26] M. Ahmadlou, H. Adeli, and A. Adeli, "Fractality and a Wavelet-chaos-Methodology for EEG-based Diagnosis of Alzheimer Disease," (in English), *Alzheimer Disease & Associated Disorders,* Article vol. 25, no. 1, pp. 85-92, JAN-MAR 2011 2011, doi: 10.1097/WAD.0b013e3181ed1160.

[27] Z. Sankari and H. Adeli, "Probabilistic neural networks for diagnosis of Alzheimer's disease using conventional and wavelet coherence," (in English), *Journal of Neuroscience Methods,* Article vol. 197, no. 1, pp. 165-170, APR 15 2011 2011, doi: 10.1016/j.jneumeth.2011.01.027.

[28] Z. Sankari, H. Adeli, and A. Adeli, "Intrahemispheric, interhemispheric, and distal EEG coherence in Alzheimer's disease," (in English), *Clinical Neurophysiology,* Article vol. 122, no. 5, pp. 897-906, MAY 2011 2011, doi: 10.1016/j.clinph.2010.09.008.

[29] M. Ahmadlou, A. Adeli, R. Bajo, and H. Adeli, "Complexity of functional connectivity networks in mild cognitive impairment subjects during a working memory task," (in English), *Clinical Neurophysiology,* Article vol. 125, no. 4, pp. 694-702, APR 2014 2014, doi: 10.1016/j.clinph.2013.08.033.

[30] J. Amezquita-Sanchez, A. Adeli, and H. Adeli, "A new methodology for automated diagnosis of mild cognitive impairment (MCI) using magnetoencephalography (MEG)," (in English), *Behavioural Brain Research,* Article vol. 305, pp. 174-180, MAY 15 2016 2016, doi: 10.1016/j.bbr.2016.02.035.

[31] D. Lopez-Sanz, P. Garces, B. Alvarez, M. Delgado-Losada, R. Lopez-Higes, and F. Maestu, "Network Disruption in the Preclinical Stages of Alzheimer's Disease: From Subjective Cognitive Decline to Mild Cognitive Impairment," (in English), *International Journal of Neural Systems,* Article vol. 27, no. 8, DEC 2017 2017, Art no. ARTN 1750041, doi: 10.1142/S0129065717500411.

[32] L. Khedher, I. Illan, J. Gorriz, J. Ramirez, A. Brahim, and A. Meyer-Baese, "Independent Component Analysis-Support Vector Machine-Based Computer-Aided Diagnosis System for Alzheimer's with Visual Support," (in English), *International Journal of Neural Systems,* Article vol. 27, no. 3, MAY 2017 2017, Art no. ARTN 1650050, doi: 10.1142/S0129065716500507.

[33] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, "Alzheimer's disease and models of computation: Imaging, classification, and neural models," (in English), *Journal of Alzheimers Disease,* Review vol. 7, no. 3, pp. 187-199, JUN 2005 2005.

[34] G. Mirzaei, A. Adeli, and H. Adeli, "Imaging and machine learning techniques for diagnosis of Alzheimer's disease," (in English), *Reviews in the Neurosciences,* Article vol. 27, no. 8, pp. 857-870, DEC 2016 2016, doi: 10.1515/revneuro-2016-0029.

[35] C. Li *et al.*, "A Neuroimaging Feature Extraction Model for Imaging Genetics with Application to Alzheimer's Disease," (in English), *2017 Ieee 17th International Conference on Bioinformatics and Bioengineering (Bibe),* Proceedings Paper pp. 15-20, 2017 2017, doi: 10.1109/BIBE.2017.00010.

[36] H. Braak and K. Del Tredici, "Alzheimer's disease: Pathogenesis and prevention," (in English), *Alzheimers & Dementia,* Article vol. 8, no. 3, pp. 227-233, MAY 2012 2012, doi: 10.1016/j.jalz.2012.01.011.

[37] T. GomezIsla, J. Price, D. McKeel, J. Morris, J. Growdon, and B. Hyman, "Profound loss of layer II entorhinal cortex neurons occurs in very mild Alzheimer's disease," (in English), *Journal of Neuroscience,* Article vol. 16, no. 14, pp. 4491-4500, JUL 15 1996 1996.

[38] U. Khan *et al.*, "Molecular drivers and cortical spread of lateral entorhinal cortex dysfunction in preclinical Alzheimer's disease," (in English), *Nature Neuroscience,* Article vol. 17, no. 2, pp. 304-311, FEB 2014 2014, doi: 10.1038/nn.3606.

[39] C. Chu, A. L. Hsu, K. H. Chou, P. Bandettini, C. Lin, and A. s. D. N. Initiative, "Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images," (in eng), *Neuroimage,* vol. 60, no. 1, pp. 59-70, Mar 2012, doi: 10.1016/j.neuroimage.2011.11.066.

[40] F. Liu, C. Y. Wee, H. Chen, and D. Shen, "Inter-modality relationship constrained multi-modality multi-task feature selection for Alzheimer's Disease and mild cognitive impairment identification," (in eng), *Neuroimage,* vol. 84, pp. 466-75, Jan 2014, doi: 10.1016/j.neuroimage.2013.09.015.

[41] X. Zhu, H. I. Suk, L. Wang, S. W. Lee, D. Shen, and A. s. D. N. Initiative, "A novel relational regularization feature selection method for joint regression and classification in AD diagnosis," (in eng), *Med Image Anal,* vol. 38, pp. 205-214, 05 2017, doi: 10.1016/j.media.2015.10.008.

[42] K. Ota, N. Oishi, K. Ito, H. Fukuyama, S.-J. S. Group, and A. s. D. N. Initiative, "Effects of imaging modalities, brain atlases and feature selection on prediction of Alzheimer's disease," (in eng), *J Neurosci Methods,* vol. 256, pp. 168-83, Dec 2015, doi: 10.1016/j.jneumeth.2015.08.020.

[43] C. Li *et al.*, "Greater Regional Cortical Thickness is Associated with Selective Vulnerability to Atrophy in Alzheimer's Disease, Independent of Amyloid Load and APOE Genotype," (in eng), *J Alzheimers Dis,* vol. 69, no. 1, pp. 145-156, 2019, doi: 10.3233/JAD-180231.

[44] D. Loewenstein *et al.*, "Recovery from Proactive Semantic Interference and MRI Volume: A Replication and Extension Study," (in English), *Journal of Alzheimers Disease,* Article vol. 59, no. 1, pp. 131-139, 2017 2017, doi: 10.3233/JAD-170276.

[45] C. Plant *et al.*, "Automated detection of brain atrophy patterns based on MRI for the prediction of Alzheimer's disease," (in eng), *Neuroimage,* vol. 50, no. 1, pp. 162-74, Mar 2010, doi: 10.1016/j.neuroimage.2009.11.046.

[46] F. Zhang, S. Tian, S. Chen, Y. Ma, X. Li, and X. Guo, "Voxel-Based Morphometry: Improving the Diagnosis of Alzheimer's Disease Based on an Extreme Learning Machine Method from the ADNI cohort," (in English), *Neuroscience,* Article vol. 414, pp. 273-279, AUG 21 2019 2019, doi: 10.1016/j.neuroscience.2019.05.014.

[47] X. Sun *et al.*, "A human brain tau PET template in MNI space for the voxel-wise analysis of Alzheimer's disease," (in eng), *J Neurosci Methods,* vol. 328, p. 108438, Dec 2019, doi: 10.1016/j.jneumeth.2019.108438.

[48] R. Ossenkoppele *et al.*, "Prevalence of amyloid PET positivity in dementia syndromes: a meta-analysis," (in eng), *JAMA,* vol. 313, no. 19, pp. 1939-49, May 2015, doi: 10.1001/jama.2015.4669.

[49] A. Farzan, S. Mashohor, A. Ramli, and R. Mahmud, "Boosting diagnosis accuracy of Alzheimer's disease using high dimensional recognition of longitudinal brain atrophy patterns," (in English), *Behavioural Brain Research,* Article vol. 290, pp. 124-130, SEP 1 2015 2015, doi: 10.1016/j.bbr.2015.04.010.

[50] M. Lopez *et al.*, "Principal component analysis-based techniques and supervised classification schemes for the early detection of Alzheimer's disease," (in English), *Neurocomputing,* Article|Proceedings Paper vol. 74, no. 8, pp. 1260-1271, MAR 15 2011 2011, doi: 10.1016/j.neucom.2010.06.025.

[51] R. E. Curiel Cid *et al.*, "A cognitive stress test for prodromal Alzheimer's disease: Multiethnic generalizability," (in eng), *Alzheimers Dement (Amst),* vol. 11, pp. 550-559, Dec 2019, doi: 10.1016/j.dadm.2019.05.003.

[52] D. A. Loewenstein *et al.*, "Utilizing semantic intrusions to identify amyloid positivity in mild cognitive impairment," (in eng), *Neurology,* vol. 91, no. 10, pp. e976-e984, 09 2018, doi: 10.1212/WNL.0000000000006128.

[53] M. Boye, T. Tran, N. Grabar, A. Przepiorkowski, and M. Ogrodniczuk, "NLP-Oriented Contrastive Study of Linguistic Productions of Alzheimer's and Control People," (in English), *Advances in Natural Language Processing,* Proceedings Paper vol. 8686, pp. 412-+, 2014 2014.

[54] S. Zygouris and M. Tsolaki, "Computerized Cognitive Testing for Older Adults: A Review," (in English), *American Journal of Alzheimers Disease and Other Dementias,* Review vol. 30, no. 1, pp. 13-28, FEB 2015 2015, doi: 10.1177/1533317514522852.

[55] M. Canini *et al.*, "Computerized Neuropsychological Assessment in Aging: Testing Efficacy and Clinical Ecology of Different Interfaces," (in English), *Computational and Mathematical Methods in Medicine,* Article 2014 2014, Art no. ARTN 804723, doi: 10.1155/2014/804723.

[56] I. Tarnanas, M. Tsolaki, T. Nef, R. Muri, and U. Mosimann, "Can a novel computerized cognitive screening test provide additional information for early detection of Alzheimer's disease?," (in English), *Alzheimers & Dementia,* Article vol. 10, no. 6, pp. 790-798, NOV 2014 2014, doi: 10.1016/j.jalz.2014.01.002.

[57] S. Weintraub *et al.*, "Cognition assessment using the NIH Toolbox," (in English), *Neurology,* Article vol. 80, pp. S54-S64, MAR 2013 2013, doi: 10.1212/WNL.0b013e3182872ded.

[58] M. Inoue, D. Jimbo, M. Taniguchi, and K. Urakami, "Touch Panel-type Dementia Assessment Scale: a new computer-based rating scale for Alzheimer's disease," (in English), *Psychogeriatrics,* Article vol. 11, no. 1, pp. 28-33, MAR 2011 2011, doi: 10.1111/j.1479-8301.2010.00345.x.

[59] C. Jacova *et al.*, "C-TOC (Cognitive Testing on Computer) Investigating the Usability and Validity of a Novel Self-administered Cognitive Assessment Tool in Aging and Early Dementia," (in English), *Alzheimer Disease & Associated Disorders,* Article vol. 29, no. 3, pp. 213-221, JUL-SEP 2015 2015, doi: 10.1097/WAD.0000000000000055.

[60] V. Berisha, S. Wang, A. LaCross, and J. Liss, "Tracking Discourse Complexity Preceding Alzheimer's Disease Diagnosis: A Case Study Comparing the Press Conferences of Presidents Ronald Reagan and George Herbert Walker Bush," (in English), *Journal of Alzheimers Disease,* Article vol. 45, no. 3, pp. 959-963, 2015 2015, doi: 10.3233/JAD-142763.

[61] S. Pakhomov, D. Chacon, M. Wicklund, and J. Gundel, "Computerized assessment of syntactic complexity in Alzheimer's disease: a case study of Iris Murdoch's writing," (in English), *Behavior Research Methods,* Article vol. 43, no. 1, pp. 136-144, MAR 2011 2011, doi: 10.3758/s13428-010-0037-9.

[62] P. Garrard, K. Patterson, P. Watson, and J. Hodges, "Category specific semantic loss in dementia of Alzheimer's type - Functional-anatomical correlations from cross-sectional analyses," (in English), *Brain,* Article vol. 121, pp. 633-646, APR 1998 1998, doi: 10.1093/brain/121.4.633.

[63] M. Reuter, N. Schmansky, H. Rosas, and B. Fischl, "Within-subject template estimation for unbiased longitudinal image analysis," (in English), *Neuroimage,* Article vol. 61, no. 4, pp. 1402-1418, JUL 16 2012 2012, doi: 10.1016/j.neuroimage.2012.02.084.

[64] N. Tustison *et al.*, "Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements," (in English), *Neuroimage,* Article vol. 99, pp. 166-179, OCT 1 2014 2014, doi: 10.1016/j.neuroimage.2014.05.044.

[65] A. Klein *et al.*, "Mindboggling morphometry of human brains," (in English), *Plos Computational Biology,* Article vol. 13, no. 2, FEB 2017 2017, Art no. ARTN e1005350, doi: 10.1371/journal.pcbi.1005350.

[66] "Alzheimer's Disease Big Data DREAM Challenge 1." https://www.synapse.org/#!Synapse:syn2290704/wiki/60828 (accessed August 2, 2017).

[67] "“Mindboggle Data,” Mindboggle-101." https://www.synapse.org/#!Synapse:syn2290704/wiki/60828 (accessed August 2, 2017).

[68] L. SQUIRE and S. ZOLAMORGAN, "THE MEDIAL TEMPORAL-LOBE MEMORY SYSTEM," (in English), *Science*, Article vol. 253, no. 5026, pp. 1380-1386, SEP 20 1991 1991, doi: 10.1126/science.1896849.

[69] S. SHAPIRO and M. WILK, "AN ANALYSIS OF VARIANCE TEST FOR NORMALITY (COMPLETE SAMPLES)," (in English), *Biometrika*, Article vol. 52, pp. 591-&, 1965 1965, doi: 10.1093/biomet/52.3-4.591.

[70] L. Wang *et al.*, "Changes in hippocampal connectivity in the early stages of Alzheimer's disease: Evidence from resting state fMRI," (in English), *Neuroimage*, Article vol. 31, no. 2, pp. 496-504, JUN 2006 2006, doi: 10.1016/j.neuroimage.2005.12.033.

[71] M. Xia, J. Wang, and Y. He, "BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics," (in English), *Plos One*, Article vol. 8, no. 7, JUL 4 2013 2013, Art no. ARTN e68910, doi: 10.1371/journal.pone.0068910.

[72] A. Ortiz, J. Munilla, J. Gorriz, and J. Ramirez, "Ensembles of Deep Learning Architectures for the Early Diagnosis of the Alzheimer's Disease," (in English), *International Journal of Neural Systems*, Article vol. 26, no. 7, NOV 2016 2016, Art no. ARTN 1650025, doi: 10.1142/S0129065716500258.

[73] A. Bakkour, J. Morris, and B. Dickerson, "The cortical signature of prodromal AD Regional thinning predicts mild AD dementia," (in English), *Neurology*, Article vol. 72, no. 12, pp. 1048-1055, MAR 24 2009 2009, doi: 10.1212/01.wnl.0000340981.97664.2f.

[74] C. Fang *et al.*, "A Novel Gaussian Discriminant Analysis-based Computer Aided Diagnosis System for Screening Different Stages of Alzheimer's Disease," (in English), *2017 Ieee 17th International Conference on Bioinformatics and Bioengineering (Bibe)*, Proceedings Paper pp. 279-284, 2017 2017, doi: 10.1109/BIBE.2017.00054.

[75] J. Zhang, M. Liu, L. An, Y. Gao, and D. Shen, "Alzheimer's Disease Diagnosis Using Landmark-Based Features From Longitudinal Structural MR Images," (in English), *Ieee Journal of Biomedical and Health Informatics*, Article vol. 21, no. 6, pp. 1607-1616, NOV 2017 2017, doi: 10.1109/JBHI.2017.2704614.

[76] R. E. Curiel *et al.*, "Semantic Intrusions and Failure to Recover From Semantic Interference in Mild Cognitive Impairment: Relationship to Amyloid and Cortical Thickness," (in eng), *Curr Alzheimer Res*, vol. 15, no. 9, pp. 848-855, 2018, doi: 10.2174/1567205015666180427122746.

[77] R. Duara *et al.*, "Effect of age, ethnicity, sex, cognitive status and APOE genotype on amyloid load and the threshold for amyloid positivity," (in eng), *Neuroimage Clin,* vol. 22, p. 101800, 2019, doi: 10.1016/j.nicl.2019.101800.

[78] C. Li *et al.*, "The Relationship of Brain Amyloid Load and APOE Status to Regional Cortical Thinning and Cognition in the ADNI Cohort," (in eng), *J Alzheimers Dis,* vol. 59, no. 4, pp. 1269-1282, 2017, doi: 10.3233/JAD-170286.

[79] E. Westman, C. Aguilar, J. S. Muehlboeck, and A. Simmons, "Regional magnetic resonance imaging measures for multivariate analysis in Alzheimer's disease and mild cognitive impairment," (in eng), *Brain Topogr,* vol. 26, no. 1, pp. 9-23, Jan 2013, doi: 10.1007/s10548-012-0246-x.

[80] M. Jenkinson, C. Beckmann, T. Behrens, M. Woolrich, and S. Smith, "FSL," (in English), *Neuroimage,* Review vol. 62, no. 2, pp. 782-790, AUG 15 2012 2012, doi: 10.1016/j.neuroimage.2011.09.015.

[81] C. Li *et al.*, "Pattern analysis of the interaction of regional amyloid load, cortical thickness and APOE genotype in the progression of Alzheimer's disease," (in English), *2017 Ieee International Conference on Bioinformatics and Biomedicine (Bibm),* Proceedings Paper pp. 2171-2176, 2017 2017.

[82] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," (in English), *Journal of Machine Learning Research,* Article vol. 12, pp. 2825-2830, OCT 2011 2011.

[83] S. Pei, J. Guan, and S. Zhou, "Classifying early and late mild cognitive impairment stages of Alzheimer's disease by fusing default mode networks extracted with multiple seeds," (in eng), *BMC Bioinformatics,* vol. 19, no. Suppl 19, p. 523, Dec 2018, doi: 10.1186/s12859-018-2528-0.

[84] K. Hett *et al.*, "Multimodal Hippocampal Subfield Grading For Alzheimer's Disease Classification," (in eng), *Sci Rep,* vol. 9, no. 1, p. 13845, Sep 2019, doi: 10.1038/s41598-019-49970-9.

[85] B. Jie, M. Liu, D. Zhang, and D. Shen, "Sub-Network Kernels for Measuring Similarity of Brain Connectivity Networks in Disease Diagnosis," (in eng), *IEEE Trans Image Process,* vol. 27, no. 5, pp. 2340-2353, May 2018, doi: 10.1109/TIP.2018.2799706.

[86] B. Jie, M. Liu, and D. Shen, "Integration of temporal and spatial properties of dynamic connectivity networks for automatic diagnosis of brain disease," (in eng), *Med Image Anal,* vol. 47, pp. 81-94, 07 2018, doi: 10.1016/j.media.2018.03.013.

[87] C. Wee *et al.*, "Cortical graph neural network for AD and MCI diagnosis and transfer learning across populations," (in English), *Neuroimage-Clinical,* Article vol. 23, 2019 2019, Art no. UNSP 101929, doi: 10.1016/j.nicl.2019.101929.

[88] P. Yang *et al.*, "Fused Sparse Network Learning for Longitudinal Analysis of Mild Cognitive Impairment," (in eng), *IEEE Trans Cybern*, Sep 2019, doi: 10.1109/TCYB.2019.2940526.

[89] T. E. Kam, H. Zhang, Z. Jiao, and D. Shen, "Deep Learning of Static and Dynamic Brain Functional Networks for Early MCI Detection," (in eng), *IEEE Trans Med Imaging*, Jul 2019, doi: 10.1109/TMI.2019.2928790.

[90] G. Peters and J. T. Woolley. "“The American Presidency Project,” Presidential News Conferences." http://www.presidency.ucsb.edu/news_conferences.php (accessed June 5, 2017).

[91] S. Bird, E. Klein, and E. Loper, *Natural Language Processing with Python*. *O'Reilly*, 2009.

[92] D. Klein, C. Manning, and ACL, "Accurate unlexicalized parsing," (in English), *41st Annual Meeting of the Association For Computational Linguistics, Proceedings of the Conference*, Proceedings Paper pp. 423-430, 2003 2003.

[93] L. GOTTSCHALK, R. ULIANA, and R. GILBERT, "PRESIDENTIAL-CANDIDATES AND COGNITIVE IMPAIRMENT MEASURED FROM BEHAVIOR IN CAMPAIGN DEBATES," (in English), *Public Administration Review*, Article vol. 48, no. 2, pp. 613-619, MAR-APR 1988 1988, doi: 10.2307/975762.

VITA

CHEN FANG

2007-2011	B.S., Applied Mathematics Huangshan University Huangshan, Anhui, China
2014-2015	M.S., Computer Engineering Florida International University Miami, Florida
2015-2020	Research Assistant Florida International University Miami, Florida
2018	Data Science Intern AT&T Chief Data Office Palo Alto, California
2018-2020	Doctoral Candidate Florida International University Miami, Florida

PUBLICATIONS AND PRESENTATIONS

C. Fang, C. Li, M. Cabrerizo, A. Barreto, J. Andrian, N. Rishe, D. Loewenstein, R. Duara, and M. Adjouadi, "Gaussian discriminant analysis for optimal delineation of mild cognitive impairment in Alzheimer's disease," *Int J Neural Syst*, vol. 28, no. 8, 1850017, May 2018.

C. Fang, C. Li, P. Forouzannezhad, M. Cabrerizo, R. E. Curiel, D. Loewenstein, R. Duara, and M. Adjouadi, "Gaussian Discriminative Component Analysis for Early Detection of Alzheimer's Disease: A Supervised Dimensionality Reduction Algorithm," *J Neurosci Meth*, under second Review (Major Revision).

C. Wu, C. Fang, X. Wu, and G. Zhu, "Health-risk assessment of arsenic and groundwater quality classification using random forest in the Yanchi region of northwest China," *Exposure and Health*, in Press. Available: https://doi.org/10.1007/s12403-019-00335-7

P. Forouzannezhad, A. Abbaspour, C. Fang, M. Cabrerizo, D. Loewenstein, R. Duara, and M. Adjouadi, "A survey on applications and analysis methods of functional magnetic resonance imaging for Alzheimer's disease," *J Neurosci Meth*, vol. 317, pp. 121-140, April 2019.

P. Forouzannezhad, A. Abbaspour, C. Li, C. Fang, U. Williams, M. Cabrerizo, A. Barreto, J. Andrian, N. Rishe, R. E. Curiel, D. Loewenstein, R. Duara, and M. Adjouadi, "A Gaussian-based model for early detection of mild cognitive impairment using multimodal neuroimaging," *J Neurosci Meth*, vol. 333, pp. 108544, March 2020.

C. Fang, C. Li, M. Cabrerizo, A. Barreto, J. Andrian, D. Loewenstein, R. Duara, and M. Adjouadi, "A novel Gaussian discriminant analysis-based computer aided diagnosis system for screening different stages of Alzheimer's disease," *in Proceedings of IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE)*, Oct. 23-25, 2017, Washington DC, pp. 279-284.

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," *in Proceedings of IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE)*, Oct. 23-25, 2017, Washington DC, pp. 15-20.

P. Janwattanapong, M. Cabrerizo, C. Fang, H. Rajaei, A. Pinzon-Ardila, S. Gonzalez-Arias, and M. Adjouadi, "Classification of interictal epileptiform discharges using partial directed coherence," *in Proceedings of IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE)*, Oct. 23-25, 2017, Washington DC, pp. 473-478.

C. Fang, C. Li, M. Cabrerizo, A. Barreto, J. Andrian, D. Loewenstein, R. Duara, and M. Adjouadi, "A Gaussian discriminant analysis-based generative learning algorithm for the early diagnosis of mild cognitive impairment in Alzheimer's disease," *in Proceedings of IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Nov. 13-16, 2017, Kansas City, MO, pp. 538-542.

C. Fang, P. Janwattanapong, H. Martin, M. Cabrerizo, A. Barreto, D. Loewenstein, R. Duara, and M. Adjouadi, "Computerized neuropsychological assessment in mild cognitive impairment based on NLP-oriented feature extraction," *in Proceedings of IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Nov. 13-16, 2017, Kansas City, MO, pp. 543-546.

C. Li, C. Fang, M. Cabrerizo, A. Barreto, J. Andrian, R. Duara, D. Loewenstein, and M. Adjouadi, "Pattern analysis of the interaction of regional amyloid load, cortical thickness and APOE genotype in the progression of Alzheimer's disease," *in Proceedings of IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Nov. 13-16, 2017, Kansas City, MO, pp. 2171-2176.

C. Fang, P. Janwattanapong, C. Li and M. Adjouadi, "A global feature extraction model for the effective computer aided diagnosis of mild cognitive impairment using structural MRI images," *31st International Conference on Neural Information Processing Systems (NIPS)*, Machine Learning for Health (ML4H) Workshop, Long Beach, CA, 2017.